Title
High-Resolution Quantitative Analysis of Genetic Interactions

Author
Braberg, Hannes

Publication Date
2013

Peer reviewed|Thesis/dissertation

High-Resolution Quantitative Analysis of Genetic Interactions

by

Hannes Braberg

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

[redacted]

Foreword

by Gerard Cagney

Hannibal. Hector. Hercules…

Strong conquering names that evoke images of pitiless battle, blood draining into the sand, ancient loyalties betrayed.

It was therefore in a state of keen anticipation that the Krogan Lab awaited the arrival of a rumoured new force from the North - a computational biologist of such immense mathematical power that it was said to be of supernatural origin, who would put to the sword all but the most stringent of data. From the land of the Viking would come a man capable of forging pure gold from a crucible of sketchy EMAP experiments and dirty TAP tags – elegant genetic and physical interactions inconceivable to the human mind, with p-values finely calibrated to delight the most exacting of scholars. Into the mythology of the Krogan Lab strode….

Hannes.

Well I wasn't present that day, but most likely he emerged from a 3rd floor service corridor sipping a Mocha Frappe Crappe Latte and wearing a pair of Merril Chameleon Arc 2 Rival Waterproof hiking boots, sat down at his desk, and worked for two or three weeks before Ellie noticed he had arrived.

For Hannes is a quiet, unassuming, utterly reliable researcher. When you see him he tends to be working, or at least looking like he is working. The most outrageous, ribald conversation may be swirling around him (or on occasion a toy helicopter controlled by Assen may be swirling around him), but Hannes will stick to the task and get it done.

Mostly that task is hierarchical clustering in some form, or bullying young students into producing more data to feed his insatiable clustering habit. Although he disguises it well, 99% of what Hannes does is clustering. Like a jazz musician playing variations on a theme, Hannes will normalize the data some fancy way, he will cluster by distribution, by correlation, by density, he will rank and evaluate, he will visualize the clusters, but it's still the same damn jazz theme and it's still just clustering Hannes!

Of course, this is not remotely true. Much as it pains me to admit it, Hannes is a skilled, perceptive and imaginative scientist, with a broad arsenal of analytical tools. His work is always high quality, and there is a purpose to each experiment or analysis he carries out - no data for the mere sake of data with Hannes. In my opinion, the literature of epistasis mapping will be seen as pioneering and important, and Hannes has made many critical contributions to this work. In a publication record still only a few years old, he has co-authored with many great

scientists – Jim Haber, Andrej Sali, Brian Strahl, Tobias Walther, Kevan Shokat, Christine Guthrie, Jonathan Weissman (but not, notably, the great Peter Uetz).

Hannes seems to mix a curiosity about the world, with a natural contentment with life. He is easy company and a good friend. He is universally liked and gets on well with any type of personality. I have never seen Hannes become angry, anxious, or even slightly perturbed. He is a good mentor, and generous and selfless with helping other students. He should be highly valued as a person one can approach for rigorous but constructive discussion of novel ideas. I wish I could say I have discovered some unspeakable private obsessions, but the worst I've come across is his secret mastery of the bandoneón.

But this is getting embarrassing. Obviously, it was no merely academic or humane talent that prompted Nevan to offer him a position in the Krogan Lab. Instead, Nevan recognized in Hannes a skill so rare and refined that it outshines his other attributes. For Nevan is an aficionado of the AWKWARD SILENCE….and Hannes is the MAGICIAN OF AWKWARD SILENCES. In the elevator with a stranger, in the minutes before a tense meeting begins, when the new person has made their first *faux pas* – these are the moments Hannes cherishes (with Nevan approvingly looking on), transforming what was merely a passing embarrassment into an excruciating lengthy trial of suspended tension – by his refusal to speak, move, or show any indication that an awkwardness exists.

Each one of us can remember with shame a personal Hannes awkward silence moment, but few realize that this moment was probably planned long in advance by Hannes, executed with precision, and later recounted behind the closed doors of Nevan's office to the muffled sound of clinking beer glasses and laughter. Over the years, Nevan has assembled a crack team of ASSes (Awkward Silence Sustainers) – Art Wuster, Mike Shales, Cathal Mahon, Dorothea Fielder, Jeff Johnson, Colm Ryan – hardened men and women resistant to any type of awkward situation. But Hannes returns their silence, amplifying it to earsplitting levels, forcing them to back down – he cannot be broken.

Of course there are many stories about Hannes it would be uncharitable to mention here – the entertaining situation before he met Ayse when a stream of giggling females would visit his desk for 'advice'; the moment he finally surpassed Gian Luca's tolerance of trivial banter with one outrageous claim too many; the fact that he is Norwegian – I will not mention them.

Instead, I'd like thank him for the many hours of intellectual engagement and fun, enjoyed by Krogan Lab members and passersbys in QB3 over the years. My apologies to Hannes' many many friends and colleagues I haven't mentioned here by name.

And so the time has come Hannes, when you must defend your thesis, leave the comforting milk of the Krogan Lab and step into the world. You have learnt all you

can from the two great masters…. the Breaking Bad counterculture genius of Assen Roguev…the nefarious 'bioinformatics of the night' practised by Pedro Beltrao…..which way will you choose?

Den som lever får se

Best of luck buddy

ger

Acknowledgements

Right off the bat, I'd like to thank my advisor, Nevan Krogan. You've been a great mentor and friend and I could not imagine a better experience than I've had during my time in your lab. From the truly engaging science, to the paper writing sessions at a café that shall remain nameless, and the countless beers in your office and at other venues, I would not change a thing… Well, except the time you strong-armed me into carrying your thick wool sweater during a 107-mile bike ride on a hot summer's day, or the time you convinced me to see if I could fit into a cupboard that subsequently collapsed under my weight. Objectionable events such as those aside, I truly doubt I could have found a better advisor. Thank you!

Then there are the many other PIs I have had the pleasure and good fortune to work with. My dissertation committee – Nevan Krogan, David Agard, Tanja Kortemme and Andrej Sali – I truly appreciate all the support and time you've given me. My rotation PIs – Carol Gross and Kevan Shokat – I much enjoyed my time in your labs and I am grateful that I have had the opportunity to keep interacting with you both. Christine Guthrie and Jim Haber, thank you for all the time you have put into our collaborations. My qualifying exam committee – Yifan Cheng, Tanja Kortemme, Geeta Narlikar and Andrej Sali – thanks for a truly engaging and valuable exam, as well as the many meetings and advice leading up to it. My mentors prior to joining the Biophysics program at UCSF – Andrej Sali, M.S. Madhusudhan and Cyrus Chothia – first of all, I very much enjoyed working with each of you, and appreciate your support and friendship outside of lab as well. Second, I would never have been

able to come to UCSF for my PhD if it was not for the three of you… Suffice it to say I have had an absolutely amazing time here and it is much thanks to you.

I have gained many collaborators and friends during my time at UCSF. Assen Roguev… a genius, a loose cannon, an outcast, and a great friend. Thanks for teaching me all I ever wanted to know about genetics and being so generous with your time whenever my clumsy hands needed guidance in the wet lab. Then there's all the times outside of lab, be it planned excursions or chance encounters of you comfortably asleep in a bus stop or banding pigeons at 16th and Mission. It would not have been the same without you, but I did not appreciate the time you took off on the I-80, leaving me chasing you for 30 minutes to guide you back to our sailing lesson that had already started. Cathal Mahon and Grace Li, I wanted to bundle you with Assen here, but since you're neither loose cannons nor outcasts, that would be unfair. Weekends would have been so much less fun without you two around. In particular I would like to thank you for making Sunday the best day of the week. This is quite a feat, mind you, given that Sundays are immediately succeeded by the dreaded Mondays. A special thanks to Grace for always being the driver. Gerard Cagney, you were there from the very beginning and spent a lot of time mentoring me. I really appreciate that and am grateful for your friendship. I'd also like to thank you for not putting an end to one of my early dates with Ayse, even though I had offered you $50 to do so. Turns out she was a keeper… Finally, thanks for teaching me everything I ever wanted to know (actually, a lot more) about ospreys. Erik Verschueren and Anne Campagna, I'm really glad you two have joined us here in SF. I knew you were a force to be reckoned with as soon as Erik challenged the great

Pedro ‘P-P’ Beltrao for his desk, while Anne distracted Silvia. It was a gutsy move, but it paid off and Erik ousted P-P, who had been my cubicle neighbor for almost 5 years. P-P, you’re a brilliant colleague and friend, and I was truly fortunate to sit next to you, asking you for advice over and over again. Also, you are a walking encyclopedia of hair products. You did once ruin what would have possibly been the best practical joke of mankind, but for that I forgive you. Jeff ‘Jeffie’ Johnson, thanks for being a wonderful neighbor and always such a good sport with my craziness (which I know you enjoy). Mike Shales, thanks for being so dependable and always there to help with whatever we ask for, be it on weekdays, weekends or Christmas Day (but possibly not on Canada Day). Colm “Rebel Boy” Ryan, thanks for being a great help with everything statistics. I’m no statistics slouch myself, but you always seem to take it up a notch. Gian Luca Negri, Laura Kapitzky, Marie Fahey & Natali Gulbahce… Come back guys, our cubicle misses you (as do I)! Stefan Bohn, you always invoke emotion, be it fear, joy or simply confusion. I enjoy this much, keep it up. Peter and Klavdija Cimermancic, I really enjoy the company of you guys, and I bet it will be even more exciting when Adrian learns to say more than “car”. Art Wuster, the wearer of classy shoes and purchaser of last year’s calendars, thanks for your company both here and in Cambridge. Erica Moehle, thanks for being such a great collaborator. I am happy that our project is successfully wrapped up, but I will much miss the many meetings over coffee or at the Ramp. Dorothea Fiedler, Craig Kaplan, Sean Collins, Shuyi Wang, Jiewei Xu, Koy Saeteurn, Ada Cheng, Paivand Jalalian, Richard Alexander, Tomas Linder, Nassos Typas, Scooter Morris, Kathy Franks-Skiba, Rebecca Brown, Aileen Paterson… I’ve enjoyed working with all of

you and appreciate the help… I would also like to thank all the members of the Krogan Lab I didn't mention here, as well as my colleagues in the Biophysics and Bioinformatics Graduate Groups.

I owe much to all my friends that took care of me and made me feel so welcome during my first times in SF, prior to my PhD. Madhu and Andrej, I would have been unemployed if not for you. Kasia Bernacki and Petri Fast, I would have been homeless if not for you, and I won't forget your kindness and hospitality. Kingshuk Ghosh, I would have behaved a lot better if not for you, and I am very fortunate to keep seeing you so often even after you left SF. Banu Ozkan, thanks for keeping Kings under control. And if it wasn't for all of you, I would not nearly have enjoyed myself as much, nor would I have come back to pursue my PhD. Thanks so much!

I'd also like to thank the many other friends I have made in San Francisco during my time here. Steve and Lin Woodward, thanks for providing us with such wonderful company and for countless dinners. David Hoatson, you almost broke me the first time I met you, forcing me to run a dozen miles at a ludicrous pace. I much prefer you as a friend to a running partner. Berta Rodriguez, thanks for the many enjoyable weekday dinners. Eric Brecht, my mountain bike buddy, thanks for supporting my action sports lifestyle and for always driving to the trails. Christopher Magan, thanks for being the only person I know who looks forward to calling me "Doctor".

Finally, I want to thank family and old friends who mean a lot to me. Jonas Carlstedt, thanks for always being such a great friend and for being so welcoming

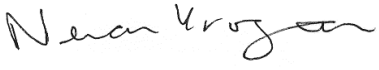
whenever I go back to visit. My father, Lennart, and Katherine, thanks so much for all support and for many good times together. I've enjoyed it very much and am happy I get to see you so often. Thanks to my aunt, Inger, for providing housing and great company during my studies in Uppsala, and for being an extraordinary aunt in general. Torgil, thanks for always having been a wonderful big brother and not having beaten me up too much. Thanks to my mother, Ingrid, for being the best mother anyone could ever wish for and for being so supportive of all my choices. For driving me to ice-hockey training as a kid, for spending your weekends at the go-kart racing track with Torgil and me as teens, for supporting my motorcycle classes as a young adult, and for encouraging my choice to move all the way to San Francisco for my PhD… Thank you! I would like to thank Mocha the dog for guarding the silence all those nights and days I worked from home, but he didn't, so I won't. However, he does keep Ayse company when she needs more attention than I have, so I thank him for that. And finally, Ayse (who in all honesty does not ask for much attention)… Thank you for being such a good sport, never stopping my crazy ideas, for being so sweet and selfless, and for making my life so much fun!

Contributions

Several chapters of this dissertation describe studies that have been published in the journals cited herein. The chapters do not necessarily represent the final published form. These studies were carried out with the collaborators listed as co-authors for each chapter. The study described in Chapter 2, "High-throughput, quantitative analyses of genetic interactions in *E. coli"*, was published in September 2008 in *Nature Methods*, Vol. 5(9), pp. 781-787. In this study, Hannes Braberg developed an extension of the scoring scheme for genetic interactions in yeast to function with E. coli, and carried out data analysis. The study described in Chapter 3, "A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing", was published in December 2008 in *Molecular Cell*, Vol. 32(5), pp. 735-746, and that described in Chapter 4, "A lipid E-MAP identifies Ubx2 as a critical regulator of lipid saturation and lipid bilayer stress", is in press at *Molecular Cell*. In these studies, Hannes carried out computational analyses of genetic interaction data. The study described in Chapter 5, "Functional organization of the *S. cerevisiae* phosphorylation network", was published in March 2009 in *Cell*, Vol. 136(5), pp. 952-963. Hannes carried out the computational analysis of this study and wrote the paper with co-authors. The study described in Chapter 6, "A plasma-membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking", was published in June 2010 in *Nature Structural & Molecular Biology*, Vol. 17, pp. 901-908. Hannes carried out data analysis of the plasma-membrane E-MAP. The study described in Chapter 7, "Systematic Triple-Mutant Analysis Uncovers Functional Connectivity

between Pathways Involved in Chromosome Regulation", was published in June 2013 in *Cell Reports*, Vol. 3. Hannes designed the Triple-Mutant Analysis (TMA) experiments with Jim Haber and Nevan Krogan, carried out TMA experiments with Richard Alexander, analyzed data, and wrote the paper with co-authors. The study described in Chapter 8, "Individual lysine acetylations on the N terminus of *Saccharomyces cerevisiae* H2A.Z are highly but not differentially regulated" was published in December 2010 in the *Journal of Biological Chemistry*, Vol. 285(51), pp. 39855-39865. Here, Hannes performed genetic interaction assays of *htz1* alleles together with Shuyi Wang and carried out data analysis. The study described in Chapter 9, "RNA polymerase II carboxyl-terminal domain phosphorylation regulates protein stability of the Set2 methyltransferase and histone H3 di- and trimethylation at lysine 36" was published in January 2012 in the *Journal of Biological Chemistry*, Vol. 287(5), pp. 3249-3256. Hannes carried out the genetic interaction analysis of *set2* alleles in this study. The study described in Chapter 10, "From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II", is in press at *Cell*. Hannes designed and directed the study, generated the pE-MAP, carried out growth rate assays and all computational analyses, managed collaborations, and wrote the paper.

The work of Hannes Braberg described herein was done under the supervision of Nevan Krogan and exceeds the requirements for a standard PhD dissertation.



High-Resolution Quantitative Analysis of Genetic Interactions

Hannes Braberg

Abstract

Genetic interactions describe how the presence of one gene affects the function of a second gene. Negative genetic interactions arise when two mutations together cause a stronger growth defect than expected based on the individual single mutations, while positive genetic interactions correspond to cases where the double mutant is either no sicker (epistatic) or healthier (suppressive) than the sickest single mutant.

Our lab has developed a high-throughput technology in budding yeast for quantitative analysis of the entire spectrum of genetic interactions, termed epistatic miniarray profile (E-MAP). The genetic profiles generated by this method provide highly specific readouts that can be used to identify genes that are functionally related. In this dissertation I describe my work related to the employment and extension of the E-MAP technology. With collaborators, I have interrogated several biological processes using E-MAPs, including the phosphorylation network, RNA processing, lipid metabolism, and the plasma membrane. We further extended the technology to facilitate analysis of higher-order genetic interactions by examination of triple mutants, and we developed an adaptation of E-MAP to function in E. coli.

My main focus was on the development of an important advance of the E-MAP technology, which allows us to address higher levels of complexity by examining genetic interactions of point mutant alleles of multi-functional genes. The technique, termed point mutant E-MAP (or pE-MAP), greatly increases the

achievable resolution as it allows assignment of genetic relationships to individual residues and domains. We applied this system to genetically dissect RNA polymerase II (RNAPII) in budding yeast, and generated a pE-MAP comprising ~60,000 quantitative genetic interactions. Using these data, we assigned functions to RNAPII sub-domains and uncovered connections to protein complexes. The pE-MAP further allowed us to characterize connections related to RNAPII activity. These include an inverse relationship between *in vitro* transcription rate and *in vivo* splicing efficiency, classification of fast and slow mutants that shift transcription start upstream and downstream, respectively, and identification of Sub1 as a positive transcription factor that regulates start site selection and influences splicing. The pE-MAP approach provides a powerful strategy to understand other multi-functional machines at amino acid resolution.

Table of Contents

List of Figures and Tables

List of Supplementary Materials

Chapter 1

Introduction

Introduction

Genetics is the science of heredity and variation, and provides a gene-based toolbox to study biological problems. Budding yeast has long been a model organism for genetics and since the identification of its genes in the genome sequencing project (Goffeau et al., 1996) a large effort has been carried out to map their functions. By deleting a gene and studying the resulting phenotype we can gain information about its function. Perhaps the most fundamental phenotype to study for this purpose is growth rate, and this was completed for budding yeast on a genome-wide scale in 2002 (Giaever et al., 2002). Shortly thereafter, two proteome-wide studies were carried out to characterize the protein complexes in yeast (Gavin et al., 2006; Krogan et al., 2006). A natural complement to these studies is the characterization of functional connections and pathways in yeast. To this end, an effort was undertaken to map genetic interactions, which describe how the function of a given gene is affected by the presence or absence of a second gene.

These genetic interactions are invaluable for determining functional relationships between pairs of genes. Negative genetic interactions (synthetic sick/lethal interactions, SSL) arise when two mutations together cause a stronger growth defect than expected based on the individual single mutations. For example, removal of two factors in parallel pathways that produce the same essential product would result in cell death, whereas removing each alone would at most result in cell sickness (Figure 1). Conversely, positive genetic interactions correspond to pairs of mutations where the double mutant is either no sicker (epistasis) or healthier (suppression) than the sickest single mutant. In context of the above example, we

would observe this if the two factors were in the same serial pathway; removal of the first factor would abolish the function of the pathway, and removing a factor upstream or downstream in the same pathway would add no harm to the cell (Figure 1).

Two approaches have been developed to identify SSL interactions on a large-scale in budding yeast; synthetic genetic arrays (SGA) (Tong et al., 2001; Tong et al., 2004) and diploid based synthetic lethality analysis on microarrays (dSLAM) (Pan et al., 2004). SSL interactions however only represent a subset of the genetic interaction spectrum. This is a major limitation and precludes the collection of complete genetic interaction patterns, which are often more informative than single interactions as they report on the phenotype in a large number of mutant backgrounds. These genetic profiles provide highly specific readouts that allow precise identification of sets of genes that have similar effects on cell physiology.

Shortly prior to my joining Nevan Krogan's lab, he developed an extension of the SGA approach, together with colleagues Sean Collins, Maya Schuldiner and Jonathan Weissman. This technique, termed E-MAP, allows for quantitative identification of the entire spectrum of genetic interactions (Collins et al., 2006; Schuldiner et al., 2005; Schuldiner et al., 2006). The resulting genetic profiles are of much higher resolution and specificity than previous attempts that were based on qualitative SSL identifications. The method is based on the array of double mutants in high-density grids on yeast plates and relies on imaging software to determine colony sizes from images of the plates. The first two E-MAPs, focused on the early secretory pathway (Schuldiner et al., 2005) and chromosome biology (Collins et al.,

2007), were completed shortly before my joining the lab. The method was at this point only applicable to budding yeast, but an extension to work in fission yeast was soon developed (Roguev et al., 2007). However, at the time, prokaryotes were not amenable to large-scale genetic interaction studies, and one of my first projects involved the extension of the E-MAP technique to prokaryotes. We thus developed a method for high-throughput generation of double mutants in *Eschericia coli*, and designed a modified genetic interaction scoring system to function in *E. coli*. We named this system Genetic Interaction ANalysis Technology for *E. coli* (GIANT-coli), and it is described in Chapter 2.

During the development of GIANT-coli, I grew increasingly interested in further extending and exploiting the E-MAP technology. I decided to focus on the original budding yeast E-MAP system, as it was most thoroughly benchmarked and budding yeast is a well-studied model organism. I thus continued onto the analysis of an E-MAP focused on genes involved in RNA processing, including factors that generate or regulate mRNAs, tRNAs, snRNAs, snoRNAs and rRNAs. We predicted and verified several functional connections within and between RNA related processes. Furthermore, by combining the genetic interactions with physical interaction data, we discovered two previously uncharacterized protein complexes involved in splicing and mRNA export. Finally, the data revealed several proposed functions for poorly annotated genes. This study is described in Chapter 3. Using a similar approach, we generated an E-MAP focused on lipid metabolism and trafficking, which revealed that the protein Ubx2 is involved in fatty acid

desaturation and membrane homeostasis. More details on this connection, and several other findings from the study, are provided in Chapter 4.

Because the E-MAP technology efficiently identifies functional connections between proteins, complexes and pathways, it provides a powerful framework for systematic network analysis. Harnessing this potential, we applied it to map the phosphorylation network in budding yeast. Our genetic interaction data added new levels of detail to known kinase pathways and revealed a link between kinases Cak1 and Fus3, and a chromatin integrity pathway. In this analysis, described in Chapter 5, we also extended the previous view of pairwise genetic interactions to a higher-order map of sets of three genes that interact strongly with one another. This triplet motif analysis proved highly useful, and we later re-employed it to interrogate the genetics of plasma membrane biology. Detailed in Chapter 6, we used this map and analysis to identify a new component of the eisosome and link a poorly characterized gene to endocytic and eisosome function.

While the triplet motifs involve sets of three genes, each interaction is still between a single pair of genes, and we were thus only exploring double mutant interactions. However, we believed that greater mechanistic insight could be gained through higher-level genetic interaction analyses. To this end, we developed an approach termed Triple Mutant Analysis (TMA), which allows for the generation and quantitative analysis of triple mutants in budding yeast. This method is particularly useful to discover novel genetic interactions when cells are deleted for pairs of genes known to act redundantly. In Chapter 7 I describe the TMA analysis

and how we used it to uncover complex relationships that could not have been gleaned simply from double mutant analysis.

To this point, the strains screened in E-MAPs had either deletions of non-essential genes or hypomorphic alleles of essential genes. There are however limitations to this approach. First, by screening different alleles of the same gene, we can characterize functional differences between these and hence determine the regions of the gene that are most important for its function. I utilized this methodology on a limited scale in Chapter 8, where we compared a range of unacetylable *htz1* alleles, and in Chapter 9, where we examined a truncated form of Set2.

However, since many important proteins are multi-functional, the most significant limitation to the 'all or nothing' approach of deletions and hypomorphic alleles is that secondary functions tend to get masked. I spent the majority of my PhD developing an important advance of the E-MAP technology, termed point mutant E-MAP (or pE-MAP), which allows us to address higher levels of complexity by examining point mutants of multi-functional and essential genes. This approach greatly augments the precision of gene function analysis as it reports on underlying genetic relationships at an amino acid resolution. I applied the pE-MAP technique to genetically dissect RNA polymerase II (RNAPII), an essential multifunctional protein complex that plays a central role in the dynamic control of gene expression. The approach allowed us to functionally characterize RNAPII sub-domains and to uncover connections to protein complexes. In addition, the pE-MAP classified groups of RNAPII mutants with increased and decreased transcription elongation

rate *in vitro* that favor upstream and downstream start-site selection respectively. Analyzing these rate-altering mutants on splicing microarrays, we further discovered an inverse relationship between RNAPII catalytic activity *in vitro* and splicing efficiency *in vivo*. Finally, the pE-MAP identified the protein Sub1 as a positive transcription factor that controls start-site selection and affects splicing. These results, detailed in Chapter 10, highlight the pE-MAP technique as a powerful tool for the functional characterization of multi-functional molecular machines.

Figures

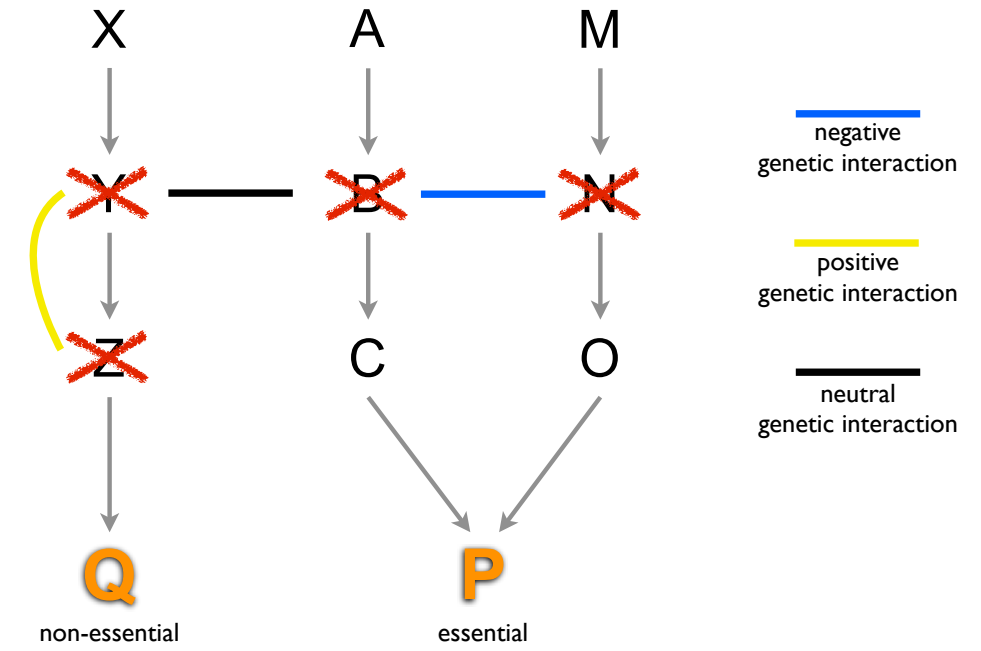


Figure 1. Negative and positive genetic interactions. Three pathways are shown. The two pathways on the right function in parallel and produce the same essential product, **P**. If the gene encoding either protein **B** or **N** is deleted, the corresponding pathway is abolished, but the parallel pathway maintains production of **P**. However, if both **B** and **N** are deleted, the production of **P** is completely abolished, resulting in cell death. The effect on the cell is more severe than for two unrelated genes (e.g. **Y** and **B**) and we observe a negative genetic interaction (specifically synthetic lethal). The pathway on the left is non-essential, and deleting either **Y** or **Z** results in a growth defect due to lack of product **Q**. However, deletion of **Y** and **Z** together has the same effect as deletion of each separately. The double mutant cell growth is thus better than expected, and this is an example of a positive genetic interaction (specifically epistasis).

References

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M*., et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol *7*, R63.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B*., et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B*., et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387-391.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M*., et al.* (1996). Life with 6000 genes. Science *274*, 546, 563-547.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P*., et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F., and Boeke, J.D. (2004). A robust toolkit for functional profiling of the yeast genome. Mol Cell *16*, 487-496.

Roguev, A., Wiren, M., Weissman, J.S., and Krogan, N.J. (2007). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat Methods *4*, 861-866.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F.*, et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Schuldiner, M., Collins, S.R., Weissman, J.S., and Krogan, N.J. (2006). Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344-352.

Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M.*, et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Chapter 2

High-throughput, quantitative analyses of genetic interactions in *E. coli*

High-throughput, quantitative analyses of genetic interactions in *E. coli*

Athanasios Typas[1], Robert J. Nichols[1], Deborah A. Siegele[2], Michael Shales[3], Sean Collins[3], Bentley Lim[1], Hannes Braberg[3], Natsuko Yamamoto[4], Rikiya Takeuchi[4], Barry L. Wanner[5], Hirotada Mori[4]*, Jonathan S. Weissman[3,6], Nevan J. Krogan[3]* and Carol A. Gross[1]*

1 Department of Microbiology and Immunology, University of California at San Francisco, USA

2 Department of Biology, Texas A&M University, USA

3 Department of Cellular and Molecular Pharmacology & The California Institute for Quantitative Biomedical Research, University of California at San Francisco, USA

4 Graduate School of Biological Sciences, Nara Institute of Science and Technology, Ikoma, Japan

5 Department of Biological Sciences, Purdue University, West Lafayette, IN 47907 USA

6 Howard Hughes Medical Institute, University of California at San Francisco, USA

*corresponding authors; e-mail addresses: cgross@cgl.ucsf.edu, krogan@cmp.ucsf.edu & hmori@gtc.naist.jp

.

Summary

Large-scale genetic interaction studies provide the basis for defining gene function and pathway architecture. Recent advances in the ability to generate double mutants on a genome-wide basis in the budding yeast, *S. cerevisiae*, have dramatically accelerated the acquisition of such genetic interaction information and the biological inferences that follow. Here, we describe a method based on F-driven conjugation, which allows for high-throughput generation of double mutants in *E. coli*. This method, termed Genetic Interaction ANalysis Technology for *E. coli* (GIANT-coli), permits us to systematically generate and array double mutant cells on solid media, in a variety of formats, including high-density arrays (1536 colonies/plate). We show that colony size provides a robust and quantitative output of cellular fitness and that GIANT-coli can recapitulate known synthetic interactions and identify new negative (synthetic sickness/lethality) and positive (suppressive/epistatic) relationships, either when applied to a limited gene set or when used for genome-wide screens. Finally, we describe a complementary strategy for identification of suppressor mutants on a genome-wide level. Together, these methods permit rapid, large-scale genetic interaction studies in *E. coli*.

Introduction

Genetic interactions report on the extent to which the function of one gene depends on the presence of a second gene and have a long history of facilitating the identification and characterization of cellular pathways. In *Saccharomyces cerevisiae*, recent advances in the ability to systematically create double mutant strains has resulted in new technologies enabling genome-wide genetic interaction screens (Boone et al., 2007). For example, the plate based synthetic genetic array (SGA) (Tong et al., 2001; Tong et al., 2004) and diploid based synthetic lethality analysis on microarrays (dSLAM) (Pan et al., 2004) approaches have been extensively used to identify synthetic sick/lethal (SSL) interactions for a given query deletion strain. However, SSL interactions can be often difficult to interpret as they tend to occur between genes involved in different pathways (Kelley and Ideker, 2005). Recently, we exploited the SGA strategy to develop an approach called E-MAP (Epistatic Mini Array Profiles) (Collins et al., 2007; Schuldiner et al., 2005), which allows for identification of not only negative interactions, such as SSL pairs, but also positive interactions in which the double mutant is healthier than would be expected based on the growth of the two single mutants. E-MAPs contain comprehensive and quantitative information about genetic interactions between genes that are logically connected either because they participate in common processes or are constrained in cellular location. Together, these three approaches have led to a dramatic increase in the number of genetic interactions reported, and have provided functional insights into numerous cellular networks (Collins et al., 2007; Denic and Weissman, 2007; Fujita et al., 2006; Keogh et al., 2006; Keogh et al.,

2005; Krogan et al., 2003; Mayer et al., 2004; Schuldiner et al., 2005). Analogous approaches have been developed in *C. elegans* using RNAi technology (Byrne et al., 2007; Lehner et al., 2006) and we have recently described a system that allows for rapid generation of quantitative genetic interaction data in the fission yeast *Schizosaccharomyces pombe* (Roguev et al., 2007).

In contrast to the central place of genetic interaction screens for the phenotypic analyses of eukaryotic organisms, such screens are relatively rare in prokaryotic organisms and when employed, the query gene-set has been usually very limited (Campbell and Brown, 2008; Elena and Lenski, 1997; Onufryk et al., 2005; Ullers et al., 2007; Wolfe et al., 2008; Wu et al., 2005). Additionally, by definition, synthetic lethal double mutants are not viable in the haploid state, meaning that their identification in prokaryotes is difficult using conventional genetic tools. Although conditional alleles are sometimes used, strategies must be developed on a case-by-case basis, preventing high-throughput approaches that facilitate the functional characterization of unknown genes. As a consequence reports of genetic interactions in *E. coli* are rare; we are aware of fewer than 200 reported synthetic lethal interactions, as compared to almost 20,000 in *S. cerevisiae* (Breitkreutz et al., 2008), even though the two genomes are of roughly comparable size. Lack of genetic interaction data undoubtedly contributes to the lagging functional annotation of bacterial genomes. Even in *E. coli* and *B. subtilis*, arguably the best-studied prokaryotes, one third of the genes have no known function (Ara et al., 2006; Ito et al., 2005) and for less well-studied bacterial species, the situation is even worse. The immense amount of information generated by genomic and

metagenomic approaches exacerbates this problem. To-date, 626 bacterial genomes have been sequenced, and 961 are in progress (NCBI database), and direct DNA sequencing from uncultured microbial communities has identified several thousand additional genes (Rusch et al., 2007). In addition, the Human Microbiome, a new NIH roadmap project, aims to identify human microbial flora and provide sequence information for over 1000 organisms. Next-generation sequencing technologies (Schuster, 2008) ensure that the pace of discovering new genes will accelerate. The development of high-throughput genetic interaction screens applicable to bacterial species will be invaluable in utilizing this information to understand gene function and pathway organization.

Here, we report a method, termed GIANT-Coli (Genetic Interaction ANalysis Technology for *E. coli*) that allows for rapid construction of double mutant strains for the first time in a prokaryotic organism. This method is based on the well-characterized Hfr conjugation gene transfer system, and uses two comprehensive *E. coli* mutant libraries of ~4000 single-gene deletions, the Keio collection (Baba et al., 2006) and the newly developed ASKA single-gene deletion set (N. Yamamoto et al. unpublished data), to systematically generate double mutant strains arrayed in high density on a solid surface. Colony sizes of the resultant double mutant strains are quantitatively assessed using a modified version of the algorithm originally developed for *S. cerevisiae* (Collins et al., 2006; Schuldiner et al., 2005), together with normalization procedures that are tailored for *E. coli* growth. As a proof of principle experiment, we generated a 12 by 12 genetic interaction matrix, which demonstrated the capacity of the methodology to recapitulate previously described

genetic interactions and to discover novel negative and positive interactions. We also show that we can use the same methodology to quantitatively identify genetic interactions on a genome-wide scale. Finally, we describe a complementary strategy for rapid identification of suppressor mutants.

Results and Discussion

Rationale for Method of Gene Transfer

The major bottleneck in producing genetic interaction data is development of an efficient and robust method for the mass generation of double gene deletion mutants. We chose to transfer marked deletions from one *E. coli* strain to another by Hfr mating (conjugation), rather than P1 phage generalized transduction, because Hfr mating is more efficient in transferring markers and works well on solid agar surfaces, which is more amenable to high throughput technology. The Hfr donor (male) has a chromosomally integrated conjugative F plasmid. Upon contact with a recipient cell lacking an F (F^-; female), the Hfr donor is nicked at the origin of transfer (*oriT*) within the integrated F plasmid and oriented transfer of a single strand of the circular *E. coli* chromosome proceeds from *oriT*. The transferred single-stranded DNA is replicated in the recipient and can be maintained in the cell only after integration via double crossover recombination. The Hfr itself is almost never transferred to the daughter cell chromosome. Existing Hfr strains (Low, 1996) were isolated after multiple rounds of mutagenesis, have poorly documented genotypes, and would thus introduce uncontrolled heterogeneity into a genetic interaction analysis. Therefore, we created a new Hfr strain isogenic to our single

gene-knockout libraries. We transduced the "pseudo Hfr" made by integrating the transfer region of F into the *trp* operon (Francois et al., 1990) into our wild-type background to create an isogenic Hfr (see methods). The "pseudo-Hfr" transferred as efficiently as the classic high mating Hfr 3000 strain (Bachmann, 1996) in our high-throughput mating methodology (Fig. S1A, B) and was used in all subsequent experiments in this report.

Development of GIANT-coli, a methodology for rapid creation of double mutants using Hfr mating

We devised a high-throughput mating system that robustly generates double mutants in *E. coli* (Fig. 1). In Step 1, a derivative of the pseudo-Hfr donor strain, containing a single gene deletion marked with kanamycin (Kan^R; Keio collection (Baba et al., 2006)), was mated on agar plates to ASKA "recipient strains", a set of single-gene knockouts marked with chloramphenicol (Cm^R; N. Yamamoto et al. unpublished data) or vice versa. To perform mating in a high-throughput format, recipient strains were robotically arrayed on agar plates in the desired format (384, 768 or 1536 colonies per plate), grown overnight and then transferred onto an agar plate previously inoculated with a lawn of an isogenic donor strain. These "mating plates" were incubated overnight to allow growth and mating of the parental strains. In Step 2, cells from the mating plates were transferred robotically onto agar plates containing one antibiotic (kanamycin). The rationale for this "intermediate selection" is explained in (A) below. In Step 3, cells from the intermediate selection plate were transferred robotically to a plate with both antibiotics to select for

double recombinants. The double recombinants were imaged after growing for an experimentally determined time that ensures low background growth of synthetic lethal pairs and permits easy differentiation between sick and healthy mutants (see Fig. S2 for a more detailed protocol).

Three critical parameters were optimized so that this procedure performed robustly: 1) efficiency of mating, 2) efficient transfer of most of the chromosome and 3) low recovery of strains in which the locus examined has been duplicated, leading to recovery of "false positives" such as strains having gene A marked both with kan^R and Cm^R.

First, to obtain high and reproducible mating, the ratio between donor and recipient cells proved critical. The efficiency of mating is influenced by the growth phase and number of the donor cells spread on the lawn, by the length of time the lawn has been grown prior to transfer of the arrayed recipients to it, and by the type of pinner head used to robotically transfer recipients. The exact conditions used in this study are described in Fig. S2. Second, we examined whether markers far from *oriT* were transferred with lower efficiency than those close to *oriT*. Transfer of the entire chromosome from *oriT* to the F^- occurs across a fragile pilus-bridge and takes ~ 90 minutes. Breakage of the pilus aborts transfer, leading to a gradient of transfer depending upon distance of the marker from *oriT*, which is ultimately reflected in recovery of fewer recombinants for distal markers. Longer matings on a solid surface partially obviate this problem. Even when markers far away from *oriT* were used, i.e. 4.2 MB from *trp::oriT* (90% coverage of the chromosome), there was no significant drop in the recovery of recombinants, as measured by colony size (data

not shown). Note that these colonies are the visible manifestation of the cells arrayed on the plate and do not arise from a single cell. To ensure efficient transfer of the furthest markers from *trp::oriT* (close to 100% chromosome coverage), we designed a new plasmid that can target Hfr transfer genes to specified chromosomal positions, thereby allowing us to position an *oriT* to transfer those regions early (see "Using GIANT-coli at the genomic level" section below). Third, we developed two strategies, the intermediate selection step and the minimal media protocol to minimize recovery of false positives, created by duplication of a region of the chromosome. The effectiveness of our strategy was monitored by the ability to prevent generation of strains having the same gene marked by both the donor and recipient antibiotics (e.g. when an Hfr strain with gene A::Kan^R is mated with an F^- with gene A::Cm^R), which we refer to as "self-mating".

(A) The intermediate selection step. Rapidly growing *E. coli* in LB initiates multiple (up to 16) rounds of replication to ensure complete duplication of its chromosome before division (Nielsen et al., 2007), allowing for facile generation of strains that contain some duplicated regions of the chromosome. Plating directly on the double antibiotic selects for and maintains such strains, thereby increasing the fraction of "self-mating" recombinants. Because the "intermediate selection" is performed on plates with only the bacteriocidal antibiotic (kanamycin), strains with duplicated regions are not selected. Instead, they are eliminated both by spontaneous resolution of this unstable state and by competition from growth of the Kan^R parent on these plates. The intermediate selection also ensures that only one parent is transferred to the subsequent double antibiotic plate, thereby eliminating

any new rounds of mating and further generation of rare duplications. Finally, this step magnifies small differences in growth of the daughter double mutants, allowing for easier detection of genetic interactions. As a consequence, the intermediate selection virtually eliminates the background from "self-mating" (Fig. S1C&D) as well as growth observed with truly synthetic lethal pairs.

(B) The minimal media protocol. A second way to reduce generation of duplicated regions of the chromosome is to minimize the occurrence of multiple genomes within a single cell, which is the primary source of duplication events. When *E. coli* cells grow on a poor carbon source e.g. M9-glycerol minimal medium, division is sufficiently slow that cells have ~1 genome per cell (Helmstetter, 1996; Nielsen et al., 2006). As expected, when mating was performed on M9 glycerol medium, the background growth of self-mating and synthetic lethal pairs was significantly decreased (data not shown). The intermediate selection decreased this background even further.

Validation of GIANT-Coli

As a proof of principle experiment to assess our strategy for mapping genetic interactions in *E. coli*, we performed a 12 by 12 genetic cross, which provided 66 distinct, pair-wise double mutant strains as well as 12 self-matings. Our choice of genes (*surA*, *ybaY*, *ycbS*, *ompC*, *yraI*, *cpxR*, *degP*, *pal*, *ompA*, *yfgL*, *yraP* and *basR*) was based on the following criteria: i) spatial compartmentalization to increase interaction probability; all gene-products studied here have roles associated with the envelope compartment of the cell, ii) gene pairs with previously detected

interactions as a positive control (*surA*/*degP*, *surA*/*yfgL*, *degP*/*yfgL*), iii) gene pairs close together in the chromosome to provide information about the linkage cut-off of our methodology (*yraP*/*yraI*, *ompA*/*ycbS*), iv) a marker far from *oriT* to evaluate whether chromosomal position affects recombination efficiency (*ompC*), and v) mutants with significant growth defects (*surA* or *pal*) to facilitate identification of double mutants exhibiting positive interactions.

Choosing a relatively small subset of genes allowed us also to array each recipient multiple times on the same plate so that we could directly assess reproducibility, compare different plate formats and media, and examine the growth differences between reciprocal mutant pairs. Specifically, we compared genetic interaction results when recipients were arrayed in 384 or 1536 format and when the screens were carried out in rich (LB) or minimal (M9-glycerol) media. Representative plates are shown for crosses of pseudo-Hfr *pal*::*Kan*^{R} with the 12 Cm^R-marked recipients arrayed in 1536 (128 replicas of each mutant; Fig. 2A & S3A) and 384 (32 replicas of each mutant; Fig. S3C & E) formats. Qualitative examination of these plates indicated there was almost no growth in the self-mating pair (*pal::Kan*^{R} x *pal::Cm*^{R}). In fact, the growth in these regions is comparable to that seen in the sterility control (see red box in Fig 2A), arguing that false positives were rare. Several new synthetic lethal interactions were apparent: *pal/ompA* and *pal/yfgL* on M9 plates (Fig. 2A & S3C), and *pal/surA* on LB plates (Fig. S3A, E*)*. Note that the *surA::Cm*^{R} clone used as recipient did not grow in minimal medium (Fig. 2A & S3C), so no inference can be made about this interaction from the M9 glycerol plates.

We quantified growth by obtaining an image of the plate and measuring pixel counts of the colony within a defined boundary. A key challenge is to distinguish differences that result from the growth properties of the double mutants from those arising from properties of the parental single mutants and from plate to plate variation. To do this, we used a dual normalization procedure (Fig. S4 & Methods) (Collins et al., 2006), which adjusts both for growth differences between plates (in-plate, horizontal normalization) and growth differences or mating defects of the recipient single mutants (strain or vertical normalization). The raw and normalized data are displayed in red and gray, respectively (Fig. 2B, S3B, S3D & S3F). The importance of strain normalization is demonstrated by its ability to filter out both the observed growth defect of single *pal* mutants in M9 and the known deficiency of *ompA* mutants to act as recipients in mating (Hoekstra and Havekes, 1979). Thus, the *pal/ycbS and ompA/ycbS* interactions that are scored as negative in the raw data were essentially neutral after normalization (Fig. S5). ~70% of the normalized scores exhibited acceptable standard deviations, i.e. within 25% of the plate mean (Fig. 2B, S3B, S3D & S3F). The larger error bars apparent in a few cases resulted from either pinning problems (*basR*/*pal* pair in LB1536; Fig. S3A&B) or from suppressors of synthetic sick and lethal interactions (see *ompA/pal* pair in LB384 and LB1536; Fig. S3A, B, E & F); such suppressors were never evident in more than 10% of the colonies pinned. We note that *ompC* (3.64MB away from *trp::oriT*; 78% chromosome coverage) displayed no noticeable effect on its efficiency to form double mutants.

To visualize all genetic interaction data simultaneously we generated heat maps. The colony-size scores that deviated the most from the median colony-size score of all double mutants after normalization within each 12 by 12 genetic matrix (Fig. 2C,D and S6) were denoted either as black (slower growing, negative genetic interactions) or red (faster growing, positive genetic interactions). These heat maps were comprised of data from both the 384 and 1536 plate formats as they gave almost identical data (Fig. S6, S7) as well as from reciprocally constructed double mutants, which also gave concordant data (Fig. S6). We independently verified the 8 most reproducibly negative interactions, ranging from synthetic sick to synthetic lethal, and 4 most reproducibly positive genetic interactions that were seen in both plate formats and in the reciprocally constructed double mutants using established methodologies (Table 1). Each of these interactions exhibited at least 20% impaired/ enhanced growth for both reciprocally constructed double mutants, Two negative interactions, *surA*/*degP* (Rizzitello et al., 2001) and *degP/yfgL* (Charlson et al., 2006), were previously described , and therefore provided validation for the efficacy of the method.. Inability to reproduce the synthetic lethal genetic interaction previously observed for *surA/yfgL* (Onufryk et al., 2005) was not due to failure of our methodology as it was also not reproduced by standard P1 transduction (data not shown), and may be due to strain/allele differences.

The 3 new synthetic lethal interactions (*pal/surA*, *pal/ompA*, *pal/yfgL*) were validated by examining co-transduction of a linked marker. In this technique, the P1 donor has a selectable marker closely linked to the first gene deletion being tested, and the recipient is either the wild-type strain or has the second gene deletion.

When two mutations are synthetically lethal, co-transduction of the linked markers from the donor strain is never observed when the recipient is deleted for the second gene because all events bringing in the first gene deletion are lethal (see also Table S1). New synthetic sick phenotypes (*degP/pal*, *cpxR/pal*, *ompA/yraP*) were recapitulated by reconstructing the double mutants with P1 transduction and then examining either colony size on plates or growth rate in liquid (see Fig. 2F-H & S8). Potential positive interactions (*pal/yraP*, *surA/ompA*, *degP/ompA, cpxR/ompA*) were reconstructed by P1 transduction and validated by monitoring growth on plates (Fig. S8) or by performing competition experiments in liquid (Lac^+ versions of the double mutants were competed with their parental Lac^- single mutants over several generations of growth in liquid medium). These verification experiments confirmed all positive and negative interactions (Table 1) except for the two negative interactions that resulted from linkage effects: *yraP/yraI* separated by only 7.7 kb (synthetic lethal), and *ompA/ycbS* separated by 17.2 kb (synthetic sick). This verified an expected limitation of our methodology: recombination between closely linked markers is less frequent than that of markers that are far apart (see also linkage section below and Fig 4C).

There were several differences between results in M9 glycerol and LB (Fig. 2C, D), some of which have been validated (Fig. 2 E-H), arguing that screening double mutants in a variety of conditions will be a worthwhile endeavor. Where tested, LB-specific effects appear to reflect growth rate rather than effects of the rich medium per se, as decreasing growth rate by lowering the temperature to 30°C partially restored growth of these double mutants in LB (see *degP-pal* in Fig. S8),

consistent with previous work indicating that slower growth partially compensates for outer membrane defects (Charlson et al., 2006; Wu et al., 2005).

Using GIANT-coli at the genomic level

A critical prerequisite for extending GIANT-coli to larger datasets was to be able to rapidly convert single-gene deletion mutants into Hfr donor strains. To accomplish this, we created a "double male" strain (Fig. 3A), which transfers from two origins: 1) an upstream *oriT* in a pseudogene linked to a tetracycline resistance marker (Tet^R; see Methods); and 2) from the pseudo-Hfr (Amp^R or Gen^R). When transfer initiates from the upstream *oriT*, the downstream pseudo-Hfr and its *oriT* are transferred efficiently, as assessed by acquisition of the adjacent antibiotic resistance marker (Amp^R or Gen^R). We mated the double male with the entire Keio collection arrayed in 384 format and selected Hfr versions of the Keio collection on Amp/Kan or Gen/Kan plates. The same methodology was successfully used to generate Hfr derivatives of several Cm^R-marked ASKA deletion mutants (data not shown).

An alternative approach employed a Chromosomal Integration Plasmid (CIP) to deliver the F-transfer region to various chromosomal locations. This plasmid is replicated from the Pi dependent R6K g-ori and has the F transfer region, a ~300 bp of chromosomal homology, and a streptomycin-spectinomycin cassette for selection. CIPs are carried in a *pir*$^+$ *recA*$^-$ host where they replicate as plasmids but do not integrate. Upon transfer to a *pir*$^-$, *recA*$^+$ F$^-$ strain, the plasmid integrates into the chromosome by homologous recombination (Fig. 3B & Methods for more details),

simultaneously converting the F^- strains into recombinant Hfr donors in a high-throughput manner (Takeuchi *et al*., unpublished data).

The use of a much larger recipient gene-set means that each strain is present only a few times on the plate. Therefore, we tested whether we could recapitulate results from our 12x12 matrix by crossing several recombinant Hfr Cm^R donors against the entire Kan^R Keio collection, both in rich and minimal media. To our satisfaction, important quality controls were met: i) the self-mating control was always one of the most negative interactions on the plate (for an example see Fig. 4A); ii) we could recapitulate a majority of the genetic interactions discovered in our 12x12 matrix (e.g., *pal/ompA* produced a synthetic lethal interaction and *pal/degP* a neutral interaction on minimal medium; Fig. 4A). Apart from those few previously validated interactions, many new ones were identified and verified by reconstructing the double mutants with P1 transduction and quantitative assessment and comparison of their growth with that of the parental single mutants on agar surfaces (see Fig. 4B and Table S2 for the results of the Pal screen on M9). We note that double mutants with a mucoid/slimy phenotype gave false readings in our screen (6 cases in the Pal screen on M9), since their colony size did not reflect their actual fitness. We are currently developing a program that will identify such colonies, exclude them from our initial analysis and assign them a score reflective of extent of mucoidy.

Crossing several donors against the entire Kan^R single-gene Keio collection gave us sufficient data to make an initial assessment of the linkage cut-off of our methodology. To define linkage biases, we plotted interaction scores from 9x3985

(LB) and 14x3985 (M9) crosses as a function of the distance between genes in kilobases (kb) (Fig. 4C). Gene-pairs separated by less than 60 kb in M9 and 30 kb in LB displayed slight to severe negative interactions that results from a decrease in recombination efficiency. The lower cut-off in LB may reflect greater recombination resulting from the multiple genomes present under those conditions. As many functionally related genes are arranged in operons in bacteria, it is important to determine double mutant phenotypes of genes that are in close proximity to each on the chromosome. To do this, we will normalize the interaction scores of closely linked genes to a standard linkage curve generated from the hundred thousand gene-pairs tested (similar to Fig. 4C). This will allow us to distinguish whether the observed colony sizes differ significantly from the size predicted from the decreased recombination efficiency of closely linked genes.

Interestingly, the moderately separated *ompA*/*ycbS* pair (17.2 kb) gave more recombinants when the donor transferred the gene of the pair that is closer to *oriT* rather than the converse (i.e. more recombinants when the donor transferred *ycbS* than *ompA;* Fig. S6). At present, we lack sufficient data to determine whether this asymmetry is a general rule.

The methodology described in this report can be modified to identify suppressors of conditionally lethal mutants. This new application can be broadly used as it can be accomplished with hand pinning the Keio collection, does not rely on quantitative analysis of large datasets, requires very few plates and can be completed in less than one week. Here, the conditional lethal mutant is used as a donor strain and mated to the entire library of deletion mutants with the protocol

described in Figure 1, except that the final double antibiotic selection plates contains an inhibitory compound or is incubated in the condition that unmasks the lethality of the donor. As a proof of principle experiment, we mated the conditional lethal *yraP::Cm*^{R} strain with the Keio collection, selecting double recombinants able to grow in 3% SDS, the condition causing *yraP*$^{-}$ lethality (Onufryk et al., 2005) (Fig. 4D and Table S3). The same approach can be used to find suppressors of an essential gene. In this case the donor stain has an inactivated copy of the essential gene in the chromosome and a functional copy of this gene on a non-mobilizable plasmid that cannot be transferred to the recipient; double recombinants that grow identify suppressors that compensate for loss of the essential gene function.

Conclusions and Perspectives

We have developed and validated a method for high-throughput generation of double mutant strains in *E. coli*. This method can be used for large-scale quantitative analyses of genetic interactions and is compatible with both the 384 and 1536 high-density arrays. In addition, we have developed facile methods to convert F^{-} female (recipient) strains into Hfr males (donors) *en masse* and have devised a simple variant of our genetic interaction methodology that identifies single-gene knockouts able to suppress deletions that cause a conditional lethal phenotype. These methods will permit rapid exploration of the genetic interaction landscape in *E. coli*.

The small data set that we have generated illustrates the great potential of this method to provide new information: quantitative genetic interaction analysis of

only 66 reciprocal double mutants revealed 10 significant synthetic interactions (out of which 8 are new), ranging from lethal to suppressive growth phenotypes. Genome-wide screens provided many additional candidates for genetic interactions. These results are in strong accord with the view that the paucity of documented genetic interactions in *E. coli* as compared to *S. cerevisiae* reflects differences in methodologies employed to search for such interactions, rather than fundamental differences in the genetic interaction networks of the these organisms.

Our initial findings only hint at the rich biology that remains to be uncovered by systematic exploration of synthetic relationships. Some of the interactions can be rationalized, such as the three positive interactions identified for the *ompA* mutant. SurA and to a lesser degree DegP are required for proper OmpA folding (Lazar and Kolter, 1996; Rouviere and Gross, 1996; Sklar et al., 2007), while CpxR activates expression of DegP and DsbA (Danese et al., 1995; Pogliano et al., 1997), which are both required for proper OmpA folding (Bardwell et al., 1991; Sklar et al., 2007). In all three backgrounds (*cpxR*$^-$, *surA*$^-$ and *degP*$^-$) removing OmpA could decrease envelope stress caused by accumulation of misfolded OmpA. Likewise, some suppressors of the *yraP* mutant are also suggestive of compensation. Skp, a periplasmic chaperone, and DsbA and DsbB, which orchestrate disulfide-bond formation in the periplasm, may suppress because misfolding of a downstream target of these chaperones buffers the envelope defect of *yraP*$^-$ cells, possibly by relieving an imbalance in membrane composition. As no known DsbA/DsbB and Skp substrates (Leichert and Jakob, 2004; Qu et al., 2007) were identified as suppressors, an essential outer membrane protein targeted by those chaperones

(e.g. Imp or YaeT), not present in our query set, may be responsible for this phenotype.

In contrast, the importance of Pal was unanticipated because its precise function(s) is unknown. Pal is an outer-membrane (OM)-associated lipoprotein that tethers the OM to the peptidoglycan (PG) (Parsons et al., 2006), an interaction believed to stabilize the OM. Pal is also the OM anchor of the 5-protein Tol-Pal system that bridges the inner and outer membranes (Cascales and Lloubes, 2004) and is energized by proton motive force (Cascales et al., 2000; Cascales et al., 2001; Germon et al., 2001). This protein complex is highly conserved among gram-negative bacteria and has recently been implicated in ensuring proper OM invagination during division (Gerding et al., 2007). Six newly identified synthetic interactions in our 12X12 matrix involve *pal*, several in a growth-conditional manner (*pal/ompA*; *pal/yfgL; pal/surA*; *pal/degP*; *pal/cpxR; pal/yraP*). Furthermore the fact that our genome-wide screen with the *pal* mutant identifies many additional genetic interactions including the entire pathways for lipid-A core and ADP-L-glycero-b-manno-heptose (a precursor of the inner core lipopolysaccharide-LPS) biosynthesis, the biosynthetic pathway for the enterobacterial common antigen and several additional envelope proteins of known and unknown function (Table S2) suggests the possibility that the Tol/Pal system may be one of the central organizers of envelope functionality. The abundance of interactions identified for Pal illustrates the degree of information GIANT-coli can generate and gives a taste of its potential to provide insights into the function and integration of different cellular processes, when coherent data are collected and analyzed.

The addition of GIANT-coli to the genetic toolbox of *E. coli* ensures that systematic genetic interaction data can be rapidly accumulated and then interfaced with information from studies of individual pathways and regulatory systems as well as with large datasets including global phenotypic screens (Bochner, 2003; Inoue et al., 2007; Ito et al., 2005; Niba et al., 2007), protein-protein interaction data (Arifuzzaman et al., 2006; Butland et al., 2005), protein-DNA interaction studies (Chen et al., 2008; Wade et al., 2007), gene-expression profiling (Gene Expression Omnibus) and other forward genetics screens (Girgis et al., 2007; Mazurkiewicz et al., 2006). It will be relatively straightforward to adapt this methodology to give additional readouts other than growth such as promoter activity (e.g., by use of *gfp* or *lacZ* fusions), biofilm formation (Congo-red or calcofluor plates, crystal violet absorption), siderophore production (chrome azurol S plates), and growth inhibition (halo assays). Moreover, our goal is to extend GIANT-coli to other gram-negative and -positive bacteria. Hfr's have been successfully used in *Salmonella* and *Shigella*; slight variations of GIANT-coli can be employed in cases where Hfr's are not applicable, e.g. the marked gene deletion can be carried on a conjugative plasmid. We also want to develop a similar technology for naturally competent organisms (e.g. *B. subtilis*, *S. pneumoniae*, *N. gonorrhoeae*); preliminary experiments verify that DNA uptake can be used as a high-throughput approach to generate double mutants *en masse* in such organisms (M. Winkler, personal communication). Together, these methodologies should lead to rapid progress in discovery of gene function and network connections in the bacterial kingdom.

Methods

Strains, plasmids and growth conditions

All single mutants used in this study are part of the *E. coli* single-gene Keio collection (Baba et al., 2006) or the newly developed ASKA single-gene deletion library (Yamamoto *et al*., unpublished data) that were constructed in BW25113 and BW38029, respectively, which are closely related to MG1655 (Table S4A). The pseudo-Hfr strain carries an ~43 kb HindIII fragment of F, which includes the entire transfer region, integrated into the *trp* locus and linked to ampicillin (Francois et al., 1990). To generate an Hfr donor isogenic with the *E. coli* single-gene deletion libraries, the "pseudo-Hfr" was transduced with P1 phage into BW25113 and BW38029. A transductant was purified and retransduced to BW25113 and BW38029 to minimize DNA heterogeneity in the vicinity of *trp*. Additionally, the ampicillin marker was changed to gentamycin.

To construct a double male strain, a ~300-bp *oriT* fragment and *tetAR* (Tet^R) from pCP16 (Cherepanov and Wackernagel, 1995) were cloned next to each other in pBR322 (using Pst I/EcoR I and Sma I/ Hind III respectively); the *oriT*-*tetAR* region was amplified and recombined into two pseudogenes, *intQ* and *ycdN* (Grosse et al., 2006), using recombineering (Datsenko and Wanner, 2000; Sawitzke et al., 2007).

pNTM3 was made by engineering F. First, a 41.7 kb derivative carrying *oriT*, all transfer functions, and *oriR* was made by replacing a ~60 kb fragment of F with a gentamicin resistance gene from pAH143 using the phage l Red helper plasmid pKD46 (Datsenko and Wanner, 2000). A ~300 bp region of *rhaM* was cloned into pAH144 (Haldimann and Wanner, 2001) between oriRg (the Pi-dependent

replication origin from plasmid R6K) and *aadA*, conferring streptomycin and spectinomycin resistance, using Not I. The *oriR*g-*rhaM*-*aadA* region was then PCR-amplified and used to replace *oriR* and nearby gentamicin resistance using the phage l Red helper plasmid pKD46 (Datsenko and Wanner, 2000). All primers and new strains are listed in Table S4A&B.

Cells were routinely grown in LB medium or M9-minimal mediium supplemented with 0.4% glycerol and, when necessary, 50 μg/ml L-tryptophan at 37°C; L-tryptophan was omitted from the media when the strain was not a pseudo-Hfr derivative. Antibiotics were added when necessary in the following concentrations; 30 μg/ml kanamycin; 20 μg/ml or 30 μg/ml chloramphenicol (in LB and M9-glycerol respectively (Le Grice and Matzura, 1981)) 10 μg/ml tetracycline; 15 μg/ml gentamicin; 100 μg/ml ampicillin; 30 μg/ml streptomycin; 50 μg/ml spectinomycin.

Array manipulations

We arrayed colonies on LB or M9-glycerol with agar in rectangular plates (Singer Instruments). Single-mutant libraries were kept as glycerol stocks in 384 micro-titer plates and were always thawed and transferred to agar plates using the Singer RoToR HDA station. Arrays were also replicated using the Singer RoToR HDA station. Incubation times between different steps were different and are indicated in the detailed protocol in Fig S2.

Analysis of the data from the 12 by 12 genetic cross

High resolution images of each plate were obtained using a 6 megapixel camera under controlled lighting conditions[29]. Image contrast was enhanced through a batch processing Levels Adjustment in Adobe Photoshop (Composite channel, Input 30, 200, gamma 0.35). Raw colony size data for each location on the plate array was obtained by automated image analysis software, HT Colony Grid Analyzer (http://sourceforge.net/project/showfiles.php?group_id=163953).

Since colonies on the periphery of the array were oversized relative to interior colonies, data from the 2 outer rows around the 1536 plates, and the outermost row around the 384 plates were discarded. Data was processed manually for the 12 x 12 experiment using Excel spreadsheets. The four major conditions (M9-384, M9-1536, LB-384 & LB-1536) were processed as separate groups using the following protocol.

First, we normalized individual plates to permit facile comparison between plates. The average size of all colonies on the plate larger than a threshold level (taken as10 pixels for 1536 plates and 75 pixels for 384 plates) was scaled to be equal to 500. Colony sizes below the threshold level were averaged; each "empty" or "trace" spot received this average value rather than the measured size of the empty spot. Normalized sizes for each plate were further adjusted by subtracting the average size of colonies in the sterility control zone (n=4 for 384; n=16 for 1536). Second, we normalized for the growth of each strain across all matings. Here, common recipient strains were grouped together (as opposed to the grouping by donor strain used in the plate normalization), and the median colony size of the

recipient was scaled to be equal to 500. Self-mating cases where the donor and recipient strains were the same were excluded from this adjustment. For each donor-recipient mating these doubly normalized colony sizes were averaged (and standard deviations calculated) to provide the representative values plotted in Figures 2, S3, S5, S6, S7.

Analysis of genome-wide screens for extraction of linkage data

14 genome-wide screens were performed in M9-glycerol (*cpxR*, *degP*, *degQ*, *djlA*, *hlpA*, *lpp*, *nlpB*, *ompC*, *pal*, *rcsD*, *smpA, ybaY, ycbS, yfeY*) and 9 in LB (*cpxA*, *cpxR*, *degP*, *djlA*, *fkpA*, *hlpA*, *ompA*, *pal*, *smpA*) and analyzed for linkage using a new iteration of the publicly available software used for E-MAPs (Collins et al., 2006). The new software is adapted for 1536 arrays (S Collins & JS Weismann; unpublished data). Images of *E. coli* arrays were captured and subjected to artificial background darkening as described above, prior to being fed to the new analysis software. Note here that "S interaction scores" are calculated for each double mutant by the analysis software (Collins et al., 2006).

Acknowledgements

We wish to thank Sueyoung Lee, Angela Wong and Jeanyoung Lee for excellent technical assistance. We also thank C. J. Ingles, P.J. Kiley, C. Squires, S. A. Johnson, A. Hochschild & T. J. Silhavy for critically reading this manuscript and offering useful suggestions. This work was supported by Sandler Family Funding to CAG and to NJK, NIH GM036278 to CAG, NIH GM62662 to BLW, CREST-JST, Grant-in-

Aid for Scientific Research & Grant-in-Aid for Scientific Research on Priority Areas to HM. AT is a recipient of an EMBO fellowship.

Figures

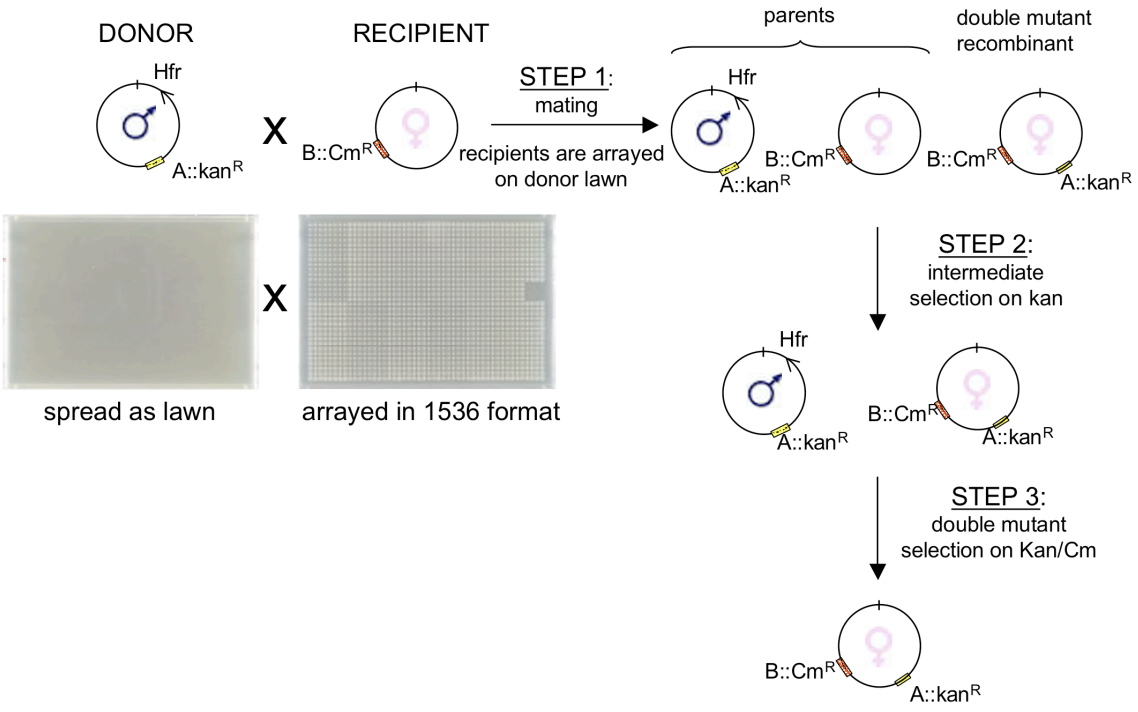


Figure 1. A flowchart depicting the different steps used in GIANT-coli. An Hfr donor (male) strain carrying a selectable marker (Kan^R) replacing an ORF is mated on agar plates with arrayed F^- recipients (females; 1536/plate) carrying a different selectable marker (Cm^R) replacing another ORF (Step1). Following mating, cells are subjected to an intermediate selection on the bacteriocidal antibiotic Kan (Step 2) and then to a final selection for double mutants using both antibiotics (Step 3). Images of two representative plates used for generating a mating plate are shown below the cartoon.

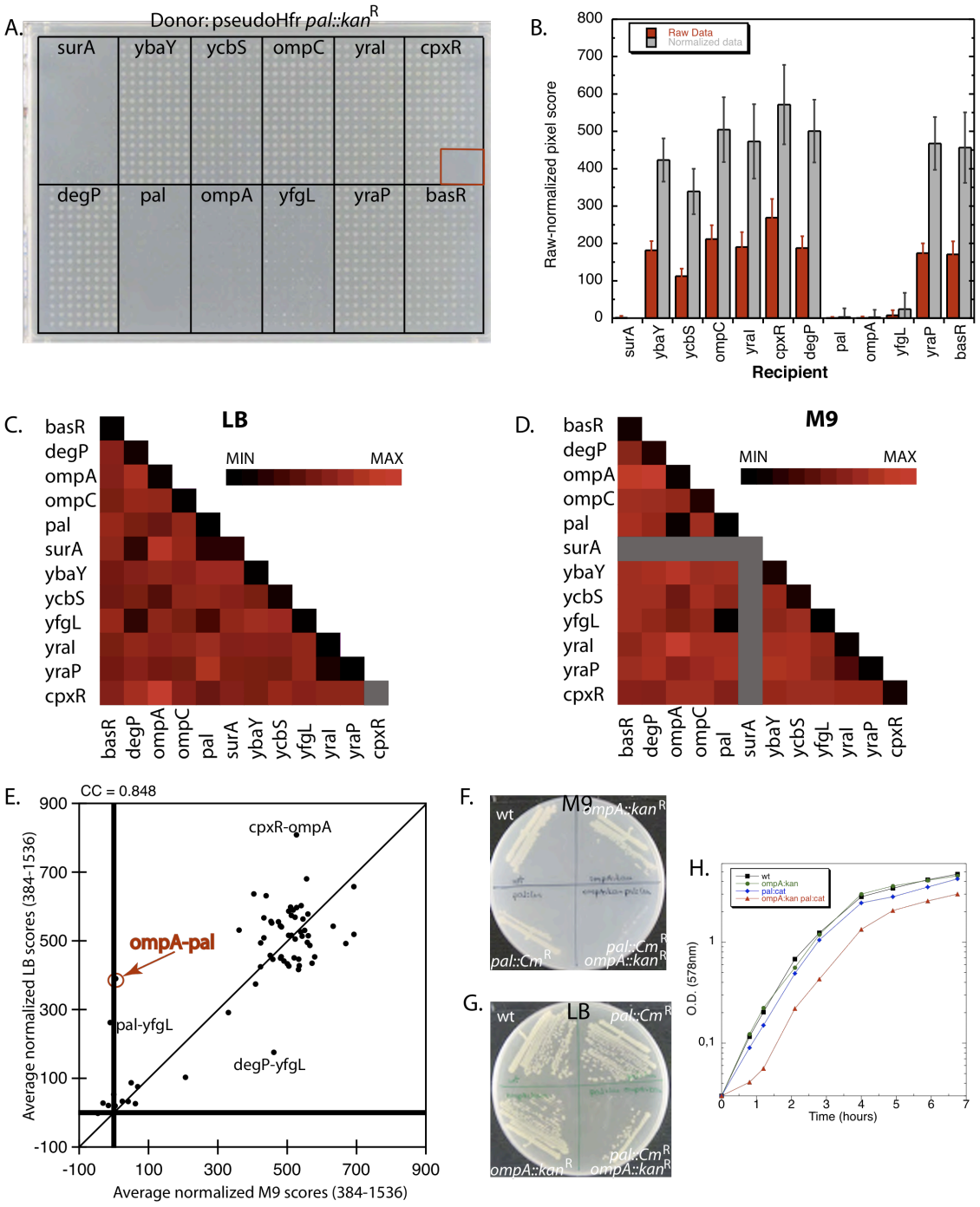


Figure 2. A 12x12 genetic interaction matrix to validate GIANT-coli. **A.** A representative 1536-format, M9-glycerol plate showing the double mutants resulting from crossing pseudo-Hfr *pal*::*kan*^{R} with 12 Cm^R-marked ASKA recipients arrayed in boxes of 16x8=128 replicas. The red box is a sterility control, since no

recipients are arrayed in this spot. **B.** Quantification of (A). Error bars depict standard deviations. **C-D.** Heat maps representing 12x12 crosses in LB (**C)** and M9-glycerol **(D)** based on the combined data from the 384 and 1536 plate formats and averaged results of reciprocal genetic interactions. The gray lines indicate that no results were extracted in M9-glycerol from *surA*::*kan*^{R} and *surA*::*Cm*^{R}, as these clones grew very poorly in this medium. The color-coded bar ranges from a minimum size score (noted as MIN) to a maximum (noted as MAX) calculated for each dataset separately. **E.** Scatter plot of averaged normalized colony-size scores comparing growth in M9-glycerol vs LB for the 65 pairs (of 78 total) that grew in both media. Double mutants with significantly different growth in the two media are identified by name. The differential phenotype of the *palompA* double mutant (red) is further analyzed in F-H. **F-H.** The *palomp*A double mutant was reconstructed by P1 transduction and the conditional interaction identified by GIANT-coli was recapitulated by lethality on M9-glycerol plates **(F),** smaller colony size in LB plates (**G),** and longer lag-phase and slightly slower growth rate in LB medium (**H**; doubling times of wild-type, *ompA::Cm*^{R}, *pal::Kan*^{R} and *ompA::Cm*^{R} *pal::Kan*^{R} are approximately 28.5', 30', 31.5' and 38', respectively).

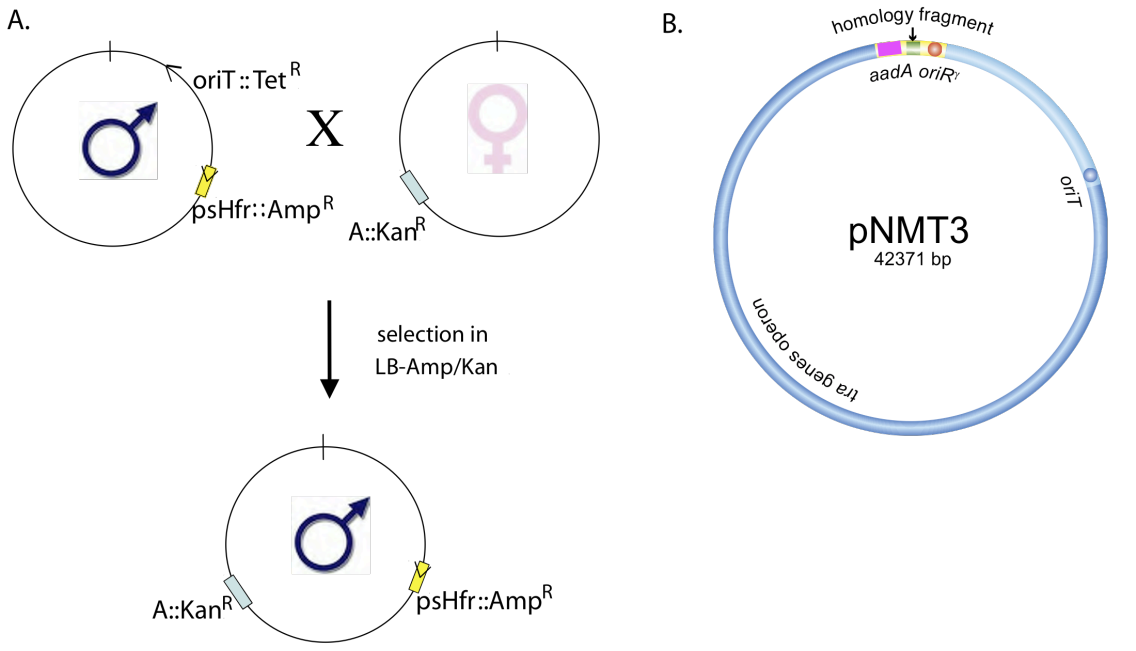


Figure 3. A toolkit that facilitates the use of GIANT-coli in genome-wide analyses. **A.** High-throughput conversion of an entire single-gene knockout F- library to an Hfr donor library. A double male strain is crossed with the Keio deletion library and a male (Hfr) Keio library is isolated by selecting on Amp/Kan or Gen/Kan plates (the pseudo-Hfr locus is linked to Gen^R or Amp^R). The entire process is carried out on agar plates. Transfer capabilities of a number of the newly generated pseudo-Hfr's have been validated. **B.** Targeted integration of F-transfer functions at different chromosomal loci. Conditionally-replicating CIP vectors contain *oriR*g (red circle), a ~40 kb BamH1 fragment of F (blue) including its 33 kb transfer region (dark blue) and *oriT* (blue circle), *aadA*, conferring streptomycin and spectinomycin resistance (violet) and ~300 bp of chromosomal homology (green). For pNTM3, the chromosomal region is from *rhaM*. CIPs are carried in a strain expressing the P protein, which allows the plasmid to replicate from *oriR*g. Upon mating to the F- Keio

or ASKA deletion mutants, which lack P, CIPs are unable to replicate; selection for streptomycin/spectinomycin resistance results in chromosomal integration dictated by the particular homology region present on the CIP.

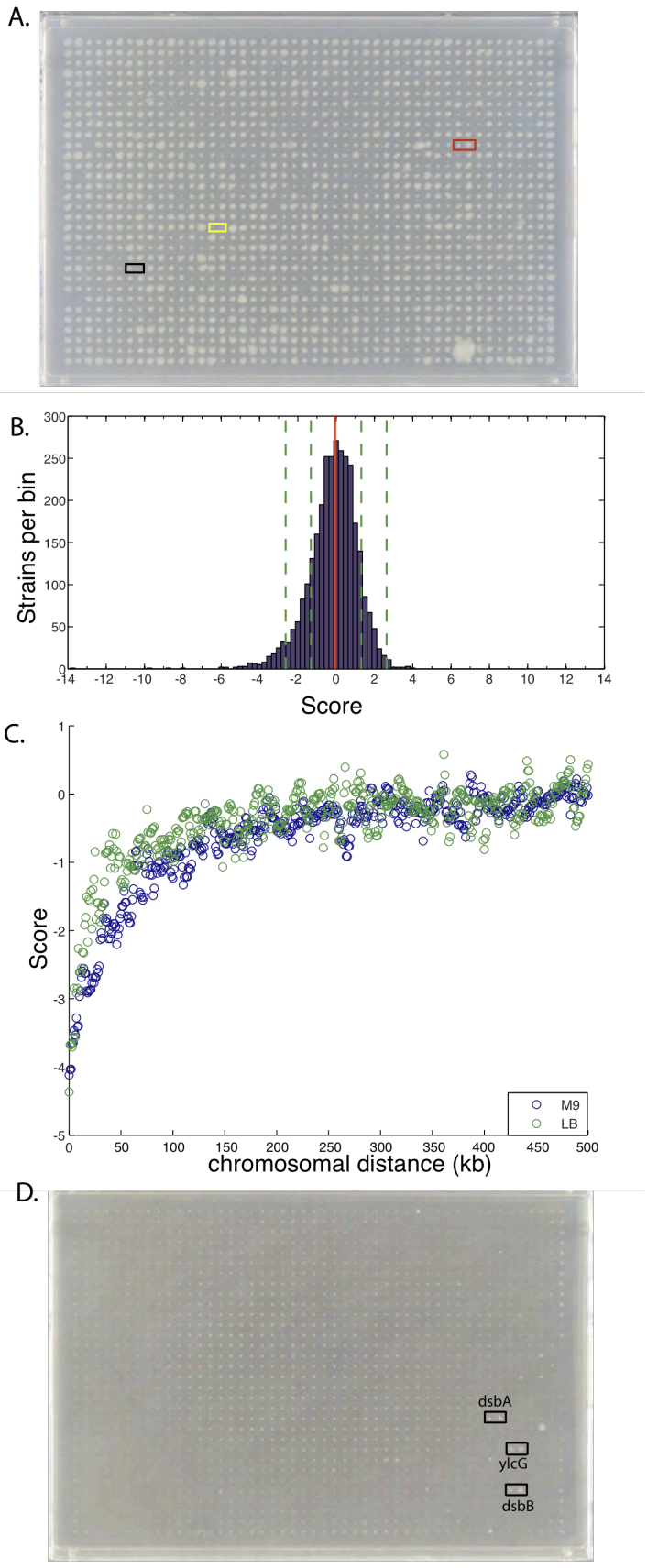


Figure 4. Genome-wide screens using GIANT-coli. **A, B.** Cross of pseudo-Hfr *pal*::*Cm*^{R} ASKA mutant with the entire F^{-} Kan^{R} Keio collection (3985 mutants) arrayed in 1536 format (1536 colonies per plate); each Keio mutant is present twice as adjacent duplicates (768 unique recipients/plate). A representative image of one M9 plate (out of six total) is shown in panel (**A**). Interactions identified in the 12x12 matrix recapitulated here are marked with differently colored boxes. Black: the self-mating pair, *pal::Cm*^{R} *pal::Kan*^{R}; yellow: the synthetic lethal pair, *pal::Cm*^{R} *ompA::Kan*^{R}; red: *pal::Cm*^{R} *degP::Kan*^{R}, a neutral

interaction in minimal media. The distribution of all interaction scores of *pal::Cm*^{R} with the Keio collection is shown as a histogram in **(B)**. In the histogram the number of Keio mutants within a bin of 0.25 are plotted against the interaction scores; the mean of the distribution (red dotted line) and 1 and 2 standard deviations (green dotted lines) are shown.. **C.** Linkage biases of Hfr mating in LB (green) and M9 (blue). Median interaction scores, extracted using a variation of the E-MAP analysis software (Collins et al., 2006), are aggregated into 5kb bins sliding with 1kb steps and are plotted as a function of chromosomal distance in kbs between genes. The analysis is based on 14 genome-wide screens in M9 glycerol and 9 in LB (data not shown). **D.** Suppression analysis of *yraP*::*Cm*^{R} lethality in 3% SDS. Pseudo-Hfr *yraP*::*Cm*^{R} was mated with the Keio collection as in (A above) except that double mutants were selected on both antibiotics in the presence of 3% SDS. A representative image of one plate is shown; complete data are shown in Table S3.

Tables

Table 1. Reproduction of synthetic interactions detected in the 12X12 genetic interaction experiment. **A**: colony size; **B**: growth curves in liquid media; **C**: competition experiment; double mutant versus single mutants in liquid media; **D**: co-transduction of a linked marker (see also text and Table S1); **lit**: reported earlier in literature. Negative and positive interactions are indicated in black and blue fonts respectively. *yraP/yraI* and *ompA/ycbS* were both recorded as synthetic sick-lethal in mating experiments, but this was due to linkage. Nd stands for not-determined.

Pairs	LB		M9	
	interaction	**verification**	**interaction**	**verification**
degP/surA	lethal	D, lit	nd	-
pal/surA	lethal	D	nd	-
pal/yfgL	lethal	D	lethal	D
pal/ompA	sick	A, B (30 &37°C)	lethal	A, D
degP/yfgL	sick	A, B (mostly at 37°C), lit	-	-
degP/pal	slightly sick	A, B (both only at 37°C)	-	-
cpxR/pal	sick*	A, B	slightly sick	A
ompA/yraP	slightly sick	-	sick	A
pal/yraP	positive	A	-	A
ompA/degP	slightly positive	-	positive	A
ompA/surA	positive	(A), C	nd	-
cpxR/ompA	positive*	A	slightly positive	A

* only one way available; *cpxR*$^{-}$ is very poor recipient in LB and therefore that data were not evaluated

Supplementary Figures

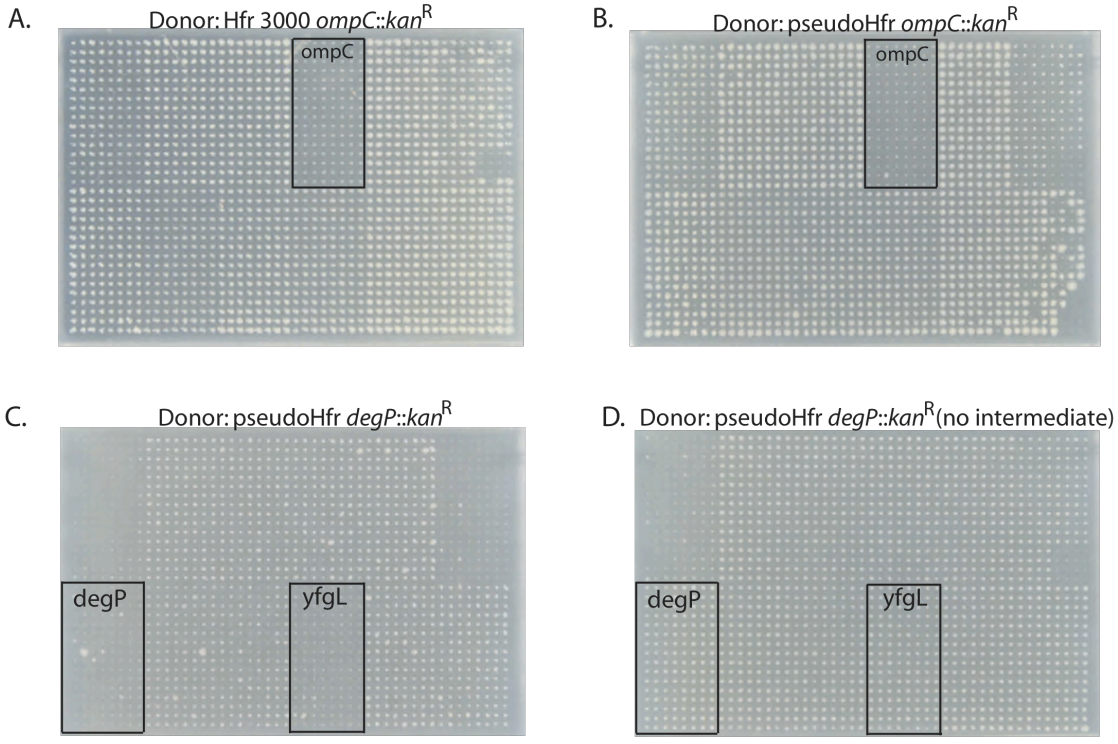


Figure S1. Optimization of GIANT-coli. **A-B.** Comparison of mating efficiencies of pseudo-Hfr and Hfr 3000 using our high-throughput mating technology. A representative mating of Hfr 3000 *ompC*::*kan*^{R} (**A**) or pseudo-Hfr *ompC*::*kan*^{R} (**B**) with 12 Cm-marked recipients arrayed in box format. Note that colony sizes, an indication of robustness of transfer and mating, are roughly comparable, indicating equivalent performance. **C-D.** Comparison of a representative mating of pseudo-Hfr *degP*::*kan*^{R} with (**C**) and without (**D**) an intermediate step. The intermediate selection step increases the data quality by eliminating recovery of false positives (*degP::kan*^{R} *degP::Cm*^{R}) and facilitating detection of synthetic interactions (*degP::kan*^{R} *yfgL::Cm*^{R}).

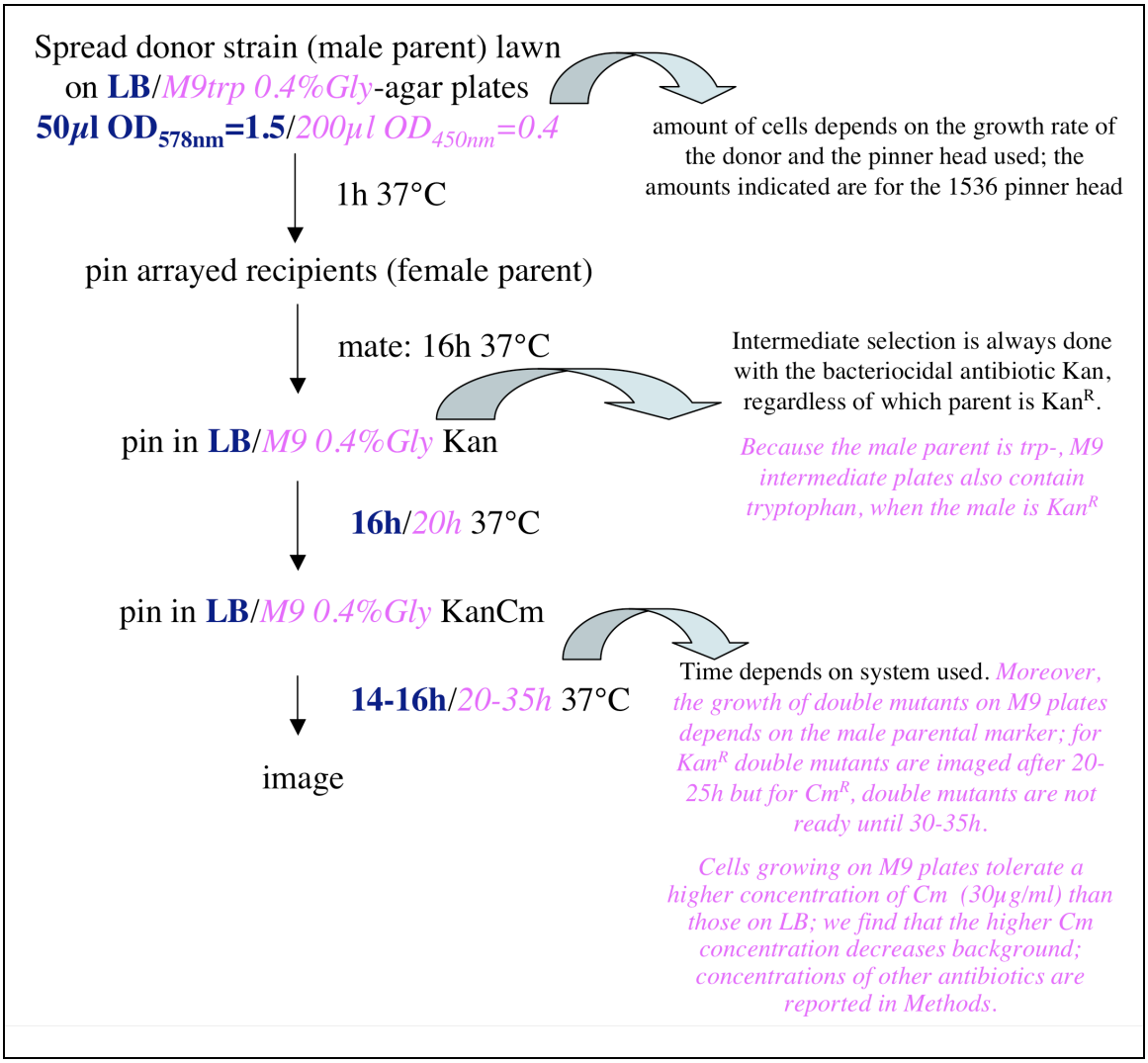


Figure S2. Flow diagram of the detailed GIANT-coli protocol. Bold blue and magenta italic fonts represent specific steps of the rich and minimal media protocols respectively. Notes for each step are on the right of the flow diagram.

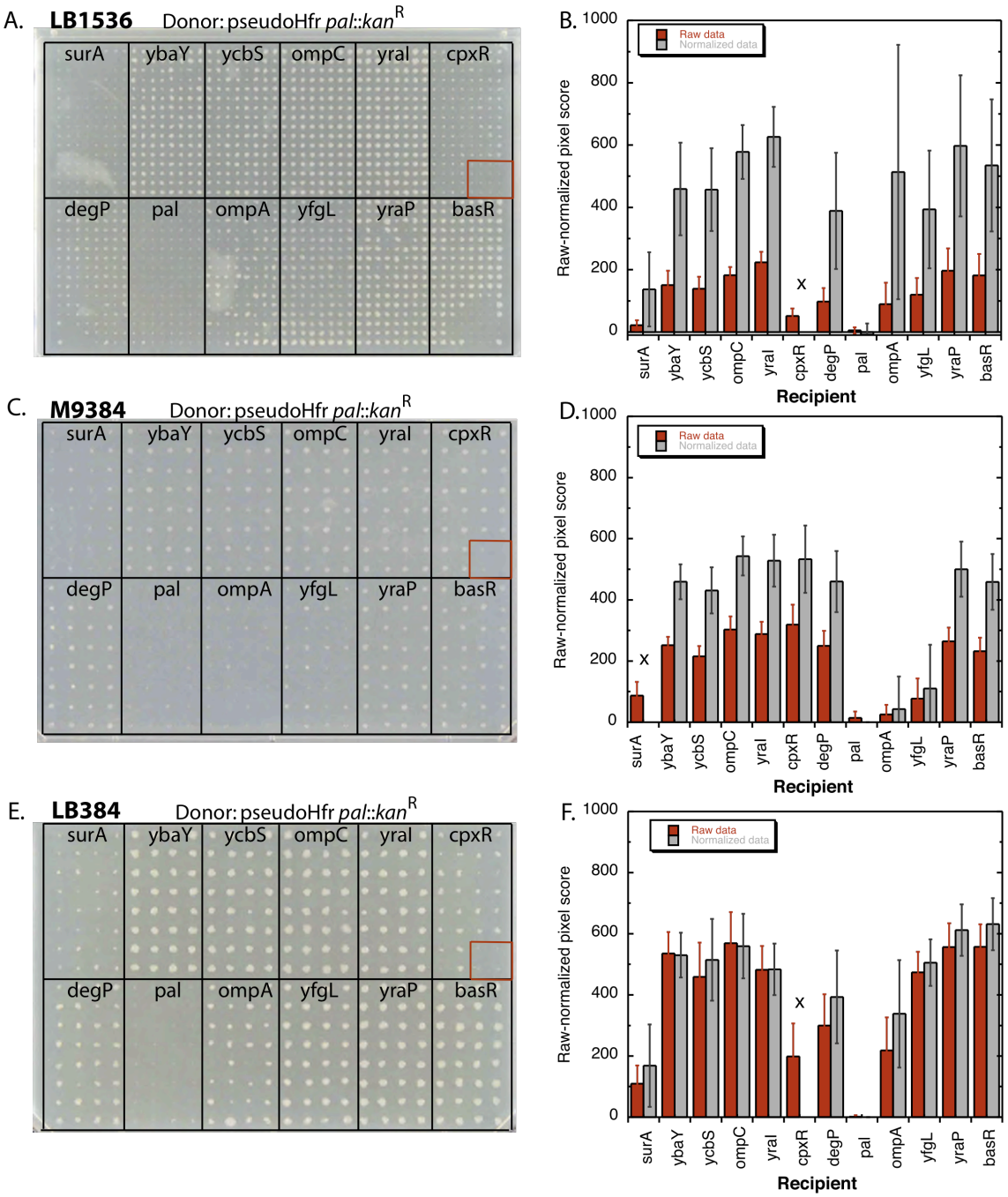


Figure S3. Comparison of mating of pseudo-Hfr *pal*::*kan*^{R} with 12 Cm-marked recipients arrayed in the 1536-platform on LB agar plates (**A**), with those arrayed in the 384-platform on M9-glycerol (**C**) or LB agar plates (**E**). The red box is a sterility

control as no recipients are arrayed in this spot. Quantification of each plate is shown in the panel on the right of each plot (**B**, **D**, **F**).

Error bars depict standard deviations. Bars marked with an x were not be taken into account, since *surA*::*Cm*^{R} grew very poorly in M9-glycerol and *cpxA*::*Cm*^{R} mated very poorly in LB.

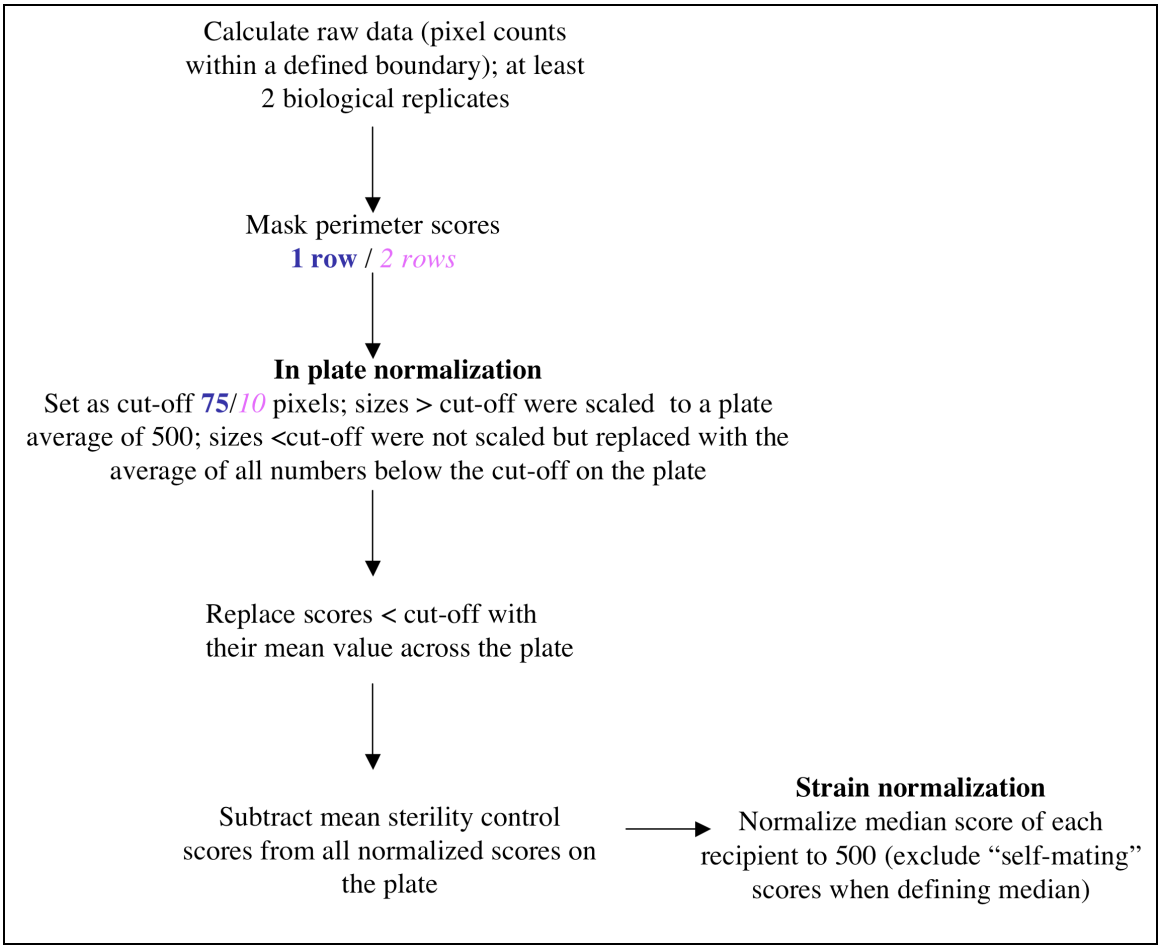


Figure S4. Flow diagram of the normalization procedure employed for the analysis of the 12x12 genetic matrix. Discrepancies between the normalizations applied to the two different array platforms used in this study are color-coded (bold blue for 384 and magenta italics for 1536) and include: i) masking the outer two rows of the higher-density arrayed plates (1536) rather than only one in the 384 platform (colonies in the outer rows grow significantly more than colonies in the rest of the plate) and ii) a lower pixel cut-off for the 1536 array (10 vs 75), which reflects the smaller average colony size in the higher-density arrayed plates.

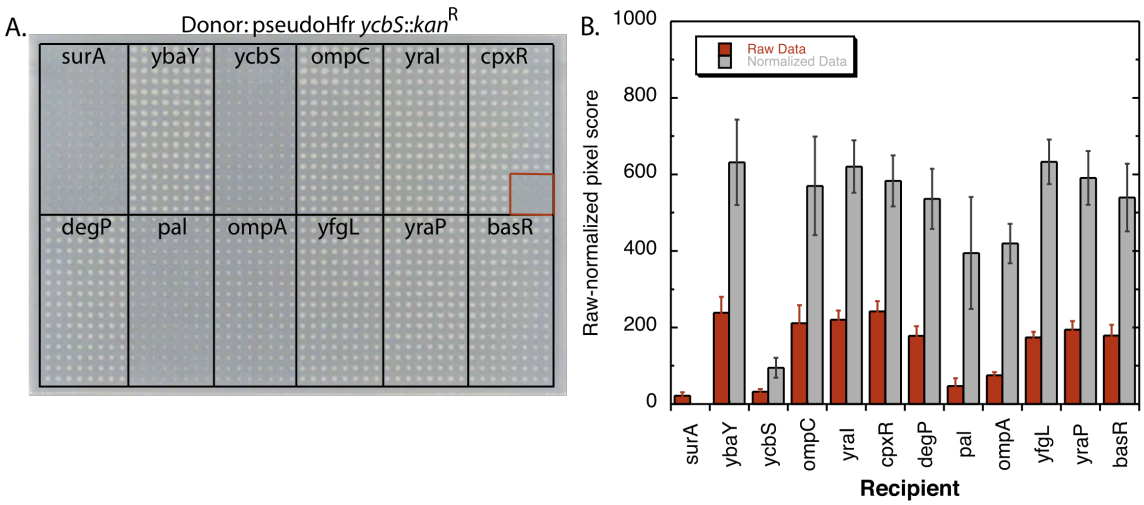


Figure S5. Strain (vertical) normalization adjusts for growth deficiencies or mating defects of the recipient parents. Mating of pseudo-Hfr *ycbS*::*kan*^{R} with 12 Cm^{R}-marked recipients arrayed in boxes using the 1536-platform on M9-gylcerol (**A**) and its quantification (**B**). *pal*::*Cm*^{R}, which is a slow grower, and *ompA*::*Cm*^{R}, which is impaired in mating, both generate low colony-size scores when combined with *ycbS*::*kan*^{R}. Strain normalization takes into account the deficiencies of the parental single-mutants and corrects the colony-size scores, which are in the same range as the rest of the neutral gene-combination after normalization. In sharp contrast, the self-mating control, *ycbS*::*kan*^{R} *ycbS*::*Cm*^{R}, retains its low colony-size score after normalization.

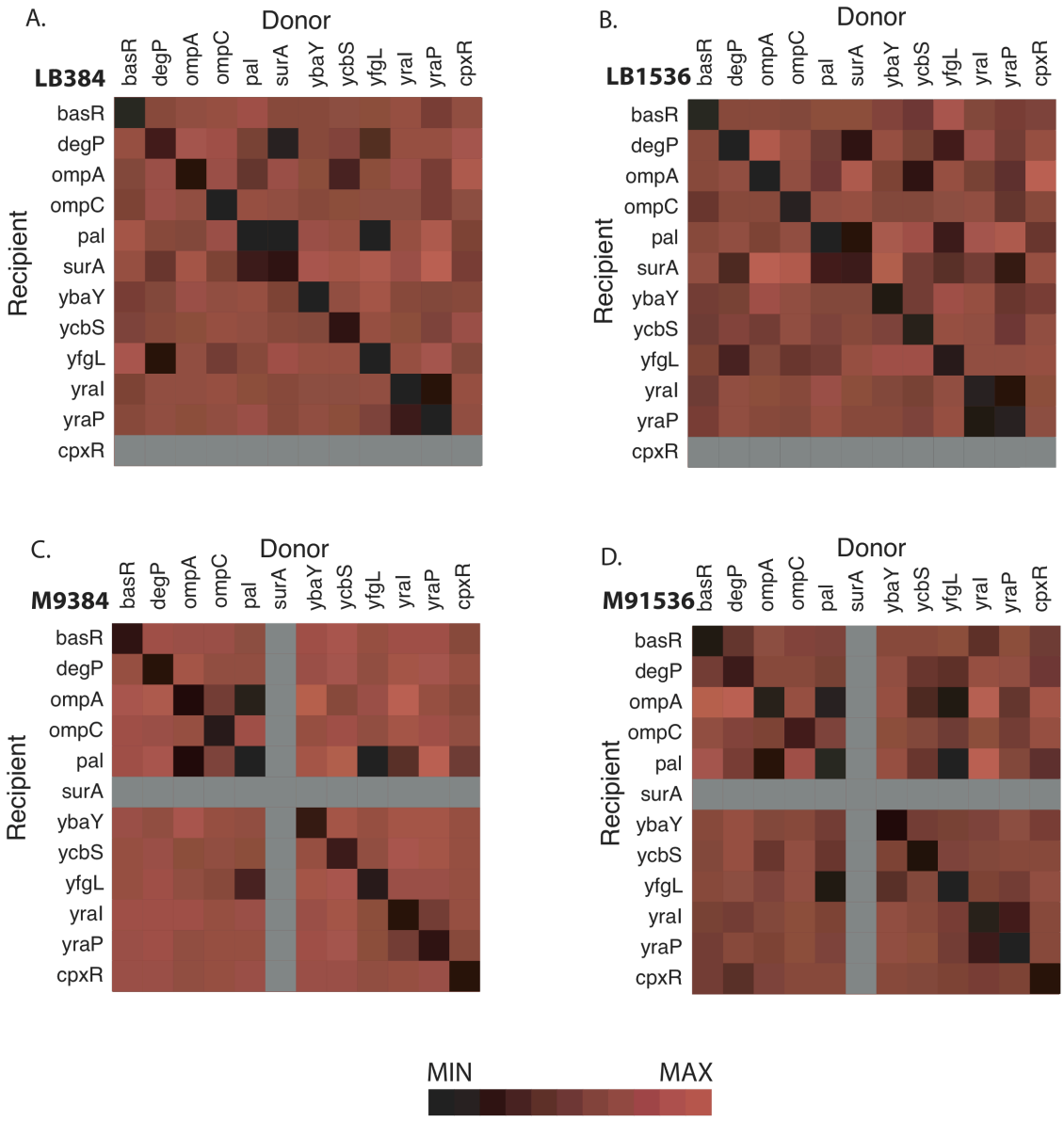


Figure S6. Heat maps representing 12x12 crosses in all four different datasets: **A**. LB-384 **B**. LB-1536; **C**. M9-384; **D**. M9-1536. In these heat maps, reciprocal genetic interactions are not averaged and are located in symmetrical positions relative to the diagonal. The gray lines refer to data that were not evaluated because of strain problems: *cpxR::Cm*^{R} is a very poor recipient in LB and *surA*::*kan*^{R} and *surA*::*Cm*^{R} clones used in this study grow very poorly in M9-glycerol. The color-coded bar

ranges from the minimum interaction score (noted as MIN) to the maximum interaction score (noted as MAX) for each dataset separately.

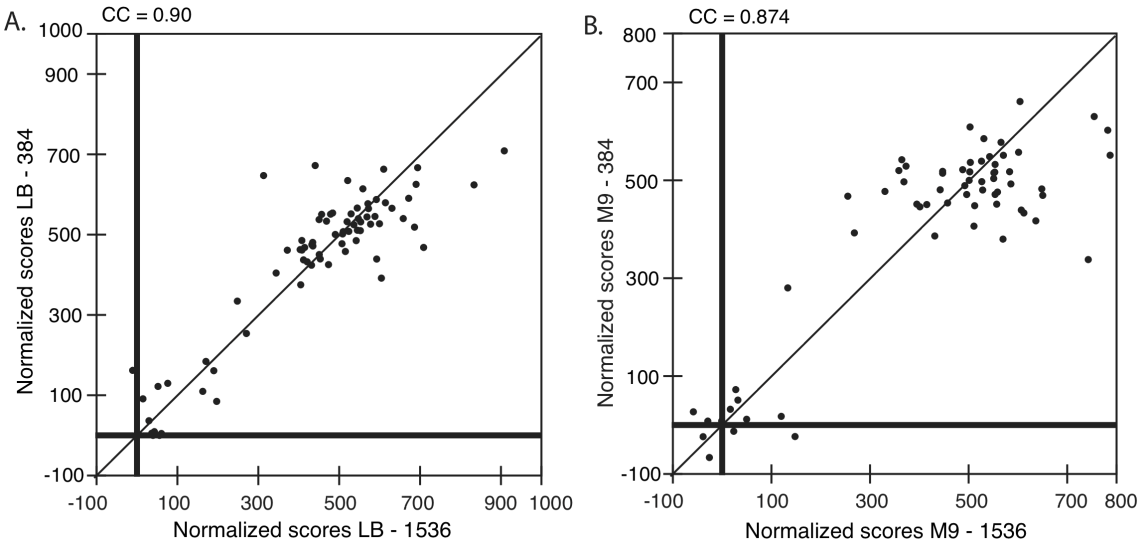


Figure S7. 384- and 1536-platforms generate similar quality of data. Scatter plots of averaged normalized colony-size scores for all 78 gene pair combinations in the two platforms using either LB (**A)** or the M9-glycerol (**B)** data indicate that there is no systematic bias.

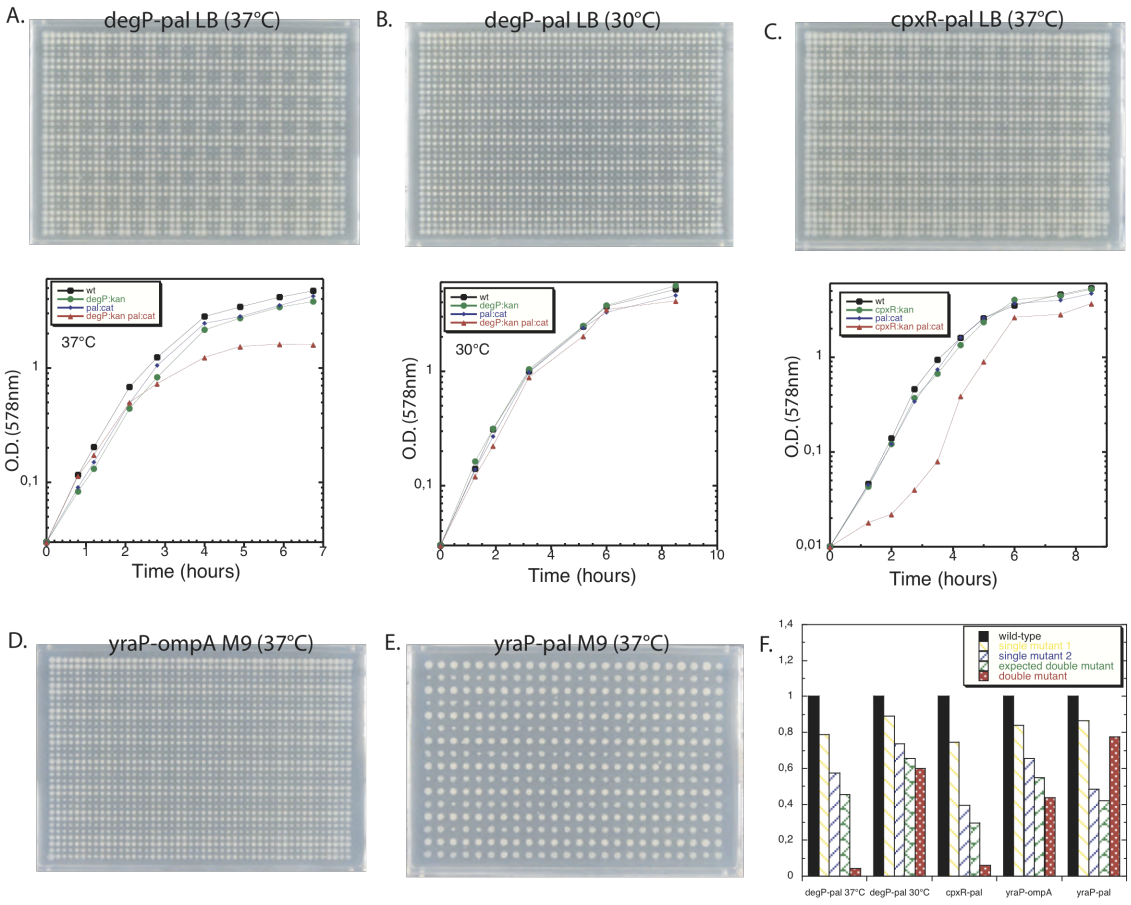


Figure S8. Independently reconstructed double mutants verify interactions detected by GIANT-coli. Wild-type (wt), parental single mutants and double mutants are arrayed in boxes of 4x4 replicas on the 1536-format plates, or boxes of 2x2 (each mutant once) on the 384-format plates. Each box consists of wt: left top-corner, single-mutant 1: right top-corner, single-mutant 2: left bottom-corner and double mutant on the right bottom-corner. **A**: *degP/pal* (LB-37°C); **B**: *degP/pal* (LB-30°C); **C**: *cpxR/pal* (LB-37°C); **D**: *ompA/yraP* (M9-37°C); **E**: *yraP/pal* (M9-37°C). Quantification of all 5 plates is shown in panel **F**. The average wild-type colony-size is set to 1 and the mutant colony sizes are plotted relative to wt; the expected colony size of the double mutant is the product of the single-mutant colony sizes. For A-C,

growth curves of the 4 strains in liquid media are shown below the plates. Growth differences between single and double mutants are very similar in liquid media and solid agar surfaces.

Supplementary Tables

Available online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2700713/

References

Ara, K., Ozaki, K., Nakamura, K., Yamane, K., Sekiguchi, J., and Ogasawara, N. (2006). *Bacillus* minimum genome factory - effective utilization of microbial genome information. Biotechnol Appl Biochem.

Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.C., Hirai, A.*, et al.* (2006). Large-scale identification of protein-protein interaction of *Escherichia coli* K-12. Genome Res *16*, 686-691.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol *2*, 2006 0008.

Bachmann, B.J. (1996). Derivations and Genotypes of Some Mutant Derivatives of *Escherichia coli* K-12. In *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, F.C.e.a. Neidhardt, ed. (ASM), p. chapter 133.

Bardwell, J.C., McGovern, K., and Beckwith, J. (1991). Identification of a protein required for disulfide bond formation in vivo. Cell *67*, 581-589.

Bochner, B.R. (2003). New technologies to assess genotype-phenotype relationships. Nat Rev Genet *4*, 309-314.

Boone, C., Bussey, H., and Andrews, B.J. (2007). Exploring genetic interactions and networks with yeast. Nat Rev Genet *8*, 437-449.

Breitkreutz, B.J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D.H., Bahler, J., Wood, V.*, et al.* (2008). The BioGRID Interaction Database: 2008 update. Nucleic Acids Res *36*, D637-640.

Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N.*, et al.* (2005). Interaction network containing conserved and essential protein complexes in *Escherichia coli*. Nature *433*, 531-537.

Byrne, A.B., Weirauch, M.T., Wong, V., Koeva, M., Dixon, S.J., Stuart, J.M., and Roy, P.J. (2007). A global analysis of genetic interactions in Caenorhabditis elegans. J Biol *6*, 8.

Campbell, T.L., and Brown, E.D. (2008). Genetic interaction screens with ordered overexpression and deletion clonesets implicate the *Escherichia coli* GTPase YjeQ in late ribosome biogenesis. Journal of bacteriology.

Cascales, E., Gavioli, M., Sturgis, J.N., and Lloubes, R. (2000). Proton motive force drives the interaction of the inner membrane TolA and outer membrane pal proteins in *Escherichia coli*. Mol Microbiol *38*, 904-915.

Cascales, E., and Lloubes, R. (2004). Deletion analyses of the peptidoglycan-associated lipoprotein Pal reveals three independent binding sequences including a TolA box. Mol Microbiol *51*, 873-885.

Cascales, E., Lloubes, R., and Sturgis, J.N. (2001). The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. Mol Microbiol *42*, 795-807.

Charlson, E.S., Werner, J.N., and Misra, R. (2006). Differential effects of yfgL mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. Journal of bacteriology *188*, 7186-7194.

Chen, C.S., Korobkova, E., Chen, H., Zhu, J., Jian, X., Tao, S.C., He, C., and Zhu, H. (2008). A proteome chip approach reveals new DNA damage recognition activities in *Escherichia coli*. Nature methods *5*, 69-74.

Cherepanov, P.P., and Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene *158*, 9-14.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature.

Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol *7*, R63.

Danese, P.N., Snyder, W.B., Cosma, C.L., Davis, L.J., and Silhavy, T.J. (1995). The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP. Genes Dev *9*, 387-398.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Nat Acad Sci USA *97*, 6640-6645.

Denic, V., and Weissman, J.S. (2007). A molecular caliper mechanism for determining very long-chain fatty acid length. Cell *130*, 663-677.

Elena, S.F., and Lenski, R.E. (1997). Test of synergistic interactions among deleterious mutations in bacteria. Nature *390*, 395-398.

Francois, V., Conter, A., and Louarn, J.M. (1990). Properties of new *Escherichia coli* Hfr strains constructed by integration of pSC101-derived conjugative plasmids. Journal of bacteriology *172*, 1436-1440.

Fujita, M., Umemura, M., Yoko-o, T., and Jigami, Y. (2006). PER1 is required for GPI-phospholipase A2 activity and involved in lipid remodeling of GPI-anchored proteins. Molecular biology of the cell *17*, 5253-5264.

Gerding, M.A., Ogata, Y., Pecora, N.D., Niki, H., and de Boer, P.A. (2007). The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in *E. coli*. Mol Microbiol *63*, 1008-1025.

Germon, P., Ray, M.C., Vianney, A., and Lazzaroni, J.C. (2001). Energy-dependent conformational change in the TolA protein of *Escherichia coli i*nvolves its N-terminal domain, TolQ, and TolR. Journal of bacteriology *183*, 4110-4114.

Girgis, H.S., Liu, Y., Ryu, W.S., and Tavazoie, S. (2007). A comprehensive genetic characterization of bacterial motility. PLoS Genet *3*, 1644-1660.

Grosse, C., Scherer, J., Koch, D., Otto, M., Taudte, N., and Grass, G. (2006). A new ferrous iron-uptake transporter, EfeU (YcdN), from *Escherichia coli*. Mol Microbiol *62*, 120-131.

Haldimann, A., and Wanner, B.L. (2001). Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. Journal of bacteriology *183*, 6384-6393.

Helmstetter, C.E. (1996). Timing of synthetic activities in the cell cycle. In Escherichia coli and Salmonella Cellular, Molecular Biology, F.C.e.a. Neidhardt, ed. (Washington, DC, American Society for Microbiology Press), pp. 1627–1639.

Hoekstra, W.P., and Havekes, A.M. (1979). On the role of the recipient cell during conjugation in *Escherichia coli*. Antonie Van Leeuwenhoek *45*, 13-18.

Inoue, T., Shingaki, R., Hirose, S., Waki, K., Mori, H., and Fukui, K. (2007). Genome-wide screening of genes required for swarming motility in *Escherichia coli* K-12. Journal of bacteriology *189*, 950-957.

Ito, M., Baba, T., and Mori, H. (2005). Functional analysis of 1440 *Escherichia coli* genes using the combination of knock-out library and phenotype microarrays. Metab Eng *7*, 318-327.

Kelley, R., and Ideker, T. (2005). Systematic interpretation of genetic interactions using protein networks. Nature biotechnology *23*, 561-566.

Keogh, M.C., Kim, J.A., Downey, M., Fillingham, J., Chowdhury, D., Harrison, J.C., Onishi, M., Datta, N., Galicia, S., Emili, A.*, et al.* (2006). A phosphatase complex that dephosphorylates gammaH2AX regulates DNA damage checkpoint recovery. Nature *439*, 497-501.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V.*, et al.* (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell *12*, 1565-1576.

Lazar, S.W., and Kolter, R. (1996). SurA assists the folding of *Escherichia coli* outer membrane proteins. Journal of bacteriology *178*, 1770-1773.

Le Grice, S.F., and Matzura, H. (1981). Binding of RNA polymerase and the catabolite gene activator protein within the cat promoter in *Escherichia coli*. Journal of molecular biology *150*, 185-196.

Lehner, B., Tischler, J., and Fraser, A.G. (2006). RNAi screens in *Caenorhabditis elegans* in a 96-well liquid format and their application to the systematic identification of genetic interactions. Nat Protoc *1*, 1617-1620.

Leichert, L.I., and Jakob, U. (2004). Protein thiol modifications visualized in vivo. PLoS Biol *2*, e333.

Low, K.B. (1996). Hfr Strains of *Escherichia coli* K-12. In Escherichia coli and Salmonella: Cellular and Molecular Biology, F.C.e.a. Neidhardt, ed. (ASM), p. chapter 127.

Mayer, M.L., Pot, I., Chang, M., Xu, H., Aneliunas, V., Kwok, T., Newitt, R., Aebersold, R., Boone, C., Brown, G.W.*, et al.* (2004). Identification of protein complexes required for efficient sister chromatid cohesion. Molecular biology of the cell *15*, 1736-1745.

Mazurkiewicz, P., Tang, C.M., Boone, C., and Holden, D.W. (2006). Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. Nat Rev Genet *7*, 929-939.

Niba, E.T., Naka, Y., Nagase, M., Mori, H., and Kitakawa, M. (2007). A Genome-wide Approach to Identify the Genes Involved in Biofilm Formation in *E. coli*. DNA Res *14*, 237-246.

Nielsen, H.J., Li, Y., Youngren, B., Hansen, F.G., and Austin, S. (2006). Progressive segregation of the Escherichia coli chromosome. Mol Microbiol *61*, 383-393.

Nielsen, H.J., Youngren, B., Hansen, F.G., and Austin, S. (2007). Dynamics of *Escherichia coli* chromosome segregation during multifork replication. Journal of bacteriology *189*, 8660-8666.

Onufryk, C., Crouch, M.L., Fang, F.C., and Gross, C.A. (2005). Characterization of six lipoproteins in the s^E regulon. Journal of bacteriology *187*, 4552-4561.

Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F., and Boeke, J.D. (2004). A robust toolkit for functional profiling of the yeast genome. Mol Cell *16*, 487-496.

Parsons, L.M., Lin, F., and Orban, J. (2006). Peptidoglycan recognition by Pal, an outer membrane lipoprotein. Biochemistry *45*, 2122-2128.

Pogliano, J., Lynch, A.S., Belin, D., Lin, E.C., and Beckwith, J. (1997). Regulation of *Escherichia coli* cell envelope proteins involved in protein folding and degradation by the Cpx two-component system. Genes Dev *11*, 1169-1182.

Qu, J., Mayer, C., Behrens, S., Holst, O., and Kleinschmidt, J.H. (2007). The trimeric periplasmic chaperone Skp of *Escherichia coli* forms 1:1 complexes with outer membrane proteins via hydrophobic and electrostatic interactions. Journal of molecular biology *374*, 91-105.

Rizzitello, A.E., Harper, J.R., and Silhavy, T.J. (2001). Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. Journal of bacteriology *183*, 6794-6800.

Roguev, A., Wiren, M., Weissman, J.S., and Krogan, N.J. (2007). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat Methods *4*, 861-866.

Rouviere, P.E., and Gross, C.A. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. Genes Dev *10*, 3170-3182.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K.*, et al.* (2007). The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biol *5*, e77.

Sawitzke, J.A., Thomason, L.C., Costantino, N., Bubunenko, M., Datta, S., and Court, D.L. (2007). Recombineering: in vivo genetic engineering in *E. coli*, *S. enterica*, and beyond. Methods in enzymology *421*, 171-199.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F.*, et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Schuster, S.C. (2008). Next-generation sequencing transforms today's biology. Nature methods *5*, 16-18.

Sklar, J.G., Wu, T., Kahne, D., and Silhavy, T.J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. Genes Dev *21*, 2473-2484.

Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M.*, et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Ullers, R.S., Ang, D., Schwager, F., Georgopoulos, C., and Genevaux, P. (2007). Trigger Factor can antagonize both SecB and DnaK/DnaJ chaperone functions in *Escherichia coli*. Proc Natl Acad Sci U S A *104*, 3101-3106.

Wade, J.T., Struhl, K., Busby, S.J., and Grainger, D.C. (2007). Genomic analysis of protein-DNA interactions in bacteria: insights into transcription and chromosome organization. Mol Microbiol *65*, 21-26.

Wolfe, A.J., Parikh, N., Lima, B.P., and Zemaitaitis, B. (2008). Signal integration by the two-component signal transduction response regulator CpxR. Journal of bacteriology.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J., and Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. Cell *121*, 235-245.

Chapter 3

A Genetic Interaction Map of RNA Processing Factors Reveals Links Between Sem1/Dss1-Containing Complexes and mRNA Export and Splicing

A Genetic Interaction Map of RNA Processing Factors Reveals Links Between Sem1/Dss1-Containing Complexes and mRNA Export and Splicing

Gwendolyn M. Wilmes[1,*], Megan Bergkessel[1,*], Sourav Bandyopadhyay[2], Michael Shales[3], Hannes Braberg[3], Gerard Cagney[3], Sean R. Collins[3,4], Gregg B. Whitworth[1], Tracy L. Kress[1], Jonathan S. Weissman[3,4], Trey Ideker[2], Christine Guthrie[1], Nevan J. Krogan[3]

1 Department of Biochemistry and Biophysics, University of California, San Francisco, 600 16th Street, Genentech Hall, San Francisco, CA 94143-2200, USA

2 Program in Bioinformatics, Department of Bioengineering, University of California San Diego, La Jolla, CA, USA

3 Department of Cellular and Molecular Pharmacology, California Institute for Quantitative Biomedical Research, University of California, San Francisco, 1700 4th Street, San Francisco, California, USA, 94158

4 Howard Hughes Medical Institute

Co-corresponding authors: Email: christineguthrie@gmail.com, Email: krogan@cmp.ucsf.edu

*equal contributing authors

Summary

We have used a quantitative, high-density genetic interaction map to interrogate functional relationships within and between RNA processing pathways. Due to their complexity, and the essential roles of many of the components, these pathways have previously been difficult to dissect. Here we report the results for 107,155 individual interactions involving 552 mutations, 166 of which are hypomorphic alleles of essential genes. Our data allowed for the discovery of links between components of the mRNA export and splicing machineries and Sem1/Dss1, a component of the 19S proteasome. In particular, we demonstrate that Sem1 has a proteasome-independent role in mRNA export as a functional component of the Sac3-Thp1 complex. We further show that Sem1 interacts with Csn12, a component of the COP9 signalosome. Finally, we show that Csn12 plays an unexpected role in pre-mRNA splicing, and that this role is independent of other signalosome components. Thus Sem1 is involved in three separate and functionally distinct complexes.

Introduction

Following the initial identification of yeast genes by the genome sequencing project (Goffeau et al., 1996), the phenotypic analysis of gene deletion mutations (Giaever et al., 2002), and the characterization of cellular protein complexes (Gavin et al., 2006; Krogan et al., 2006), an important next step toward a comprehensive understanding of cellular function is to place these complexes within specific *in vivo* pathways. Genetic interactions, which measure the extent to which the function of one protein depends on the presence of a second, can not only help place complexes in pathways but also provide information about the functional cross-talk between pathways and processes. Two approaches, synthetic genetic array (SGA) (Tong et al., 2001) and diploid based synthetic lethality analysis on microarrays (dSLAM) (Pan et al., 2004), have been developed to identify synthetic sickness/lethal (SSL) relationships on a genome-wide scale in *S. cerevisiae.* These approaches systematically reveal cases where two mutations in combination cause a more severe growth defect than that due to either mutation alone.

However, SSL interactions only represent a subset of the total possible genetic interactions. We recently exploited the SGA methodology to develop the E-MAP (epistatic miniarray profile) approach, which uses colony size of double mutants as a quantitative read-out. This allows for the characterization of the entire range of epistatic relationships, including both negative (i.e. SSL) and positive genetic interactions (Collins et al., 2007b; Schuldiner et al., 2005; Roguev et al., 2008). Positive interactions include cases where the double mutant is inferred to be either no sicker or actually healthier than the expected fitness of the sicker single

mutant. The E-MAP method highlights the broad effects of perturbing biochemically distinct processes, and so gives a picture of how these cellular processes interact, while also resolving functional interactions at the level of the protein complex or pathway. It is ideally suited for analysis of RNA processing functions, which are biochemically complex, being separated temporally and compartmentally within the cell, but linked by many physically interacting protein complexes.

In this study, we used the E-MAP approach to generate a high density genetic interaction map for a large set of genes (~10% of the genome) that are collectively involved in virtually all aspects of RNA metabolism, including regulating and/or processing rRNAs, mRNAs, snoRNAs, snRNAs and tRNAs. Since many of the factors we genetically analyzed are required for cell viability, we created hypomorphic alleles, which allowed for an unprecedented view of the epistatic architecture of essential genes. We employed recently developed analytical tools for visualization of the quantitative genetic interaction data. This analysis provides a global view of the epistatic behavior within and between RNA-related processes as well as the genetic "cross-talk" between individual protein complexes that function within these processes. The network view allowed for the discovery of a number of previously uncharacterized functional connections. For example, this E-MAP allowed us to place the 19S proteasome subunit Sem1/Dss1 into two other functionally and physically distinct complexes: one involved in mRNA export and the other linked to mRNA splicing. We fully anticipate that this dataset will be a useful resource to the RNA community both on its own and in conjunction with previously published E-MAP datasets (Collins et al., 2007b; Schuldiner et al., 2005) and protein-protein

interaction datasets (Gavin et al., 2006; Ho et al., 2002; Krogan et al., 2006; Collins et al., 2007a) (http://interactome-cmp.ucsf.edu).

Results and Discussion

Generation of an RNA Processing Genetic Interaction Map

To genetically interrogate RNA processing in *Saccharomyces cerevisiae*, we created an E-MAP comprised of 552 mutations, each within a different gene involved in one of various RNA-related processes, including mRNA splicing and export, tRNA modification, translation, rRNA processing and RNA degradation (Figure 1A, Supplemental Figure 1). To facilitate comparison with other EMAPs, we also included sets of genes involved in other processes that indirectly impinge on RNA metabolism, including transcription, chromatin remodelling, and ubiquitin/protein degradation. Finally, we included a number of previously uncharacterized genes that have been functionally linked to RNA-related processes through unbiased, large-scale screens (Gavin et al., 2006; Krogan et al., 2006; Krogan et al., 2004; Tong et al., 2004). Since a large fraction of genes involved in these processes are essential for viability, we employed the DAmP (decreased abundance by mRNA perturbation) strategy (Schuldiner et al., 2005) to create 166 hypomorphic (partial loss of function) alleles of essential genes (Figure 1A). In all, this genetic map contains 107,155 distinct, pair-wise genetic interaction measurements (http://interactome-cmp.ucsf.edu).

We subjected the entire dataset to hierarchical clustering, an approach that groups genes with similar patterns of genetic interactions. Several known protein

complexes, such as the nuclear pore, Prefoldin, and a complex involved in nonsense mediated mRNA decay (NMD) are represented by distinct clusters (Supplemental Figure 1). Similarly, all six known components of the Elongator complex, which was initially linked to transcriptional elongation (Otero et al., 1999), cluster together. Furthermore, components of Elongator also cluster next to *TRM10*, a tRNA methyltransferase (Jackman et al., 2003), and display strong negative interactions with several other tRNA modifying enzymes (e.g. *ARC1* and *DEG1*), consistent with work implicating this complex in the regulation of tRNA modification (Esberg et al., 2006). Interestingly, the ubiquitin-like modifier *URM1* (Furukawa et al., 2000) also clusters with the components of Elongator, and recent work has suggested that Urm1p functions as a thiol group carrier important for thiol modification of tRNAs (Nakai et al., 2008; Schmitz et al., 2008) (Supplemental Figure 1).

Components of known spliceosomal subcomplexes, including the U1 snRNP (*MUD1*, *NAM8*, *PRP42* and *SNU56*) (Gottschalk et al., 1998; Neubauer et al., 1997) and the Prp19 associated Nineteen Complex (NTC) (*NTC20*, *ISY1*, *SYF2*, *PRP19*) (Chen et al., 2002) satisfyingly cluster together (Supplemental Figure 1). Two largely uncharacterized factors, *CWC21* and *CWC27*, also cluster with the NTC, consistent with work demonstrating that they physically associate with the NTC component Cef1 (Ohi and Gould, 2002). Additionally, we identified many positive genetic interactions between the NTC and the U1 snRNP (Supplemental Figure 1), perhaps reflecting the fact that the U1 snRNP departs the spliceosome approximately when the NTC joins (Chan et al., 2003). Further investigation will be

required to understand the implications of these and the many other genetic interactions contained within this E-MAP (http://interactome-cmp.ucsf.edu).

Global Analysis of Genetic and Physical Interaction Data

While inspection of the E-MAP genetic interaction datasets in isolation have already yielded many interesting observations, analysis of these data in conjunction with published high throughput protein-protein interaction datasets promises to be especially powerful. We have shown in other E-MAPs that genes exhibiting positive genetic interactions are more likely than expected by chance to encode proteins that have physical interactions (Collins et al., 2007b; Roguev et al., 2007), and inspection of the data from the RNA processing E-MAP revealed a similar overrepresentation of factors with protein-protein interactions (PPIs) among pairs of factors with positive genetic interactions. Protein-protein interactions were defined by tandem affinity purification (TAP)-derived purification enrichment (PE) scores of > 1.0, as described previously (Collins et al., 2007a). However, in contrast to what has been observed in other EMAPS, there is also an enrichment of PPIs among factors displaying negative genetic interactions (Figure 1B). Notably, the RNA processing E-MAP greatly expands the number of components of essential complexes that have been interrogated by the E-MAP technique, both by including many non-essential factors that participate in several large essential complexes, and by including a large number of hypomorphic alleles of essential factors. This difference from previous E-MAPs may at least in part account for the more prevalent negative genetic interactions observed among physically interacting factors. Indeed, a recent analysis

of a previously published E-MAP focused on chromosome biology showed that complexes dominated by negative genetic interactions are more likely to contain at least one essential gene (Bandyopadhyay et al., 2008).

Interestingly, the hypomorphic alleles of essential genes that are included in this E-MAP do not generally have more negative genetic interactions than the non-essential genes when each of these subsets is considered as a whole. We compared the ratio of negative ($S \leq -2.5$) to positive ($S \geq 2.0$) genetic interactions (N to P) involving either pairs of non-essential genes or pairs where at least one gene is essential (Figure 1C, see Experimental Procedures). In the entire dataset, we found an N to P ratio of ~1, regardless of whether the mutant pairs contained at least one essential gene (1.00) or not (1.29). However, when we restricted the analyses to genes whose corresponding proteins are physically associated (as defined above), we found that the genetic interactions observed with essential genes are over three times as likely to be negative (N to P =2.84, p=.0003) (Figure 1C; Supplemental Table 2) compared to interactions between null mutations of non-essential genes (N to P =0.93).

These observations suggest that the introduction of two mutations in the same essential complex is very likely to have a synthetic negative effect on growth, whereas a second mutation in a non-essential complex is unlikely to have a further detrimental effect above that caused by the first mutation, and may indeed improve growth. Thus, a positive interaction observed between two factors that physically interact within an essential complex may suggest that these factors form a non-essential module within the complex. One example of this is the positive interaction

between IST3 and BUD13, which have been suggested to physically interact with both the non-essential RES complex and with the essential SF3b complex of the spliceosome (Dziembowski et al., 2004; Wang et al., 2005). Our data support the model that IST3 and BUD13 function as a non-essential module in these complexes.

While factors with protein-protein interactions are overrepresented among the strong genetic interactions observed in the RNA processing E-MAP, there are also certainly many genetic interactions between factors that have functional, but not physical, interactions. Indeed, these interactions comprise a powerful tool for discovering previously uncharacterized pathways connecting protein complexes. However, visualizing the relative contributions of within-complex and between-complex genetic interactions to the overall interaction landscape is a complicated problem. To address this, we used an approach we have recently described (Bandyopadhyay et al., 2008) that classifies protein pairs as either operating within the same module or between functionally related modules given their genetic and physical interactions (see Experimental Procedures for a more detailed description of the analyses). In the resulting representation, blue and yellow denote a statistically significant enrichment of negative and positive interactions, respectively, whereas green corresponds to cases where there are significant enrichments in total interactions, but the numbers of positives and negatives are roughly equal. Nodes (boxes) correspond to complexes and edges (lines), the genetic cross-talk between the complexes (Figure 2A).

Several interesting connections become evident when the data are analyzed in this way. For example, there are negative genetic interactions (blue line) between

the chromatin remodeling complex, SWR-C, which incorporates the histone H2A variant, Htz1, into chromatin (Kobor et al., 2004; Krogan et al., 2003; Mizuguchi et al., 2004), and the nuclear pore factors Mex67-Nup120, consistent with recent work demonstrating that Htz1 is involved in tethering genes to the nuclear periphery (Brickner et al., 2007). Also, we found positive genetic interactions (yellow line) between the complex containing both casein kinase II (CKII) and the chromodomain-containing protein Chd1, and the spliceosome. Consistent with this finding, Reinberg and colleagues found a physical and functional connection in human cells between Chd1 and the spliceosome (Sims et al., 2007). There are many testable hypotheses that can be extracted from this data set in this format and we have created an interactive website that allows for searching of individual interactions as well as the connections between complexes (http://www.cellcircuits.org/complexes/RNAProcessing/html/).

In an effort to gain insight into the functional organization of RNA-related processes at an even broader level, we created a map that highlights strong genetic trends both within and between these processes, without respect to known physical interactions (Figure 2B). Nodes (boxes) now correspond to the interactions within distinct functional processes (defined by a combination of GO term analysis and manual annotation) whereas edges (lines) represent the genetic cross-talk between processes.

A striking finding is that several sets of genes that are known to function in the same biochemical processes contain predominantly positive or predominantly negative interactions, as was also observed in an analysis of the computed epistatic

behavior among yeast metabolic genes (Segre et al., 2005). For example, genes classed as involved in cytoplasmic rRNA biogenesis and mitochondrial rRNA biogenesis are significantly enriched in positive genetic interactions (Figure 2B yellow nodes). In contrast, factors involved in mRNA splicing, mRNA export and the nuclear pore have strongly negative interactions (blue nodes). Finally, not all groups show consistent patterns of interactions; for instance, RNA degradation, tRNA biogenesis and transcription show mixed but still significant genetic interactions (green nodes). The broad trends that emerge when the data are viewed in this way suggest that the patterns of genetic interactions are influenced not only by direct physical interactions between factors and direct functional interactions between complexes, but also probably by higher-level interactions within and between processes.

The Sem1/Dss1 Component of the Proteasome is Involved in mRNA Export Via the Sac3-Thp1 Complex

Our E-MAP uncovered positive genetic interactions between the proteasome and the Sac3-Thp1 (TREX-2) complex involved in mRNA export (Figure 2A), suggesting that a direct functional relationship may exist between these two complexes. We confirmed that these positive interactions represented suppressive relationships through tetrad analysis and dilution growth assays at 16°C (Supplemental Figure 2). The proteasome is a large molecular machine that degrades ubiquitylated substrates (Collins and Tansey, 2006) and is concentrated at the nuclear periphery in yeast (Enenkel et al., 1998; Wilkinson et al., 1998). During

the process of mRNA export, multiple components of the Mex67-mediated export pathway interact with the nuclear pore basket via the Sac3-Thp1 complex (Luna et al., 2008).

To investigate the significance of the genetic connections between the proteasome and the Sac3-Thp1 mRNA export complex, we compared their global genetic interaction patterns from the E-MAP. The genetic interaction profiles we obtained for *SAC3* and *THP1* are extremely similar (correlation coefficient = 0.7), consistent with their known biochemical and functional interactions (Fischer et al., 2002; Gallardo et al., 2003). After each other, the genetic interactions of *SAC3* and *THP1* are most similar to those of the nuclear pore components *NUP60* and *NUP120*, *THP2* (a component of the THO/TREX complex that co-transcriptionally associates with the mRNP), and, interestingly, the proteasome lid component *SEM1* (Figure 3A and http://interactome-cmp.ucsf.edu/). In contrast, the genetic profiles of other components of the proteasome (e.g. *RPN10*, *PRE9* and *RPN4*) do not correlate with these export factors. *SEM1* differs from other proteasome mutants in sharing many genetic interactions with *THP1-SAC3*, including negative ones with the exosome (*LRP1*, *RRP6*), nuclear pore (*NUP100*, *NUP120*, *NUP60*, *NUP57*) and other factors implicated in mRNA export (such as *MEX67*, *APQ12*, *DBP5*) (Figure 3B and Supplemental Figure 1).

These genetic data suggest that Sem1 may act as an export factor in *S. cerevisiae,* and so we directly tested whether deleting *SEM1* affected mRNA export. Deletion of either *SAC3* or *THP1* results in a substantial increase of polyadenylated mRNA in the nucleus (Fischer et al., 2002; Gallardo et al., 2003; Lei et al., 2003). We

found that deletion of *SEM1* also results in a moderate mRNA export defect at 25°C, 30°C and 37°C (Figure 3C and data not shown), suggesting a direct role for Sem1 in mRNA export. This is consistent with previous work in fission yeast (Mannen et al., 2008; Thakurta et al., 2005). Interestingly, we found that deletion of *SEM1* in the *thp1*Δ background not only partially suppresses its growth defect at 16°C (Supplemental Figure 2) but also reduces the intensity of the mRNA export phenotype compared to *thp1*Δ alone (Figure 3C).

However, deletion of proteasome factors other than *SEM1* (*RPN10*, *PRE9*, *UBP6*, *DOA1*, *RPN4*), or inhibition of the proteasome by the drug MG132 in a *pdr5Δ* background (which allowed for more efficient uptake of the drug), do not cause an mRNA export defect in this assay (Figure 3C and data not shown). Furthermore, although *RPN10* deletion exacerbates the growth, proteasome integrity and DNA damage response defects of a *sem1Δ* strain (Funakoshi et al., 2004; Sone et al., 2004), we found no significant effect on *sem1Δ*'s moderate mRNA export defect (Figure 3C). Together, these data argue that Sem1 is directly involved in mRNA export, but other components of the proteasome are not.

We next asked whether Sem1 is physically associated with the Sac3-Thp1 complex. Since we found that introducing a tag on the C-terminus of Sem1 results in a defect in mRNA export (data not shown), we used a polyclonal antibody (Funakoshi et al., 2004) to immunoprecipitate or detect Sem1 from cellular extracts of a panel of GFP-tagged strains. Both Thp1-GFP and Sac3-GFP reciprocally co-immunoprecipitate with Sem1 (Figure 3D and Supplemental Figure 3). Consistent with previous results, GFP-tagged Rpt4, a 19S proteasome component, co-

immunoprecipitates both Sem1 and Rpt6, another known component of the proteasome (Funakoshi et al., 2004; Krogan et al., 2004; Sone et al., 2004). In contrast, Sac3-GFP and Thp1-GFP are not associated with Rpt6 (Figure 3D), strongly suggesting that Sem1 interacts with Sac3-Thp1 independent of the rest of the proteasome. Previous work in *S. pombe* suggested physical associations between Dss1 (*S. cerevisiae* Sem1) and other mRNA export factors, including Nup159, Rae1 (*S. cerevisiae* Gle2), and Mlo3 (*S. cerevisiae* Yra1) (Thakurta et al., 2005); however, we do not detect these physical connections in budding yeast (Figure 3E and data not shown). We also do not find Sem1 associated with TREX (e.g. Thp2) or the transcriptional activation complex SAGA (e.g. Ada2), both of which are complexes that have been reported to interact with Sac3-Thp1 (Fischer et al., 2002; Rodriguez-Navarro et al., 2004). Finally, the physical interaction between Sac3-Thp1 and Sem1 is insensitive to RNAse A treatment (Figure 3D and data not shown), suggesting that this interaction is not mediated by RNA.

To gain further insight into the physically interacting partners of Sem1, we analyzed data from our earlier large-scale protein-protein interaction project that used a systematic affinity tagging, purification and mass spectrometry approach to identify protein complexes (Krogan et al., 2006). After components of the proteasome, the proteins that most significantly immunoprecipitated Sem1 as prey were Csn12, a component of the CSN (COP Nine Signalosome, another large complex with structural similarity to the proteasome lid), and Ypr045c, an uncharacterized protein, both of which also co-purified with each other (http://tap.med.utoronto.ca/). We confirmed the Sem1-Csn12 and Sem1-Ypr045C

interactions by co-immunoprecipitation (Figure 3D). Together with our Sac3-Thp1-Sem1 results, these physical interactions suggest that Sem1 may be part of at least three functionally distinct complexes in *S. cerevisiae*: the 19S proteasome (Funakoshi et al., 2004; Krogan et al., 2004; Sone et al., 2004), Sac3-Thp1, and a complex containing Csn12 and Ypr045c (see Figure 5). This is consistent with previous glycerol gradient fractionation data showing that a significant fraction of the Sem1 in the cell is contained in complexes smaller than the proteasome (Funakoshi et al., 2004). These complexes are all likely to be conserved, as a purification of protein complexes associated with human Dss1 identified two metazoan-specific complexes and components of all three *S. cerevisiae* complexes that we describe here (Baillat et al., 2005). Furthermore, *S. cerevisiae* or human Sem1/Dss1 can complement a *dss1Δ* strain for mRNA export in *S. pombe* (Thakurta et al., 2005).

Examining the domain structure of Sem1-interacting proteins suggests a model for how Sem1 associates with each of these complexes. Rpn3, Thp1, and Csn12 are the only three proteins in *S. cerevisiae* that harbor the PAM (PCI-Associated Module) domain (Ciccarelli et al., 2003), a highly conserved region upstream of a subset of PCI domains (protein interaction domains named after three complexes rich in these domains: the proteasome, the CSN, and eIF3) (Hofmann and Bucher, 1998; Scheel and Hofmann, 2005). Similarly, Rpn12, Sac3, and Ypr045c are the only three proteins in *S. cerevisiae* that contain a highly conserved variant of the PCI domain, called the SAC3 domain (PFAM database (Finn et al., 2006) # PF03399) (Figure 5). Indeed, the region of Sac3 that interacts with Thp1 was mapped to a

small section containing the SAC3 domain (Fischer et al., 2002). Furthermore, a number of studies have placed Rpn3, Rpn12, and Sem1 in a subcomplex within the proteasome lid (Fu et al., 2001; Isono et al., 2005; Sharon et al., 2006). Sem1 crosslinks directly to Rpn3 *in vivo* (Sharon et al., 2006), and binds to Rpn3 *in vitro* (Wei et al., 2008). Thus, these domains suggest a model that Sem1 may be interacting with PAM and/or SAC3 domains in multiple large protein complexes.

Indeed, our results raise the intriguing possibility that the SAC3 and PAM domains may in fact bind to one another, and in turn, recruit Sem1. Consistent with this notion, we found that both Sac3 and Thp1 are required for the other to interact with Sem1 (Figure 3D and Supplemental Figure 3). Previous studies have suggested that Sem1 is involved in stabilizing protein complexes. One of the metazoan-specific Sem1/Dss1 containing complexes, that containing BRCA2, is unstable unless Dss1 is co-purified (Li et al., 2006; Yang et al., 2002). The proteasome is also unstable in cells lacking Sem1 (Funakoshi et al., 2004; Sone et al., 2004), although a recent study suggests that this instability does not stem from abrogation of the PAM-domain (Rpn3)- SAC3-domain (Rpn12) interaction (Wei et al., 2008). In the presence of 1M salt, the proteasome falls apart in a *sem1Δ* strain (Sone et al., 2004). In contrast, we find that in the absence of Sem1, Sac3-GFP still remains associated with immunoprecipitated Thp1-HA in the presence of at least 1 M NaCl (Figure 3E). Consistent with these data, we find that the localization of Thp1 to the nuclear envelope is not strongly affected by *sem1Δ*, in contrast to *sac3Δ* , which results in its complete mislocalization ((Fischer et al., 2002) and data not shown). Thus, these data suggest that Sem1 has roles other than simply stabilizing protein complexes.

The Sac3-Thp1 mRNA export complex is implicated in the pathway that loads the export receptor Mex67 (or TAP in higher eukaryotes) onto mRNAs, and our genetic, functional, and physical data suggest that Sem1 may act as part of the Sac3-Thp1 complex. Sub2 and Yra1 are co-transcriptionally loaded onto mRNAs as part of the THO/TREX complex, and Yra1 acts as an adaptor protein to load Mex67 onto the mRNP to making it competent for export (Luna et al., 2008). Because Yra1 can interact with either Sub2 or Mex67, but not both simultaneously, Sub2 may be exchanged for Mex67 (Strasser and Hurt, 2001) (Figure 3G). To investigate the role of Sac3-Thp1-Sem1 in RNP formation, we UV crosslinked proteins to RNA *in vivo*, made cell extracts, purified the poly(A) RNA over an oligo dT cellulose column, and then examined protein components in the resulting poly(A) eluates in wild-type and mutant backgrounds. In a wild-type strain, low levels of Sub2, Yra1, and Mex67 can be detected crosslinked to poly(A) mRNA in a UV-dependent manner (Figure 3F). Deletion of *SEM1*, *THP1* or *SAC3* all result in a significant accumulation of Sub2 and Yra1 on polyA mRNA, whereas Mex67 binding increases only modestly in all three cases (Figure 3F and data not shown). In contrast, large quantities of Mex67, but not Yra1, accumulate on mRNA when it is blocked at the nuclear periphery in a *dbp5,rrp6* double mutant (Lund and Guthrie, 2005). These data suggest that the Thp1-Sac3-Sem1 complex facilitates the exchange of Sub2 for Mex67 (Figure 3G). The modest increase in Mex67 association may reflect a partial block during Mex67 loading, or may reflect aberrant binding of Mex67 to the particularly large amounts of Yra1 that appear to be accumulating on the mRNA. Although further work will be

required to identify the exact role that this complex plays in the export pathway, our data argue that Sem1 is acting in concert with Sac3-Thp1 to load Mex67 onto mRNA.

The Csn12 Component of the COP9 Signalosome is Involved in mRNA Splicing

The functions of two of the three Sem1-containing complexes (the proteasome and Sac3-Thp1) were known, so we used the E-MAP, in conjunction with physical interaction data, to gain insight into the function of the Csn12-Ypr045c complex. Csn12 was previously described as a subunit of the COP9 signalosome (CSN), a complex responsible for the removal of the Rub1 modifier (known as Nedd8 in higher eukaryotes) from cullin proteins, which are components of various E3 ubiquitin ligase complexes (Maytal-Kivity et al., 2003). However, deletion of *CSN12* fails to cause accumulation of neddylated Cdc53p, a phenotype observed in all other mutants of CSN subunits (Wee et al., 2002). Inspection of the quantitative genetic interaction data revealed that *CSN12* is most similar genetically to components of the spliceosome. For example, the profile of genetic interactions of *csn12*Δ is highly correlated with that of *isy1*Δ and also several other mRNA splicing factor mutants, including *syf2*Δ, *prp11*-DAmP, *ist3*Δ, *snu66*Δ and *prp19-DAmP*, whereas deletions of other components of the CSN display very different genetic interaction profiles (Figure 4A). Consistent with a functional link between Csn12 and the spliceosome, large-scale protein TAP studies identified a physical link between only this component of the CSN and many factors involved in mRNA splicing (Collins et al., 2007a; Gavin et al., 2006; Krogan et al., 2006).

To test whether Csn12 plays a functional role in mRNA splicing, we used a splicing-specific microarray to measure the levels of pre-mRNA, mature mRNA, and total mRNA for most of the intron-containing genes in the yeast genome (Pleiss et al., 2007). In general, a decrease in splicing efficiency for a given transcript results in an accumulation of its pre-mRNA, and often in a measurable loss of mature mRNA (Clark et al., 2002; Pleiss et al., 2007). We examined the splicing efficiencies of *isy1*Δ, *csn12*Δ, and *ypr045c*Δ (the SAC3 domain-containing protein that physically interacts with Csn12) as well as deletions of *CSN9* and *PCI8*, two other components of the CSN. As expected, deletion of *ISY1* causes a substantial accumulation of pre-mRNA for many intron-containing genes (Figure 4B). Interestingly, deletion of either *CSN12* or *YPR045c* causes a pre-mRNA accumulation in many of the same transcripts adversely affected in the *isy1*Δ strain (Figure 4B). In contrast, *csn9*Δ and *pci8*Δ strains fail to show a splicing defect (Figure 4B). We also failed to detect a splicing defect in the *sem1*Δ strain (data not shown). It may be that under the conditions tested, the loss of Sem1p does not significantly impair the function of the Csn12-Ypr045c complex in splicing.

In order to more quantitatively assess the similarity of the splicing microarray profiles observed with *csn12*Δ and *ypr045c*Δ to those obtained from well characterized splicing factors, we calculated pairwise Pearson correlation coefficients between a number of mutant pairs using the ratios from all of the microarray features reporting on intron-containing genes (Figure 4C). As previously reported, various splicing factor mutations result in a wide range of splicing phenotypes (Pleiss et al., 2007). The splicing microarray profiles of the *csn12*Δ and

*ypr045c*Δ strains are highly correlated with each other as well as with the profiles of a number of well-characterized splicing factor mutants (e.g. *prp2-1*, *prp16-302*, *prp8-1*, *ist3*Δ, *isy1*Δ), whereas *csn9*Δ and *pci8*Δ exhibit distinct behavior on the splicing array (Figure 4C). Thus, Csn12 and Ypr045c, but not other components of the CSN, are involved in maintaining normal levels of mRNA splicing.

Perspective

In this study, we generated a quantitative genetic interaction map in budding yeast that is focused on virtually all aspects of RNA processing and includes factors implicated in the generation and regulation of tRNAs, mRNAs, rRNAs, snoRNAs and snRNAs. We have applied novel analytical tools that allow for the visualization of the genetic crosstalk between the different RNA-related processes and the protein complexes that function within these processes. Using these representations and other data exploration tools such as hierarchical clustering and searchable databases, we predicted and experimentally verified several novel biological connections as examples of how this data can be mined. Along with the examples explored here, we also found a strong link between Npl3, an SR domain-containing protein involved in mRNA export, and mRNA splicing (see Kress etal., in press).

The large amount of genetic data from the E-MAP is even more useful when analyzed in combination with other large data sets, including other E-MAPs (Collins et al., 2007b; Schuldiner et al., 2005) and physical interaction data sets such as the TAP database (Collins et al., 2007a; Gavin et al., 2006; Ho et al., 2002; Krogan et al., 2006) (http://interactome-cmp.ucsf.edu). This approach allowed us to identify two

novel protein complexes containing the 19S proteasome component Sem1, one of which (Sac3-Thp1) acts in mRNA export, and one of which (Csn12-Ypr045c) acts in mRNA splicing (Figure 4). We failed to detect an mRNA export defect in both *csn12*Δ and *ypr045c*Δ, and we failed to detect a splicing defect in both *sac3*Δ and *thp1*Δ (data not shown), suggesting that these two complexes are functionally distinct despite sharing Sem1 as a common component.

SEM1 has positive genetic interactions with *SAC3*, *THP1*, *CSN12*, and *YPR045C*, consistent with the global pattern observed when we interrogated the entire data set, whereby factors that share physical interactions in non-essential complexes tend to have positive genetic interactions. By contrast, *SEM1* has mainly negative genetic interactions with components of the proteasome, which is essential. *SAC3* and *THP1* have positive genetic interactions with each other, as do *CSN12* and *YPR045C*, again consistent with the observed pattern. Interestingly, the interactions between components of different Sem1-containing complexes were largely positive, possibly reflecting crosstalk mediated by the shared Sem1 component (Supplemental Figure 1). Both the genetic and functional data confirmed a role for Sem1 in mRNA export, but we did not observe a splicing defect in the *sem1Δ* strain (data not shown). The genetic interactions of *SEM1* with *CSN12* and *YPR045C* are weaker than those observed with *SAC3* and *THP1*, and the splicing defect caused by deletion of *CSN12* or *YPR045C* is mild compared to the dramatic export defect caused by deletion of *SAC3* or *THP1*. It may be that the rich growth conditions under which these studies were performed do not strongly necessitate the function of the Csn12-Ypr045c-Sem1 complex. Alternatively, Sem1's function in

this complex may be related to a still undiscovered splicing-independent role. Further studies will be required to better understand the role of this complex in splicing, and the conditions under which its function is most relevant.

The links reported here between Sem1 and multiple mRNA processes suggest a model for Sem1 action in multiple distinct complexes. We favor a model in which Sem1/Dss1 acts as a nucleic acid analog, modulating access between nucleic acids and the protein complexes with which they interact. The proteasome has been suggested to play roles in transcription activation, DNA repair, and RNA degradation (Ezhkova and Tansey, 2004; Ferdous et al., 2001; Gautier-Bert et al., 2003; Gonzalez et al., 2002), all of which would involve interactions with nucleic acids, which could be facilitated at least in part by Sem1. Indeed, inspection of a co-crystal of human Dss1 with the metazoan22 specific breast cancer protein BRCA2 and ssDNA led to a similar model, in which Dss1 might modulate access of ssDNA to BRCA2 during double strand break repair.

Sem1/Dss1 is a small (10 kD) extremely acidic protein, with physical characteristics strikingly similar to those of nucleic acids (Yang et al., 2002). It has recently been shown that a small patch of conserved very acidic residues in Sem1 is necessary and sufficient to mediate interaction with a PAM domain-containing protein (Rpn3) (Wei et al., 2008). Therefore, it is tempting to speculate that it could act as both an RNA and DNA chaperone, facilitating the association of nucleic acid with several complexes. Further work will be required to elucidate the biophysical details of such interactions.

Finally, combining the unbiased genetic map with physical interaction data suggests functions for several previously uncharacterized proteins. For example, *YNR004W* encodes a small protein that has no identifiable domains or homology, but deletion of this non-essential gene causes a similar genetic interaction profile to that seen with *tgs1*Δ, a protein that trimethylates snoRNA and spliceosomal snRNA caps (Mouaikel et al., 2002). Both *TGS1* and *YNR004W* cluster with components of the spliceosome (Supplemental Figure 1), and a two-hybrid interaction has been reported for the corresponding pair of proteins (Uetz et al., 2000). Strikingly, we have found that deletion of *YNR004W* causes a nearly identical splicing defect to that observed with *tgs1*Δ (Supplemental Figure 4). These data suggest that Ynr004w functions directly with Tgs1 to ensure proper tri-methylation of snRNA caps. A number of other currently unnamed open reading frames included in the E-MAP display strong patterns of genetic interactions suggestive of roles in specific processes, providing numerous directions for further experimentation.

In order to facilitate further data-mining by the RNA community, we have developed an interactive and searchable web-based database (http://interactome-cmp.ucsf.edu). One can query single mutants for individual genetic interactions or for other factors that share similar genetic interaction profiles. This database is integrated with the results from previous E-MAPs, allowing a wealth of information from different fields to be cross-referenced. Additionally, an interface for navigating the higher-level analysis of interactions between complexes has been developed, and is linked to from within the database (http://www.cellcircuits.org/complexes/RNAProcessing/html/). We anticipate that

this dataset will be used as a resource by others to launch more detailed investigations into the biology of RNA processing.

Experimental Procedures

E-MAP analysis

Strains were constructed and E-MAP experiments were performed as previously described (Collins et al., 2007b; Collins et al., 2006; Schuldiner et al., 2005).

Functional connections between processes and complexes

For analysis of connections between related processes, 446 genes were classified with one of 12 functions, using a combination of GO annotations and manual curation. The median genetic interaction score of protein pairs spanning two processes or within the same process was compared to equal sized random samples of genetic interactions. A cutoff of $p<0.001$ was used for the between-process connections and a cutoff of $p<0.05$ was used for connections within a process. Enrichments for strong (both positive and negative) interactions were computed by comparison of the median of the absolute values of the S-scores to the absolute value of a random distribution with a $p<0.001$. Complexes involved in RNA processing were identified using previously described methods (Bandyopadhyay et al., 2008). This analysis categorizes interactions supported by both strong genetic and physical evidence as operating within a module, or complex. Interactions with a strong genetic but weak physical signal are better characterized as operating

between two functionally related modules. Given within- and between-module likelihoods for individual interactions, an agglomerative clustering procedure seeks to merge these interactions into increasingly larger modules and to identify pairs of modules interconnected by many strong genetic interactions. Pairs of functionally related complexes and genetic interactions within a complex were then categorized through analysis of their pairwise genetic interaction distributions as described above.

Yeast Genetics

Yeast were grown and double deletions made according to standard methods (Guthrie and Fink, 2002). For serial dilution growth tests, cells were grown to log phase and diluted to OD 0.1, and then serially diluted by one fifth. Serial dilutions were transferred onto YPD and incubated at 16°C for 7 days.

In Situ Hybridizations and Microscopy

In situ hybridizations were performed as described previously with a few modifications (Duncan et al., 2000). Briefly, cells were grown to OD 0.15-0.25 in YEPD, and fixed for one hour in 5% formaldehyde, spheroplasted with zymolyase 100T (Seikagaku Corporation), fixed briefly with paraformaldehyde and permeabilized with cold methanol. Bulk polyadenylated RNA was detected by in situ hybridization with approximately 10 ng of digoxigenin-11-dUTP (Roche) labeled oligo dT50 probe followed by staining with anti-digoxigenin-fluroescein Fab fragments (Roche) (1:25). In parallel, cells were treated with secondary antibody

and no probe. Cells were mounted in SlowFade Gold antifade reagent with DAPI (Invitrogen). Images were collected on an Olympus BX-60 microscope with a 100x Olympus UPlanFI NA 1.30 objective and appropriate filters with a Sensys CCD camera (Photometrics) and analyzed using iVision imaging Software (Biovision Technologies). All strains were analyzed at least three times, and all images within a set were treated identically.

Immunoprecipitations

Twenty-five mLs log phase cultures of cells from the GFP tagged collection (Huh et al., 2003) were washed and resuspended in lysis buffer containing 50 mM Tris HCl pH 7.5, 150 mM NaCl, 0.1 % NP40, and protease inhibitors, and lysed by vigorous shaking with glass beads. Immunoprecipitations between Sem1 and Thp1GFP were stable in buffer containing up to at least 1 M NaCl. Extracts were clarified by centrifugation at 10,000 g for 8 minutes, an input sample was removed and the extracts were immunoprecipitated overnight with 1 µL polyclonal antibody to Sem1 (Funakoshi et al., 2004) (generous gift of Hideki Kobayashi), or 2 µg monoclonal antibody to GFP (Roche) followed by a one hour incubation with 30 µL of a 50% slurry of IgG beads (Amersham). The beads were washed 5 times with lysis buffer and resuspended in 1X SDS sample buffer. Western blots were then probed for Sem1 (1:1000), GFP (1:2000), and Rpt6/Sug1 (1:1000) (generous gift of Thomas Kodadek), or HA (12CA5 Roche). For RNAse A treatment, either 25 µg of RNAse A was added to extracts for 10 min at room temperature, or the extracts were mock treated, and a total RNA sample was analyzed by agarose gel electrophoresis to

check degradation of total RNA. All immunoprecipitations were performed at least three times.

UV Crosslinking Assay

Cells were grown to OD 0.8 and subjected to UV crosslinking, extract preparation, and poly(A) RNA purification performed as in (Gilbert et al., 2001). Western blots were then probed for Mex67 (1:10000, generous gift of Catherine Dargement), Sub2 (1:2000, polyclonal antisera specific to Sub2 produced in the lab), and Yra1 (1:2000, generous gift of Douglas Kellogg).

Splicing Microarrays

Splicing microarray sample collection, data analysis, and arrays are as previously described with the following modifications (Pleiss et al., 2007). Briefly, total RNA was purified from wild type and mutant yeast cultures growing in log phase at 30° C, and converted to animo-allyl labeled cDNA, which was then coupled to dyes and competitively hybridized to splicing microarrays. Values represent averaged data collected from dye-flipped technical replicates performed on two biological replicates for the *csn12Δ*, *ypr045cΔ*, *csn9Δ*, and *pci8Δ* arrays. The *isy1Δ* values are derived from a single representative biological replicate. Similarities between splicing microarray profiles were measured by Pearson correlation, calculated on the complete observations (Becker et al., 1988). Experimental factors were then symmetrically arranged along the X- and Y-axis of Figure 4C using hierarchical clustering (de Hoon et al., 2004) of euclidean distances between

correlation values with unweighted centroid linkages. Array data has been deposited in the GEO database (http://www.ncbi.nlm.nih.gov/geo/), accession number GSE11634.

Acknowledgements

We would like to thank H. Kobayashi, T. Kodadek, D. Kellogg, and Catherine Dargement for antibodies, D. Cameron for strains, M. Dinglasin for technical assistance, M. Schuldiner for help with the screens, P. Kemmerman for help with the website, L. Booth for work on *ynr004w*Δ, and V. Panse for sharing unpublished data. We would like to thank members of the Guthrie lab, A. Frankel and J. Gross for reading and advice on the manuscript. G.M.W. was supported by postdoctoral fellowships from the American Cancer Society and the Sandler Foundation. M.B. was supported by an HHMI predoctoral fellowship. C.G. is an American Cancer Society Research Professor of Molecular Genetics. This research was funded by NIH grant GM21119 (C.G.) and from Sandler Family Funding (N.J.K.).

Figures

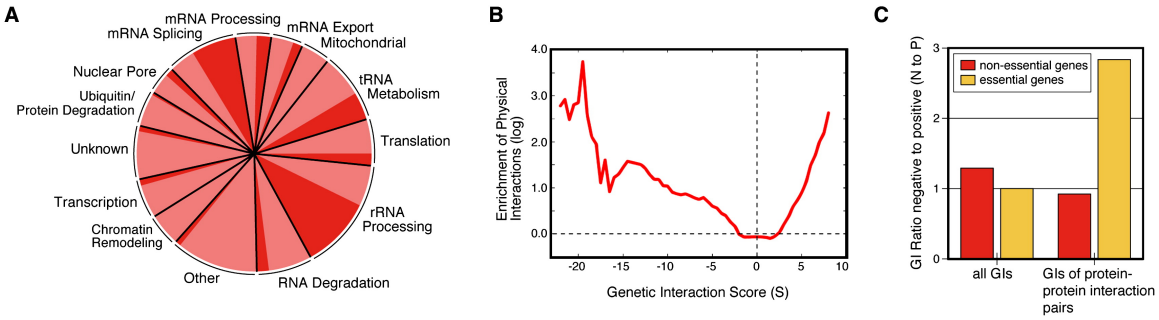


Figure 1. Description of the RNA processing E-MAP. A) Composition of the RNA processing E-MAP. The genetic map is comprised of 552 genes (166 of which are essential) and represents virtually all processes involved in RNA metabolism, including rRNA, tRNA, snoRNA, snRNA and mRNA processing. The dark red portions of the pie chart represent the proportions of the processes that correspond to essential genes. **B)** Comparison of genetic and physical interactions. The graph compares pairs of proteins that are physically associated (PE > 1.0, (Collins et al., 2007a)) and the genetic interaction scores from the corresponding mutants (S-score, (Collins et al., 2006)). Based on a comparison to randomized genetic interaction scores, protein pairs with strong positive and negative genetic interactions are more likely to be physically associated. **C**) Comparisons of ratios of negative (S ≤ -2.5) to positive (S ≥ 2.0) (N to P) genetic interactions. The graph is divided into two parts: 1) all genetic interactions from the RNA processing E-MAP and 2) only those from pairs of genes whose corresponding proteins are physically associated (PE > 1.0, (Collins et al., 2007a)). Red bars correspond to genetic interactions derived entirely from deletions of pairs of non-essential genes (n=

88,879 total (left), 110 physically interacting (right)), while yellow bars are derived from pairs including one (n=49,380 total ,73 physically interacting) or two (n= 639 total, 4 physically interacting) DAmP alleles of essential genes. The overrepresentation of negative genetic interactions among pairs of genes that include an essential gene and have physical interactions was found to be highly significant using Fisher's exact test, with a two-tailed p-value of 3x10-4. The trend was not strongly dependent on using various different thresholds for defining negative and positive interactions (data not shown). GIs, genetic interactions.

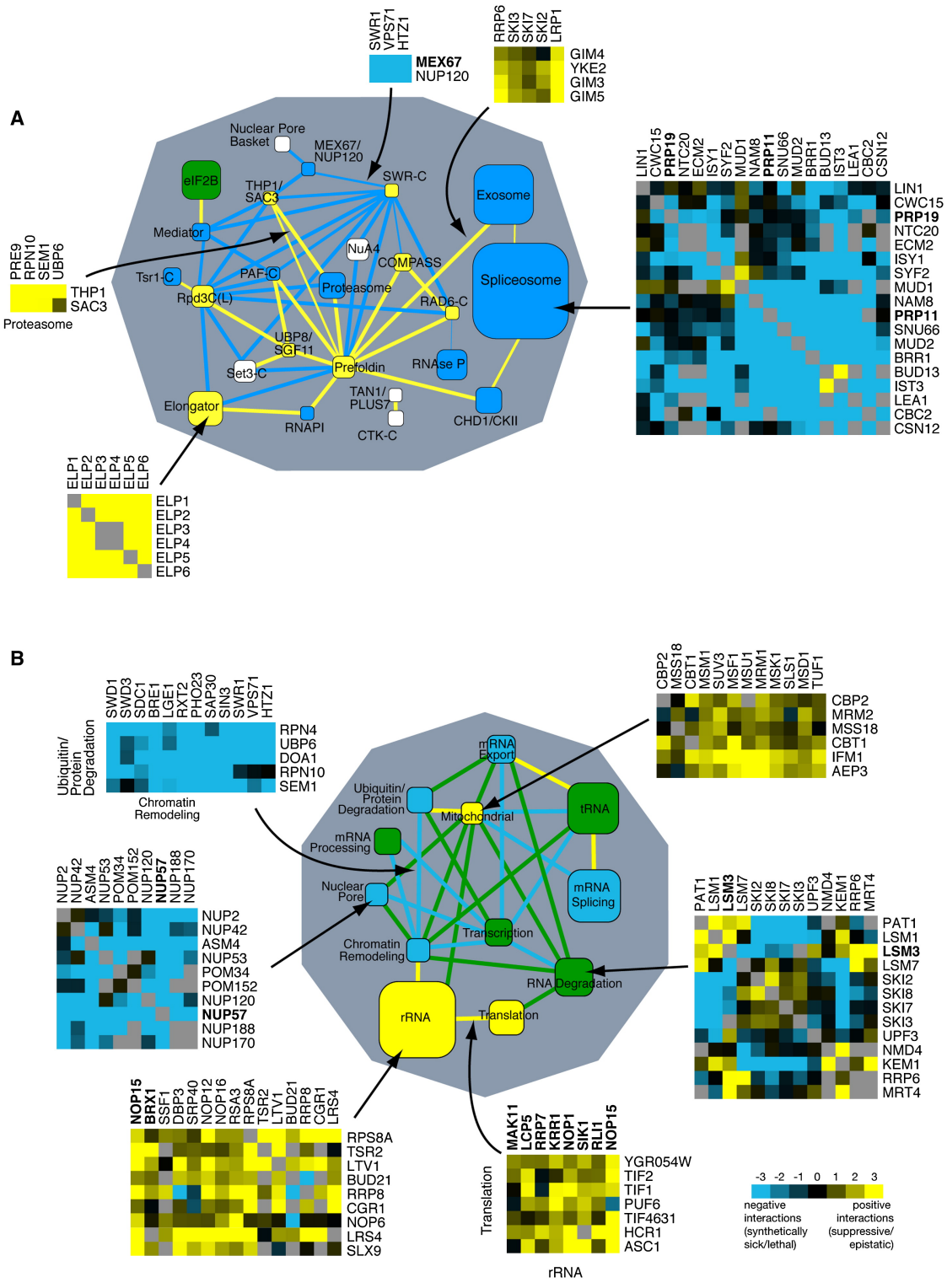


Figure 2. Functional cross-talk between biological processes and protein complexes. Global views of the genetic cross-talk between different RNA-related

protein complexes (**A**) and processes (**B**). Blue and yellow represent a statistically significant enrichment of negative and positive interactions, respectively, whereas green corresponds to cases where there are roughly equal numbers of positive and negative genetic interactions. White corresponds to a lack of significant interactions or lack of data. Nodes (boxes) correspond to distinct protein complexes (**A**) or functional processes (**B**) whereas edges (lines) represent how the complexes and processes are genetically connected. Line thickness represents the significance of the connection. Node size is proportional to the number of genes in the process or complex. Essential genes (DAmP alleles) are in bold. Representative genetic interactions which contributed to the overall enrichment for interactions are shown for sample nodes and edges, according to the scale shown, with grey boxes representing missing data points. See Experimental Procedures for a description of how the networks are generated.

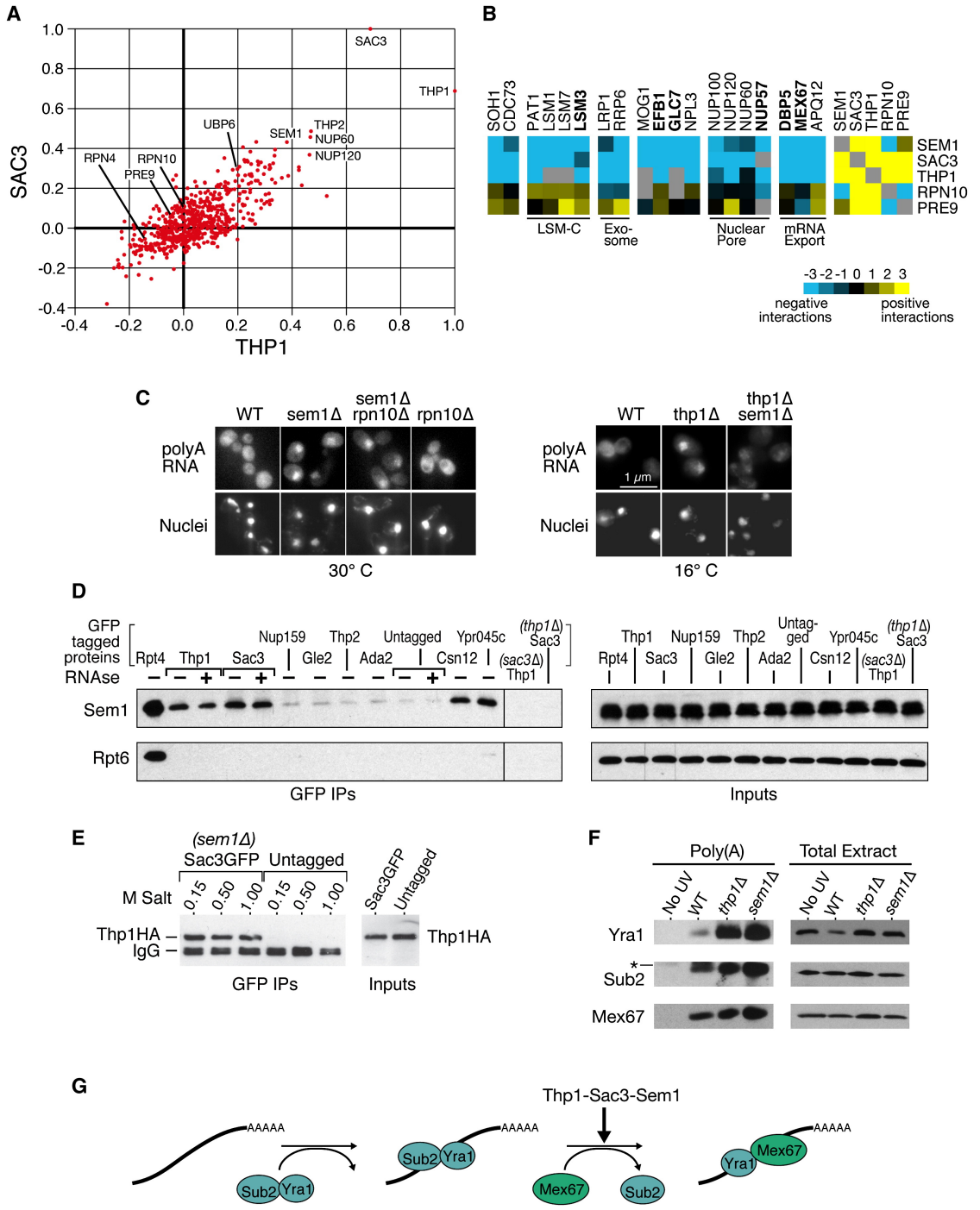


Figure 3. Sem1 is involved in mRNA export via the Sac3-Thp1 complex. A) Scatter plot of correlation coefficients for each mutant compared to the profiles generated from *SAC3* (y-axis) and *THP1* (x-axis). **B)** Representative genetic interactions from the EMAP that differentiate *SEM1*, *SAC3*, and *THP1* from *RPN10*

and *PRE9*. Blue and yellow indicate negative and positive genetic interactions, respectively. **C)** In situ hybridizations with a dT50 probe to detect accumulation of poly(A) RNA (top row). The bottom row is DAPI staining to detect nuclei. Cells were either kept at permissive temperature (30°C, right) or shifted to 16°C for 2 hours before fixation (left). **D)** Co-immunoprecipitations (IPs) of Sem1 and Rpt6 with GFP-tagged proteins. The indicated strains were IPed with a monoclonal GFP antibody, either with or without prior RNAse A treatment of the extract, and the blot was cut and probed with polyclonal antibodies against Sem1 or Rpt6. The right panel is 1/200th of the sample for the GFP IPs exposed identically to the left. **E)** Co- IPs of Thp1-6HA with Sac3GFP in a *sem1Δ* background. The IPs were performed in buffer with 150 mM NaCl, and then washed 4 times at room temperature for 10 minutes each wash in buffers containing the indicated amount of salt. The right panel represents1/20th the sample for the GFP IPS. **F)** In vivo UV crosslinking of proteins to poly(A) RNA. Each lane of the poly(A) eluates contains equal amounts of purified poly(A) RNA. The "no UV" strain contains Thp1-6HA. The * in the Sub2 blot represents a nonspecific cross-reacting band. The left and right panels were exposed differently, and the Yra1 poly(A) blot was overexposed to allow visualization of the Yra1 band in the wild-type strain. **G)** Depiction of the export block identified by the poly(A) crosslinking in part F. Sac3-Thp1-Sem1 could facilitate an exchange of Sub2 for Mex67. See text for more details.

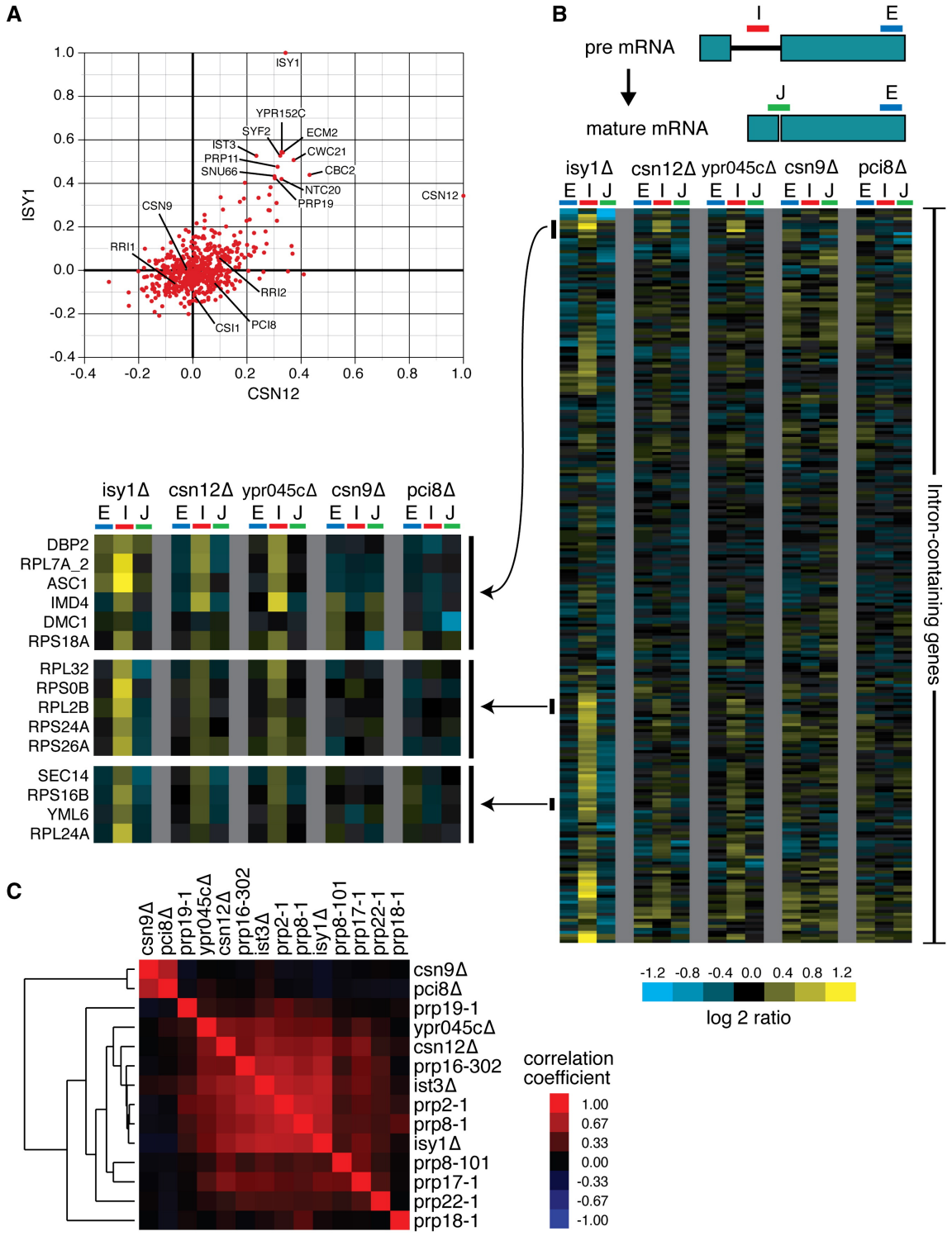


Figure 4. Csn12 is involved in mRNA splicing. (**A**) Plot of correlation coefficients generated from comparison of the genetic interaction profiles from *csn12*Δ or *isy1*Δ to all other profiles in the E-MAP. **B**) Splicing-specific microarray profiles for several

mutant strains. The schematic displays the positions of the microarray probes that report specifically on the levels of pre-mRNA (in the Intron), mature mRNA (at the Junction), and total mRNA (in the second Exon) for each intron-containing transcript. The relative levels of exon, intron, and junction for a single intron-containing gene are displayed as log2 ratios for the indicated mutant strains compared to a wild type strain, across each row. The ordering of genes was determined by hierarchical clustering. For selected clusters of genes, the splicing profiles across the mutants tested are displayed at higher resolution to the left of the full splicing profiles. **C**) Pairwise Pearson correlation coefficients were calculated between each of the mutants tested, as well as between these mutants and several previously characterized splicing mutants. The matrix of correlations was subjected to hierarchical clustering.

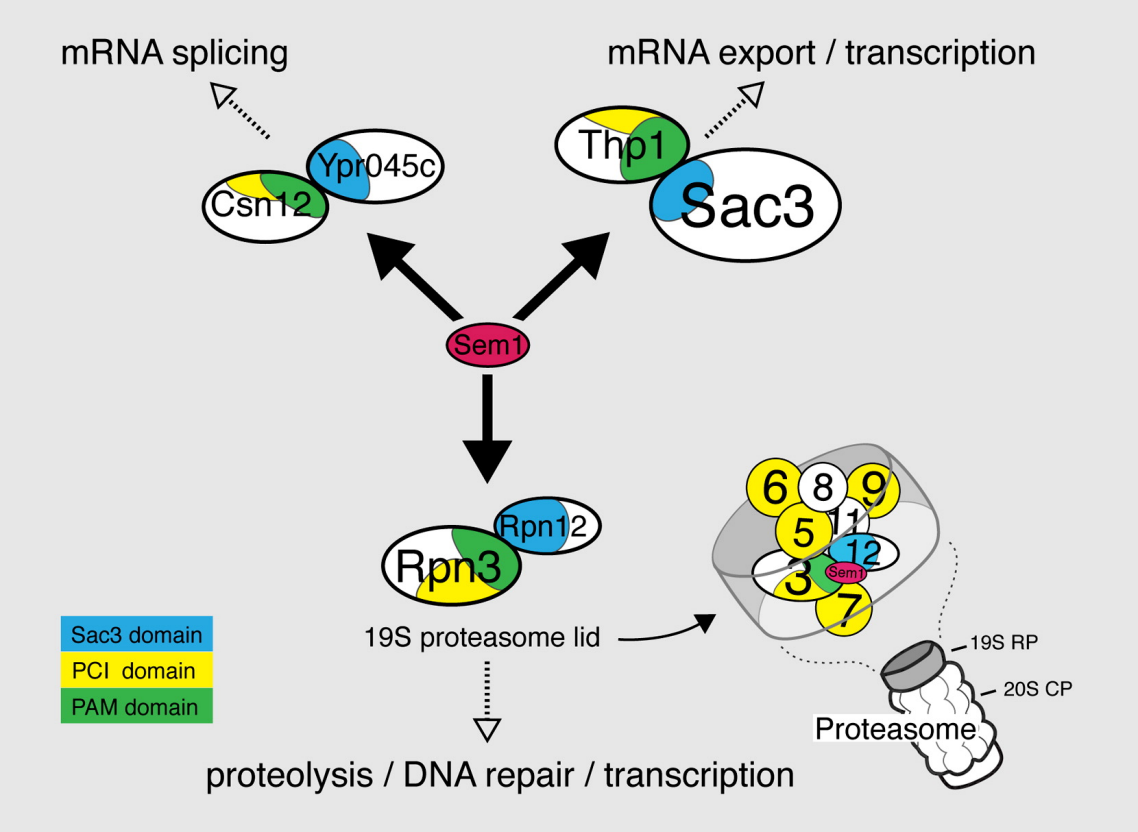


Figure 5. Model for three Sem1-containing complexes. 19S RP= Regulatory Particle of the proteasome, which includes both the lid (diagrammed here), and the base. 20S CP=Core Particle of the proteasome. The schematic of protein organization in the proteasome lid is based on the model from Sharon et al (Sharon et al., 2006). The "SAC3" domains (PFAM PF03399) are amino acids: Rpn12 20-211, Ypr045c 205-408, and Sac3 248-443, and the PAM domains (Ciccarelli et al., 2003): Rpn3 204-378, Csn12 136-343, and Thp1 151-332. Also shown are the PCI domains (PFAM PF01399) associated with the PAM domains: Rpn3 343-447, Csn12 298-415, and Thp1 300-430.

Supplementary Figures

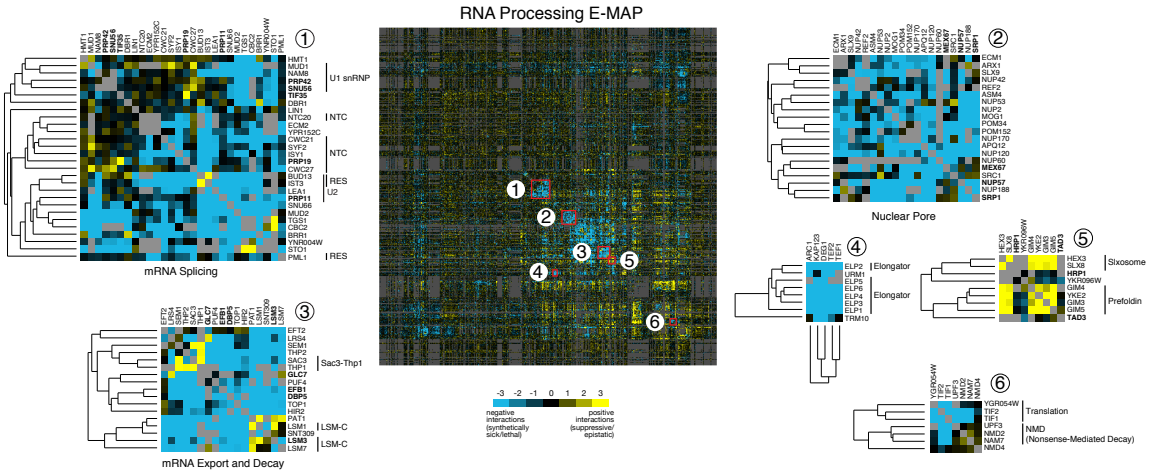


Figure S1. Hierarchical clustergram of all 552 genes in the RNA Processing E-MAP. Numbered pullouts represent subclusters with processes and protein complexes labeled. Essential genes are in bold. Blue and yellow represent negative and positive genetic interactions, respectively, whereas grey corresponds to data that was not collected.

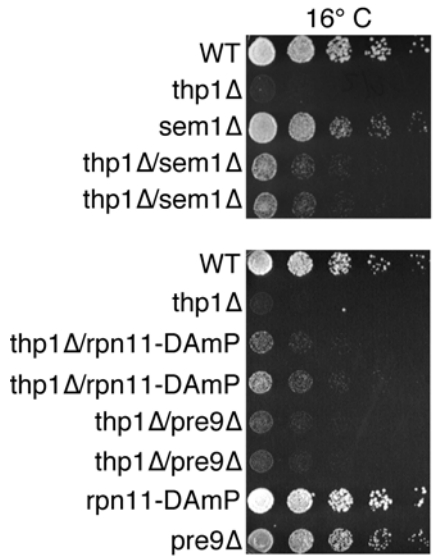


Figure S2. Verification of positive genetic interactions derived from the E-MAP. Five-fold serial dilutions of wild-type, single and double deletion cells spotted onto YPD and grown at 16°C for 7 days.

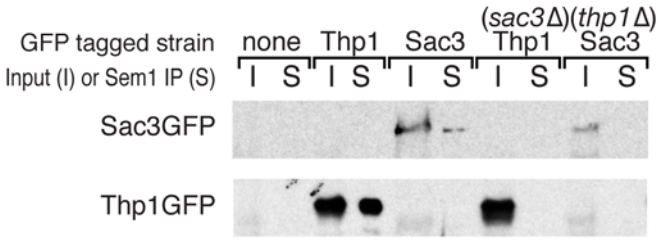


Figure S3. Sac3GFP and Thp1GFP co-immunoprecipitate with Sem1 if both Sac3 and Thp1 are present in the strain. Sem1 was immunoprecipitated with a polyclonal antibody in the indicated strains and the blots were probed for GFP. Inputs represent 1/20th of the IP sample.

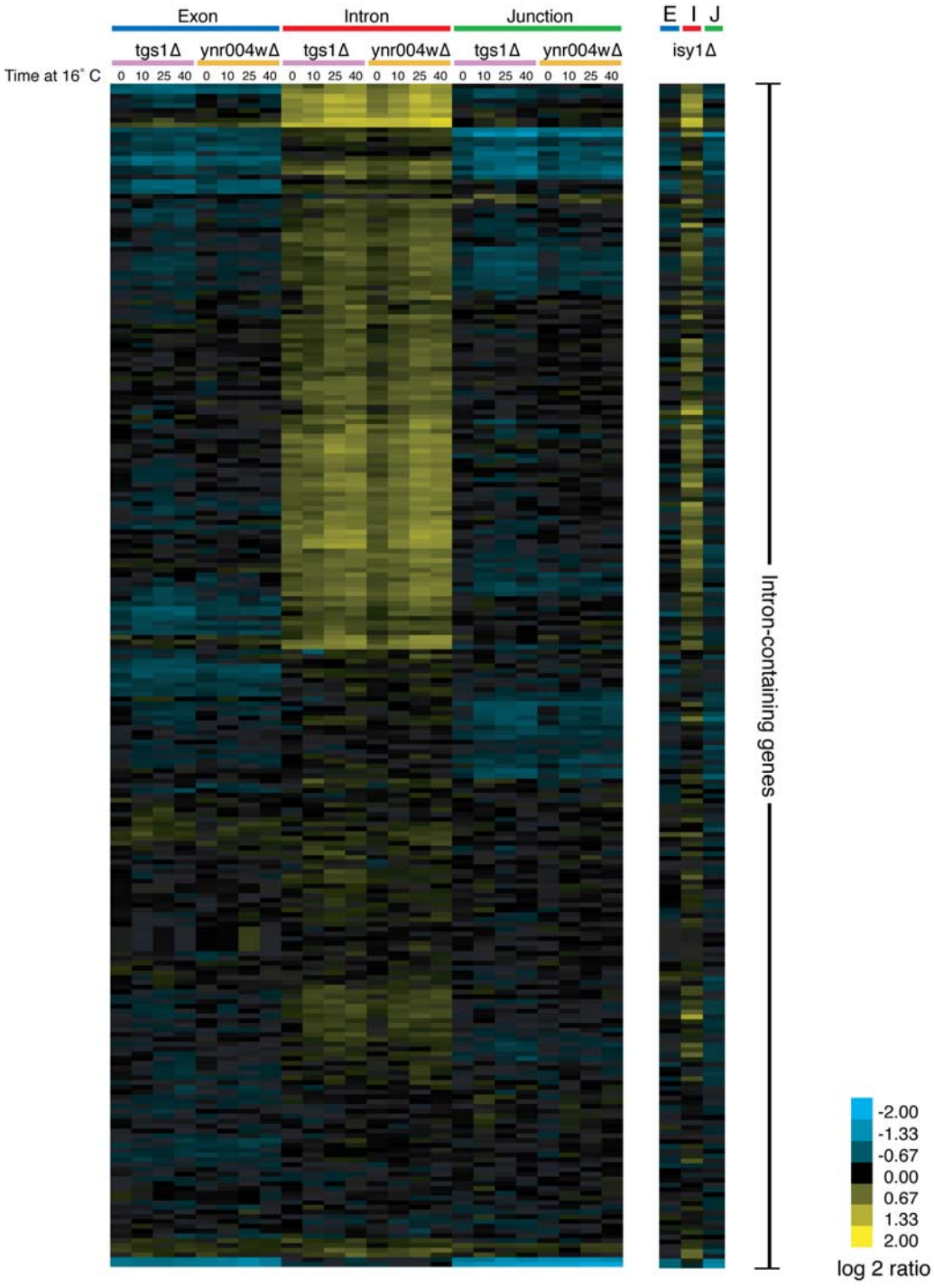
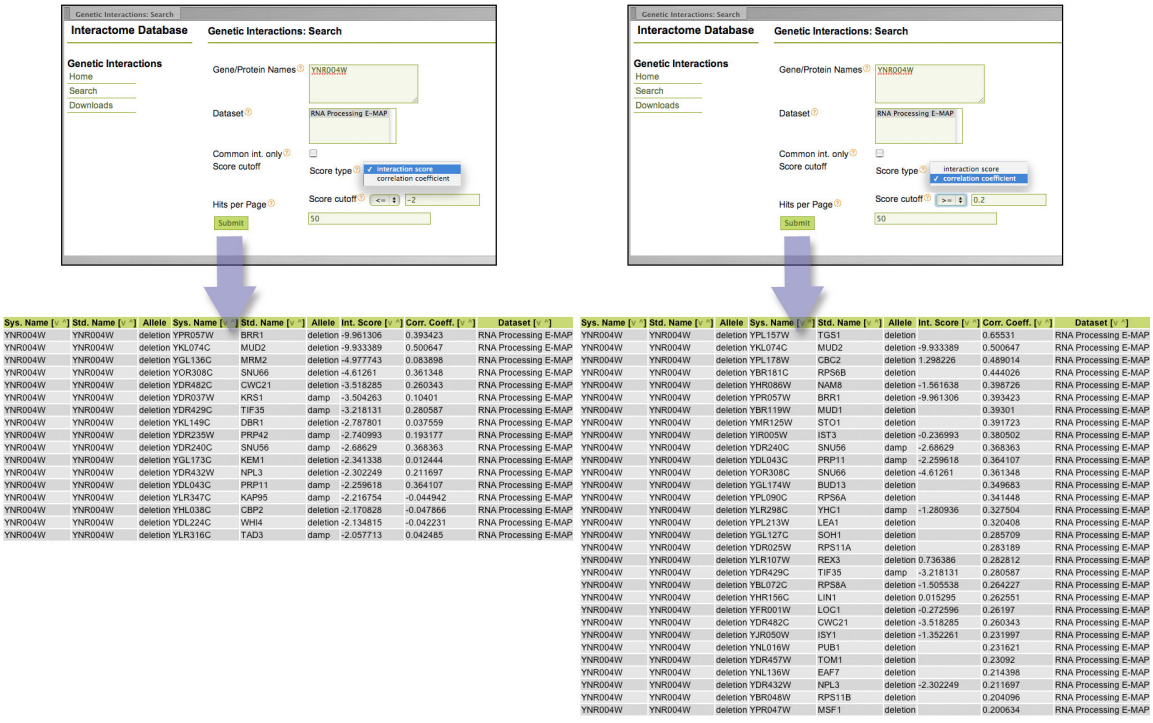


Figure S4. Tgs1 and Ynr004w are required for normal mRNA splicing. Splicing-specific microarray profiles were generated for *tgs1*Δ and *ynr004w*Δ, as described in figure 4 and Experimental Procedures, with the exception that this experiment

involves a temperature shift. *tgs1*Δ and *ynr004w*Δ cultures grown at 30° C were shifted to 16°C and samples were collected at the indicated times (in minutes) and hybridized against samples collected from a wild type strain also shifted to 16°C. Genes are ordered by hierarchical clustering of this dataset. The splicing profile for the splicing mutant *isy1*Δ strain in log phase at 30°C (as described in figure 4) is included here for reference, with the genes ordered to match the order in the *tgs1*Δ and *ynr004w*Δ profile.



Sys. Name	Std. Name	Allele	Sys. Name	Std. Name	Allele	Int. Score	Corr. Coeff.	Dataset
YNR004W	YNR004W	deletion	YPR057W	BRR1	deletion	-9.961306	0.393423	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YKL074C	MUD2	deletion	-9.933389	0.500647	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YGL136C	MRM2	deletion	-4.977743	0.083898	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YOR308C	SNU66	deletion	-4.61261	0.361348	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR482C	CWC21	deletion	-3.518285	0.260343	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR037W	KRS1	damp	-3.504263	0.10401	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR429C	TIF35	damp	-3.218131	0.280587	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YKL149C	DBR1	deletion	-2.787801	0.037559	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR235W	PRP42	damp	-2.740993	0.193177	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR240C	SNU56	damp	-2.68629	0.368363	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YGL173C	KEM1	deletion	-2.341338	0.012444	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR432W	NPL3	deletion	-2.302249	0.211697	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDL043C	PRP11	damp	-2.259618	0.364107	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YLR347C	KAP95	deletion	-2.216754	-0.044942	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YHL038C	CBP2	deletion	-2.170828	-0.047866	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDL224C	WHI4	deletion	-2.134815	-0.042231	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YLR316C	TAD3	damp	-2.057713	0.042485	RNA Processing E-MAP

Sys. Name	Std. Name	Allele	Sys. Name	Std. Name	Allele	Int. Score	Corr. Coeff.	Dataset
YNR004W	YNR004W	deletion	YPL157W	TGS1	deletion		0.65531	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YKL074C	MUD2	deletion	-9.933389	0.500647	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YPL178W	CBC2	deletion	1.298226	0.489014	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YBR181C	RPS6B	deletion		0.444026	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YHR086W	NAM8	deletion	-1.561638	0.398726	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YPR057W	BRR1	deletion	-9.961306	0.393423	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YBR119W	MUD1	deletion		0.39301	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YMR125W	STO1	deletion		0.391723	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YIR005W	IST3	deletion	-0.236993	0.380502	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR240C	SNU56	damp	-2.68629	0.368363	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDL043C	PRP11	damp	-2.259618	0.364107	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YOR308C	SNU66	deletion	-4.61261	0.361348	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YGL174W	BUD13	deletion		0.349683	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YPL090C	RPS6A	deletion		0.341448	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YLR298C	YHC1	damp	-1.280936	0.327504	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YPL213W	LEA1	deletion		0.320408	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YGL127C	SOH1	deletion		0.285709	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR025W	RPS11A	deletion		0.283189	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YLR107W	REX3	deletion	0.736386	0.282812	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR429C	TIF35	damp	-3.218131	0.280587	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YBL072C	RPS8A	deletion	-1.505538	0.264227	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YHR156C	LIN1	deletion	0.015295	0.262551	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YFR001W	LOC1	deletion	-0.272596	0.26197	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR482C	CWC21	deletion	-3.518285	0.260343	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YJR050W	ISY1	deletion	-1.352261	0.231997	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YNL016W	PUB1	deletion		0.231621	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR457W	TOM1	deletion		0.23092	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YNL136W	EAF7	deletion		0.214398	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YDR432W	NPL3	deletion	-2.302249	0.211697	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YBR048W	RPS11B	deletion		0.204096	RNA Processing E-MAP
YNR004W	YNR004W	deletion	YPR047W	MSF1	deletion		0.200634	RNA Processing E-MAP

Figure S5. Representative screenshots from the web-based database that allows for the query of individual genetic interactions and/or profiles from specific mutants derived from the genetic data from the RNA Processing E-MAP.

Supplementary Tables

Available online at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2644724/

References

Baillat, D., Hakimi, M. A., Naar, A. M., Shilatifard, A., Cooch, N., and Shiekhattar, R. (2005). Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell *123*, 265-276.

Bandyopadhyay, S., Kelley, R., Krogan, N. J., and Ideker, T. (2008). Functional maps of protein complexes from quantitative genetic interaction data. PLoS Comput Biol *4*, e1000065.

Becker, R. A., Chambers, J. M., and Wilks, A. R. (1988). The new S language : a programming environment for data analysis and graphics (Pacific Grove, Calif., Wadsworth & Brooks/Cole Advanced Books & Software).

Brickner, D. G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P. C., Widom, J., and Brickner, J. H. (2007). H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol *5*, e81.

Chan, S. P., Kao, D. I., Tsai, W. Y., and Cheng, S. C. (2003). The Prp19p-associated complex in spliceosome activation. Science *302*, 279-282.

Chen, C. H., Yu, W. C., Tsao, T. Y., Wang, L. Y., Chen, H. R., Lin, J. Y., Tsai, W. Y., and Cheng, S. C. (2002). Functional and physical interactions between components of the Prp19p-associated complex. Nucleic Acids Res *30*, 1029-1037.

Ciccarelli, F. D., Izaurralde, E., and Bork, P. (2003). The PAM domain, a multi-protein complex-associated module with an all-alpha-helix fold. BMC Bioinformatics *4*, 64.

Clark, T. A., Sugnet, C. W., and Ares, M., Jr. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science *296*, 907-910.

Collins, G. A., and Tansey, W. P. (2006). The proteasome: a utility tool for transcription? Curr Opin Genet Dev *16*, 197-202.

Collins, S. R., Kemmeren, P., Zhao, X. C., Greenblatt, J. F., Spencer, F., Holstege, F. C., Weissman, J. S., and Krogan, N. J. (2007a). Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics *6*, 439-450.

Collins, S. R., Miller, K. M., Maas, N. L., Roguev, A., Fillingham, J., Chu, C. S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M*., et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S. R., Schuldiner, M., Krogan, N. J., and Weissman, J. S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol *7*, R63.

de Hoon, M. J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.

Duncan, K., Umen, J. G., and Guthrie, C. (2000). A putative ubiquitin ligase required for efficient mRNA export differentially affects hnRNP transport. Curr Biol *10*, 687-696.

Dziembowski, A., Ventura, A. P., Rutz, B., Caspary, F., Faux, C., Halgand, F., Laprevote, O., and Seraphin, B. (2004). Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. Embo J *23*, 4847-4856.

Enenkel, C., Lehmann, A., and Kloetzel, P. M. (1998). Subcellular distribution of proteasomes implicates a major location of protein degradation in the nuclear envelope-ER network in yeast. Embo J *17*, 6144-6154.

Esberg, A., Huang, B., Johansson, M. J., and Bystrom, A. S. (2006). Elevated levels of two tRNA species bypass the requirement for elongator complex in transcription and exocytosis. Mol Cell *24*, 139-148.

Ezhkova, E., and Tansey, W. P. (2004). Proteasomal ATPases link ubiquitylation of histone H2B to methylation of histone H3. Mol Cell *13*, 435-442.

Ferdous, A., Gonzalez, F., Sun, L., Kodadek, T., and Johnston, S. A. (2001). The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II. Mol Cell *7*, 981-991.

Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R.*, et al.* (2006). Pfam: clans, web tools and services. Nucleic Acids Res *34*, D247-251.

Fischer, T., Strasser, K., Racz, A., Rodriguez-Navarro, S., Oppizzi, M., Ihrig, P., Lechner, J., and Hurt, E. (2002). The mRNA export machinery requires the novel Sac3p- Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. Embo J *21*, 5843 5852.

Fu, H., Reis, N., Lee, Y., Glickman, M. H., and Vierstra, R. D. (2001). Subunit interaction maps for the regulatory particle of the 26S proteasome and the COP9 signalosome. Embo J *20*, 7096-7107.

Funakoshi, M., Li, X., Velichutina, I., Hochstrasser, M., and Kobayashi, H. (2004). Sem1, the yeast ortholog of a human BRCA2-binding protein, is a component of the proteasome regulatory particle that enhances proteasome stability. J Cell Sci *117*, 6447-6454.

Furukawa, K., Mizushima, N., Noda, T., and Ohsumi, Y. (2000). A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes. J Biol Chem *275*, 7462-7465.

Gallardo, M., Luna, R., Erdjument-Bromage, H., Tempst, P., and Aguilera, A. (2003). Nab2p and the Thp1p-Sac3p complex functionally interact at the interface between transcription and mRNA metabolism. J Biol Chem *278*, 24225-24232.

Gautier-Bert, K., Murol, B., Jarrousse, A. S., Ballut, L., Badaoui, S., Petit, F., and Schmid, H. P. (2003). Substrate affinity and substrate specificity of proteasomes with RNase activity. Mol Biol Rep *30*, 1-7.

Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dumpelfeld, B.*, et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B.*, et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387-391.

Gilbert, W., Siebel, C. W., and Guthrie, C. (2001). Phosphorylation by Sky1p promotes Npl3p shuttling and mRNA dissociation. Rna *7*, 302-313.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M.*, et al.* (1996). Life with 6000 genes. Science *274*, 546, 563-547.

Gonzalez, F., Delahodde, A., Kodadek, T., and Johnston, S. A. (2002). Recruitment of a 19S proteasome subcomplex to an activated promoter. Science *296*, 548-550.

Gottschalk, A., Tang, J., Puig, O., Salgado, J., Neubauer, G., Colot, H. V., Mann, M., Seraphin, B., Rosbash, M., Luhrmann, R., and Fabrizio, P. (1998). A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. Rna *4*, 374-393.

Guthrie, C., and Fink, G. R. (2002). Guide to yeast genetics and molecular and cell biology. Part B (San Diego, Calif., Academic Press).

Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K.*, et al.* (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180-183.

Hofmann, K., and Bucher, P. (1998). The PCI domain: a common theme in three multiprotein complexes. Trends Biochem Sci *23*, 204-205.

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., and O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686-691.

Isono, E., Saito, N., Kamata, N., Saeki, Y., and Toh, E. A. (2005). Functional analysis of Rpn6p, a lid component of the 26 S proteasome, using temperature-sensitive rpn6 mutants of the yeast Saccharomyces cerevisiae. J Biol Chem *280*, 6537-6547.

Jackman, J. E., Montange, R. K., Malik, H. S., and Phizicky, E. M. (2003). Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9. Rna *9*, 574-585.

Kobor, M. S., Venkatasubrahmanyam, S., Meneghini, M. D., Gin, J. W., Jennings, J. L., Link, A. J., Madhani, H. D., and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol *2*, E131.

Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P.*, et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Krogan, N. J., Keogh, M. C., Datta, N., Sawa, C., Ryan, O. W., Ding, H., Haw, R. A., Pootoolal, J., Tong, A., Canadien, V.*, et al.* (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell *12*, 1565-1576.

Krogan, N. J., Lam, M. H., Fillingham, J., Keogh, M. C., Gebbia, M., Li, J., Datta, N., Cagney, G., Buratowski, S., Emili, A., and Greenblatt, J. F. (2004). Proteasome involvement in the repair of DNA double-strand breaks. Mol Cell *16*, 1027-1034.

Lei, E. P., Stern, C. A., Fahrenkrog, B., Krebber, H., Moy, T. I., Aebi, U., and Silver, P. A. (2003). Sac3 is an mRNA export factor that localizes to cytoplasmic fibrils of nuclear pore complex. Mol Biol Cell *14*, 836-847.

Li, J., Zou, C., Bai, Y., Wazer, D. E., Band, V., and Gao, Q. (2006). DSS1 is required for the stability of BRCA2. Oncogene *25*, 1186-1194.

Luna, R., Gaillard, H., Gonzalez-Aguilera, C., and Aguilera, A. (2008). Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus. Chromosoma.

Lund, M. K., and Guthrie, C. (2005). The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. Mol Cell *20*, 645-651.

Mannen, T., Andoh, T., and Tani, T. (2008). Dss1 associating with the proteasome functions in selective nuclear mRNA export in yeast. Biochem Biophys Res Commun *365*, 664-671.

Maytal-Kivity, V., Pick, E., Piran, R., Hofmann, K., and Glickman, M. H. (2003). The COP9 signalosome-like complex in S. cerevisiae and links to other PCI complexes. Int J Biochem Cell Biol *35*, 706-715.

Mizuguchi, G., Shen, X., Landry, J., Wu, W. H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonne, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. Mol Cell *9*, 891-901.

Nakai, Y., Nakai, M., and Hayashi, H. (2008). Thio-modification of Yeast Cytosolic tRNA Requires a Ubiquitin-related System That Resembles Bacterial Sulfur Transfer Systems. J Biol Chem *283*, 27469-27476.

Neubauer, G., Gottschalk, A., Fabrizio, P., Seraphin, B., Luhrmann, R., and Mann, M. (1997). Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A *94*, 385-390.

Ohi, M. D., and Gould, K. L. (2002). Characterization of interactions among the Cef1p-Prp19p-associated splicing complex. Rna *8*, 798-815.

Otero, G., Fellows, J., Li, Y., de Bizemont, T., Dirac, A. M., Gustafsson, C. M., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. (1999). Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. Mol Cell *3*, 109-118.

Pan, X., Yuan, D. S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J. S., Hieter, P., Spencer, F., and Boeke, J. D. (2004). A robust toolkit for functional profiling of the yeast genome. Mol Cell *16*, 487-496.

Pleiss, J. A., Whitworth, G. B., Bergkessel, M., and Guthrie, C. (2007). Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS Biol *5*, e90.

Rodriguez-Navarro, S., Fischer, T., Luo, M. J., Antunez, O., Brettschneider, S., Lechner, J., Perez-Ortin, J. E., Reed, R., and Hurt, E. (2004). Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. Cell *116*, 75-86.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S. R., Qu, H., Shales, M., Park, H. O., Hayles, J.*, et al.* (2008). Conservation and Rewiring of Functional Modules Revealed by an Epistasis Map in Fission Yeast. Science.

Roguev, A., Wiren, M., Weissman, J. S., and Krogan, N. J. (2007). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat Methods *4*, 861-866.

Scheel, H., and Hofmann, K. (2005). Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. BMC Bioinformatics *6*, 71.

Schmitz, J., Chowdhury, M. M., Hanzelmann, P., Nimtz, M., Lee, E. Y., Schindelin, H., and Leimkuhler, S. (2008). The sulfurtransferase activity of Uba4 presents a link between ubiquitin-like protein conjugation and activation of sulfur carrier proteins. Biochemistry *47*, 6479-6489.

Schuldiner, M., Collins, S. R., Thompson, N. J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J. F.*, et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Segre, D., Deluna, A., Church, G. M., and Kishony, R. (2005). Modular epistasis in yeast metabolism. Nat Genet *37*, 77-83.

Sharon, M., Taverner, T., Ambroggio, X. I., Deshaies, R. J., and Robinson, C. V. (2006). Structural organization of the 19S proteasome lid: insights from MS of intact complexes. PLoS Biol *4*, e267.

Sims, R. J., 3rd, Millhouse, S., Chen, C. F., Lewis, B. A., Erdjument-Bromage, H., Tempst, P., Manley, J. L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and premRNA splicing. Mol Cell *28*, 665-676.

Sone, T., Saeki, Y., Toh-e, A., and Yokosawa, H. (2004). Sem1p is a novel subunit of the 26 S proteasome from Saccharomyces cerevisiae. J Biol Chem *279*, 28807-28816.

Strasser, K., and Hurt, E. (2001). Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. Nature *413*, 648-652.

Thakurta, A. G., Gopal, G., Yoon, J. H., Kozak, L., and Dhar, R. (2005). Homolog of BRCA2-interacting Dss1p and Uap56p link Mlo3p and Rae1p for mRNA export in fission yeast. Embo J *24*, 2512-2523.

Tong, A. H., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C. W., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Tong, A. H., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G. F., Brost, R. L., Chang, M.*, et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P.*, et al.* (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature *403*, 623-627.

Wang, Q., He, J., Lynn, B., and Rymond, B. C. (2005). Interactions of the yeast SF3b splicing factor. Mol Cell Biol *25*, 10745-10754.

Wee, S., Hetfeld, B., Dubiel, W., and Wolf, D. A. (2002). Conservation of the COP9/signalosome in budding yeast. BMC Genet *3*, 15.

Wei, S. J., Williams, J. G., Dang, H., Darden, T. A., Betz, B. L., Humble, M. M., Chang, F. M., Trempus, C. S., Johnson, K., Cannon, R. E., and Tennant, R. W. (2008). Identification of a Specific Motif of the DSS1 Protein Required for Proteasome Interaction and p53 Protein Degradation. J Mol Biol.

Wilkinson, C. R., Wallace, M., Morphew, M., Perry, P., Allshire, R., Javerzat, J. P., McIntosh, J. R., and Gordon, C. (1998). Localization of the 26S proteasome during mitosis and meiosis in fission yeast. Embo J *17*, 6465-6476.

Yang, H., Jeffrey, P. D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N. H., Zheng, N., Chen, P. L., Lee, W. H., and Pavletich, N. P. (2002). BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. Science *297*, 1837-1848.

Chapter 4

A lipid E-MAP identifies Ubx2 as a critical regulator of lipid saturation and lipid bilayer stress

A lipid E-MAP identifies Ubx2 as a critical regulator of lipid saturation and lipid bilayer stress

Michal A. Surma[1,2], Christian Klose[1], Debby Peng[2], Michael Shales[2], Caroline Mrejen[3], Adam Stefanko[4], Hannes Braberg[2], David Gordon[2], Daniela Vorkel[1], Christer S. Ejsing[4], Robert Farese Jr.[3], Kai Simons[1], Nevan J. Krogan[2,3+] and Robert Ernst[1,5+]

1 Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

2 Department of Cellular and Molecular Pharmacology, California Institute for Quantitative Biosciences, University of California, San Francisco, CA, 94158, USA

3 J. David Gladstone Institutes, San Francisco, CA, 94158,USA

4 Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark

5 Goethe University Frankfurt, Institute of Biochemistry, 60438 Frankfurt, Germany

[+] To whom correspondence should be addressed.

nevan.krogan@ucsf.com and ernst@em.uni-frankfurt.edu

Summary

Biological membranes are complex and the mechanisms underlying their homeostasis are incompletely understood. Here, we present a quantitative genetic interaction map (E-MAP) focused on various aspects of lipid biology including their metabolism, sorting, and trafficking. This E-MAP contains ~250,000 genetic interaction scores, both negative and positive, and identifies a molecular crosstalk of protein quality control pathways with lipid bilayer homeostasis. Ubx2p, a component of the ER-associated degradation (ERAD) pathway, surfaces as a key upstream regulator of the essential fatty acid desaturase Ole1p. Loss of Ubx2p affects the transcriptional control of *OLE1* resulting in impaired fatty acid desaturation and a severe shift towards more saturated membrane lipids. Both the induction of the unfolded protein response and aberrant nuclear membrane morphologies observed in cells lacking *UBX2* are suppressed by supplementation of unsaturated fatty acids. Our results point towards the existence of dedicated bilayer stress responses for membrane homeostasis.

Introduction

A biological membrane is a defining feature of all cellular life forms. Cell membranes possess a staggering complexity of lipids and proteins that cooperate to compartmentalize biochemical processes and allow for selective exchange of molecules and information (van Meer et al., 2008). Lipids play a central role in shaping the physicochemical environment of the membrane, which in turn affects the membrane proteins (Phillips et al., 2009). The molecular packing of lipids is the underlying principle of phase behavior and membrane fluidity. Cellular lipids exhibit a large fraction of kinked and 'poorly' packing unsaturated fatty acids (UFAs) to increase fluidity of the membrane at physiological temperatures (van Meer et al., 2008). Generally, saturated fatty acid (SFA)-containing lipids can pack with higher order and have higher melting temperatures. Not surprisingly, lipid saturation is subjected to crucial regulation in response to changes of temperature and other conditions for cellular growth (Klose et al., 2012).

Membrane proteins evolved with the surrounding membrane lipids to achieve optimal biological activity but also contribute to the general structural properties of biological membranes (Kaiser et al., 2011). Systematic interrogations of cellular pathways elucidated a crosstalk between lipid metabolism and protein quality control systems such as the unfolded protein response (UPR) and the ER-associated degradation (ERAD) pathway, responsible for removing aberrant proteins from the endoplasmic reticulum (Jonikas et al., 2009; Schuldiner et al., 2005). These findings have placed fresh emphasis on the regulation of lipid metabolism and provided a new perspective to the field of protein homeostasis. An

important role of fatty acids (FAs) in cellular stress responses has been recognized both in *Saccharomyces cerevisiae* and mammals (Eizirik and Cnop, 2010; Kohlwein and Petschnigg, 2007). A misbalance between SFAs and UFAs in the diet and, consequently, in cells induces ER-stress and can lead to cell death (Kim et al., 2008). Conversely, a massive induction of ER-stress is counter-balanced by increased utilization of FAs for lipid synthesis and expansion of the ER (Bernales et al., 2006; Schuck et al., 2009). This introduces a conundrum surrounding the nature by which ER-stress is modulated by lipid species.

Functional relationships between the different cellular processes involving lipids and other factors might be uncovered through genetic- interaction mapping of the genes of interest. Genetic interactions have been employed for decades to evaluate the functional relationships of genes by identifying synthetic sickness/lethality. More recently, this type of analysis achieved a new level of predictive power with the development of epistatic miniarray profiles (E-MAPs) (Roguev et al., 2008; Schuldiner et al., 2005). Here, a large set of double mutants is generated and their growth is measured using colony size. Both negative (aggravating) and positive (alleviating) pairwise epistatic relationships between genes are determined and expressed quantitatively as S-scores by calculating the deviations of the observed growth rate of each double mutant from the expected growth rate in absence of a genetic interaction. Negative genetic interactions occur if a double mutant is less fit than expected (synthetic sickness/lethality), whereas positive genetic interactions occur if a double mutant is fitter than expected

(alleviation). The frequency of strong genetic interactions between randomly chosen genes is low, but much higher for functionally related genes.

Using E-MAP data, functional groups of genes can be identified through the unbiased hierarchical clustering of genetic interaction profiles, but also patterns of interactions with all other genes in the study can be used to gain the insight into gene functions (Beltrao et al., 2010). E-MAPs have proven as a powerful technique to discover gene functions, to identify structural complexes, and to organize them into pathways (Aguilar et al., 2010; Collins et al., 2007b; Hoppins et al., 2011; Schuldiner et al., 2005).

Here we present an E-MAP focused on *S. cerevisiae* genes involved in lipid metabolism, -sorting, post-Golgi trafficking and related processes, which complements previous E-MAPs that focused on the early secretory pathway and plasma membrane/endocytosis (Aguilar et al., 2010; Schuldiner et al., 2005). Combined with quantitative lipidomics, we uncover an intriguing connection between FA desaturation, the UPR and the ERAD machinery. Ubx2, a central component of the ERAD pathway (Neuber et al., 2005; Schuberth and Buchberger, 2005), emerges as a key upstream regulator of the single, essential FA desaturase of *S. cerevisiae*, Ole1 (Stukey et al., 1990). Our results identify a molecular link between protein degradation and FA desaturation and reveal new insight into the regulation of membrane homeostasis.

Results

Generation of a lipid E-MAP

To study the inter-connections of lipid metabolism and protein homeostasis, we used a quantitative genetic interaction map, or E-MAP, to functionally interrogate a set of 741 genes (corresponding to ~13% of the *S. cerevisiae* genome), 133 of which were essential (Source data 1). The genes, selected based on gene ontology terms, represent several functional categories, including trafficking and metabolism of lipids (Figure 1A). In total, we identified 251,383 genetic interactions (Source data 2). As a quality control, we checked the correlation between replicate measurements where the query and array strains were flipped (i.e. A-B vs. B-A), hence providing a strict test, as such flipped replicates originate from independently created double mutants. Our replicate genetic scores are highly correlated (r=0.661) (Figure 1B), more so than other datasets that have been generated (Collins et al., 2007b; Costanzo et al., 2010; Schuldiner et al., 2005; Wilmes et al., 2008), implying a genetic interaction map of high quality (Figure S1A). Nevertheless, the measured interactions are highly correlated with these previous studies (Figure S1B), but not redundant. Out of the entire dataset, ~50% of the interactions (134,059) have not been measured in previous screens (Figure 1C).

Our results are consistent with previous reports of a correlation between physically associated proteins and the corresponding genes displaying similar interaction profiles (Collins et al., 2007b) (Figure S1C). With a set of 90 high confident protein-protein interactions (PPIs) (Source data 3) (Babu et al., 2012; Collins et al., 2007a), we compared the correlation of genetic profiles of the

corresponding genes from this E-MAP and from a larger screen (Costanzo et al., 2010). Even though our E-MAP covers only 741 genes compared to several thousand from the larger screen, almost all the profiles from the PPI pairs are more highly correlated within our dataset (Figure 1D; S1D), again implying a high quality genetic interaction dataset.

Hierarchical clustering of the genetic interaction profiles recapitulates known pathways and complexes and identifies previously unrecognized connections (Figure 1E, Source data 4). For example, we identified clusters of genes involved in ER protein folding and β-1-6-glucan synthesis, vacuolar delivery, and peroxisome function. Interestingly, the sphingolipid biosynthesis cluster, which includes *TSC10*, *TSC13* and *PHS1*, also contains a component of RSC chromatin remodeling complex (*RSC3*) (Lorch et al., 1998), thereby suggesting a link between chromatin biology and sphingolipid metabolism.

This wealth of genetic interaction data on diverse aspects of cell biology with many more evident connections can be harnessed for in depth interrogation and rigorous biochemical characterization (supporting material).

Uncovering potential regulators of phosphatidylcholine synthesis

Our lipid E-MAP uncovered novel candidates for regulation of the well-characterized pathways for phosphatidylcholine (PC) biosynthesis. The dataset includes genes involved in the two parallel branches of PC synthesis (Figure 2A): *PSD1*, *CHO2* and *OPI3* (phosphatidylserine (PS) to PC conversion pathway) and *PCT1*, *CKI1*, *CPT1* (Kennedy pathway), whose genetic interaction profiles clustered

separately and showed strong negative interaction between each other (Figure 2B,C). Negative interactions between genes from two branches of a biochemical pathway result from both branches contributing to the same pool of critically important products – in this case PC species.

SNF4 encodes a subunit of AMP-activated Snf1 kinase complexes, critical for gene regulation in response to glucose, but has not been implicated in the modulation of PC synthesis. Inspection of the clustered E-MAP data identified highly negative interaction scores and clustering of *SNF4* with genes from the PS to PC pathway (Figure 2C), but not with genes from the Kennedy pathway (Figure 2C). Because yeast lipidomes are remodeled in response to glucose (Klose et al., 2010), these data implicated *SNF4* specifically in the regulation of PS to PC conversion, suggesting a link between glucose sensing and glycerolphospholipid synthesis by an yet uncharacterized mechanism.

Similarly, our data connect the triacylglycerol lipase encoded by the *TGL2* gene with the Kennedy pathway, as *TGL2* specifically clustered with Kennedy pathway genes (Figure 2B). Moreover, *TGL2* exhibited highly negative genetic interactions with *CHO2* and *OPI3*, the genes from the other branch, which is typical for parallel pathways (Figure 2C). Since diacylglyceroles (DAGs) are the products of Tgl2, and DAGs are required for the Kennedy pathway, it is likely that Tgl2 serves as a DAG donor for PC- synthesis. *TGL1*, *TGL3* and *TGL4*, additional acylglycerol lipase encoding genes, did not show similar genetic interactions, suggesting a specific function of Tgl2 for the Kennedy pathway. This interpretation is substantiated by the subcellular localization of these enzymes: Tgl2 and Cpt1 (the next-in-pathway

enzyme) localize to mitochondria, while Tgl1, Tgl3 and Tgl4 localize to different organelles (Henry et al., 2012). Together, our analysis suggests that Tgl2 generates DAG for the *de novo* synthesis of PC. These examples of *SNF4* and *TGL2* illustrate the potential of this E-MAP to identify novel candidates for various cellular functions.

Genetic interaction between protein degradation and FA desaturation

Our E-MAP highlighted an intricate connection between the protein quality control machinery and FA desaturation. Specifically, we found a strong negative genetic interaction between the *UBX2* gene, encoding a component of the ERAD pathway (Neuber et al., 2005), and *OLE1*, the single and essential FA desaturase gene of *S. cerevisiae* (Figure 2D-F). In the same network of genetic interactions, we also found *MGA2* and *SPT23*, encoding two transcriptional regulators that cooperatively mediate and modulate the expression of *OLE1* (Figure 2F) (Kandasamy et al., 2004). These regulators are ER-membrane proteins and have to be activated in order to drive *OLE1* transcription (Figure 2G). To this end, Mga2 and Spt23 become ubiquitylated and processed by the action of the proteasome and the Cdc48 ATPase along with its ubiquitin binding cofactors Npl4 and Ufd1 (Hoppe et al., 2000; Rape et al., 2001; Shcherbik and Haines, 2007). Partial degradation of the membrane embedded precursor (P120) releases a soluble and active transcriptional activator (P90) from the ER that translocates to the nucleus to drive the expression of *OLE1* (Figure 2G).

The OLE pathway shares several similarities and components with the ERAD pathway, where Ubx2 acts as a membrane anchor for Cdc48 and recruits

ubiquitylated, misfolded proteins to their site of membrane extraction (Neuber et al., 2005). Therefore, a strong genetic interaction of UBX2 with IRE1 and HAC1, essential to mount an UPR, was not surprising (Ron and Walter, 2007) (Figure 2E,F). However, our genetic data imply an equally important function of Ubx2 for the modulation of FA desaturation and the pronounced negative genetic interactions of *UBX2* with *SPT23* and *OLE1* raised the question to what extend *UBX2* contributes to FA metabolism and membrane biogenesis (Figure 2D-F).

We directly tested a role of *UBX2* for *OLE1* expression and regulation of the Ole1 level (Figure 2H,I). Exponentially growing *ubx2*Δ and *mga2*Δ cells exhibited a reduced mRNA level of *OLE1* (Figure 2H), resulting in a 40% and 50% drop of the Ole1 enzyme level compared to wildtype cells (Figure 2I). While a reduction by a factor of two may appear modest, one should bear in mind that *OLE1* is an essential gene and that its products, UFAs, are abundant building blocks for membrane biogenesis.

The lipid composition of a desaturase-depleted cell

We employed quantitative, mass-spectrometry based shotgun lipidomics to study the impact of a mildly reduced Ole1 level on the membrane lipid composition (Ejsing et al., 2009). The lipidomes of *ubx2*Δ or *mga2*Δ cells were strikingly different from wild type controls (Figure 3A-C; S2). In *ubx2*Δ cells, we observed a systematic shift towards more saturated lipids manifested by a dramatic increase of glycerophospholipids with two saturated acyl chains (Double bonds (DB)=0) at the expense of glycerophospholipids with two monounsaturated acyl chains (DB=2).

This effect was even more pronounced in *mga2*Δ cells but unexpectedly absent in *spt23*Δ cells (Figure 3A). An increased fraction of saturated lipids was observed with varying degree in all glycerophospholipid classes (Figure 3B; S3). The observed effect was remarkably specific and not phenocopied by the deletion of *HRD1* or *DOA10*, the two major E3 ubiquitin ligases of the ERAD system (Vembar and Brodsky, 2008) or by the deletion of other UBX-domain encoding genes (*UBX3* to *UBX7*) (Figure S4). This is especially noteworthy because of the overlapping functions of Ubx2 and Ubx4 in the ERAD pathway (Alberts et al., 2009) and because the tested UBX-domain containing proteins all exert their diverse cellular functions via the recruitment of ubiquitylated substrates to the abundant ATPase Cdc48 (Schuberth and Buchberger, 2008). Like Ubx2, Ubx7 also localizes to the ER and interacts with ERAD components (Goder et al., 2008). This emphasizes the specific function of Ubx2 and illustrates a remarkable modularity of the ERAD system.

Cellular adaptations to perturbed FA metabolism caused by UBX2 deletion

Saturated and tightly packing glycerophospholipids have a major impact on the biophysical properties of a membrane. How does a cell cope with a highly increased fraction of saturated acyl moieties? In comparison to WT cells, *ubx2*Δ cells displayed significant changes in virtually all lipid classes (Figure 3C, Supp. Table 6), which may in part compensate for the severe changes in the FA composition. Even the levels of ergosterol and sphingolipids (MIPC, MIP_2C) were significantly perturbed, although these lipids do not incorporate UFAs, suggesting widespread homeostatic effects. Notably, the increased abundance of lipids with

relatively small head-groups, such as PA and PE at the expense of PC, could compensate for intrinsic curvature stress imposed by the reduced volume occupied by the saturated acyl chains in the hydrophobic core of the membrane (Figure 3C). Thus, the *ubx2*Δ lipidome exhibited a systematic shift towards more saturated membrane lipids, but at the same time showed major rearrangements in lipid classes to maintain membrane shape and fluidity.

A complex interplay of regulatory factors determines the level of Ole1.

Membrane homeostasis requires a tight regulation of the Ole1 level. Because we assessed the level of Ole1 expressed from its endogenous locus and promoter, we could identify Mga2 as the dominant factor determining the desaturase level in exponentially growing cells (Figure 2I). Surprisingly, the lack of Spt23, the other transcriptional regulator, did not interfere with the expression of *OLE1* as efficiently as the loss of Mga2 or Ubx2 (Figure 2H) and barely affected the resulting desaturase level (Figure 2I). Spt23 can compensate for the lack of Ubx2 or Mga2 when cells resume from the stationary phase, but upon increased demand for *de novo* lipid synthesis, Ubx2 and Mga2 become critical to yield high levels of Ole1 (Figure S5)

As a feedback control, excess dietary UFAs abrogate the expression of *OLE1* and destabilize the mRNA of *OLE1* (Kandasamy et al., 2004). In presence of dietary linoleate (18:2), the *OLE1* mRNA level dropped significantly in wild type and *ubx2*Δ, *mga2*Δ, and *spt23*Δ cells (Figure 4A). Thus, none of these three genes encodes the sole sensor for the regulation of the OLE pathway. Consistent with the UFA-induced drop of mRNA abundance, the Ole1 enzyme level was markedly reduced in presence

of linoleate (18:2) in wild type cells (Figure 4B). Intriguingly, the dynamic range of this response was damped in *ubx2*Δ cells, implying that also a post- translational mechanism might be at work. We conclude that Ubx2 is required for normal *OLE1* expression and that Mga2 is the dominant factor controlling the expression of *OLE1* in growing cell.

The processing and activation of Spt23 and Mga2 via the OLE pathway (Figure 2G) shares components and mechanistic cues with the ERAD pathway. Because Cdc48 along with its ubiquitin-binding co-factors Ufd1 and Npl4 have been shown to release active P90 fragments of Spt23 and Mga2 from the ER-membrane (Rape et al., 2001; Shcherbik and Haines, 2007) and Cdc48-Ufd1-Npl4 has been implicated in the processing in Spt23 (Hitchcock et al., 2001), we were interested if Ubx2 might assist this reaction. In *ubx2*Δ cells we observed the accumulation of the membrane-embedded, full-length Spt23 P120 as ubiquitylated species and a mildly reduced level of Spt23 P90 (Figure 4C,D), suggesting a processing defect. No accumulation of P120 was observed in cells lacking one of the two central E3-ubiquitin ligases of the ERAD system (Hrd1/Der3, Doa10) or Ubx4 (Vembar and Brodsky, 2008). A complex of Spt23 and Ubx2 was only detectable when downstream events were blocked by the temperature sensitive *cdc48-3* mutation, implying a very transient interaction in WT cells (data not shown). Together with our lipidomic survey of several UBX-domain containing proteins (Figure S4), the processing of Spt23 reveals a *specific* function of Ubx2 in the regulation of *OLE1* expression and consequently the membrane lipid saturation.

Stationary cells that do not depend on *de novo* lipid biosynthesis stalled the processing of Spt23 into the active P90 form (Figure 4E). When stationary cells resumed growth, they rapidly produced P90 and this was controlled by exogenous UFAs (Figure 4E). In *ubx2*Δ cells, however, the processing of Spt23 was significantly impaired, thereby limiting the response to exogenous UFAs (Figure 4E). Consistent with earlier findings, the Rsp5- dependent ubiquitylation of Spt23 observed in *ubx2*Δ cells was blocked by exogenous linoleate, but less so by oleate (Hoppe et al., 2000) (Figure 4E; S5B), implying that the ubiquitylation reaction *per se* or an upstream event is responsible for stalling the activation of Spt23 in presence of abundant UFAs.

Upon examination of Mga2, the dominant transcriptional regulator, we observed an increased level of both P120 and processed P90 Mga2 in *ubx2*Δ cells, suggesting a role of *UBX2* in controlling the steady state level of Mga2, but not for the processing reaction (Figure 4F). In contrast to Spt23, the proteolytic processing of Mga2 is independent of ubiquitylation by Rsp5 (Shcherbik et al., 2003) and consequently does not seem to require Ubx2 (Figure 2F). Nevertheless, our experiments provide compelling evidence that Ubx2 affects the enzyme level of Ole1 via both transcriptional activators, Spt23 and Mga2.

Like Hmg2, a key enzyme for sterol production, Ole1 is intrinsically unstable and processed via the ERAD pathway (Bays et al., 2001; Braun et al., 2002). When Ole1 was expressed from a *GAL*-promoter, the steady-state level was increased in cells lacking central ERAD components such as *UBX2* or the ubiquitin ligase *HRD1*/*DER3* (Figure 4G) (Bordallo et al., 1998; Hampton et al., 1996). Supporting

this, immunoprecipitation experiments identified a physical interaction of both Ubx2 and Ole1 with the ERAD specific ubiquitin ligase Hrd1/Der3, all expressed at their endogenous levels (Figure 4H). We conclude that Ole1 is degraded by the ERAD system in an Ubx2- dependent manner.

In summary, Ubx2 regulates the level of Ole1 via Mga2 and Spt23 but at the same time by mediating Ole1 degradation via the ERAD pathway.

UPR induction in cells with membrane lipid alternations due to deficiency of UBX2 or MGA2

The deletion of *UBX2* or *MGA2* represents a major perturbation to the lipidome (Figure 3A-C; S2). Does the increase of saturated lipids accompanied by lipidome remodeling result in cellular stress? As with proteins, accumulation of misfolded species in the lumen of the ER induces the UPR (Ron and Walter, 2007). However, perturbations of lipid metabolism can also activate UPR signaling (Han et al., 2010; Kohlwein and Petschnigg, 2007; Pineau et al., 2009; Volmer et al., 2013). Using a FACS-based assay to determine the degree of UPR activation, Jonikas and colleagues identified *MGA2* and *UBX2*/*SEL1* deletions among the most potent inducers of UPR (Jonikas et al., 2009). We adapted this strategy to identify protein-folding stress by driving the expression of GFP from a promoter with four repeats of the unfolded protein response element (UPRE) and validated the assay by titrating DTT to wild type cells (Figure 5A). Our experiments confirmed that *mga2*Δ and *ubx2*Δ cells exhibit a strong activation of the UPR (Figure 5B). Because we had identified a critical function of these genes in the regulation of lipid saturation and

because membrane lipid saturation might directly contribute to UPR signaling (Volmer et al., 2013), we wondered whether the UPR activation could be reverted by UFAs to these cells. While wild type cells and *hrd1/der3*Δ cells were mildly stressed by oleate, the UPR activation of *ubx2*Δ and *mga2*Δ cells was relieved upon UFA feeding (Figure 5B). The lipidomic analysis of wild type and *ubx2*Δ cells grown with oleate showed an indistinguishable yet higher degree of UFA incorporation in membrane and storage lipids (Figure 5C,D). Likewise, almost all differences in the lipid class profile of wild type versus *ubx2*Δ cells (Figure 3C) were abolished upon UFA feeding (Figure 5E). Exogenous oleate overcompensated the limited ability of *ubx2*Δ cells to mediate high-level expression of Ole1 and the resulting deficit of UFAs. This reversible change of the lipid class composition suggested that the lipidome change in *ubx2*Δ cells represented a direct consequence of perturbed FA desaturation, an active adaptation, or both.

Membrane whorls of the outer nuclear membrane due to deficiency of UBX2 or MGA2

The increased level of saturated lipids in *mga2*Δ and *ubx2*Δ cells, as well as the activation of the UPR under UFA-limiting conditions, prompted us to investigate the cellular anatomy of these mutants by electron microscopy. Wild type cells grown to the late-logarithmic phase exhibited normal nuclear morphology (Figure 6A, D). In *ubx2*Δ and *mga2*Δ cells we frequently observed whorl formation by the outer nuclear membrane (Figure 6B,C,E-G), a structure never spotted in wild type cells. Aberrant morphologies were observed in 70 nm sections in more than 10% of

*ubx2*Δ and 30% of *mga2*Δ cells and their size ranged from 100-600 nm (Figure 6F,G). We also observed a putatively early stage of whorl formation where the outer nuclear membrane only protruded from the nucleus (Figure 6E). To establish whether the reduced degree of FA desaturation as observed in *ubx2*Δ and *mga2*Δ was causative for NE proliferation and whorl formation, we grew the cells in presence of oleate (Figure 6H-J). Not a single whorl was observed under these conditions, but an increased content of lipid droplets as consequence of oleate-induced TAG synthesis (Hapala et al., 2011) (Figure 5E). Whorls were observed also in an *ubx2*Δ*mga2*Δ mutant (Figure 6K).

Cells challenged with DTT, a drug interfering with protein maturation in the ER, induces the UPR to cope with an increased load with misfolded proteins (Ron and Walter, 2007). Similarly, *ubx2*Δ and *mga2*Δ cells exhibited a constitutively high activation of the UPR (Figure 5B). While DTT-induced stress resulted in proliferation of the cortical ER (Figure 6L) (Bernales et al., 2006), cells stressed by an overly saturated lipid composition expanded their outer nuclear membrane into whorls. Because DTT acts in the lumen of the ER, but the deletion of *UBX2* or *MGA2* affects the ER-membrane composition, we speculate that the nature and the origin of ER-stress determine the site of ER expansion.

Finally, we examined whether the formation of whorls in cells with augmented FA desaturation was solely driven by the lipid composition or whether it might represent an active cellular process. We blocked the signal propagation of the UPR by deletion of *HAC1* to test whether UPR-signaling was required for whorl formation (Ron and Walter, 2007). Because *ubx2*Δ*hac1*Δ cells are not viable, we

inspected *mga2*Δ*hac1*Δ cells. Although *mga2*Δ*hac1*Δ cells exhibited characteristic lipidome changes as observed in *ubx2*Δ and *mga2*Δ cells (Figure S6), we could not find a single whorl in these cells (Figure 6M). Therefore, whorl formation represents an active process that required a functional UPR.

Discussion

Here, we present a genetic interaction map focused on lipid biology and post-Golgi trafficking. This dataset is accessible for browsing and interpretation (Supplementary file 1 and 2) and reveals novel connections between cellular machineries and pathways that can be used infer future directions of research. Our lipid E-MAP identified novel, putative regulators of PC synthesis. From the wealth of data obtained, we focused here on the function of Ubx2, linking protein homeostasis to lipid metabolism.

Given that the ER is the organelle of membrane protein insertion and a major site for lipid biosynthesis, one may expect that the components of the ER membrane, lipids and proteins, are co-regulated. We show that Ubx2 is a fundamental component of this connection. Its dual role in the ERAD and OLE pathways physically links these processes and emphasizes a new structure-function relationship through which the ER-bioactivity is regulated.

Modes of UFA synthesis regulation

Cells require UFAs for lipid synthesis and the resulting membrane properties are influenced by the fraction of saturated lipids. Therefore, the level of Ole1 is

regulated by several mechanisms (Martin et al., 2007). Ubx2 affects the level of Ole1 substantially by three distinct ways: via Mga2, Spt23, and by mediating the degradation of Ole1. Our data are consistent with Ubx2 acting as substrate-recruiting factor by aiding Mga2 and Spt23 to Cdc48 that releases the soluble P90 fragments from the ER-membrane (Rape et al., 2001; Shcherbik and Haines, 2007). This function of Ubx2 would be analogous to its role in the ERAD pathway (Vembar and Brodsky, 2008). Remarkably, despite that Ubx7 resides in the same membrane and also binds Cdc48 (Goder et al., 2008), it cannot compensate for the loss of Ubx2. Likewise, Ubx4 implicated in ERAD and binding Cdc48 (Alberts et al., 2009) is not involved in the activation of the OLE pathway. The role of Ubx2 to assist the detachment of the transcriptional activators from the ER-membrane is tantalizingly simple, but might not represent the whole story. We show that Mga2, not Spt23, is the dominant factor controlling the Ole1 level, but were puzzled by the observation that both loss and increased levels of Mga2 (as observed in *ubx2*Δ cells) result in reduced levels of Ole1p. Mga2, in contrast to Spt23, has a function beyond transcriptional activation: stabilizing and destabilizing the *OLE1* mRNA (Kandasamy et al., 2004). A soluble expressed P90 fragment of Mga2 could only stabilize, but not destabilize the *OLE1* mRNA, suggesting that this function required the membrane-tethered form of Mga2 (Kandasamy et al., 2004). Therefore, an increase of membrane-associated Mga2, as observed in *ubx2*Δ cells, might destabilize the *OLE1* mRNA and thus decrease the Ole1 level in this mutant. The complex pattern by which Ubx2 modulates *OLE1* expression and the resulting enzyme level illustrates

interwoven and interdependent relationships between lipid metabolism and proteostasis.

Intrinsic mechanisms of lipid bilayer quality control

The biophysical properties of a membrane are a consequence of the membrane composition. A systematic shift towards more saturated lipids results in a marked change of the intrinsic membrane curvature since the hydrophobic part of fully saturated lipids occupies a smaller volume than lipids with one or two unsaturated acyl chains. Our lipidomic data reveal that *ubx2*Δ and *mga2*Δ cells exhibit an increased abundance of lipids with relatively small head-groups, suggesting a mechanism enabling yeast to counteract intrinsic curvature stress and aberrant lipid packing by modulating the quantities of geometrically different lipid classes (Deguil et al., 2011). This type of regulation represents a reverse scenario to adaptations observed in PC- depleted cells that undergo intense acyl chains remodeling (Boumann et al., 2006) and a membrane stress response (Thibault et al., 2012). A recent study identified a role of Ubx2p for TAG-metabolism, lipid droplet size and composition (Wang and Lee, 2012). Because SFAs are generally less efficiently incorporated into TAGs as compared to UFAs (Hapala et al., 2011), the decreased abundance of TAGs in *ubx2*Δ cells might reflect the consequence of decreased FA desaturation that is counteracted by exogenous UFAs (Figure 5E). The fact that the reduced TAG levels in *ubx2*Δ cells were restored by the expression of its mammalian homologue Ubxd8, suggests a conserved function of *UBX2*/Ubxd8 for protein degradation and lipid metabolism from yeast to mammals (Wang and Lee,

2012). Intriguingly, the mammalian homologue of Ubx2, Ubxd8, has been proposed to act as a FA sensor and was implicated in the regulation of sterols by acting on insig-1, allowing a cell to coordinate FA utilization and sterol biosynthesis (Lee et al., 2010). At the same time, Ubxd8 is involved in the regulation of lipid droplet turnover by acting on ATGL, the rate-limiting enzyme for lipolysis (Olzmann et al., 2013). Our results exclude that Ubx2 act as the sole sensor of the OLE pathway. Cells deleted for *UBX2* readily attenuated the *OLE1* expression in presence of dietary UFAs (Figure 4A). Apparently, more than one sensor and interconnected regulatory circuits evolved to maintain membrane composition and functionality.

Lipid bilayer quality control by the UPR

The UPR controls the secretory capacity of a cell (Ron and Walter, 2007). To this end, the UPR reconfigures protein and lipid synthesis, protein folding and quality control pathways. The binding mode of ER-luminal, misfolded proteins to Ire1 is established at the molecular level, but the question how Ire1 senses membrane aberrancies remains enigmatic (Volmer et al., 2013). Truncation mutants of Ire1 lacking the binding domain for misfolded proteins are still modulated by lipids (Promlek et al., 2011). How do perturbations of lipid metabolism such as PC synthesis (Thibault et al., 2011) or sphingolipids induce the UPR (Han et al., 2010)? Consistent with a recent report on the mammalian UPR components IRE1a and PERK (Volmer et al., 2013) we show that an increase of saturated membrane lipids massively induces the UPR. Cells deleted for *UBX2* or *MGA2*, which we know exhibit aberrant FA desaturation, were found among the top

UPR inducing mutants in a screen by Jonikas and colleagues (Jonikas et al., 2009). Intriguingly, we show that the high level of UPR activation can be reversed by exogenous UFAs (Figure 5B). This drastic dependence of the UPR on lipid composition can be explained by the molar predominance of lipids in the ER membrane, but the molecular mechanism by which lipids induce the UPR remains to be studied in further detail (Volmer et al., 2013). Only the functional reconstitution of full-length Ire1 in defined membrane environments can reveal the molecular signal(s) underlying the membrane dependent activation of the UPR.

Whorl formation of the nuclear envelope in desaturase-deficient cells

The ER membrane is an architecturally plastic organelle. It forms a structure of two domains, the outer nuclear membrane as part of the NE and the peripheral ER. The ER can exhibit diverse morphologies: tubes, sheets, stacks, hexagonal shapes, and also whorled ER in concentric or sinusoidal arrangements (Snapp et al., 2003; Voeltz et al., 2002). Such organized smooth ER structures can be induced e.g. by membrane protein overexpression or viral infection. We demonstrated that a shortage of FA desaturation induces the formation of sinusoidal deformations (whorls) of the NE but not of the peripheral ER. This is of note because other types of ER- stress induced by tunicamycin or DTT induce a proliferation of the peripheral ER but not of the NE (Bernales et al., 2006). The nature and origin of stress seems to determine the cellular localization of membrane deformations.

The full reversal of whorls by UFA supplementation establishes a causative role of lipids for their formation, but clearly, an increased fraction of saturated lipids

is not sufficient for ER-whorl formation. Instead, the sinusoidal proliferation of the NE represents an active cellular process that requires a functional UPR (Figure 6M). This suggests that an aberrant lipid composition can induce the UPR, which in turn is required for whorl formation. It is tempting to speculate on whether whorl formation represents a quality control mechanism to sequester aberrant membrane structures for subsequent degradation by an autophagy-like mechanism (Lingwood et al., 2009). Several questions remain: How is whorl formation initiated? What stabilizes the whorl structure and what is their fate? Understanding these questions will help to establish the molecular mechanisms of lipid bilayer quality control complementing the well-known pathways of protein quality control.

Ubx2 is a linchpin for ER-homeostasis

This study uncovered novel cellular functions of Ubx2 in membrane-homeostasis and ER-bioactivity. Lack of Ubx2 resembles the loss of a linchpin and reveals intimately connected pathways as the underlying design principle of ER-homeostasis. Instead of being autonomous processes, the ERAD, FA metabolism, membrane biogenesis, UPR, and ER-morphology are interdependent and evolved to communicate.

Experimental Procedures

Strains and plasmids used in this study are listed in Supplementary file 3. Generation and analysis of the quantitative genetic interaction data using the E-MAP approach was carried out as previously described (Collins et al., 2010). Quantitation

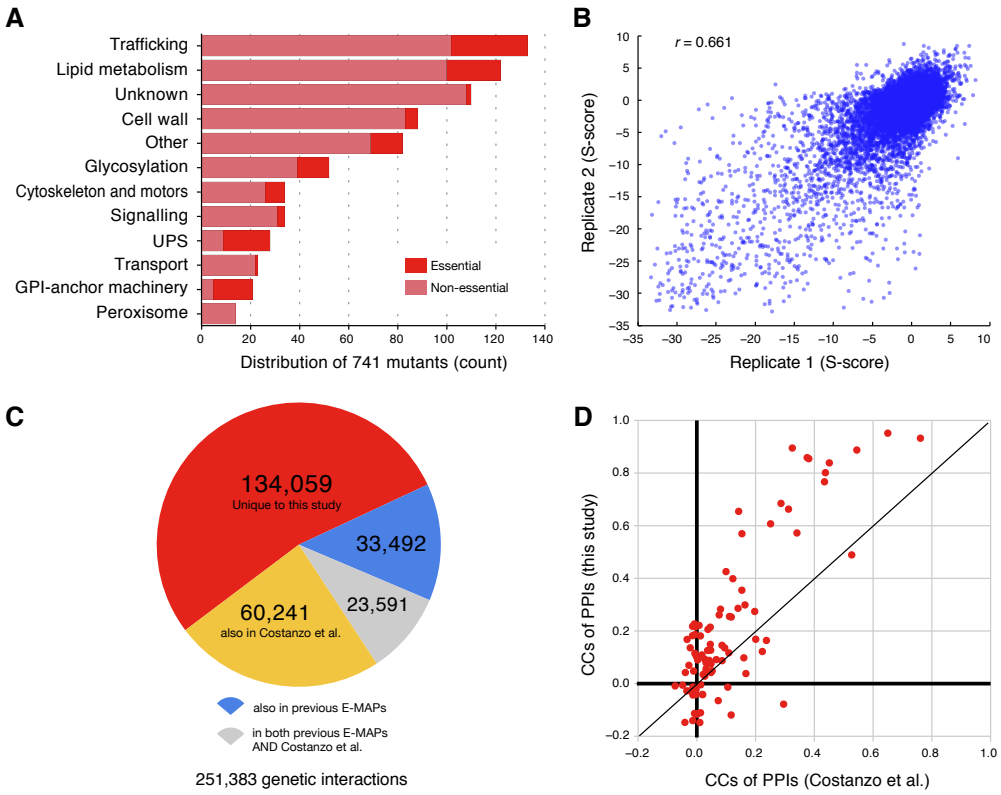
of lipids was performed essentially as described by Klose et al. (Klose et al., 2012) with modifications from the original protocol (Ejsing et al., 2009). For electron microscopy studies, cells were cryoimmobilized by high pressure freezing and, after freeze-substitution in acetone stained with osmium tetroxide und uranly acetate. 70nm sections were collected, contrasted with uranyl acetate. For more details, see supplementary materials and methods.

Acknowledgements

The authors acknowledge Pablo Aguilar, Erin Currie, Mark Hochstrasser, Robin W. Klemm, Stefan Jentsch, Thomas Sommer, Maya Schuldiner, Andrej Shevchenko, Dieter Wolf, Peter Walter, Tobias Walther and Jonathan Weissman for help and/or reagents, and James Saenz and Daniel Lingwood for critically reading the manuscript. R.E. acknowledges Jin Ye, Ray Deshaies and Pedro Carvalho for sharing unpublished data. This work was supported by the DFG (ER608/2-1 and TRR83 TP02), the NIH (GM084448, GM084279, GM081879 and GM098101) to N.J.K., the European Molecular Biology Organization (ALTF 379-2008 to R.E., ASTF 219-2009 to M.S.), the ESF "LIPIDPROD" (SI459/3-1), by Lundbeckfonden (95-310-13591, C.S.E.), the Danish Council for Independent Research | Natural Sciences (09-72484, C.S.E.) and the Klaus Tschira Foundation. N.J.K. is a Searle Scholar and Keck Young Investigator.

Figures

Figure 1



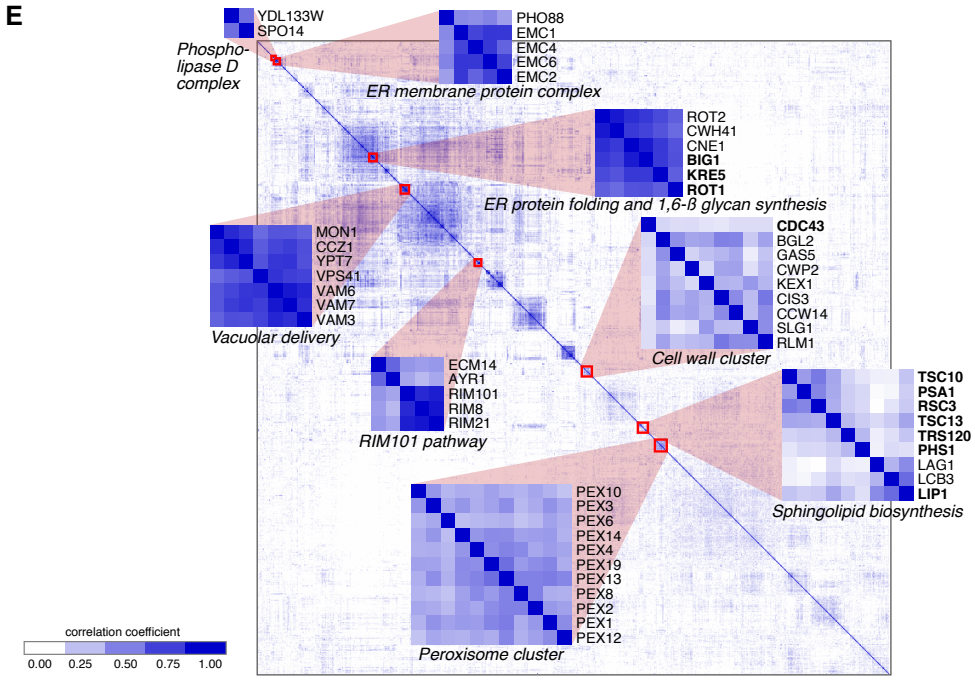


Figure 1. Overview of the E-MAP. A: Functional composition of the lipid E-MAP. Light and dark red indicate non- and essential genes, respectively. **B:** Scatter plot of the replicate scores (i.e. A-B vs. B-A pairs). **C:** Comparison of the genetic interactions identified in this work and other E-MAPs and a more global analysis (see Figure S1 for details). **D:** The correlation of genetic interaction profiles from this study and the Costanzo et al. dataset focusing on genes whose proteins participate in protein-protein interactions. **E:** Hierarchical clustering of the correlation of genetic interaction profiles for each pair of genes from the lipid E-MAP. Negative correlations were used for clustering but are omitted in the clustergram.

Figure 2

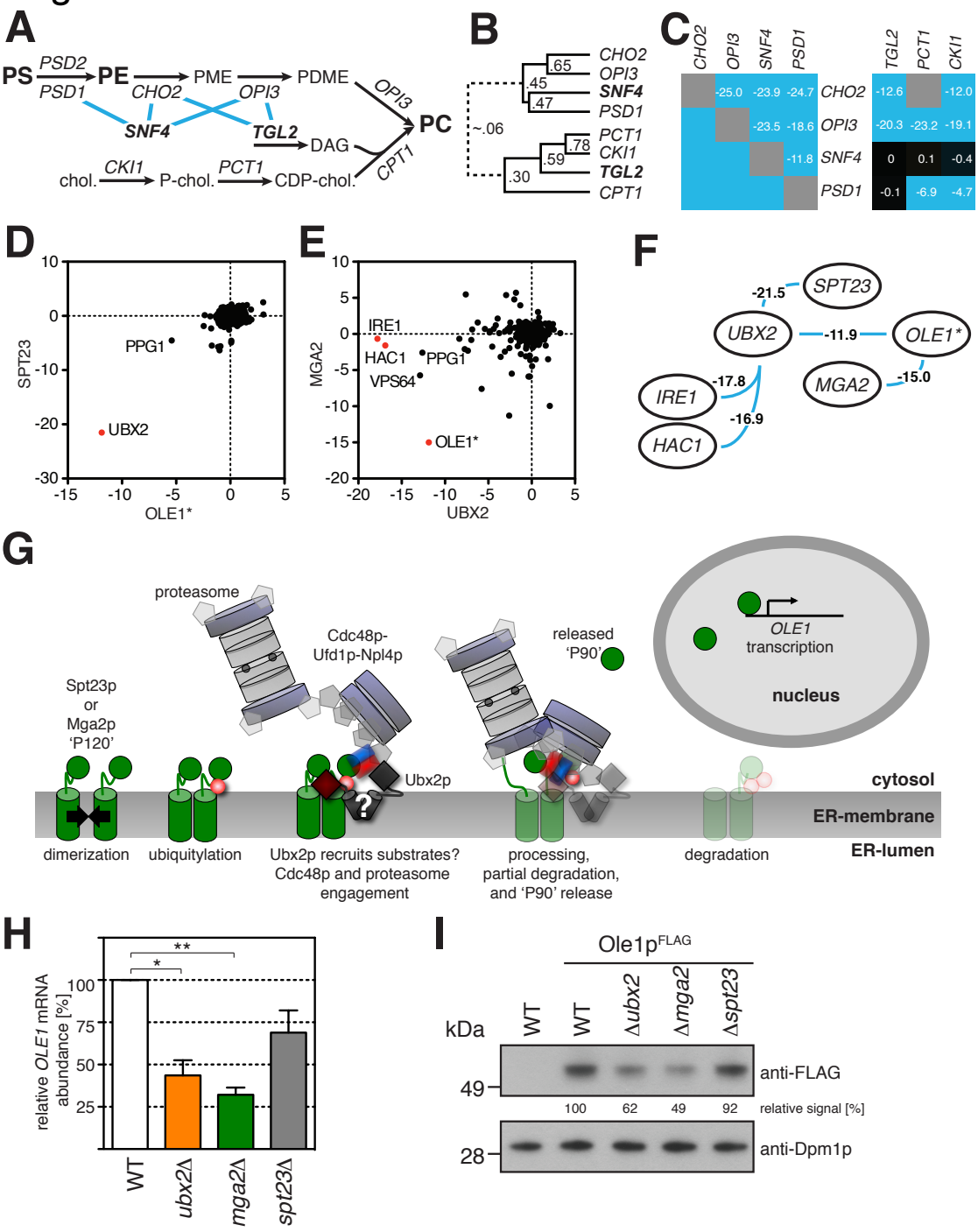


Figure 2. Genetic interactions assign gene functions to specific lipid- metabolic pathways. A: Schematic diagram of the of the PC synthesis pathway. The PS conversion is depicted on top, the *de novo* synthesis (Kennedy pathway) on the

bottom. Blue lines indicate negative genetic interactions. **B:** Clustering of *SNF4*, *TGL2* and PC synthesis-related genes. Numbers are correlation coefficients. The dotted connection indicates a non-significant correlation. **C:** The genetic interaction scores of PC synthesis-related genes with *SNF4* and *TGL2*. **D ,E:** Dot-plot of genetic interactions with genes depicted genes on the X- and Y- axis. **F:** Genetic interactions of *UBX2* highlight the crosstalk of the UPR (*IRE1* and *HAC1*), the ERAD machinery (*UBX2*) and FA-metabolism (*SPT23*, *MGA2* and *OLE1*). *OLE1*, an essential gene, was present as DAmP allele (*). Only significant interactions are shown. **G:** Schematic model of the OLE pathway. **H:** The relative mRNA levels of depicted cells were determined by qPCR. The Statistical analysis involved an unpaired, two-tailed t-test for the genotype and labeled with (*) or (**) for $p<0.05$ or $p<0.01$ ($n=3$), respectively. **I:** Wild type cells and cells deletants, with and without genomically FLAG-tagged Ole1 were grown in YPD to $OD_{600}=1$. Cell lysates were immunoblotted as indicated and the relative signal was determined by densitometry.

Figure 3

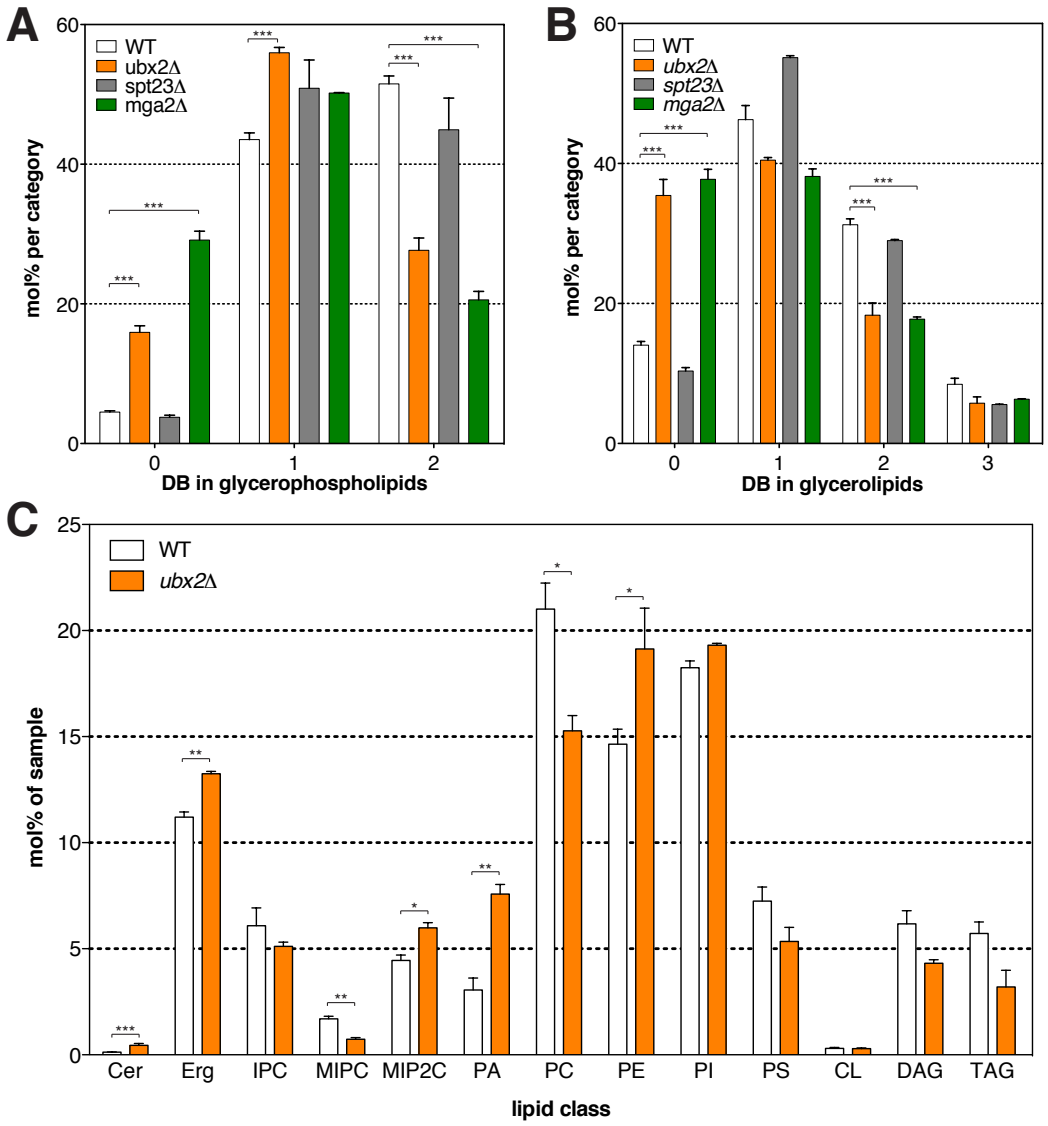


Figure 3. *UBX2* is involved in lipid metabolism. Lipidomes of the analyzed mutants. **A:** Total double bonds (DB) of glycerophospholipids (**GP**; cardiolipin omitted for clarity), given as the sum of double bonds in FAs, in mol% of all glycerophospholipids. **B:** Total double bonds (DB) of glycerolipids (**GL**; DAG and TAG), given as the sum of double bonds in FAs, in mol% of all glycerolipids. **A, B:** Highly significant deviations from WT with $p<0.001$ (unpaired, two-tailed T-test; n=6 for WT; n=2 for *ubx2*Δ and *mga2*Δ) are indicated with (***). **C:** Lipid class

composition in mol% of total lipids in the sample. (*), (**), and (***) indicate p<0.05, p<0.01 and p<0.001 in an unpaired, two-tailed t-test (n=6 for WT; n=2 for n=2 for *ubx2*Δ and *mga2*Δ). More data from lipid analyses are provided in Figure S2-S4 and Source data 5.

Figure 4

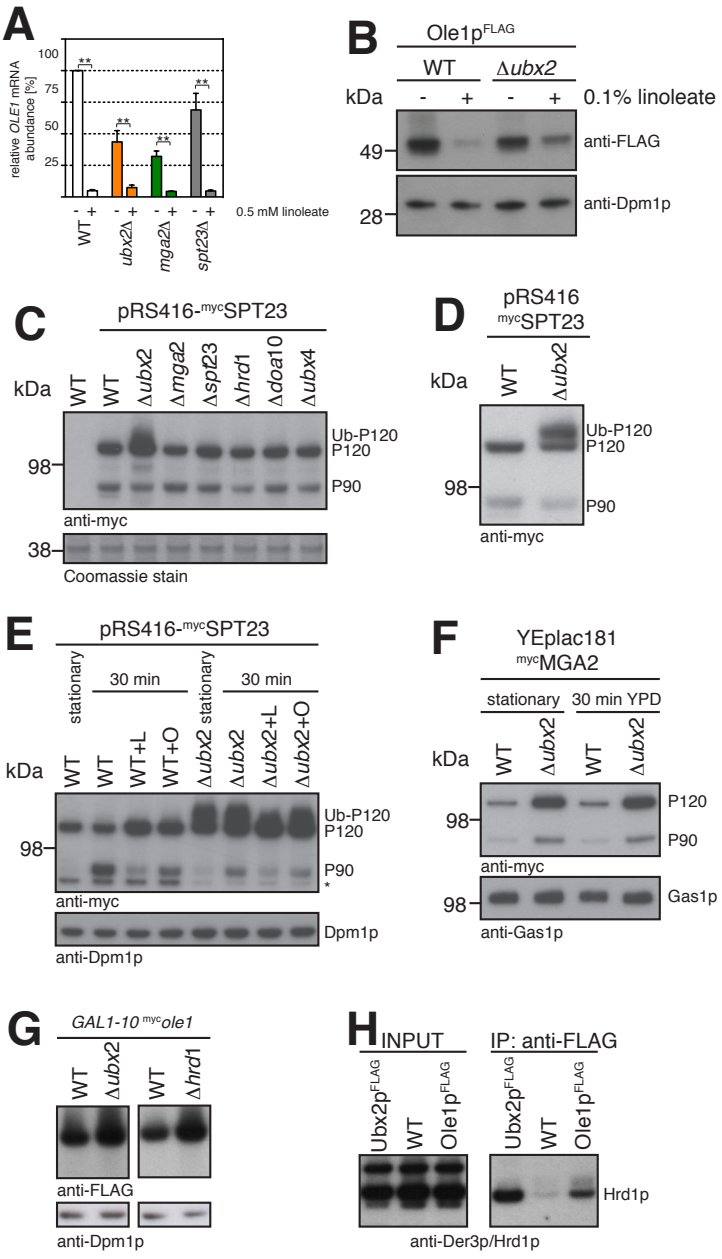


Figure 4. *UBX2* modulates the Ole1 level by distinct mechanisms. A: The relative mRNA was determined as in Figure 2I. If indicated, the medium was supplemented with 0.5 mM linoleate (UFA). (*) or (**) indicate $p<0.05$ or $p<0.01$ in an unpaired, two-tailed t-test (n=3). **B**: Wild type cells and deletants, with and without genomically FLAG-tagged Ole1 were grown in YPD to OD_{600}=1. Where indicated, the medium contained 0.1% linoleate. Cell lysates were immunoblotted as indicated. **C-E:** Wild type cells and deletants, with and without N-terminally myc-tagged Spt23 expressed from its endogenous promoter on a plasmid were grown in YPD to OD_{600}=1 (**C**,**D**) or for indicated times in YPD with and without sodium linoleate (+L, 0.1%) or sodium oleate (+O, 0.1%). Cell lysates were immunoblotted as indicated. **F:** Wild type and *ubx2*D cells carrying a plasmid encoding N-terminally myc- tagged Mga2 were grown in SC-LEU. The (*) indicates an unspecific signal that is consistently higher in wild type cells as compared to *ubx2*Δ cells. **G:** Catalytically inactive, 3xmyc-tagged Ole1 was expressed from a GAL- promoter in YP with galactose. Cell lysates were immunoblotted as indicated. **H:** Lysates of cells lysates expressing genomically tagged Ubx2-FLAG and Ole1-FLAG were subjected to immunoprecipitation. Co-precipitated material was immunoblotted as indicated. NEM was present in the lysis buffer to prevent deubiquitylation of Ole1. Cell lysates were equally loading, with wild type cells as specificity control.

Figure 5

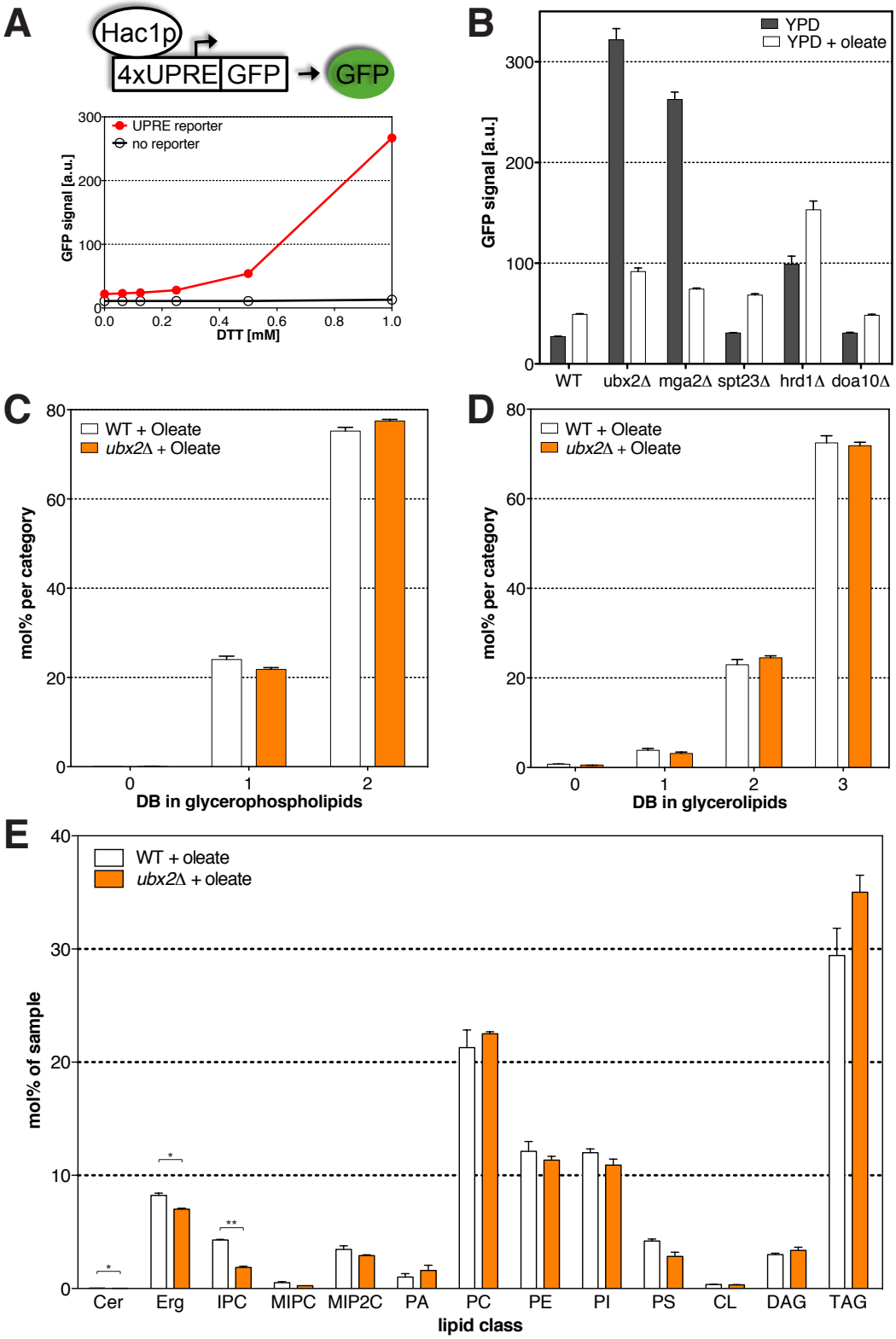

A
Hac1p
4xUPRE
GFP
GFP
UPRE reporter
no reporter
GFP signal [a.u.]
DTT [mM]
B
YPD
YPD + oleate
GFP signal [a.u.]
WT
ubx2Δ
mga2Δ
spt23Δ
hrd1Δ
doa10Δ
C
WT + Oleate
ubx2Δ + Oleate
mol% per category
DB in glycerophospholipids
D
WT + Oleate
ubx2Δ + Oleate
mol% per category
DB in glycerolipids
E
WT + oleate
ubx2Δ + oleate
mol% of sample
Cer
Erg
IPC
MIPC
MIP2C
PA
PC
PE
PI
PS
CL
DAG
TAG
lipid class

Figure 5. Misregulated FA desaturation induces the UPR and can be reversed by exogenous UFAs. A: Quantifying UPR levels in yeast with a GFP reporter as a function of DTT concentration. Cells without reporter construct served as control. **B:** The UPR reporter level in deletants exponentially grown in YPD (grey) and in presence of 0.01% oleate (white). **C-D:** Wild type cells and deletants grown to OD_{600}=1 in YPD with 0.15% Brij35 and 0.1% sodium oleate at 30ºC. **D:** Total double bonds (**DB**) of glycerophospholipids (**GP**; cardiolipin omitted for clarity), given as the sum of double bonds in FAs, in mol% of all glycerophospholipids. **E:** Total double bonds (DB) of glycerolipids (**GL**; DAG and TAG), given as the sum of double bonds in FAs, in mol% of all glycerolipids. **E:** Lipid class composition in mol% of total lipids in the sample. (*), (**), and (***) indicate $p<0.05$, $p<0.01$ and $p<0.001$ in an unpaired, two-tailed t-test (n=2).

Figure 6

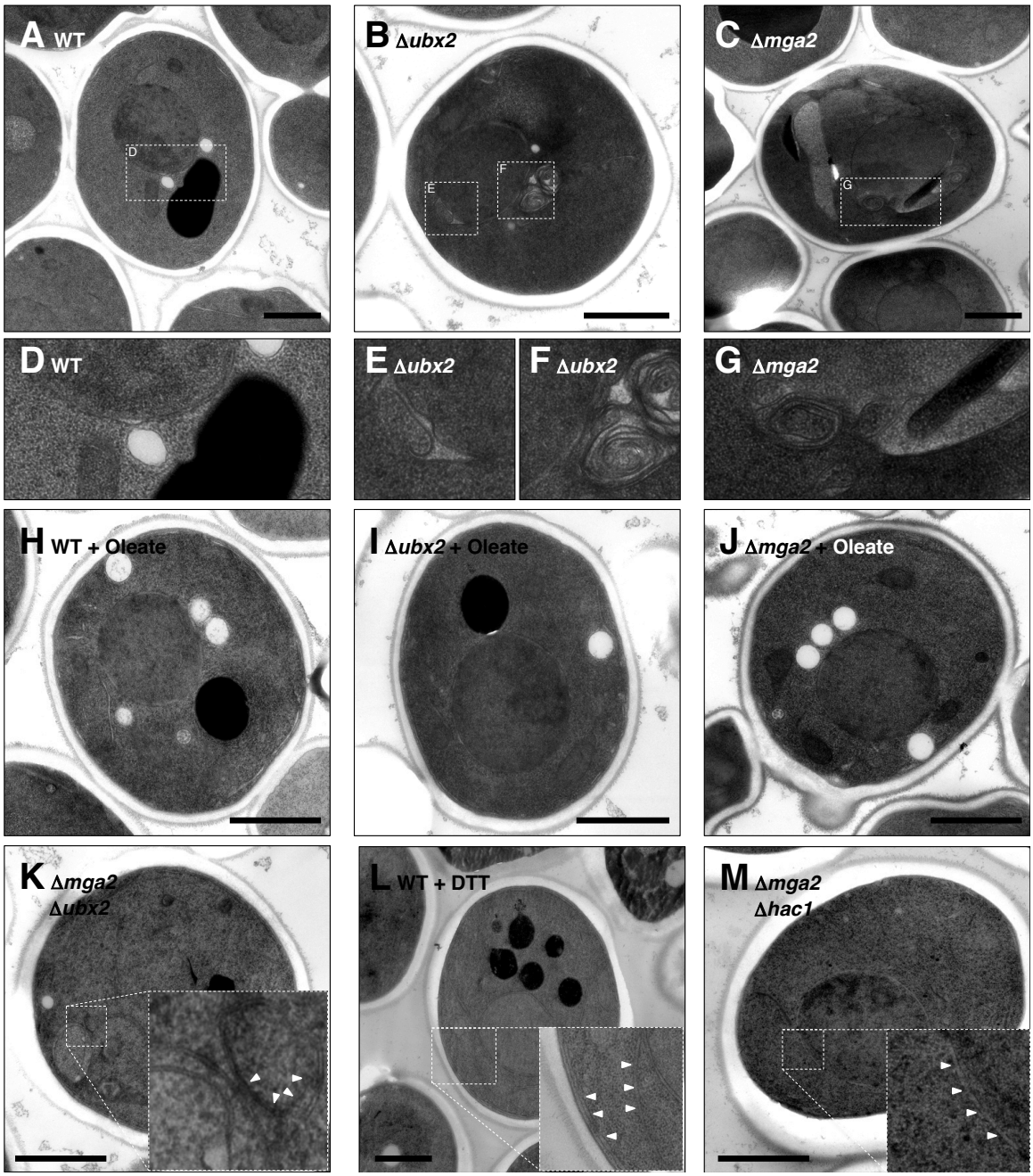


Figure 6. Misregulated FA desaturation induces expansion of the outer nuclear membrane and formation of ER whorls. A-G: Wild type cells and deletants grown in YPD to OD_{600}=3. **H-J:** Wild type cells and deletants grown to OD_{600}=3 in YPD with 0.15% Brij35 and 0.1% oleate. **K:** *mga2*Δ*ubx2*Δ double deletants were cultured

as in (A). **L:** Wild type cells grown to OD_{600}=1.5 in YPD to OD_{600}=1.5 and then stressed for 2 hours with 8 mM DTT and harvested at OD_{600}=3. **M:** *mga2*Δ*hac1*Δ double deletants cultured as in (A). Scale bar = 1 μm. **K-M**: The white arrows indicate the ER.

Supplementary Figures

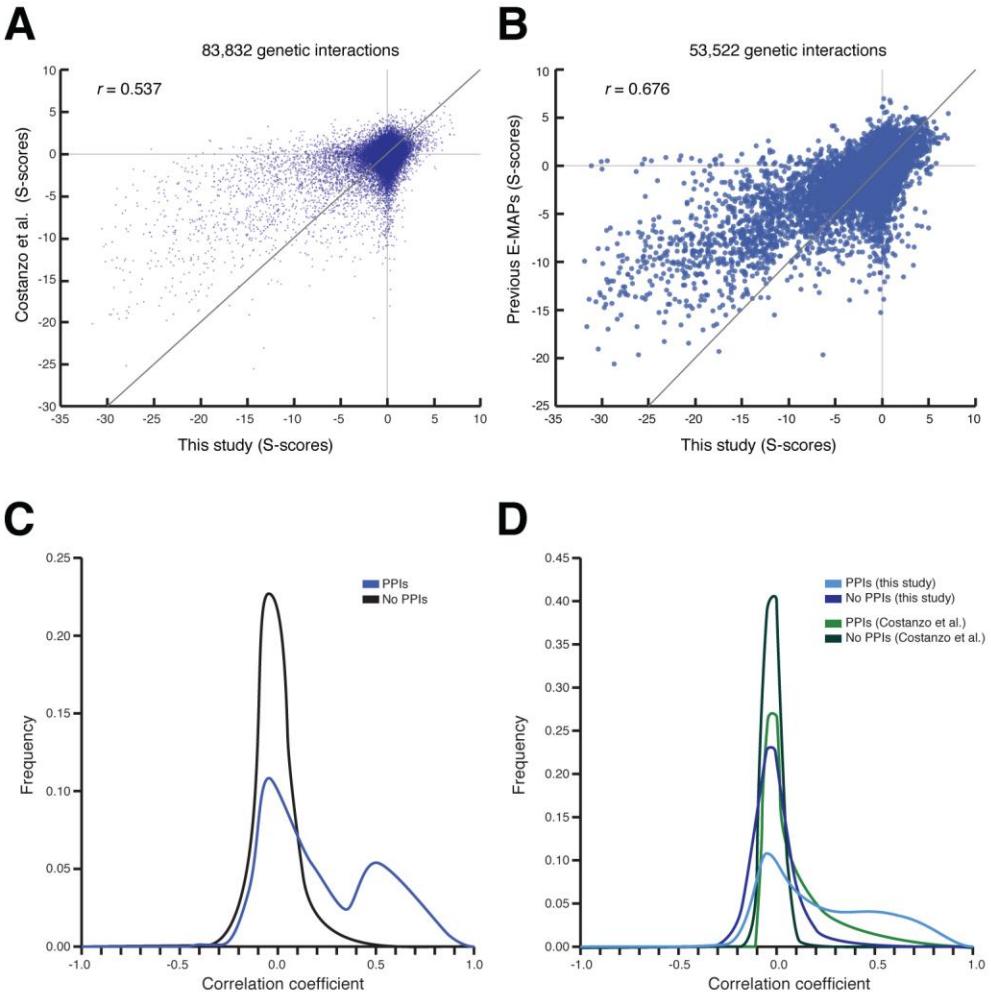


Figure S1–related to Figure 1. Benchmarking of the lipid E-MAP data. Comparison of overlapping data in the lipid E-MAP and the Costanzo et al., dataset (Costanzo et al., 2010), after it had been scaled to the S-score (Ryan et al., 2012) (**A**) or previous E-MAP data (Aguilar et al., 2010; Collins et al., 2007b; Fiedler et al., 2009; Hoppins et al., 2011; Schuldiner et al., 2005; Wilmes et al., 2008; Zheng et al., 2010) (**B**). **C:** Comparison of the lipid E- MAP data and 315 protein-protein interactions derived from previous studies (Babu et al., 2012;

Collins et al., 2007a). **D:** Comparison of the PPI data to either the lipid E-MAP data or the Costanzo et al. dataset. Protein-protein interactions were restricted to pairs that were represented in both genetic interaction datasets.

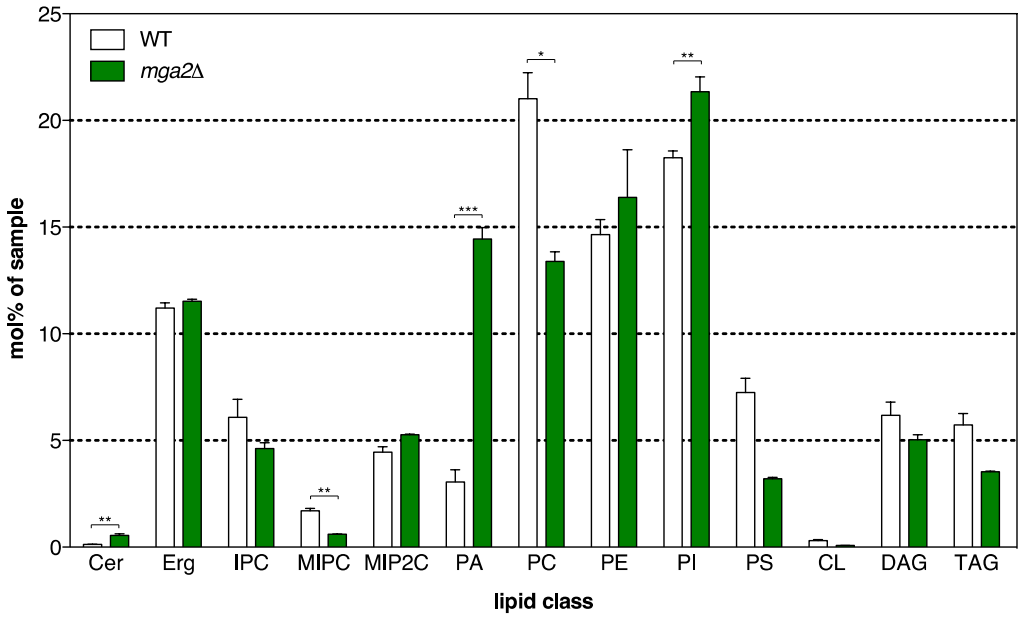
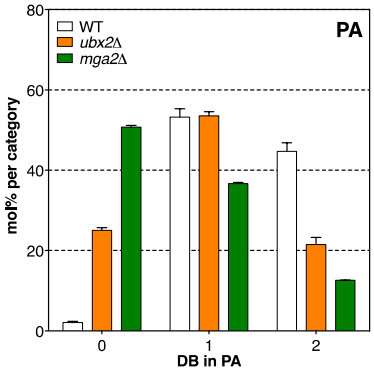
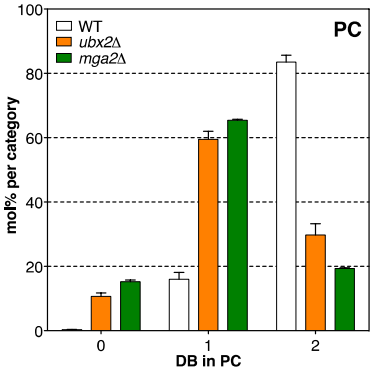
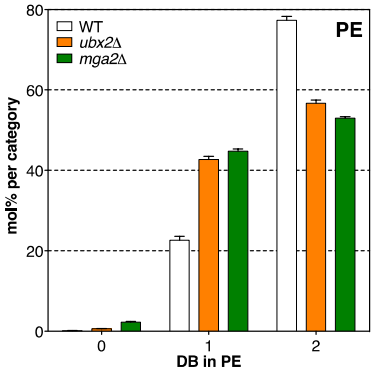
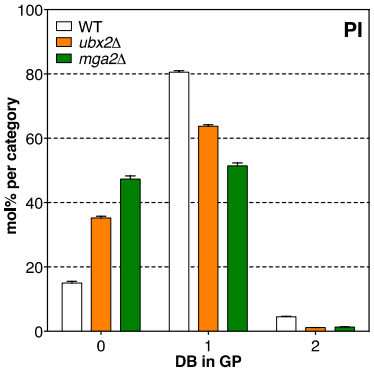


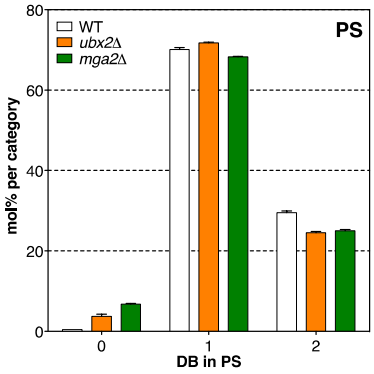
Figure S2-related to Figure 3. Lipid class profile of WT and *mga2*Δ cells. Wild type and *mga2*Δ cells were grown in YPD at 30ºC to OD_{600}=1 and their lipidomes were determined using quantitative mass-spectrometry based lipidomic platform. Illustrated is the lipid class composition in mol% of total lipids in the sample for WT and *mga2*Δ cells. Significant deviations from WT (unpaired, two-tailed T-test; n=6 for WT; n=2 for *mga2*Δ) are labeled with (*) for $p<0.05$, (**) for $p<0.01$, and (***) for $p<0.001$.

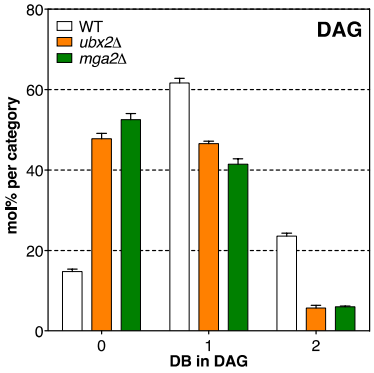

PA
WT
ubx2Δ
mga2Δ
mol% per category
DB in PA


PC
WT
ubx2Δ
mga2Δ
mol% per category
DB in PC


PE
WT
ubx2Δ
mga2Δ
mol% per category
DB in PE


PI
WT
ubx2Δ
mga2Δ
mol% per category
DB in GP


PS
WT
ubx2Δ
mga2Δ
mol% per category
DB in PS


DAG
WT
ubx2Δ
mga2Δ
mol% per category
DB in DAG

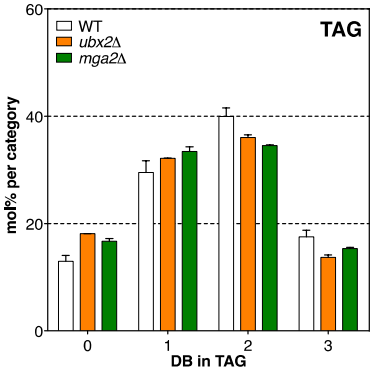

TAG
WT
ubx2Δ
mga2Δ
mol% per category
DB in TAG

Figure S3-related to Figure 3. The trend towards more saturated lipids is observed in *ubx2*Δ and *mga2*Δ cells throughout all glycerolipid classes. Cells were grown in YPD at 30ºC to an OD_{600}=1 and harvested for lipidomic analysis. For each panel the total double bonds (DB) is given as the sum of double bonds in the indicated class of glycerolipid or glycerophospholipid. The lipid classes are: PA phosphatidic acid; PC phosphatidylcholine; PE phosphatidylethanolamine; PI phosphatidylinositol; PS phosphatidylserine; DAG diacylglycerol; TAG triacylglycerol.

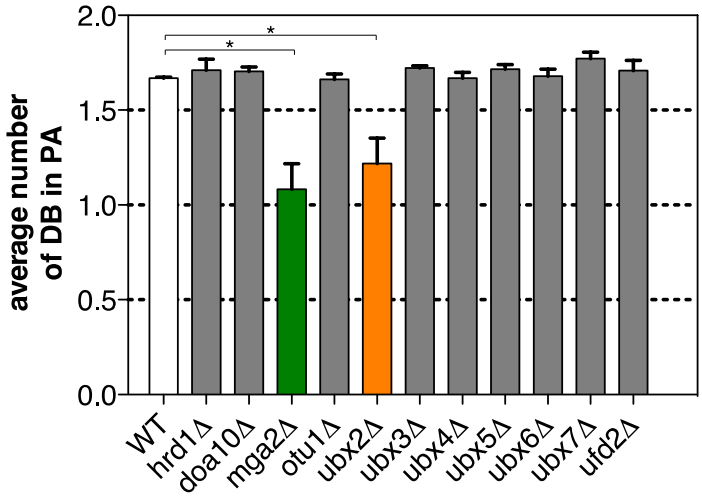


Figure S4-related to Figure 3. The trend towards more saturated lipids is specific to *ubx2*Δ cells. Wild type cells and the indicated mutants were grown at 30ºC or 22ºC in SD medium supplemented with 100 µM inositol, 100 µM choline and 2% glucose and harvested at OD_{600}=1 for lipidomic analysis. The average number of double bonds (DB) per acyl chain of phosphatidic acid (PA) is plotted for each tested mutant. A (*) indicates a significant reduction of the average number of

DB as compared to WT cells with $p<0.05$ (unpaired T- test; one-tailed; n=2). Lipidomic analysis was performed essentially as described by Ejsing et al. (Ejsing et al.,2009).

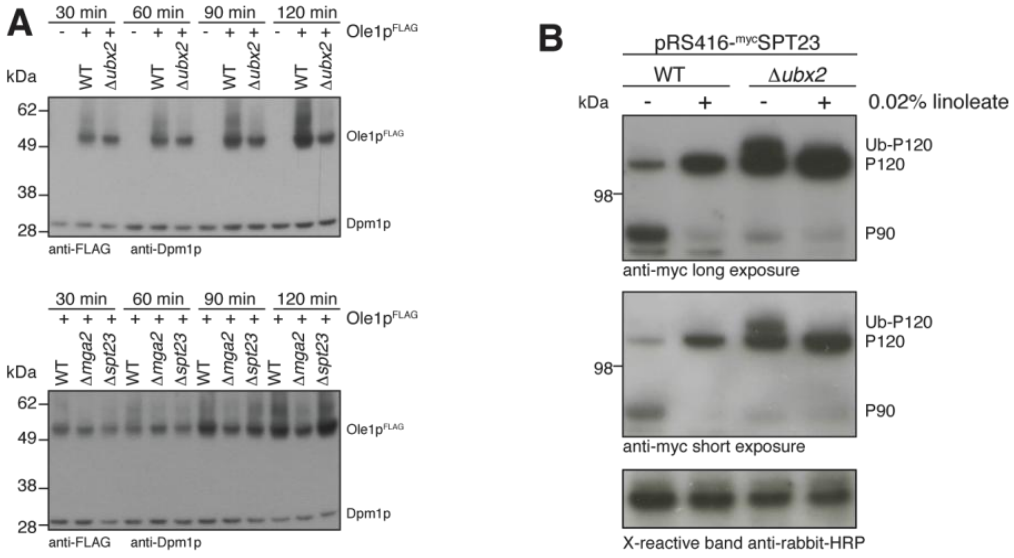


Figure S5-related to Figure 4. Regulation of Spt23 in *ubx2*Δ cells. A: Wild type cells and cells deleted for indicated genes, with and without genomically FLAG-tagged Ole1 were or cultured in fresh YPD for indicated times after inoculation to OD_{600}=0.25 from an overnight culture. Cell lysates were immunoblotted with anti-FLAG and anti-Dpm1 antibodies. **B:** Wild type and *ubx2*Δ cells expressing N-terminally myc-tagged Spt23 from its endogenous promoter on a CEN-based plasmid were transferred from an overnight culture to fresh medium containing 0.15% Brij35 with and without 0.02% linoleate to an OD_{600}=0.5 and cultured for 30ºC. Equal amounts of cells were harvested and cell lysates were immunoblotted

using a anti-myc antibodies. As loading control serves a cross-specific signal of the anti-rabbit secondary antibody.

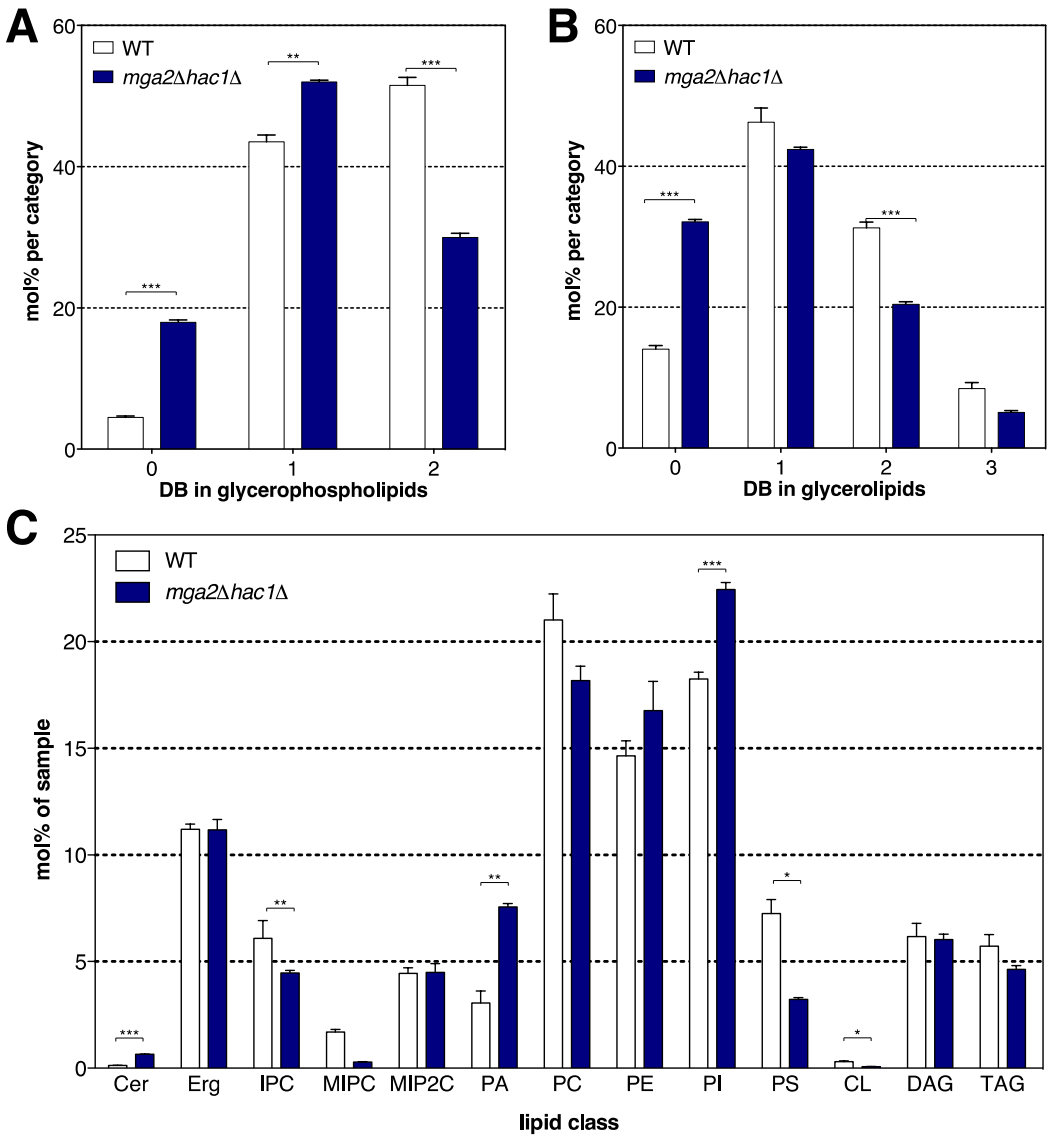


Figure S6-related to Figure 6. Lipidome of the *mga2*Δ*hac1*Δ double mutant. **A-C:** Wild type and *mga2*Δ*hac1*Δ cells were grown in YPD at 30ºC to an OD_{600}=1 and harvested for lipidomic analysis. **A:** Total double bonds (**DB**) of glycerophospholipids (**GP**; cardiolipin omitted for clarity), given as the sum of

double bonds in fatty acids, in mol% of all glycerophospholipids. **B:** Total double bonds of glycerolipids (**GL**; DAG and TAG), given as the sum of double bonds in fatty acids, in mol% of all glycerolipids. **C:** Lipid class composition in mol% of total lipids in the sample. **A-C:** Significant changes induced by the *mga2*Δ*hac1*Δ double deletion (unpaired, two-tailed T-test; n=6 for WT in YPD; n=2 for WT+DTT) are indicated with (*) for $p<0.05$, (**) for $p<0.01$, and (***) for $p<0.001$.

Supplementary Materials and Methods

Reagents and antibodies

All reagents were reagent-grade and obtained from commercial sources. Sodiumoleate and sodium linoleate (Sigma-Aldrich) were stored as 2% solution in water. The anti-FLAG M2 monoclonal antibody was from Sigma-Aldrich, a rabbit polyclonal (A-14) and a mouse monoclonal (9E10) anti-myc antibody were from Santa Cruz Biotechnology, mouse anti- Dpm1 from Invitrogen. The Der3/Hrd1 antibody was a kind gift of Dieter Wolf, rabbit anti- Gas1 was donated from H. Riezman.

Lipid extraction and mass spectrometric analysis

Quantitation of lipids was performed essentially as described by Klose et al. (Klose et al., 2012) with modifications from the original protocol by Ejsing et al. (Ejsing et al., 2009). Yeast culture were inoculated to an OD_{600} of 0.1 from an overnight culture and then grown in 30 ml YPD with or without 0.15% Brij35

at 30ºC in a 100 ml Erlenmeyer flask at 200 rpm. The presence of 0.15% Brij35 did not affect any parameter of lipidome, and was used for feeding experiments with 0.1% sodium oleate. When cells reached an OD_{600}=1, 20 OD units of cells were harvested and washed with cold 155 mM ammonium bicarbonate in presence of 10 mM azide. Cell pellets were stored for up to four weeks at -80ºC. Prior to cell lysis, pellets were thawed and resuspended in 1 ml 155 mM ammonium bicarbonate and disrupted by vigorous shaking in presence of 200 µl zirconia beads in 2 ml microfuge tubes using TissueLyser II (Qiagen) at 4ºC for 20 min. A volume equivalent to 0.2 OD_{600} units was taken and diluted with ammonium bicarbonate to 200 µl. For details of mass spectrometric analysis see Klose et al. (Klose et al., 2012). The lipid data are available as 'Supporting Table 2 – Lipid data.csv' and semicolon delimited file, with commas as decimal separator.

Electron microscopic analysis of yeast cells

The pellet of a filtered yeast culture grown to an OD_{600}=3 in YPD was cryoimmobilized by high pressure freezing (EMPACT 2, Leica Microsystems), and further processed by freeze-substitution (AFS II, Leica Microsystems) in acetone (containing 1 % osmium tetroxide, 0.1 % uranyl acetate, from -90°C up to room temperature with a slope of 4° C per hour. After short rinses with acetone, the samples got infiltrated gradually with LX112-resin. After polymerization, 70 nm sections were taken on a microtome (Ultracut UCT, Leica Microsystems), collected on copper grids, post-contrasted with 1% uranyl-acetate in water followed by lead

citrate, examined with an electron microscope (TECNAI 12, FEI). Pictures were taken using CCD camera (TVIPS, F214-Tecnai).

Flow cytometry analysis

Cells were grown in YPD with additives as indicated from OD_{600}=0.1 to OD_{600}=1 and analyzed using a FACSCalibur (Becton Dickenson).

Protein analysis and Western Blot

Yeast cells were cultured in YPD media if not otherwise noted. Between 10 and 20 OD_{600} units were collected by centrifugation, washed with phosphate buffered saline (PBS) in presence of 10 mM N-ethylmaleimide (NEM) and snap-frozen. Cell pellets were thawed on ice, resuspended in lysis buffer (PBS, 10 mM NEM, and EDTA-free protease inhibitor cocktail (Roche)) and disrupted by vigorous shaking with zirconia beads (Roche) for 5 min at 4ºC. After addition of a 5-fold concentrated sample buffer (50 mM Tris-HCl H 6.8, 8M urea, 2% SDS, 5 mM EDTA, and 0.02% Bromphenol blue and 20% glycerole) the lysates were heated to 65ºC for 10 min, cleared by centrifugation, and separated via SDS-polyacrylamide gradient gels (4-12%, Invitrogen). For immunodetection, the proteins were blotted on PVDF membranes and detected using antigen-specific antibodies.

Immunoprecipitation

10 OD_{600} units of cells grown to OD_{600}=1 were resuspended in 1.4 ml IP buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 2.5 mM $MgCl_2$, 1% NP40 substitute)

in presence of 10 mM NEM, 10 mM PMSF and EDTA-free protease inhibitor cocktail. Cell were disrupted by vigorous shaking with zirconia beads for 10 min at 4ºC and the lysate was incubated for additional 20 min at 4ºC on an overhead rotor. The non-solubilized material was removed by centrifugation at 10000 x g for 10 min. Then, the lysate was pre-cleared by 30 min incubation in presence of protein A beads. Flag-tagged proteins and their interaction partners were retrieved by 12 µl anti-FLAG M2 affinity gel (Sigma-Aldrich) and 3 hours of incubation at 4º. After extensive washes with IP buffer the immunoprecipitated proteins were eluted from the beads in sample buffer and analyzed by Western blotting.

OLE1 mRNA quantitation

Cells were grown to saturation in YPD containing 0.04% Tween-40 and 0.5 mM oleate, then diluted to OD_{600}=0.5 in YPD (-UFA) or YPD containing 0.04% Tween-40 and 0.5 mM sodium linoleate (+UFA) and cultured for 4h at 30ºC. Cells were collected by centrifugation, washed with sterile water and stored in liquid nitrogen prior to RNA extraction. The total RNA was extracted and DNAse treated using the *Turbo DNA free* kit (Invitrogen). cDNA synthesis was performed from 2 µg of DNAse-free RNA using the *SuperScript III Frist-Strand Synthesis System for RT-PCR* (Invitrogen), using OligodT primers. Quantitative Real time PCR was performed using *Power SYBR green PCR master mix* (Applied Biosystems) in an Applied Biosystems 7900HT Fast Real-Time PCR System using *OLE1*, *ACT1* and *TAF10* specific primers. The *OLE1* mRNA abundance was normalized to ACT1 mRNA levels and expressed relative to *OLE1* expression in wild type cells in absence of

exogenously added UFAs. Three independent experiments were performed and the relative mRNA levels were averaged and plotted.

References

Aguilar, P.S., Frohlich, F., Rehman, M., Shales, M., Ulitsky, I., Olivera-Couto, A., Braberg, H., Shamir, R., Walter, P., Mann, M.*, et al.* (2010). A plasma- membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking. Nature structural & molecular biology *17*, 901-908.

Alberts, S.M., Sonntag, C., Schafer, A., and Wolf, D.H. (2009). Ubx4 modulates cdc48 activity and influences degradation of misfolded proteins of the endoplasmic reticulum. J Biol Chem *284*, 16082-16089.

Babu, M., Vlasblom, J., Pu, S., Guo, X., Graham, C., Bean, B.D., Burston, H.E., Vizeacoumar, F.J., Snider, J., Phanse, S.*, et al.* (2012). Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. Nature *489*, 585-589.

Bays, N.W., Gardner, R.G., Seelig, L.P., Joazeiro, C.A., and Hampton, R.Y. (2001). Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. Nat Cell Biol *3*, 24-29.

Beltrao, P., Cagney, G., and Krogan, N.J. (2010). Quantitative genetic interactions reveal biological modularity. Cell *141*, 739-745.

Bernales, S., McDonald, K.L., and Walter, P. (2006). Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. PLoS biology *4*, e423.

Bordallo, J., Plemper, R.K., Finger, A., and Wolf, D.H. (1998). Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins. Mol Biol Cell *9*, 209-222.

Boumann, H.A., Gubbens, J., Koorengevel, M.C., Oh, C.S., Martin, C.E., Heck, A.J., Patton-Vogt, J., Henry, S.A., de Kruijff, B., and de Kroon, A.I. (2006). Depletion of phosphatidylcholine in yeast induces shortening and increased saturation of the lipid acyl chains: evidence for regulation of intrinsic membrane curvature in a eukaryote. Molecular biology of the cell *17*, 1006- 1017.

Braun, S., Matuschewski, K., Rape, M., Thoms, S., and Jentsch, S. (2002). Role of the ubiquitin-selective CDC48(UFD1/NPL4)chaperone (segregase) in ERAD of OLE1 and other substrates. EMBO J *21*, 615-621.

Collins, S.R., Kemmeren, P., Zhao, X.C., Greenblatt, J.F., Spencer, F., Holstege, F.C., Weissman, J.S., and Krogan, N.J. (2007a). Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics *6*, 439-450.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S.R., Roguev, A., and Krogan, N.J. (2010). Quantitative genetic interaction mapping using the E-MAP approach. Methods in enzymology *470*, 205-231.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S.*, et al.* (2010). The genetic landscape of a cell. Science *327*, 425-431.

Deguil, J., Pineau, L., Rowland Snyder, E.C., Dupont, S., Beney, L., Gil, A., Frapper, G., and Ferreira, T. (2011). Modulation of lipid-induced ER stress by fatty acid shape. Traffic *12*, 349-362.

Eizirik, D.L., and Cnop, M. (2010). ER stress in pancreatic beta cells: the thin red line between adaptation and failure. Sci Signal *3*, pe7.

Ejsing, C.S., Sampaio, J.L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R.W., Simons, K., and Shevchenko, A. (2009). Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. Proc Natl Acad Sci U S A *106*, 2136-2141.

Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S., et al. (2009). Functional organization of the S. cerevisiae phosphorylation network. Cell 136, 952-963.

Goder, V., Carvalho, P., and Rapoport, T.A. (2008). The ER-associated degradation component Der1p and its homolog Dfm1p are contained in complexes with distinct cofactors of the ATPase Cdc48p. FEBS letters *582*, 1575-1580.

Hampton, R.Y., Gardner, R.G., and Rine, J. (1996). Role of 26S proteasome and HRD genes in the degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum membrane protein. Mol Biol Cell *7*, 2029-2044.

Han, S., Lone, M.A., Schneiter, R., and Chang, A. (2010). Orm1 and Orm2 are conserved endoplasmic reticulum membrane proteins regulating lipid homeostasis and protein quality control. Proc Natl Acad Sci U S A *107*, 5851- 5856.

Hapala, I., Marza, E., and Ferreira, T. (2011). Is fat so bad? Modulation of endoplasmic reticulum stress by lipid droplet formation. Biol Cell *103*, 271- 285. Henry, S.A., Kohlwein, S.D., and Carman, G.M. (2012). Metabolism and regulation of glycerolipids in the yeast Saccharomyces cerevisiae. Genetics *190*, 317-349.

Hitchcock, A.L., Krebber, H., Frietze, S., Lin, A., Latterich, M., and Silver, P.A. (2001). The conserved npl4 protein complex mediates proteasome-dependent membrane-bound transcription factor activation. Mol Biol Cell *12*, 3226-3241.

Hoppe, T., Matuschewski, K., Rape, M., Schlenker, S., Ulrich, H.D., and Jentsch, S. (2000). Activation of a membrane-bound transcription factor by regulated ubiquitin/proteasome-dependent processing. Cell *102*, 577-586.

Hoppins, S., Collins, S.R., Cassidy-Stone, A., Hummel, E., Devay, R.M., Lackner, L.L., Westermann, B., Schuldiner, M., Weissman, J.S., and Nunnari, J. (2011). A mitochondrial-focused genetic interaction map reveals a scaffold- like complex required for inner membrane organization in mitochondria. The Journal of cell biology *195*, 323-340.

Jonikas, M.C., Collins, S.R., Denic, V., Oh, E., Quan, E.M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J.S.*, et al.* (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science *323*, 1693-1697.

Kaiser, H.J., Surma, M.A., Mayer, F., Levental, I., Grzybek, M., Klemm, R.W., Da Cruz, S., Meisinger, C., Muller, V., Simons, K.*, et al.* (2011). Molecular convergence of bacterial and eukaryotic surface order. The Journal of biological chemistry *286*, 40631-40637.

Kandasamy, P., Vemula, M., Oh, C.S., Chellappa, R., and Martin, C.E. (2004). Regulation of unsaturated fatty acid biosynthesis in Saccharomyces: the endoplasmic reticulum membrane protein, Mga2p, a transcription activator of the OLE1 gene, regulates the stability of the OLE1 mRNA through exosome-mediated mechanisms. J Biol Chem *279*, 36586-36592.

Kim, I., Xu, W., and Reed, J.C. (2008). Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities. Nat Rev Drug Discov *7*, 1013-1030.

Klose, C., Ejsing, C.S., Garcia-Saez, A.J., Kaiser, H.J., Sampaio, J.L., Surma, M.A., Shevchenko, A., Schwille, P., and Simons, K. (2010). Yeast lipids can phase-separate into micrometer-scale membrane domains. The Journal of biological chemistry *285*, 30224-30232.

Klose, C., Surma, M.A., Gerl, M.J., Meyenhofer, F., Shevchenko, A., and Simons, K. (2012). Flexibility of a eukaryotic lipidome--insights from yeast lipidomics. PLoS One *7*, e35063.

Kohlwein, S.D., and Petschnigg, J. (2007). Lipid-induced cell dysfunction and cell death: lessons from yeast. Curr Hypertens Rep *9*, 455-461.

Lee, J.N., Kim, H., Yao, H., Chen, Y., Weng, K., and Ye, J. (2010). Identification of Ubxd8 protein as a sensor for unsaturated fatty acids and regulator of triglyceride synthesis. Proc Natl Acad Sci U S A *107*, 21424- 21429.

Lingwood, D., Schuck, S., Ferguson, C., Gerl, M., and Simons, K. (2009). Morphological homeostasis by autophagy. Autophagy *5*, 1039-1040.

Lorch, Y., Cairns, B.R., Zhang, M., and Kornberg, R.D. (1998). Activated RSC-nucleosome complex and persistently altered form of the nucleosome. Cell *94*, 29-34.

Martin, C.E., Oh, C.S., and Jiang, Y. (2007). Regulation of long chain unsaturated fatty acid synthesis in yeast. Biochimica et biophysica acta *1771*, 271-285.

Neuber, O., Jarosch, E., Volkwein, C., Walter, J., and Sommer, T. (2005). Ubx2 links the Cdc48 complex to ER-associated protein degradation. Nat Cell Biol *7*, 993-998.

Olzmann, J.A., Richter, C.M., and Kopito, R.R. (2013). Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. Proceedings of the National Academy of Sciences of the United States of America *110*, 1345-1350.

Phillips, R., Ursell, T., Wiggins, P., and Sens, P. (2009). Emerging roles for lipids in shaping membrane-protein function. Nature *459*, 379-385.

Pineau, L., Colas, J., Dupont, S., Beney, L., Fleurat-Lessard, P., Berjeaud, J.M., Berges, T., and Ferreira, T. (2009). Lipid-induced ER stress: synergistic effects of sterols and saturated fatty acids. Traffic *10*, 673-690.

Promlek, T., Ishiwata-Kimata, Y., Shido, M., Sakuramoto, M., Kohno, K., and Kimata, Y. (2011). Membrane aberrancy and unfolded proteins activate the endoplasmic reticulum stress sensor Ire1 in different ways. Molecular biology of the cell *22*, 3520-3532.

Rape, M., Hoppe, T., Gorr, I., Kalocay, M., Richly, H., and Jentsch, S. (2001). Mobilization of processed, membrane-tethered SPT23 transcription factor by CDC48(UFD1/NPL4), a ubiquitin-selective chaperone. Cell *107*, 667-677.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J.*, et al.* (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science *322*, 405-410.

Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol *8*, 519-529.

Ryan, C.J., Roguev, A., Patrick, K., Xu, J., Jahari, H., Tong, Z., Beltrao, P., Shales, M., Qu, H., Collins, S.R., et al. (2012). Hierarchical modularity and the evolution of genetic interactomes across species. Molecular cell 46, 691-704.

Schuberth, C., and Buchberger, A. (2005). Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER- associated protein degradation. Nat Cell Biol *7*, 999-1006.

Schuberth, C., and Buchberger, A. (2008). UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97. Cell Mol Life Sci *65*, 2360-2371.

Schuck, S., Prinz, W.A., Thorn, K.S., Voss, C., and Walter, P. (2009). Membrane expansion alleviates endoplasmic reticulum stress independently of the unfolded protein response. J Cell Biol *187*, 525-536.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F.*, et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Shcherbik, N., and Haines, D.S. (2007). Cdc48p(Npl4p/Ufd1p) binds and segregates membrane-anchored/tethered complexes via a polyubiquitin signal present on the anchors. Molecular cell *25*, 385-397.

Shcherbik, N., Zoladek, T., Nickels, J.T., and Haines, D.S. (2003). Rsp5p is required for ER bound Mga2p120 polyubiquitination and release of the processed/tethered transactivator Mga2p90. Current biology : CB *13*, 1227- 1233.

Snapp, E.L., Hegde, R.S., Francolini, M., Lombardo, F., Colombo, S., Pedrazzini, E., Borgese, N., and Lippincott-Schwartz, J. (2003). Formation of stacked ER cisternae by low affinity protein interactions. The Journal of cell biology *163*, 257-269.

Stukey, J.E., McDonough, V.M., and Martin, C.E. (1990). The OLE1 gene of Saccharomyces cerevisiae encodes the delta 9 fatty acid desaturase and can be functionally replaced by the rat stearoyl-CoA desaturase gene. J Biol Chem *265*, 20144-20149.

Thibault, G., Ismail, N., and Ng, D.T. (2011). The unfolded protein response supports cellular robustness as a broad-spectrum compensatory pathway. Proceedings of the National Academy of Sciences of the United States of America *108*, 20597-20602.

Thibault, G., Shui, G., Kim, W., McAlister, G.C., Ismail, N., Gygi, S.P., Wenk, M.R., and Ng, D.T. (2012). The membrane stress response buffers lethal effects of lipid disequilibrium by reprogramming the protein homeostasis network. Molecular cell *48*, 16-27.

van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008). Membrane lipids: where they are and how they behave. Nature reviews. Molecular cell biology *9*, 112-124.

Vembar, S.S., and Brodsky, J.L. (2008). One step at a time: endoplasmic reticulum-associated degradation. Nat Rev Mol Cell Biol *9*, 944-957.

Voeltz, G.K., Rolls, M.M., and Rapoport, T.A. (2002). Structural organization of the endoplasmic reticulum. EMBO reports *3*, 944-950.

Volmer, R., van der Ploeg, K., and Ron, D. (2013). Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains. Proceedings of the National Academy of Sciences of the United States of America *110*, 4628-4633.

Wang, C.W., and Lee, S.C. (2012). The ubiquitin-like (UBX)-domain- containing protein Ubx2/Ubxd8 regulates lipid droplet homeostasis. Journal of cell science *125*, 2930-2939.

Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S.*, et al.* (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Molecular cell *32*, 735-746.

Zheng, J., Benschop, J.J., Shales, M., Kemmeren, P., Greenblatt, J., Cagney, G., Holstege, F., Li, H., and Krogan, N.J. (2010). Epistatic relationships reveal the functional organization of yeast transcription factors. Mol Syst Biol 6, 420.

Chapter 5

Functional Organization of the *S. cerevisiae* Phosphorylation Network

Functional Organization of the *S. cerevisiae* Phosphorylation Network

Dorothea Fiedler[1,2], Hannes Braberg[1,3*], Monika Mehta[4*], Gal Chechik[5*], Gerard Cagney[1,3,6*], Paromita Mukherjee[4], Andrea C. Silva[4], Michael Shales[1,3], Sean R. Collins[1,2,3], Sake van Wageningen[7], Patrick Kemmeren[7], Frank C. P. Holstege[7], Jonathan S. Weissman[1,2,3], Michael-Christopher Keogh[4], Daphne Koller[5], Kevan M. Shokat[1,2], Nevan J. Krogan[1,3]

1 Department of Cellular and Molecular Pharmacology, University of California-San Francisco, San Francisco CA 94158 USA

2 Howard Hughes Medical Institute, San Francisco CA 94158 USA

3 The California Institute for Quantitative Biomedical Research, University of California, San Francisco, California 94158 USA

4 Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461 USA

5 Computer Science Department, Stanford University, Stanford, CA 94305 USA

6 Conway Institute, University College Dublin, Belfield, Dublin 4, IRELAND

7 Department of Physiological Chemistry, University Medical Center Utrect, Utrecht, The Netherlands

*These authors contributed equally.

Summary

Reversible protein phosphorylation is a signaling mechanism involved in all cellular processes. To create a systems view of the signaling apparatus in budding yeast, we generated an E-MAP (epistatic miniarray profile) comprised of 100,000 pair-wise, quantitative genetic interactions, including virtually all protein kinases and phosphatases and key cellular regulators. Quantitative genetic interaction mapping reveals factors working in compensatory pathways (negative genetic interactions; e.g. synthetic lethality) or those operating in linear pathways (positive genetic interactions; e.g. suppression). Within kinases, phosphatases, and their substrates, we found an enrichment of positive genetic interactions. To develop a global view of the signaling apparatus, we isolated "triplet genetic motifs" and assembled these into a higher-order map. The resulting network view provides new insights into signaling pathway regulation, and revealed a link between the cell cycle kinase, Cak1, the Fus3 MAP kinase, and a pathway that regulates chromatin integrity during transcription by RNA polymerase II.

Introduction

Phosphate-based signaling is critical to almost all major cellular processes and is ubiquitously present across archea, prokaryota and eukaryota (Kannan et al., 2007). Systems-wide studies in the post-genome era have provided unprecedented information about the activities of signaling proteins (Johnson and Hunter, 2005) and several thousand sites of protein phosphorylation have been mapped using mass spectrometry (Ficarro et al., 2002; Green and Pflum, 2007; Lee et al., 2006; Matsuoka et al., 2007; Olsen et al., 2006). The knowledge of kinase-substrate relationships has been expanded by both *in vitro* protein chip analysis (Ptacek et al., 2005) and large-scale genetic screens using a kinase over-expression strategy (Sopko et al., 2006). Bioinformatic efforts have focused on network-level analyses of phosphorylation, providing database resources for phosphorylation sites and signaling pathways (Diella et al., 2008; Lee et al., 2006; Zanzoni et al., 2007). In addition, the integration of context-dependent information (including protein interactions and cell-specific kinase expression) has helped to improve the specificity of phospho-consensus site assignments (Linding et al., 2007).

Despite these achievements, signaling networks have remained difficult to study. While phosphoproteomic datasets illuminate the magnitude and diversity of protein phosphorylation, the functional relevance for the majority of these phosphorylation sites remains unknown (Johnson and Hunter, 2005). Focused studies can elucidate the function of one specific kinase or one particular pathway, but often overlook important connections to components that do not directly participate in the pathway. In particular, two hallmarks of kinase signaling are the

linear cascades of kinases important for signal amplification and the abundant cross-talk between these pathways in order to coordinate multiple cellular inputs/outputs.

Genetic interactions report on the extent to which the function of one gene depends on the presence of a second gene and can illuminate the functional organization of protein networks. Negative genetic interactions (synthetic sick/lethal interactions) describe cases where two mutations in combination cause a stronger growth defect than expected from the two single mutations. In contrast, positive genetic interactions correspond to cases where the double mutant is either no sicker (epistatic) or healthier (suppressive) than the sickest single mutant (Figure 1a). Negative genetic interactions are often found for proteins that work in compensatory pathways, while positive interactions can identify pairs of proteins that are in complex and/or function in the same pathway (Collins et al., 2007b; Kelley and Ideker, 2005; Roguev et al., 2008; Schuldiner et al., 2005; St. Onge et al., 2006). Two approaches, synthetic genetic array (SGA) technology (Tong et al., 2001) and diploid based synthetic lethality analysis on microarrays (dSLAM) (Pan et al., 2004), have been developed to identify synthetic sick/lethal (SSL) relationships on a genome-wide scale in *S. cerevisiae*. SSL interactions however, only represent a subset of the full epistatic interaction spectrum. Our interest in mapping the pathway architectures involved in phosphate-based signaling led us to use an SGA-based methodology that measures the entire range of epistatic relationships, termed epistatic mini array profile (or E-MAP) (Figure 1a) (Collins et al., 2007b; Collins et al., 2006; Roguev et al., 2008; Schuldiner et al., 2005; Schuldiner et al., 2006; Wilmes

et al., 2008). The E-MAP method does not rely on the detection of individual phosphoproteins, but it rather provides the functional connections between signaling molecules, whether these are direct or indirect. In addition, genetic interactions are not limited to the detection of proteins in physical contact and are thus particularly useful in identifying more transient interactions, and those controlled by post-translational modification. Here we use our E-MAP approach to provide an unbiased, genetic overview of the functional connections within the signaling machinery.

Results and Discussion

Composition of the Signaling E-MAP

An E-MAP quantitatively records genetic interactions between pairs of mutations for a defined subset of genes. The data is generated by systematically constructing double mutant strains, measuring their growth rates, and converting this information to individual genetic interaction scores, both positive and negative (Collins et al., 2006; Schuldiner et al., 2006). To explore the signaling apparatus of *S. cerevisiae*, our genetic approach targeted almost all known protein kinases (121) and phosphatases (38) and their regulatory proteins (45 and 39, respectively), as well as non-protein kinases (47) and phosphatases (38) (Figure 1b, Supplementary Table 1). To include those components of the signaling machinery that are essential for viability, we employed the DAmP (Decreased abundance by mRNA perturbation) strategy to create 48 hypomorphic alleles. This method relies on the insertion of the antibiotic-resistance marker to disrupt the natural 3'UTR, which leads to

destabilization of the mRNA transcripts, and, subsequently lowered protein levels (Schuldiner et al., 2005). In all, the signaling E-MAP contains 483 genes, including 135 genes that serve as reporters of a variety of major biological processes (Figure 1b, Supplementary Table 1). We comprehensively evaluated the pair-wise genetic interactions of these 483 alleles, resulting in approximately 100,000 distinct, pair-wise genetic interaction measurements (Supplementary Figure 1, Supplementary Table 2). All data is available at http://interactome-cmp.ucsf.edu in an interactive and searchable format.

Superposition of Literature-Derived Network with Signaling E-MAP

To evaluate the genetic interaction data, we first examined how the data recapitulated known phosphorylation/dephosphorylation events. To this end, we manually curated from the literature the majority of well-characterized kinase- and phosphatase-substrate relationships (654 and 141, respectively) (Figure 2a, Supplementary Table 3). Of these 795 pair-wise connections, our E-MAP contained data for 252 (32%) of them (Supplementary Table 4). We analyzed the genetic interaction patterns for different subsets of the manually curated set of known phosphorylation events: kinases and their substrates, phosphatases and their substrates, two kinases that share the same substrate, two phosphatases that share the same substrate, and kinase-phosphatase pairs that target a common substrate (Figure 2b; Supplementary Tables 4, 5). As a metric to describe the trends within these different subsets, we employed a ratio of highly positive ($S \geq 2.0$) to negative ($S \leq -2.5$) genetic interactions (P to N ratio). Importantly, the relative ratios obtained

for the different subsets were independent of the thresholds used to define the positive and negative genetic interactions (see Methods). The E-MAP data revealed a significant enrichment of positive genetic interactions for known kinase- and phosphatase-substrate pairs. While the E-MAP as a whole displayed a P to N ratio of 0.53, this ratio was almost inverted for characterized kinase- and phosphatase-substrate relationships (P to N ratio: 1.78; $p= 1.8 \times 10^{-3}$) (Figure 2b, Supplementary Tables 3, 4, 5). For example, we detected positive genetic interactions between the *CAK1* kinase and two of its substrates, *CTK1* and *SMK1* (Figure 2a (1)) (Ostapenko and Solomon, 2005; Schaber et al., 2002). These trends might be rationalized as follows: if the phosphorylated form of a substrate is important for cell growth, then either mutation of the substrate, or loss of its phosphorylation through mutation of the cognate kinase, will result in impaired cell growth. Once the kinase has been deleted (and the level of substrate phosphorylation has decreased), additional deletion of the substrate will result in a less severe growth defect than expected from the two single mutations, thus giving rise to a positive genetic interaction. Similar logic can be applied to phosphatase-substrate relationships that display positive genetic interactions (e.g. *SIT4-GCN2* (Cherkasova and Hinnebusch, 2003) or *SIT4-SAP155* (Luke et al., 1996), Figure 2a (2)) if the dephosphorylated form of a substrate is important for cell growth. Conversely, pairs of kinases (or phosphatases) that act on the same substrate can display negative interactions amongst themselves, if they are partially or completely redundant. Indeed, we observed an enrichment of negative genetic interactions in these cases (P to N ratio: 0.21; $p= 2.5 \times 10^{-3}$) (Figure 2b, Supplementary Table 5), including a negative

interaction between the *CLA4* and *STE20* kinases, both of which phosphorylate Myo3 (Wu et al., 1997), and between *PPH21* and *PPH22*, duplicated PP2A phosphatase family members known to act redundantly on a number of different substrates (Figure 2a (3, 4)) (Oficjalska-Pham et al., 2006).

The most striking P to N ratio was observed for kinase-phosphatase pairs that target a common substrate. These pairs of counterbalancing enzymes proved to be much more likely to display positive genetic interactions (P to N ratio: 5.5; $p=2.8 \times 10^{-3}$) (Figure 2b, Supplementary Table 5). For example, a positive genetic interaction was observed between deletions of the *SNF1* kinase and the phosphatase *SIT4*, both of which regulate the phosphorylation status of the Gln3 transcription factor (Figure 2a (5)) (Bertram et al., 2002; Tate et al., 2006; Wang et al., 2003). In these cases, if optimal cell growth requires a substrate to maintain a particular level of phosphorylation, then mutation of its phosphatase (or kinase) will perturb the steady state phosphorylation levels, resulting in a growth phenotype. By additionally deleting the counterbalancing kinase (or phosphatase), wild-type phosphorylation levels may be restored, and the growth defect may be suppressed. We therefore contend that this quantitative genetic platform can be used to identify kinase-substrate pairs, phosphatase-substrate pairs, and kinase-pairs, phosphatase-pairs and kinase-phosphatase pairs that share a common substrate (see Supp. Tables 2 and 6 for complete list of genetic interactions).

While we found a significant enrichment of positive interactions within the set of kinase-substrate, phosphatase-substrate and kinase-phosphatase pairs that share a common substrate, many of the literature-curated relationships were not

captured by our genetic analysis (Supplementary Table 5). There are several reasons why we would not expect a perfect correlation between these two datasets. First, many signaling pathways are only fully activated in response to a particular stimulus (i.e. high salt/Hog1). Since our genetic interaction screen was carried out under standard growth conditions, most pathways are only operating at basal levels. As a result, their perturbation may not yield a significant growth phenotype. Generating E-MAP data under a number of different stresses would help detect many of these stimulus-specific genetic interactions. Second, there are many duplicated copies of kinases or phosphatases that exist in the yeast genome (e.g. *MKK1* and *MKK2*) and this redundancy can mask potential genetic interactions. In these cases, triple mutants could provide richer interaction profiles and help identify additional connections. Despite these limitations, we detected a significant number of genetic interactions that correlated with previously described network connections, illustrating the predictive power of this dataset.

To investigate the characteristics of genetic interactions among components of the signaling apparatus, we analyzed the P to N ratio among and between the protein kinases and phosphatases. Compared to previous data sets, including an E-MAP of chromosome function (Collins et al., 2007b) which had a P to N ratio of 0.49, we found that there was a large enrichment of positive genetic interactions within the set of protein kinases and protein phosphatases (P to N ratio: 0.78; $p=1.0 \times 10^{-3}$) and between kinases and phosphatases (P to N ratio: 0.87; $p=5.1 \times 10^{-3}$) (Figure 2b, Supplementary Figure 2, Supplementary Table 5). The trend towards positive

genetic interactions could stem from the fact that kinases and phosphatases often work together in linear pathways.

Signaling E-MAP Reveals Factors Involved in Histone Htz1 Deposition and Acetylation

To test the predictive power of the E-MAP data, we examined the positive genetic interactions we identified with the histone H2A variant, Htz1. Htz1 is incorporated into chromatin by the SWR-C chromatin-remodeling complex (Korber and Horz, 2004) and is subsequently acetylated on its amino-terminus by the histone acetyltransferase, NuA4 (Babiarz et al., 2006; Keogh et al., 2006b; Millar et al., 2006). Consistent with this, *HTZ1* displayed positive genetic interactions with *SWR1* (+3.9), the catalytic subunit of SWR-C, as well as with *VPS71* (+3.5) and *VPS72* (+3.5), two other components of the complex (Figure 3a). Significant positive interactions were also observed with other factors not previously known to function with Htz1. These included Bud14, a regulatory subunit for the phosphatase Glc7 (Cullen and Sprague, 2002) and Clb2, a B-type cyclin that regulates Cdc28 (CDK) activity (Cullen and Sprague, 2002; Rua et al., 2001) (Figure 3a). Deletion of either *BUD14* or *CLB2* resulted in a marked decrease in Htz1 lysine 14 acetylation (Htz1-$K14^{Ac}$; Figure 3b), indicating that these proteins impinge on the abundance of a chromatin-associated form of this histone variant (Keogh et al., 2006b). Fractionation experiments determined that deletion of *BUD14* resulted in almost a 3-fold decrease in nuclear Htz1, which correlated with the decrease (~4 fold) in the K14-acetylated form of the histone (Figure 3c). The Bud14-associated phosphates

Glc7 has been shown to impinge on chromatin structure by dephosphorylating serine 10 of histone H3, a mark linked to chromosome transmission (Hsu et al., 2000). Therefore, Glc7/Bud14 may regulate Htz1 incorporation by SWR-C into chromatin, either directly or indirectly, potentially through histone H3 (Figure 3d).

Conversely, the effect of *clb2*Δ was primarily seen at the level of Htz1 acetylation. Here the levels of nuclear Htz1 were unperturbed by deletion of *CLB2*, but Htz1-K14^{Ac} was decreased 3 fold (Figure 3c). This implies that Clb2 exerts a regulatory function on NuA4, the enzyme complex that adds this acetylation mark (Figure 3d). An independent study determined that Yng2, a component of NuA4 required for its histone acetyl transferase activity (Krogan et al., 2004), is a target of Clb2-depenedent Cdc28 phosphorylation on two consensus CDK sites (T185 and S188) (L. Holt, J. Villen, S. Gygi and D. Morgan, unpublished data). This connection is also in agreement with previous reports that suggested Htz1 function is cell-cycle regulated (Dhillon et al., 2006). Although further work will be required to work out the mechanistic details of Htz1 regulation, these data confirm the predictive power of individual positive genetic interactions from the E-MAP.

Mapping Genetic Interactions onto Characterized Signaling Pathways

We next explored how the genetic interactions mapped within signaling cascades, involving multiple kinases and/or phosphatases. The HOG (High Osmolarity Glycerol) pathway is a well-characterized MAP kinase cascade that controls the response to osmotic shock (Figure 4a) (Hohmann, 2002). The Sln1 branch of the pathway is activated by the Sln1 osmosensor, which leads to

activation of two MAP kinase kinase kinases (MAPKKKs), Ssk2 and Ssk22. These two kinases then phosphorylate a dedicated MAP kinase kinase (MAPKK), Pbs2, which in turn dually phosphorylates and activates the MAP kinase (MAPK) Hog1. Under iso-osmotic conditions, the pathway is down-regulated by the Ypd1 phosphotransferase, and Ptc1, a phosphatase that, together with its scaffolding protein Nbp2, controls Hog1 dephosphorylation (Figure 4a). Mutations of these negative regulators (*YPD1*, *PTC1* and *NBP2*) primarily displayed positive genetic interactions with the activating proteins in the pathway (Figure 4b), presumably suppressive relationships that counterbalance the levels of phosphorylation in the pathway. In contrast, the genes coding for protein phosphatases that control Hog1 basal and maximal activity (*PTC1*, *PTC2*, *PTC3*, *PTP2*, *PTP3*) show largely negative genetic interactions among themselves (Figure 4b), suggesting a certain degree of redundancy.

In addition to the direct genetic interaction score between gene pairs, each gene possesses a genetic interaction profile which describes its interactions with all other genes in the E-MAP (see http://interactome-cmp.ucsf.edu (2008)). The genetic interaction profile provides a high resolution phenotype, and functionally related genes often have similar interaction profiles. For example, *hog1*Δ and *pbs2*Δ showed the most highly correlated profiles among the HOG pathway genes, indicating they are the most functionally related pair of proteins in this pathway (Figure 4c). Also correlated, but to a lesser extent, are *ssk1*D and *ssk2*D, which we believe is due to the fact that Pbs2 can also be activated by an alternative MAPKKK, Ste11 (O'Rourke and Herskowitz, 2004). Interestingly, *ssk2*D is more highly

correlated with *hog1*Δ and *pbs2*Δ when compared to *ssk22*D (Figure 3c), which suggests that Ssk2 is the predominant MAPKKK for pathway activation under the growth conditions of our assay. Consistent with that the interactions of *ssk2*Δ with genes encoding the negative pathway regulators *YPD1*, *PTC1*, and *NBP2* are much more positive than those seen with *ssk22*Δ (Figure 4b). Overall, the information encapsulated within both the individual pair-wise scores and the correlation coefficients between the genetic interaction profiles can be used to resolve pathway architectures more powerfully than any single observation.

The signaling E-MAP also contains the majority of small molecule kinases and phosphatases (Figure 1b, Supplementary Table 1), including several enzymes involved in inositol polyphosphate metabolism. Inositol phosphate signaling is initiated by the phospholipase Plc1, resulting in the formation of IP_3, which is converted to IP_5 by Ipk2, then to IP_6 by Ipk1 (Figure 4d) (Alcazar-Roman and Wente, 2008). IP_6 can subsequently become diphosphorylated by Kcs1 or Vip1 to respectively yield the IP_7 metabolites, 5-PP-IP_5 or 4/6-PP-IP_5 (Figure 4d) (Mulugu et al., 2007; Saiardi et al., 2004). Genetic analysis of these enzymes (except for *ipk2D*, which had mating and sporulation defects) revealed that all members of the pathway displayed positive genetic interactions with each other, except for *VIP1* and *KCS1* (Figure 4e). This is consistent with a linear pathway for the synthesis of IP_6, which then branches to yield the different IP_7 metabolites (Figure 4d) (Alcazar-Roman and Wente, 2008). Although the overall genetic interaction profiles for these four enzymes are similar, there are almost no examples where all four enzymes display the same interactions with a given mutant. For example, all components

except *VIP1* displayed strong negative interactions with the cyclin *CLB2* and *PPH3*, a phosphatase linked to cell cycle progression (Keogh et al., 2006a). Since Plc1, Ipk1 and Kcs1 – but not Vip1 – are involved in the synthesis of the 5-PP-IP_5 metabolite, these data may imply a functional connection of the 5-PP-IP_5 metabolite with *CLB2* and *PPH3*. Furthermore, only *PLC1* and *IPK1* showed positive genetic interactions with *ELP2* and *ELP3*, two proteins linked to transcriptional elongation and tRNA function (Svejstrup, 2007), which suggests that *PLC1* and *IPK1*, or the metabolites they synthesize, impinge on those processes (Figure 4e). Our results concur with recent evidence that specific inositide metabolites are involved in distinct cellular functions (Alcazar-Roman and Wente, 2008) and thus give rise to a wide range of genetic interaction patterns.

Overall, the inositol polyphosphate and MAP kinase cascades are two examples where the signaling E-MAP supports previously described network connections, and provides new functional insight into these well-characterized pathways.

Analysis of "Triplet Genetic Motifs" (TGMs)

With the goal to extract new pathway connections, we further analyzed the genetic interaction data by extracting triplet motifs. In these "triplet genetic motifs" (or TGMs), sets of three genes were mutually connected to each other by positive ($S \geq 2.0$) or negative ($S \leq -2.5$) genetic interactions. Triplet motifs are the simplest motifs, apart from binary pairings, and we propose that gene triplets displaying genetic interactions are very likely to be functionally connected. There are four

distinct types of TGMs: all three genetic interactions are positive (Type I); two positive, one negative (Type II); one positive, two negative (Type III); and all three negative (Type IV) (Figure 5a). Using only the genes coding for protein kinases, we generated each type of TGM (Supplementary Table 7) and assembled a network diagram where the motifs were connected if they shared one or two nodes (Figure 5b). We posit that TGMs comprised of three positive genetic interactions (Type I) will be enriched for factors functioning in the same pathway whereas all negative TGMs (Type IV) would potentially correspond to genes functioning in three distinct, redundant pathways (Figure 5a). Type II (++-) and Type III (+--) TGMs represent intermediate cases. For example, a Type II TGM was observed for *PTK2*, *HAL5*, and *SAT4*, three kinases that are all involved in modulating the major electrogenic transporters of yeast (Pma1, Trk1, Trk2) (Supplementary Figure 3) (Goossens et al., 2000; Mulet et al., 1999). Ptk2 phosphorylates and activates the H^+-ATPase Pma1, thereby increasing the electrical membrane potential, while the Trk1,2 K^+ transport system is positively regulated by Hal5 and Sat4 kinases, and is a major consumer of electrical potential. Together, these two systems determine the steady-state membrane potential and thereby regulate secondary active transport systems. Due to their partial redundancy, *HAL5* and *SAT4* showed a negative genetic interaction, however, both genes displayed positive genetic interactions with *PTK2*, likely a suppressive relationship between the generator and the consumers of the membrane potential. Interestingly, the *PTK2*-*HAL5* edge displayed negative genetic interactions with *PBS2* (resulting in a Type III TGM) (Supplementary Figure 3), which illustrates how cells with a compromised ability to control ion homeostasis

are unable to tolerate an additional mutation in a separate/parallel pathway that is important for responding to salt stress.

The three kinases most frequently found in TGMs were *CLA4*, which has been functionally linked to cytokinesis (12 TGMs); *BUB1*, which is essential for proper spindle checkpoint function (13 TGMs); and the AMP-activated kinase *SNF1*, required for glucose-repressed transcription and thermotolerance (14 TGMs) (Figure 5b). These three kinases were also the most genetically promiscuous when we extended this analysis to the other genes on the signaling E-MAP (Supplementary Figure 4), suggesting they occupy central signaling nodes in budding yeast.

Consistent with the pattern observed with the pair-wise genetic interactions (Figure 2b, Supplementary Figure 2), we found an overall enrichment of positive genetic interactions among and between kinases and phosphatases participating in the TGMs. When the TGMs were restricted to contain only kinases and phosphatases, we found an approximately equal number of the four types of motifs (Figure 5c). As non-kinase/phosphatase genes were introduced, the proportion of TGMs with positive genetic interactions decreased (Figure 5c).

Cak1 and Fus3 Kinases Function in the Set2/Rpd3C(S) Pathway

To uncover previously uncharacterized signaling pathways, we next focused on the Type I (+++) TGM, where only one protein was required to be either a kinase or phosphatase. We assembled a Type I TGM list, restricted to factors that were not physically associated (purification enrichment scores < 0.2 (Collins et al., 2007a)),

since we wanted to enrich for pathways rather than protein complexes. Taking the 47 most significant Type I TGMs, as defined by the product of their three individual genetic interaction scores, we assembled a network diagram (Supplementary Figure 5) to identify regions containing high-density, positive genetic interactions. Several such clusters were discovered, including the *SET2-CAK1* pair, which showed positive genetic interactions with *CTK1*, *CTK2*, *EAF3*, *ADO1* and *FUS3* (Figure 6a, Supplementary Figure 5). Several of these proteins have previously been demonstrated to act together in a pathway that is required for efficient transcriptional elongation by RNA polymerase II (RNAPII). In this cascade, the kinase Ctk1 (with its regulatory proteins, Ctk2 and Ctk3), phosphorylates Rpb1, the largest subunit of RNAPII, on its C-terminal domain (Cho et al., 2001). This modification recruits the methyltransferase Set2 to methylate Lys36 of histone H3 during transcriptional elongation (Hampsey and Reinberg, 2003). This mark, in turn, recruits Eaf3, a component of the Rpd3C(S) histone deacetylation complex, which then deacetylates histones H3 and H4 in the coding region of genes, resulting in chromatin compaction, an event that ultimately suppresses spurious cryptic initiation (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005).

Since *CAK1* displayed positive genetic interactions with multiple factors implicated in this pathway, we predicted that it would be an integral component of the RNAPII transcriptional elongation control process. The positive genetic interactions of *CAK1* with *SET2* and *EAF3* were confirmed by tetrad analysis and spot testing, which showed that *eaf3D* and *set2D* suppressed the slow growth phenotype observed in a *cak1*-DAmP background (Figure 6b). In contrast, the

positive genetic interaction between *CAK1* and *CTK1* corresponded to an epistatic interaction since the double mutant was no sicker than either of the two single mutants (Supplementary Figure 6b). To validate the genetic interaction observed between *CAK1* and *CTK1*, we expression-profiled both mutants using DNA microarrays. This revealed a very significant overlap in mRNA expression of genes affected in either mutant, implying that these two kinases function together *in vivo* (Supplementary Figure 6c, Supplementary Table 8). Furthermore, out of a large collection of kinase mutant expression-profiles, the *ctk1D* profile was most similar to that of the *cak1*-DAmP strain (S. van Wageningen and F. Holstege, unpublished data).

Finally, the *cak1*-DAmP strain, as well as other previously characterized *CAK1* mutants (*cak1-22*, *cak1-23*, *cak1-95*) (Espinoza et al., 1998), showed a significant increase of transcription initiation from a cryptic internal TATA site within *FLO8* (Figure 6c). A similar phenotype has previously been observed with *set2*D and deletions of the chromatin assembly factors *SPT2* and *HIR1* (Figure 6c) (Carrozza et al., 2005; Kaplan et al., 2003; Nourani et al., 2006). In agreement with these data, Ostapenko and Solomon reported that Cak1 phosphorylates Ctk1 *in vivo* (Ostapenko and Solomon, 2005), and we found that deletion of *CTK1* (or *CTK2* and *CTK3*) also results in a cryptic initiation defect (Figure 6e) (Cheung et al., 2008; Youdell et al., 2008).

Interestingly Bur1, a kinase known to suppress spurious transcription initiation, is another well characterized Cak1 substrate (Yao and Prelich, 2002). We previously showed that deletion of *SET2* or *EAF3* suppressed the lethality associated

with *bur1*Δ (Keogh et al., 2005), suggesting cross-talk between the two pathways (Figure 6d). The negative genetic interaction between *CAK1* and *BUR1* (Supplementary Figure 6d) suggests Bur1 is involved in another, parallel pathway. Overall these data demonstrate that Cak1 acts as a key regulator controlling two kinases cascades involved in regulating intergenic chromatin integrity (Figure 6d).

Both *CAK1* and *SET2* also displayed positive genetic interactions with *FUS3*, a MAP kinase involved in the mating pathway (Figure 6a). Further testing revealed that *fus3*Δ gave rise to a cryptic initiation defect as well (Figure 6e). Other factors known to work with *FUS3* in the mating pathway (*KSS1*, *STE50*, *DIG2*) (Chen and Thorner, 2007) or other MAP kinases (*HOG1* and *SLT2*) did not display positive genetic interactions with *CAK1* and *SET2*, and accordingly did not have the same transcriptional defect (Figure 6e). Work by Young and Colleagues demonstrated that Fus3 localizes to actively transcribed genes by chromatin immunoprecipitation (Pokholok et al., 2006). Taken together with the genetic interaction data, these findings imply that Fus3 may phosphorylate one or more chromatin factors involved in suppressing cryptic initiation. Further work will be required to reveal the mechanistic details of this intriguing connection between the mating pathway and transcriptional chromatin integrity

Perspective

In summary, we have systematically and quantitatively mapped the pair-wise genetic interactions within the yeast signaling machinery. The entire genetic interaction matrix of 100,000 pair-wise interactions is available in a web-based

searchable format, as illustrated in Supplementary Figure 7, at http://interactome-cmp.ucsf.edu. We have analyzed the quantitative genetic interaction data using triplet genetic motifs (TGMs) which we show can be a powerful approach for globally studying signaling pathways, and uncovering mechanistic details of how specific proteins function within these cascades. Our data set, when integrated with other genome-wide data sets, will help to understand and model the phosphate-based signaling behaviour of yeast cells. Furthermore, this signaling E-MAP also provides a launching platform for other analyses. While the current screen was carried out on rich media, many signaling pathways depend on specific stimuli. Condition-dependent screens will, therefore, reveal dynamic changes in the signaling apparatus. We also plan to compare the genetic interaction profiles of complete kinase gene deletions with those of catalytically dead mutants and analog-sensitive kinase alleles (Bishop et al., 2000). Such analyses should allow us to separate the catalytic activities from the scaffolding functions of these proteins. Comparison of the genetic architecture of the signaling networks between different yeast species (Roguev et al., 2008) and higher organisms (Bakal et al., 2008) will be instructive in understanding how regulation via reversible phosphorylation is conserved and how it has evolved. Finally, epistasis analyses have significant potential, not only for functionally connecting proteins into pathways and complexes, but also for identifying combinations of genes, or genes operating at critical nodes at the intersection of signaling pathways, that can serve as appropriate therapeutic targets.

Methods

Strains were constructed and E-MAP experiments were performed as previously described (Collins et al., 2007b; Collins et al., 2006; Schuldiner et al., 2005).

To highlight strong genetic interactions, we introduced a cutoff of S ≤ -2.5 for negative interactions and S ≥ 2.0 for positive interactions. Positive to negative (P to N) ratios were computed for the subsets depicted in Figure 2a. The sets consisting of kinase and phosphatase pairs that act on the same substrates were compared with a subset of the signaling E-MAP that contains only kinases, phosphatases and regulatory subunits. Two-tailed P-values were computed using Fisher's exact test. The remaining sets in Figure 2a were compared with the chromosome function E-MAP. Here, two-tailed P-values were computed using chi-square with Yates' correction. Importantly, the relative ratios obtained were independent of the thresholds used to define the positive and negative genetic interactions (data not shown). Kinases, phosphatases, kinase-substrate pairs and phosphatase-substrate pairs are listed in Supplementary Tables 3,4,5

The preparation of yeast whole cell extracts by TCA (Figure 3b) and cell fractionation experiments (Figure 3c) were performed essentially as described (Keogh et al., 2006b). The *GAL1-FLO8-HIS3* reporter (Figure 6c) consists of: (i) the *FLO8* promoter replaced by that of galactose-regulated *GAL1/10*, and (ii) the *HIS3* gene inserted out of frame into *FLO8* (*flo8*D (+1729-2505)::*HIS3* (+1-663)) such that *HIS3* product is only translated when transcription initiates from a cryptic promoter within *FLO8*(Nourani et al., 2006). In the almost identical *GAL1*$_{Sp}$*-FLO8-HIS* reporter

(Figure 6e), $GAL1_{Sp}$ is an attenuated version of the GAL1 promoter, missing 1.5 copies of the 5' GAL4-UAS (Mumberg et al., 1994). Reporter strains (Supplementary Table 9) were created by direct transformation, or mating followed by tetrad dissection. *HIS3* expression was monitored by spotting the appropriate strains onto synthetic complete (SC) medium +/- histidine with either glucose (2%) or galactose / raffinose (2% / 1% respectively) as the carbon source.

Gene expression profiling was carried out as previously described (van de Peppel et al., 2005). *P*-values for the overlap of the Venn diagrams were obtained using the hypergeometric test with a population size of 6389 genes.

Acknowledgements

The authors thank A. Roguev, C. Kaplan, C. J. Ingles, P. Aguilar, T. Walther, J. H. Morris and members of the Krogan, Shokat and Keogh labs for advice and comments. We also thank W. A. Lim for discussion on MAP kinase cascades. We are grateful to A. Chan, E. Cheng, Y. Nijati, and Li Jieying for technical assistance. We thank F. Winston and S. Buratowski for providing yeast strains and B. Strahl, S. Hahn, and D. Morgan for sharing unpublished data. The work was supported by an Ernst Schering Postdoctoral Fellowship (D.F.), the Biophysics Graduate Group at UCSF (H.B.), the National Science Foundation (G.C. and D.K.), the New York Speakers Fund for Biomedical Research (M-C.K.), the Howard Hughes Medical Institute (K.M.S.) and the NIH (N.J.K. and K.M.S.), and from Sandler Family Funding (N.J.K.).

Author information

Correspondence and requests for materials should be addressed to Nevan J. Krogan (krogan@cmp.ucsf.edu) or Kevan Shokat (shokat@cmp.ucsf.edu)

Figures

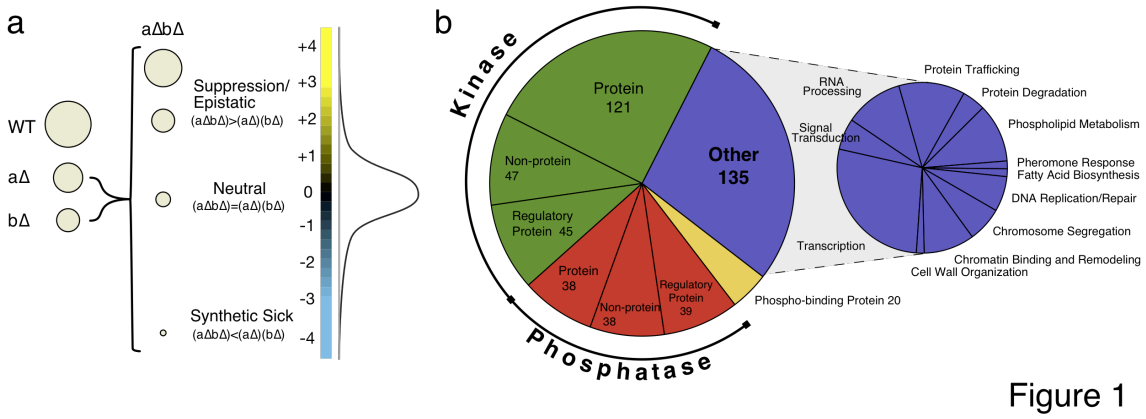


Figure 1

Figure 1. Epistasis analysis of the yeast signaling machinery. a) The entire spectrum of genetic interactions. Genetic interactions range from negative (e.g. synthetic sick) to positive (e.g. suppression) where the growth rate of the double mutant is either less ((aΔbΔ) < (aΔ)(bΔ)) or greater ((aΔbΔ) > (aΔ)(bΔ)) than the product of the growth rates of the corresponding single mutants, respectively. As shown in the distribution curve of our data set, significantly negative ($S \leq -2.5$) and positive ($S \geq 2.0$) genetic interactions are rare. b) Composition of the signaling E-MAP. For a full list of genes analyzed in this study, see Supplementary Table 1.

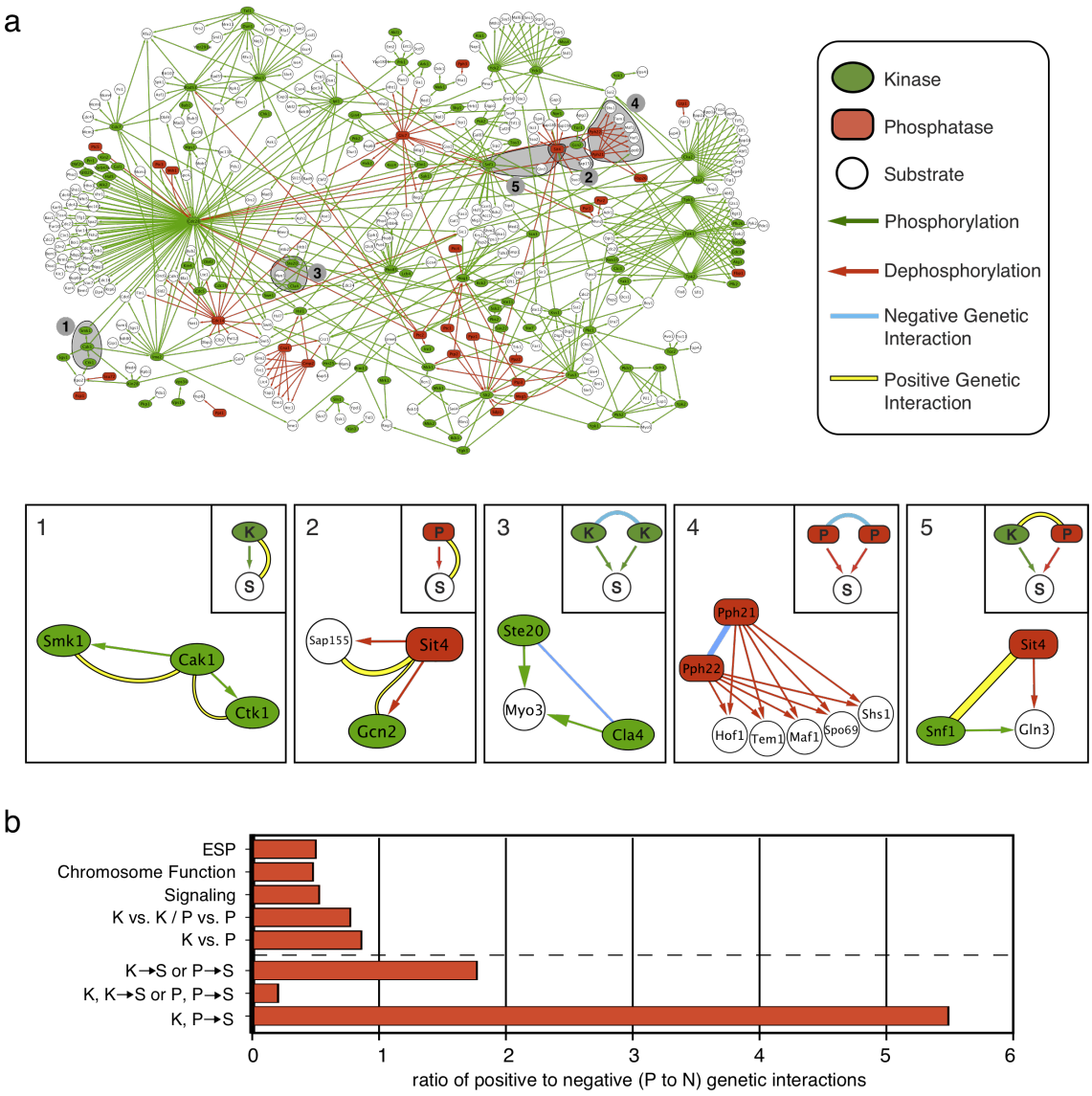


Figure 2

Figure 2. Comparison of the literature-derived signaling network to the genetic interaction data. a) A network diagram of characterized phosphorylation and dephosphorylation events between kinases and their substrates (green arrows) and phosphatases and their substrates (red arrows), manually curated from the literature (see Supplementary Table 2). Below this are specific examples of kinase-

substrate (1) and phosphatase-substrate relationships (2), and cases where two kinases (3), two phosphatases (4) or one kinase and one phosphatase (5) target one or more substrates. Blue and yellow edges correspond to negative and positive genetic interactions, respectively. The thickness of the edge correlates with the strength of the genetic interaction. b) Ratios of highly positive ($S \geq 2.0$) to negative ($S \leq -2.5$) genetic interactions (P to N ratio). Above dashed line: The P to N ratios of previously published datasets (early secretory pathway (ESP)(Schuldiner et al., 2005) and chromosome function(Collins et al., 2007b)), the signaling E-MAP, only protein kinases and protein phosphatases (K vs. K or P vs. P), and protein kinases versus protein phosphatases (K vs. P) are presented. Below dashed line: The P to N ratios for known kinase-substrate and phosphatase-substrate pairs (K → S, P → S) (out of these 252 relationships that we genetically tested, we observed significant genetic interactions (positive and negative) for 25 of them (10%)), as well as kinase-kinase, phosphatase-phosphatase (K,K → S; P,P → S) and kinase-phosphatase pairs (K,P → S) that share one or more substrates are shown. To obtain these ratios and subsequent p-values, we also included regulatory subunits known to direct kinase/phosphatase activity towards specific substrates (see Methods and Supplementary Table 3).

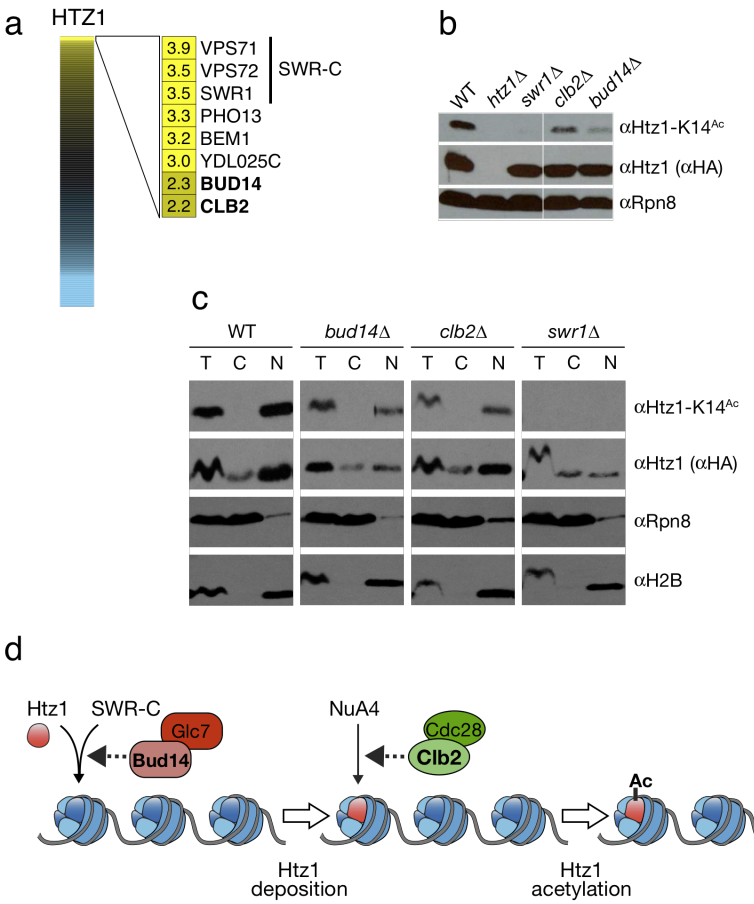


Figure 3

Figure 3. Identification of factors involved in histone Htz1 depostion and acetylation. a) Full spectrum of *HTZ1* genetic interactions. Genes with the strongest positive genetic interactions are highlighted. b) *clb2*Δ or *bud14*Δ profoundly reduce the total pool of Htz1-K14^{Ac}. Whole cell extracts were isolated from isogenic strains containing Htz1 with a C-terminal HA3-tag, resolved by SDS-PAGE and immunoblotted as indicated. Rpn8 was used as a loading control. c) *bud14*Δ reduces both chromatin-associated Htz1 and Htz1-K14^{Ac}, while *clb2*Δ specifically reduces the latter pool. Isogenic strains containing Htz1.HA3 were separated into T(otal), C(ytoplasmic) or N(uclear) fractions and immunoblotted with the indicated

antibodies. Enrichment of the proteasome component Rpn8 and the chromatin component H2B in appropriate fractions (soluble cytoplasm and insoluble nucleus) demonstrate efficient separation. Using the H2B levels as a loading control, Htz1-K14^{Ac} was decreased 3.7- and 3.0-fold in *bud14*Δ and *clb2*Δ respectively. By comparison, nuclear Htz1 was decreased 2.7-fold in *bud14*Δ but comparable to WT in *clb2*Δ. d) Schematic illustration of Htz1 deposition by SWR-C and acetylation by NuA4. The potential points of action of Bud14 and Clb2 in this pathway are indicated.

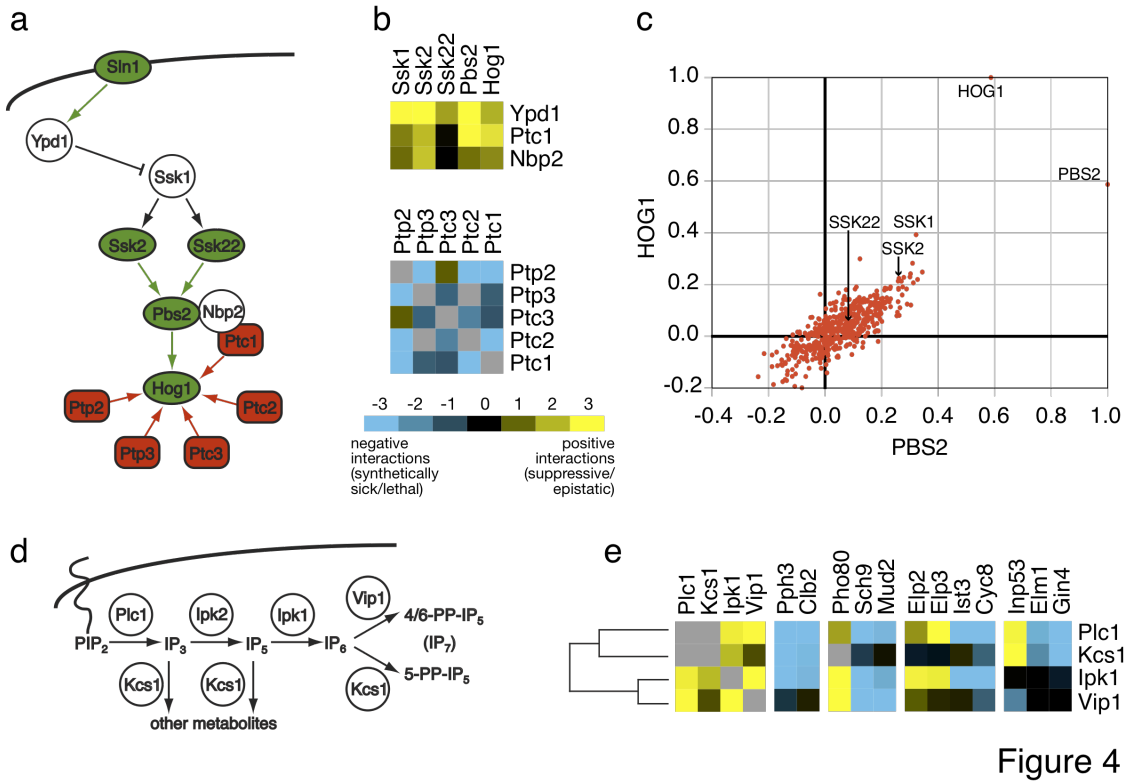


Figure 4

Figure 4. Mapping genetic interaction data onto known signaling pathways. a) General schematic of the Sln1 branch of the HOG pathway. Protein kinases and phosphatases are denoted as green and red, respectively, while green and red arrows correspond to phosphorylation and dephosphorylation actions, respectively. b) The negative pathway regulators *YPD1*, *PTC1*, and *NBP2*, show strong positive genetic interactions with the pathway activators *SSK1*, *SSK2*, *SSK22*, *PBS2*, and *HOG1*. In contrast, genes coding for the protein phosphatases (*PTC1*, *PTC2*, *PTC3*, *PTP2*, and *PTP3*) acting on *HOG1* show primarily negative genetic interactions among themselves. c) Scatter plot of the correlation coefficients of *pbs2*Δ and *hog1*Δ with all genetic profiles in the signaling E-MAP. d) Schematic illustration of the inositol polyphosphate pathway. e) A subset of genetic interactions for *PLC1*, *KCS1*, *IPK1*, and *VIP1*. In b) and e), negative and positive genetic interactions are indicated by

blue or yellow squares, respectively. Gray squares represent genetic interactions not tested.

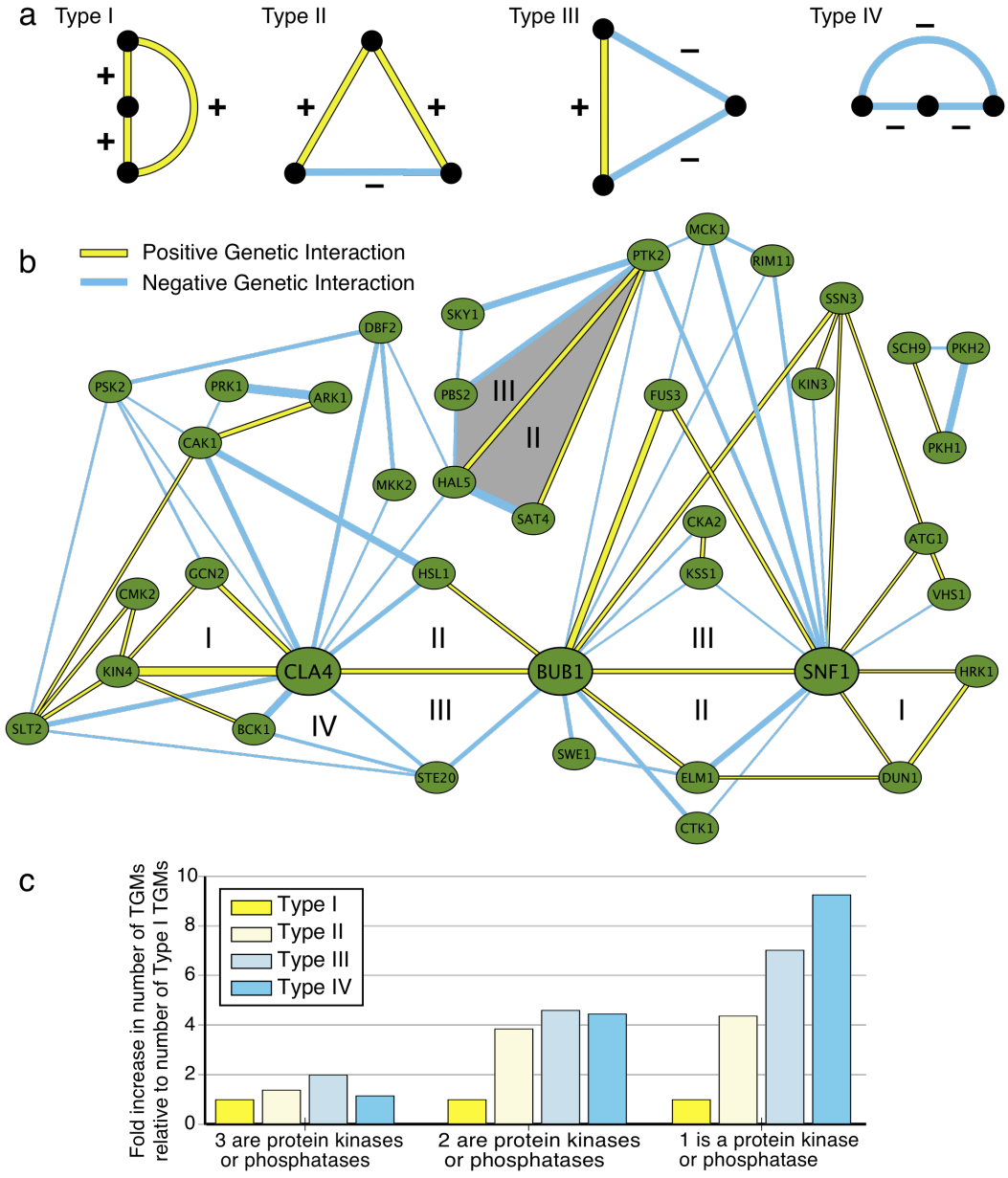


Figure 5

Figure 5. Triplet genetic motifs (TGMs) from the signaling E-MAP. a) A schematic of the four types of triplet genetic motifs: all three positive (Type I); two positive, one negative (Type II); one positive, two negative (Type III); and all three negative (Type IV). b) A network diagram of all TGMs involving three kinases using highly negative ($S \leq -2.5$) and positive ($S \geq 2.0$) genetic interactions. The motifs were connected if they shared one or two nodes; blue and yellow edges correspond to negative and positive genetic interactions, respectively. The thickness of the edges is correlated with the strength of the genetic interaction. TGMs highlighted in grey are discussed in the text. c) Comparison of the ratio of the four types of TGMs that contain one or more kinase or phosphatase. The ratios are obtained by normalizing the number of genetic triplet motifs within each set (i.e. Types II, III, and IV) to the number of TGMs that are all positive (Type I). For a complete list of all TGMs, see Supplementary Table 5.

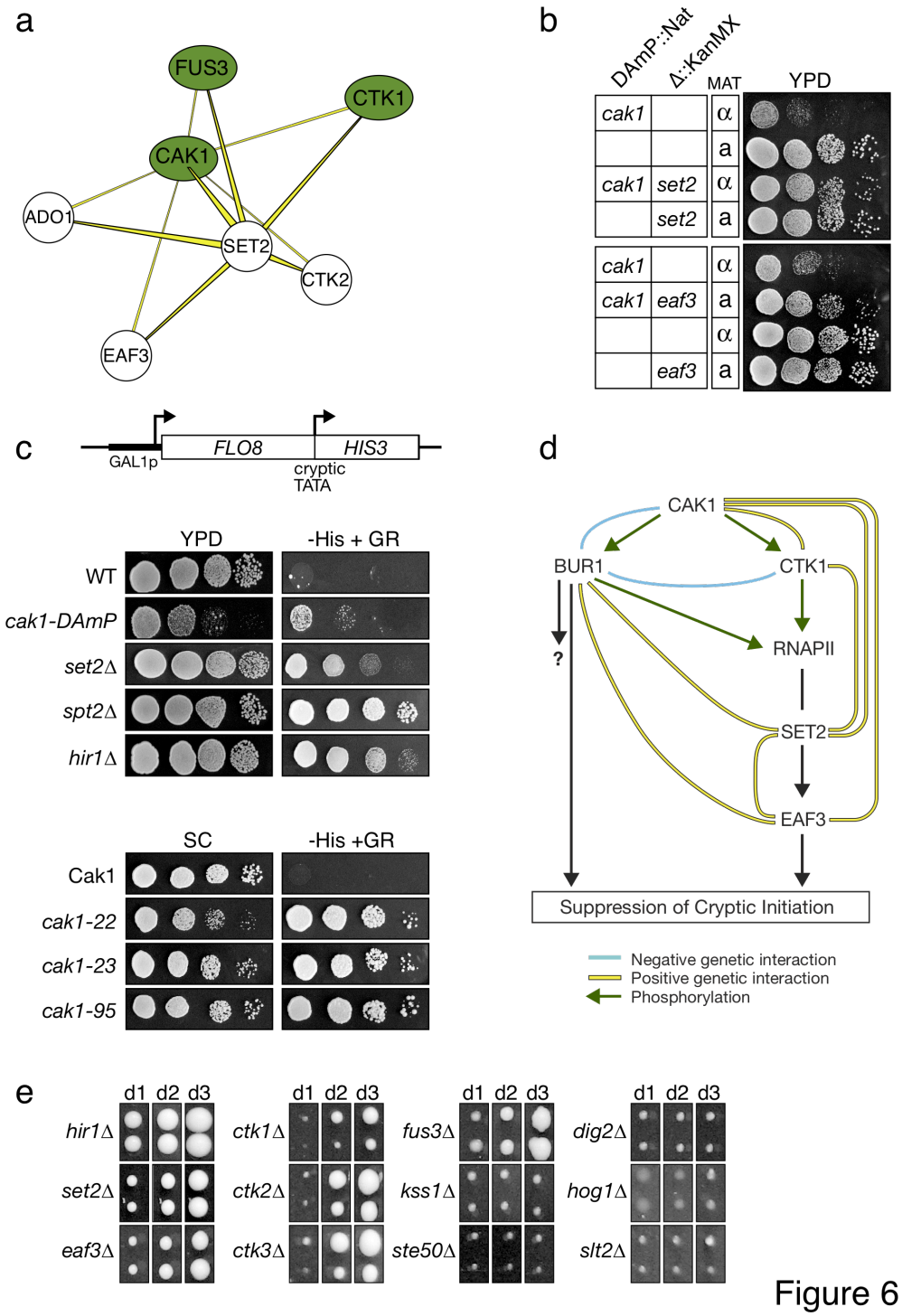


Figure 6

Figure 6. The Cak1 kinase functions in the Ctk1/Set2/Rpd3C(S) pathway that suppresses cryptic initiation by RNA polymerase II. a) Subset of Type I TGMs that share the Cak1-Set2 positive genetic interaction. For a complete picture of all

Type I TGMs see Supplementary Figure 4. b) Tetrad analysis and serial 10-fold dilution spot testing demonstrates that the slow growth of the *cak1*-DAmP allele is strongly suppressed by *set2*D or *eaf3*D. c) *CAK1* suppresses cryptic transcription initiation. The *GAL1-FLO8-HIS3* reporter (Nourani et al., 2006) is described in Methods. The *HIS3* gene product is only produced when transcription aberrantly initiates from a cryptic promoter within *FLO8*, as when chromatin structure in the transcribing gene is disrupted (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). His3 expression was monitored by 10-fold dilution spotting of the indicated strains onto synthetic complete (SC) medium +/- histidine with galactose / raffinose (2% / 1% respectively) as the carbon source. The lower panels employed a series of *cak1*-temperature sensitive alleles (Espinoza et al., 1998) at their semi-permissive temperature (SC, 32^{o}C, 48hrs; -HIS, 32^{o}C, 96hrs). d) Possible schematic of this pathway controlling intergenic chromatin fidelity. e) The MAP kinase Fus3 regulates cryptic initiation. All strains contain a *GAL1*$_{Sp}$*-FLO8::HIS* reporter, similar to that in Figure 6c (see Methods). Reporter strains were pinned in duplicate onto SC-HIS with galactose / raffinose as the carbon source, incubated at 30^{o}C and photographed on days indicated. Panels are standardized to facilitate cross-comparison of reporter expression in each strain / time-point.

Supplementary Figures

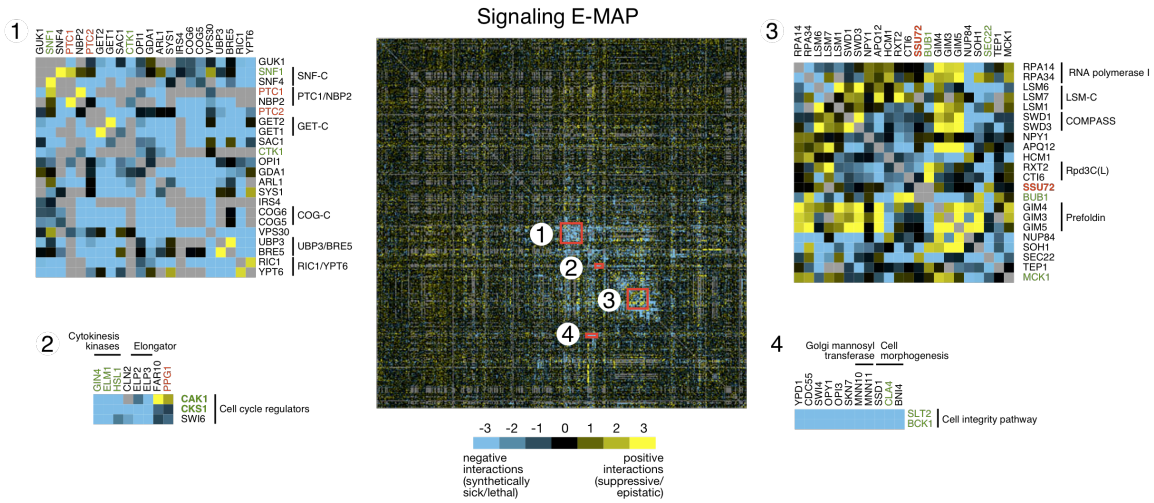


Figure S1. Clustergram of the signaling E-MAP. The entire genetic interaction dataset (483 genes) was subjected to hierarchical clustering and four regions are highlighted. Several protein complexes are shown and protein kinases and phosphatases are labeled in green and red, respectively. Negative and positive genetic interactions are indicated by blue or yellow squares, respectively. Gray squares represent genetic interactions not tested.

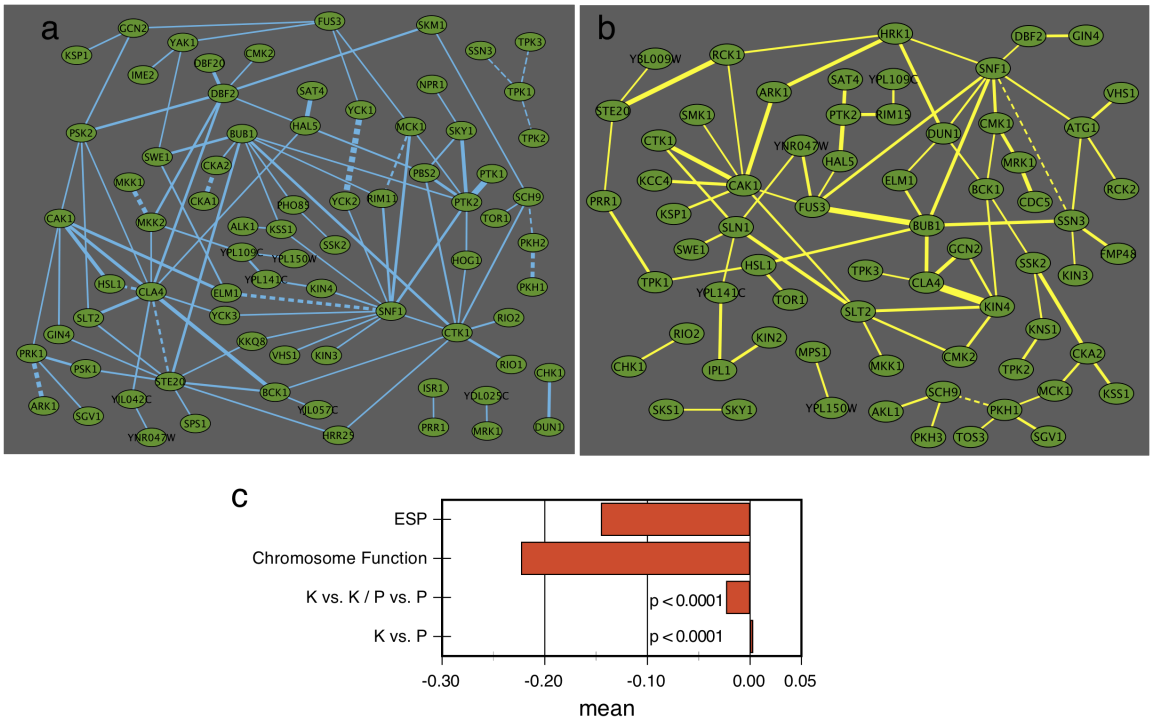


Figure S2. Epistatic architecture within the set of protein kinases and phosphatases. a) Net- work diagram of all highly negative (S ≤ -2.5) interactions between protein kinases. b) Network diagram of all highly positive (S ≥ 2.0) interactions between protein kinases. The thickness of the lines corresponds to the strength of the genetic interactions. Dashed lines indicate pairs of kinases that are known to share common substrates. c) Comparison of the mean scores in different data sets: early secretory pathway E-MAP (ESP), chromosome function E-MAP (Chromosome Function), protein kinases and protein phosphatases (K vs. K / P vs. P), and protein kinases versus protein phosphatases (K vs. P) (Supplementary Tables 4, 5).

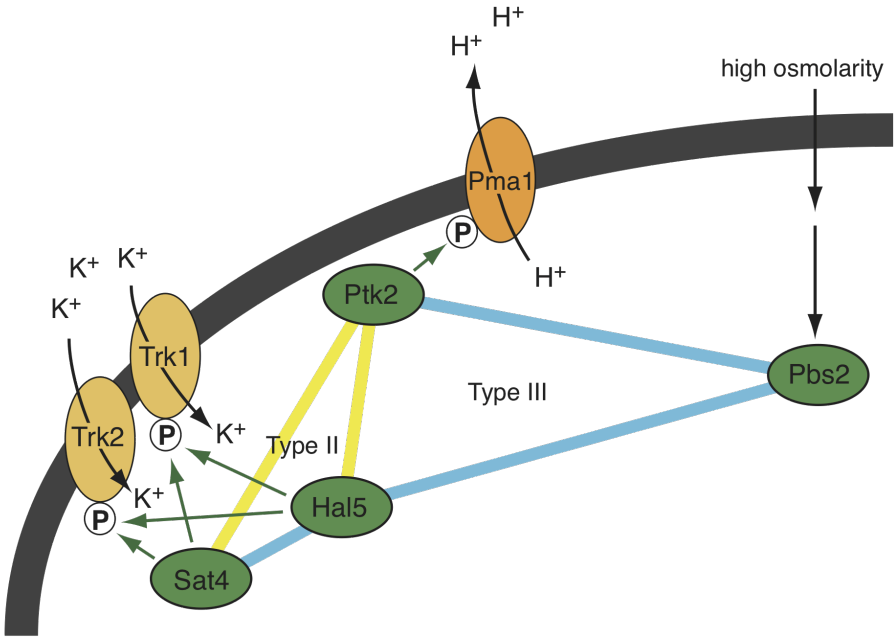


Figure S3. General schematic illustrating the roles of the Ptk2, Hal5 and Sat4 kinases. Ptk2 phosphorylates and activates the H+-ATPase Pma1, thereby increasing the elec- trical membrane potential. The Trk1,2 K+ transport system is positively regulated by Hal5 and Sat4 kinases, and is a major consumer of electrical potential. The relative activity of these two systems determines the steady-state value of the membrane potential and thereby modulates the activity of secondary active transport systems (i.e. nutrient uptake, toxic cation uptake). The kinase Pbs2 functions in a parallel pathway, which regulates the cell's response to high osmolarity. Blue and yellow edges correspond to negative and positive genetic interactions, respec- tively. *SAL4/HAL5/PTK2* and *PTK2/HAL5/PBS2* form Type II (++-) and Type III (+--) TGMs, respectively.

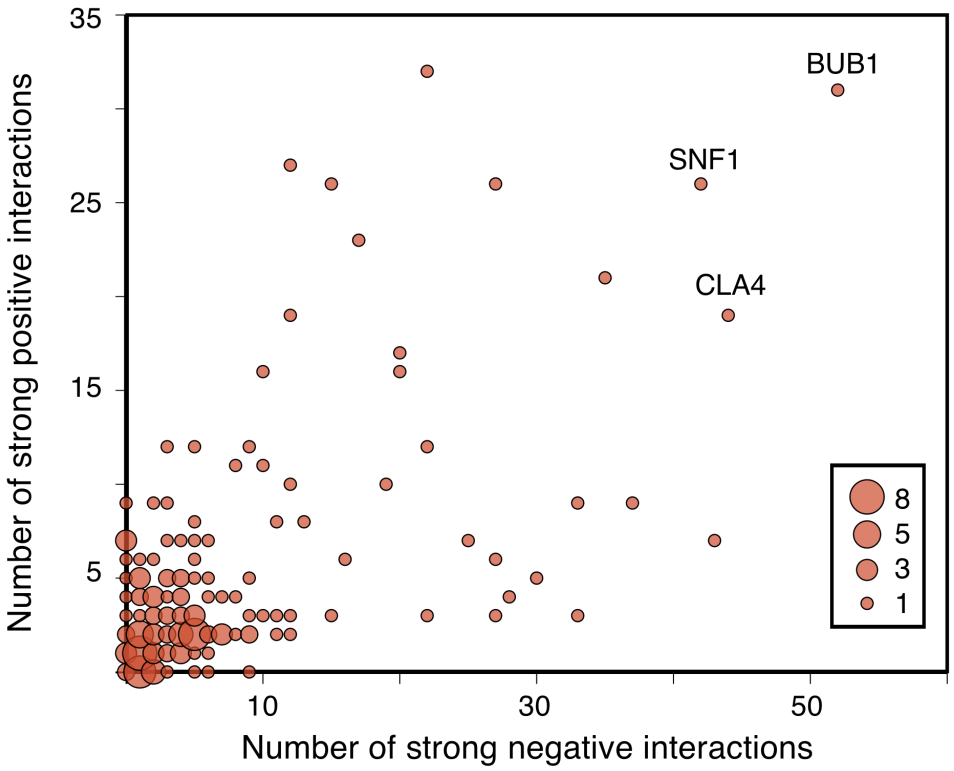


Figure S4. Number of genetic interaction partners among the protein kinases and protein phosphatases. The number of highly posi- tive (S ≥ 2.0) and negative (S ≤ -2.5) genetic interactions of all protein kinases and phosphatases with all other genes in the signaling E-MAP is displayed. Node size corresponds to how many kinases or phosphatases have the same number of positive and negative interactions. *BUB1*, *CLA4* and *SNF1* are the kinases with the greatest number of strong genetic interactions.

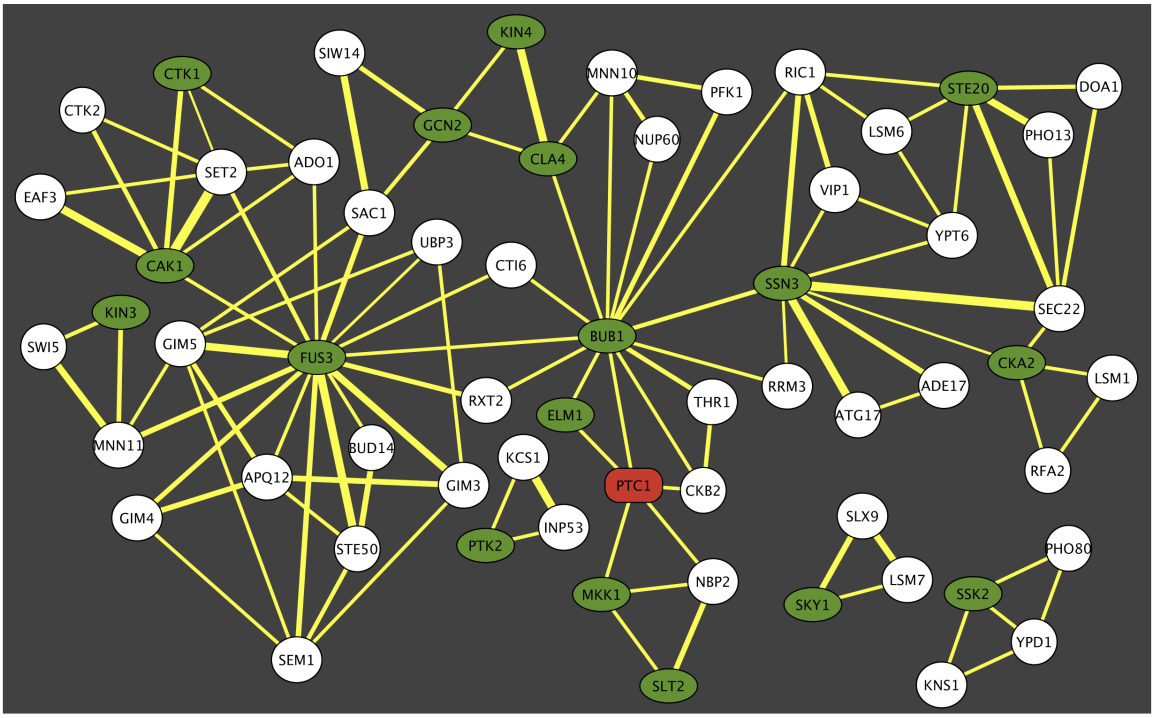


Figure S5. Network diagram of Type I TGMs. Using the product of the three genetic interac- tion scores, the 47 most positive Type I (+++) TGMs are presented in the form of a network diagram where the triplets are merged together if they have one or two nodes are that are shared. Protein kinases are illustrated in green, protein phosphatases in red, and all other proteins in white.

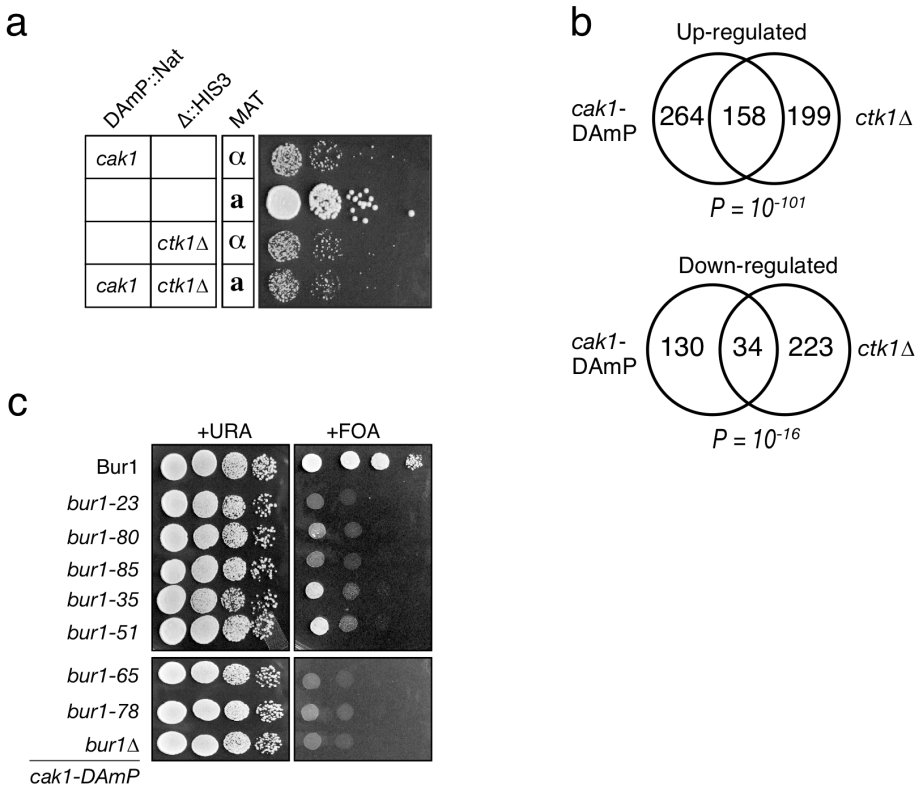


Figure S6. a) The *ctk1Δ* and *cak1*-DAmP are epistatic. Depicted are the products of a representative dissected tetrad (where all tested markers segregated 2:2). Spots are 10-fold dilutions of ODλ 600 ≈ 0.5 cultures, and plates were photographed after 72hrs at 30°C. **b**) Comparison of the *cak1*-DAmP expression profile to the *ctk1Δ* expression profile. Venn diagrams are given of the genes that show a significant change ($p<0.05$, 1.7 fold change cut off) in the *cak1*-DAmP *ctk1Δ* strains. Significance for the overlap is given. **c**) The *bur1/cak1*-DAmP double mutants have strong synthetic defects. *cak1*-DAmP was created by direct transformation of PCR product to a *bur1Δ*::KanMX shuffle strain (genomic *bur1Δ* covered by a Bur1-Cen/Ars-URA3 plasmid). The resulting strain (KLY54) was transformed with a panel of Bur1 plasmids (WT, seven mutant alleles, or empty vector, all in low-copy vector pRS315;

Leu, Cen/ARS) and plated on FOA to remove the Bur1 covering plasmid. Spots are 10-fold dilutions of ODλ600 ≈ 0.5 cultures, and plates were photographed after 72hrs at 30°C, the permissive temperature for the *bur1* mutant alleles in a *CAK1*-background (Keogh et al (2003) MCB 23:7005).

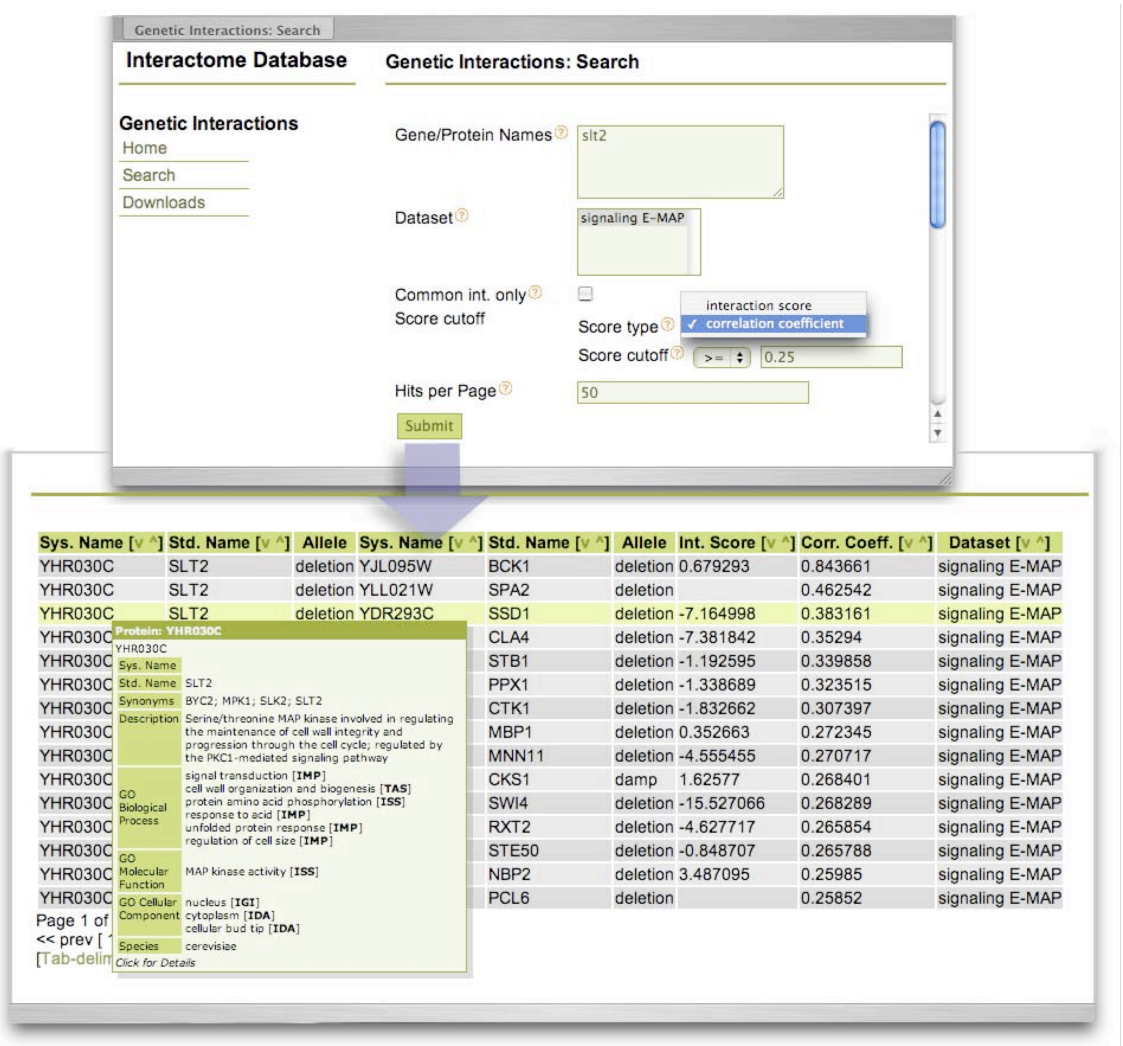


Figure S7. Overview of the interactive and searchable interactome database. Top panel: For a gene of interest, the highest scoring interaction partners can be found (score type: interaction score), or the genes with the most highly correlated genetic interaction profiles (score type: correlation coefficient). Bottom panel: The results of the query are summarized. All genes are annotated with their SGD descriptions (pop-up box) to facilitate easy browsing.

Supplementary Tables

Available online at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2856666/

References

Alcazar-Roman, A. R., and Wente, S. R. (2008). Inositol polyphosphates: a new frontier for regulating gene expression. Chromosoma *117*, 1-13.

Babiarz, J. E., Halley, J. E., and Rine, J. (2006). Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae. Genes Dev *20*, 700-710.

Bakal, C., Linding, R., Llense, F., Heffern, E., Martin-Blanco, E., Pawson, T., and Perrimon, N. (2008). Phosphorylation Networks Regulating JNK Activity in Diverse Genetic Backgrounds. Science *322*, 453-456.

Bertram, P. G., Choi, J. H., Carvalho, J., Chan, T.-F., Ai, W., and Zheng, X. F. S. (2002). Convergence of TOR-nitrogen and Snf1-glucose signaling pathways onto Gln3. Mol Cell Biol *22*, 1246-1252.

Bishop, A. C., Ubersax, J. A., Petsch, D. T., Matheos, D. P., Gray, N. S., Blethrow, J., Shimizu, E., Tsien, J. Z., Schultz, P. G., Rose, M. D.*, et al.* (2000). A chemical switch for inhibitor-sensitive alleles of any protein kinase. Nature *407*, 395-401.

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M. P., and Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Chen, R. E., and Thorner, J. (2007). Function and regulation in MAPK signaling pathways: Lessons learned from the yeast Saccharomyces cerevisiae. Biochim Biophys Acta, Mol Cell Res *1773*, 1311-1340.

Cherkasova, V. A., and Hinnebusch, A. G. (2003). Translational control by TOR and TAP42 through dephosphorylation of eIF2alpha kinase GCN2. Genes Dev *17*, 859-872.

Cheung, V., Chua, G., Batada, N., Landry, C. R., Michnick, S. W., Hughes, T. R., and Winston, F. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS Biol *6*, 2550-2562.

Cho, E.-J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev *15*, 3319-3329.

Collins, S. R., Kemmeren, P., Zhao, X.-C., Greenblatt, J. F., Spencer, F., Holstege, F. C. P., Weissman, J. S., and Krogan, N. J. (2007a). Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics *6*, 439-450.

Collins, S. R., Miller, K. M., Maas, N. L., Roguev, A., Fillingham, J., Chu, C. S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S. R., Schuldiner, M., Krogan, N. J., and Weissman, J. S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. GenomeBiology *7*, R63.

Cullen, P. J., and Sprague, G. F., Jr. (2002). The Glc7p-interacting protein Bud14p attenuates polarized growth, pheromone response, and filamentous growth in Saccharomyces cerevisiae. Eukaryotic Cell *1*, 884-894.

Dhillon, N., Oki, M., Szyjka, S. J., Aparicio, O. M., and Kamakaka, R. T. (2006). H2A.Z functions to regulate progression through the cell cycle. Mol Cell Biol *26*, 489-501.

Diella, F., Gould, C. M., Chica, C., Via, A., and Gibson, T. J. (2008). Phospho..ELM: a database of phosphorylation sites-update 2008. Nucleic Acids Res *36*, D240-D244.

Espinoza, F. H., Farrell, A., Nourse, J. L., Chamberlin, H. M., Gileadi, O., and Morgan, D. O. (1998). Cak1 is required for Kin28 phosphorylation and activation in vivo. Mol Cell Biol *18*, 6365-6373.

Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol *20*, 301-305.

Goossens, A., De la Fuente, N., Forment, J., Serrano, R., and Portillo, F. (2000). Regulation of yeast H+-ATPase by protein kinases belonging to a family dedicated to activation of plasma membrane transporters. Mol Cell Biol *20*, 7654-7661.

Green, K. D., and Pflum, M. K. H. (2007). Kinase-Catalyzed Biotinylation for Phosphoprotein Detection. J Am Chem Soc *129*, 10-11.

Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: Phosphorylation of RNA polymerase II mediates histone methylation. Cell *113*, 429-432.

Hohmann, S. (2002). Osmotic stress signaling and osmoadaptation in yeasts. Microbiol Mol Biol Rev *66*, 300-372.

Hsu, J.-Y., Sun, Z.-W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F.*, et al.* (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. Cell *102*, 279-291.

Johnson, S. A., and Hunter, T. (2005). Kinomics: methods for deciphering the kinome. Nat Methods *2*, 17-25.

Joshi, A. A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell *20*, 971-978.

Kannan, N., Taylor, S. S., Zhai, Y., Venter, J. C., and Manning, G. (2007). Structural and functional diversity of the microbial kinome. PLoS Biol *5*, 467-478.

Kaplan, C. D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096-1099.

Kelley, R., and Ideker, T. (2005). Systematic interpretation of genetic interactions using protein networks. Nat Biotechnol *23*, 561-566.

Keogh, M.-C., Kim, J.-A., Downey, M., Fillingham, J., Chowdhury, D., Harrison, J. C., Onishi, M., Datta, N., Galicia, S., Emili, A.*, et al.* (2006a). A phosphatase complex that dephosphorylates gamma H2AX regulates DNA damage checkpoint recovery. Nature *439*, 497-501.

Keogh, M.-C., Kurdistani, S. K., Morris, S. A., Ahn, S. H., Podolny, V., Collins, S. R., Schuldiner, M., Chin, K., Punna, T., Thompson, N. J.*, et al.* (2005). Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Keogh, M.-C., Mennella, T. A., Sawa, C., Berthelet, S., Krogan, N. J., Wolek, A., Podolny, V., Carpenter, L. R., Greenblatt, J. F., Baetz, K., and Buratowski, S. (2006b). The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes Dev *20*, 660-665.

Korber, P., and Horz, W. (2004). SWRred not shaken; mixing the histones. Cell *117*, 5-7.

Krogan, N. J., Baetz, K., Keogh, M.-C., Datta, N., Sawa, C., Kwok, T. C. Y., Thompson, N. J., Davey, M. G., Pootoolal, J., Hughes, T. R.*, et al.* (2004). Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. Proc Natl Acad Sci *101*, 13513-13518.

Lee, T.-Y., Huang, H.-D., Hung, J.-H., Huang, H.-Y., Yang, Y.-S., and Wang, T.-H. (2006). dbPTM: an information repository of protein post-translational modification. Nucleic Acids Res *34*, D622-D627.

Linding, R., Jensen, L. J., Ostheimer, G. J., van Vugt, M. A. T. M., Jorgensen, C., Miron, I. M., Diella, F., Colwill, K., Taylor, L., Elder, K.*, et al.* (2007). Systematic discovery of in vivo phosphorylation networks. Cell *129*, 1415-1426.

Luke, M. M., Della Seta, F., Di Como, C. J., Sugimoto, H., Kobayashi, R., and Arndt, K. T. (1996). The SAP, a new family of proteins, associate and function positively with the SIT4 phosphatase. Mol Cell Biol *16*, 2744-2755.

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., III, Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y.*, et al.* (2007). ATM and ATR Substrate Analysis Reveals Extensive Protein Networks Responsive to DNA Damage. Science *316*, 1160-1166.

Millar, C. B., Xu, F., Zhang, K., and Grunstein, M. (2006). Acetylation of H2AZ lys 14 is associated with genome-wide gene activity in yeast. Genes Dev *20*, 711-722.

Mulet, J. M., Leube, M. P., Kron, S. J., Rios, G., Fink, G. R., and Serrano, R. (1999). A novel mechanism of ion homeostasis and salt tolerance in yeast: the Hal4 and Hal5 protein kinases modulate the Trk1-Trk2 potassium transporter. Mol Cell Biol *19*, 3328-3337.

Mulugu, S., Bai, W., Fridy, P. C., Bastidas, R. J., Otto, J. C., Dollins, D. E., Haystead, T. A., Ribeiro, A. A., and York, J. D. (2007). A Conserved Family of Enzymes That Phosphorylate Inositol Hexakisphosphate. Science *316*, 106-109.

Mumberg, D., Mueller, R., and Funk, M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res *22*, 5767-5768.

Nourani, A., Robert, F., and Winston, F. (2006). Evidence that Spt2/Sin1, an HMG-like factor, plays roles in transcription elongation, chromatin structure, and genome stability in Saccharomyces cerevisiae. Mol Cell Biol *26*, 1496-1509.

O'Rourke, S. M., and Herskowitz, I. (2004). Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. Mol Biol Cell *15*, 532-542.

Oficjalska-Pham, D., Harismendy, O., Smagowicz, W. J., Gonzalez de Peredo, A., Boguta, M., Sentenac, A., and Lefebvre, O. (2006). General repression of RNA polymerase III transcription is triggered by protein phosphatase type 2A-mediated dephosphorylation of Maf1. Mol Cell *22*, 623-632.

Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635-648.

Ostapenko, D., and Solomon, M. J. (2005). Phosphorylation by Cak1 regulates the C-terminal domain kinase Ctk1 in Saccharomyces cerevisiae. Mol Cell Biol *25*, 3906-3913.

Pan, X., Yuan, D. S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J. S., Hieter, P., Spencer, F., and Boeke, J. D. (2004). A robust toolkit for functional profiling of the yeast genome. Mol Cell *16*, 487-496.

Pokholok, D. K., Zeitlinger, J., Hannett, N. M., Reynolds, D. B., and Young, R. A. (2006). Activated Signal Transduction Kinases Frequently Occupy Target Genes. Science *313*, 533-536.

Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R.*, et al.* (2005). Global analysis of protein phosphorylation in yeast. Nature *438*, 679-684.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S. R., Qu, H., Shales, M., Park, H.-O., Hayles, J.*, et al.* (2008a). Conservation and Rewiring of Functional Modules Revealed by an Epistasis Map in Fission Yeast. Science *322*, 405-410.

Rua, D., Tobe, B. T., and Kron, S. J. (2001). Cell cycle control of yeast filamentous growth. Curr Opin Microbiol *4*, 720-727.

Saiardi, A., Bhandari, R., Resnick, A. C., Snowman, A. M., and Snyder, S. H. (2004). Phosphorylation of proteins by inositol pyrophosphates. Science *306*, 2101-2105.

Schaber, M., Lindgren, A., Schindler, K., Bungard, D., Kaldis, P., and Winter, E. (2002). CAK1 promotes meiosis and spore formation in Saccharomyces cerevisiae in a CDC28-independent fashion. Mol Cell Biol *22*, 57-68.

Schuldiner, M., Collins, S. R., Thompson, N. J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J. F.*, et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Schuldiner, M., Collins, S. R., Weissman, J. S., and Krogan, N. (2006). Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344-352.

Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S. G., Cyert, M., Hughes, T. R.*, et al.* (2006). Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell *21*, 319-330.

St. Onge, R. P., Mani, R., Oh, J., Proctor, M., Fung, E., Davis, R. W., Nislow, C., Roth, F. P., and Giaever, G. (2006). Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. Nat Genet *39*, 199-206.

Svejstrup, J. Q. (2007). Elongator complex: how many roles does it play? Curr Opin Cell Biol *19*, 331-336.

Tate, J. J., Feller, A., Dubois, E., and Cooper, T. G. (2006). Saccharomyces cerevisiae Sit4 Phosphatase Is Active Irrespective of the Nitrogen Source Provided, and Gln3 Phosphorylation Levels Become Nitrogen Source-responsive in a sit4-deleted Strain. J Biol Chem *281*, 37980-37992.

Tong, A. H. Y., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C. W. V., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

van de Peppel, J., Kettelarij, N., van Bakel, H., Kockelkorn, T. T. J. P., van Leenen, D., and Holstege, F. C. P. (2005). Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. Mol Cell *19*, 511-522.

Wang, H., Wang, X., and Jiang, Y. (2003). Interaction with Tap42 is required for the essential function of Sit4 and type 2A phosphatases. Mol Biol Cell *14*, 4342-4351.

Wilmes, G. M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S. R., Whitworth, G. B., Kress, T. L., Weissman, J. S*., et al.* (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell *32*, 735-746.

Wu, C., Lytvyn, V., Thomas, D. Y., and Leberer, E. (1997). The phosphorylation site for Ste20p-like protein kinases is essential for the function of myosin-I in yeast. J Biol Chem *272*, 30623-30626.

Yao, S., and Prelich, G. (2002). Activation of the Bur1-Bur2 cyclin-dependent kinase complex by Cak1. Mol Cell Biol *22*, 6750-6758.

Youdell, M. L., Kizer, K. O., Kisseleva-Romanova, E., Fuchs, S. M., Duro, E., Strahl, B. D., and Mellor, J. (2008). Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. Mol Cell Biol *28*, 4915-4926.

Zanzoni, A., Ausiello, G., Via, A., Gherardini, P. F., and Helmer-Citterich, M. (2007). Phospho3D: a database of three-dimensional structures of protein phosphorylation sites. Nucleic Acids Res *35*, D229-D231.

Chapter 6

A Plasma-Membrane E-MAP Reveals Links of the Eisosome with Sphingolipid Metabolism and Endosomal Trafficking

A Plasma-Membrane E-MAP Reveals Links of the Eisosome with Sphingolipid Metabolism and Endosomal Trafficking

Pablo S. Aguilar[1,#], Florian Fröhlich[2,#], Michael Rehman[2,#], Mike Shales[3,#], Igor Ulitsky[4], Agustina Olivera-Couto[1,] Hannes Braberg[3], Ron Shamir[4], Peter Walter[5], Matthias Mann[6], Christer S. Ejsing[7], Nevan J. Krogan[3*], Tobias C. Walther[2*]

1 Institut Pasteur de Montevideo, Montevideo, Uruguay

2 Max Planck Institute of Biochemistry, Organelle Architecture and Dynamics, Martinsried, Germany

3 Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, USA

4 The Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel

5 Department of Biochemistry and Biophysics, University of California and Howard Hughes Medical Institute, San Francisco, CA, USA

6 Max Planck Institute of Biochemistry, Proteomics and Signal Transduction, Martinsried, Germany

7 University of Southern Denmark, Department of Biochemistry and Molecular Biology, Odense, Denmark

[#] These authors, listed in alphabetical order, contributed equally

[*] Correspondence to: krogan@cmp.ucsf.edu or twalther@biochem.mpg.de

Abstract

The plasma membrane delimits the cell and controls material and information exchange between it and the environment. How different plasma membrane processes are coordinated and how the relative abundance of plasma membrane lipids and proteins is homeostatically maintained is not yet understood. Here, we used a quantitative genetic interaction map, or E-MAP, to functionally interrogate a set of ~400 genes involved in various aspects of plasma membrane biology, including endocytosis, signaling, lipid metabolism and eisosome function. From this E-MAP, we derived a set of 57,799 individual interactions between genes functioning in these various processes. Using triplet genetic motif analysis, we identified a new component of the eisosome, Eis1 and linked the poorly characterized gene *EMP70* to endocytic and eisosome function. Finally, we implicated Rom2, a GDP/GTP exchange factor for Rho1 and Rho2, in the regulation of sphingolipid metabolism.

Introduction

The plasma membrane is the defining feature of the cell, separating its interior from the exterior space. It controls exchange and communication processes between the cell and its environment. Delivery of cellular material to the plasma membrane or cell exterior is mediated by exocytosis. Conversely, endocytosis is used to take up plasma membrane and external components. In addition, many signaling processes occur at the plasma membrane simultaneously and are often regulated by endocytosis of receptors or delivery of messenger molecules. To coordinate these processes and to maintain cell integrity under changing conditions, both plasma membrane protein and lipid composition are regulated and adjusted to external conditions. Despite impressive advances in our understanding of these individual processes, it is not well understood how they are coordinated.

To accommodate its many functions, the plasma membrane is highly organized, both spatially and temporally. In *Saccharomyces cerevisiae*, several plasma membrane domains of different composition are distinguishable by light microscopy. This organization is at least in part mediated by eisosomes, large protein complexes that underlie one of the domains, named MCC after the marker protein Can1 found there. When *PIL1,* encoding a major eisosome component, is deleted, cells have abnormal plasma membrane structure with large invaginations and loss of MCC protein organization (Grossmann et al., 2007; Walther et al., 2006). In addition, endocytosis of several plasma membrane proteins is either accelerated or delayed (Grossmann et al., 2008; Walther et al., 2006). The molecular function of eisosomes is still unknown, but recent data show that they interact with

sphingolipid-regulated Pkh-kinases, which phosphorylate their core components and are required for efficient endocytosis (deHart et al., 2002; Friant et al., 2001; Walther et al., 2007). In addition to Pkh-kinases, TORC2 is implicated in sphingolipid metabolism regulation (Aronova et al., 2008). However, it is unclear how these different signaling pathways are controlled, how they are coordinated and what their downstream effects are. Experimental evidence supports a model in which regulation of sphingolipid, sterol and glycerophospholipid levels in the plasma membrane are coordinated, but mechanistic insights as to how this is achieved are currently lacking (Guan et al., 2009; Tabuchi et al., 2006). To reveal functional links between the different processes, we generated a quantitative genetic interaction map targeting a large set of genes implicated in plasma membrane function.

Genetic interactions have long been used to dissect functional relationships between genes. Classically, researchers have looked for strong qualitative differences between observed phenotypes of double mutants and the phenotypes of the two related single mutants. More recently, we employed the E-MAP (epistatic mini-array profile) approach, a variation on synthetic genetic arrays (Tong et al., 2004). This allows for the quantitative analysis of genetic interactions, including negative (e.g. synthetic sick or lethal) as well as positive ones (e.g. suppression) (Schuldiner et al., 2005). For this approach, a comprehensive set of double mutants is generated and their growth is measured. To determine individual genetic interactions, deviations of growth rates from the medians of all combinations with one particular gene are calculated for each combination as a quantitative interaction

score (or *S-score*) (Collins et al., 2006; Schuldiner et al., 2006). Each mutation has a genetic interaction profile, or phenotypic signature, consisting of all its *S-scores* with all other genes in the E-MAP. A particularly useful parameter to judge similarities of profiles is to compare correlations of two genes' interactions with all other genes in the set. In addition, bioinformatic extraction based on mathematical models can be applied to yield functional modules in an unbiased fashion from E-MAP datasets, and correlations and *S-scores* can be used to reveal their connections (Bandyopadhyay et al., 2008; Ulitsky et al., 2008). The E-MAP approach has been previously used to functionally interrogate several processes. Dissection of genetic interactions from these E-MAPs has led to a deluge of biological insights in a variety of processes (Collins et al., 2007b; Fiedler et al., 2009; Schuldiner et al., 2005; Wilmes et al., 2008).

Here, we report an E-MAP targeting plasma membrane functions to generate novel biological insight relating to plasma membrane functions. Using this E-MAP, we have linked two new genes (*EMP70* and *EIS1*) to eisosome function and uncovered a link between Rom2- signaling and sphingolipid metabolism.

Results

Overview of the plasma membrane E-MAP

To address functional relationships between plasma membrane processes, we systematically determined the genetic interactions among a set of 374 genes involved in plasma membrane biology. We selected candidate genes encoding proteins functioning in membrane transport and organization, especially eisosomes,

actin patches, endocytosis and exocytosis. In addition, we picked genes involved in ergosterol and sphingolipid metabolism, as these lipids are implicated in many plasma membrane processes. Our selection criteria were based on available functional annotation (GO terms) and a literature survey. We also included a diverse set of genes whose products localize to the plasma membrane and/or genetically or physically interact with previously characterized plasma membrane genes/proteins. The selected genes were categorized into functional groups presented in Figure 1a and Supplementary Table 2. We included a number of genes analyzed in previous systematic genetic studies to facilitate comparison between data sets (Collins et al., 2007b; Fiedler et al., 2009; Schuldiner et al., 2005). From this set, we quantitated a total of 57,799 genetic interactions using the E-MAP approach (~83 % of the possible interactions).

Previously, we found that gene-pairs encoding physically interacting proteins are enriched for displaying positive genetic interactions and show a higher propensity for having highly correlated genetic interaction profiles (Collins et al., 2007b; Fiedler et al., 2009; Schuldiner et al., 2005). To assess the richness and quality of the genetic interaction data of the plasma membrane E-MAP, we compared the pair-wise correlation of genetic interaction profiles to a high quality set of protein-protein interactions (PPIs) (Collins et al., 2007a) and found that the power in the genetic map to predict PPIs is comparable to previously published E-MAPs (Supplementary Figure 1). Furthermore, comparison of interaction scores or correlation coefficients of gene pairs encoding physically interacting proteins (Collins et al., 2007a; Gavin et al., 2006; Krogan et al., 2006) (see Supplementary

Table 3) among all plasma membrane E-MAP gene pairs revealed that they have a higher likelihood to interact positively and to have correlated genetic interaction profiles (Figure 1b and 1c, yellow area under the green graph). Conversely, gene pairs with highly correlated interaction profiles and positive interactions are likely to physically interact.

To better visualize groups of interacting genes and their relationships, we used a previously developed algorithm that defines functional modules from quantitative genetic and protein-protein interaction (PPI) data (Ulitsky et al., 2008) (Supplementary Figure 2). This method identified 18 modules encompassing 53 genes (Supplementary Figure 2 and Supplementary Table 4). Genes in each module have similar genetic interaction profiles and form a connected sub-network in the PPI network. These modules corresponded to known protein complexes, such as the F-actin capping protein complex and the AP-3 adaptor or to known pathways, such as sphingolipid metabolism, the HOG osmosensory pathway and ergosterol biosynthesis (Supplementary Figure 2). To identify novel modules for which PPI data is not available, we performed the modular analysis without requiring PPI connectivity (Supplementary Figure 3). This identified 29 modules encompassing 190 genes (Supplementary Table 5 and http://acgt.cs.tau.ac.il/pmemap). This analysis yielded similar amounts of modules for the plasma membrane and the previously reported E-MAP on the early secretory pathway (Supplementary Table 6; [11]). Additional information can be extracted by considering interactions of single genes with modules (data not shown).

Insights from Hierarchical Clustering of the Genetic Interaction Data

Each mutant engenders a genetic interaction profile, or phenotypic signature, representing how it genetically interacts with all other mutants tested. Comparison of these profiles using hierarchical clustering (Figure 2a, Supp. Material and http://interactome-cmp.ucsf.edu/plasma_membrane/) is a powerful and unbiased approach to identify genes of the same pathway. In the following, we provide a brief summary of several functional connections revealed by such gene clustering.

RVS161 and *RVS167* encode proteins that operate together in membrane remodeling during endocytosis (Ren et al., 2006). As expected from their overlapping functions, *rvs161*Δ and *rvs167*Δ clustered together with high correlation (correlation = 0.54; Figure 2a, insert 2). Consistent with previous reports, both share positive genetic interactions with a number of genes involved in fatty acid elongation for sphingolipid synthesis, such as *FEN1* and *SUR4* (Figure 2a, insert 2d, (Revardel et al., 1995)). Interestingly, we observed positive interactions with genes encoding components of the Hog1 MAP-kinase cascade and the ergosterol biosynthesis pathway (*erg3*Δ , *erg5*Δ , *erg6*Δ, Figure 2a, insert 2). In additions to changes in their sterols, these *erg*-mutants have altered sphingolipid composition (Guan et al., 2009). Thus, defects resulting from deletion of *RVS*- genes could be compensated by *erg*-mutants by changes of sphingolipids. Also in line with previous work, both *rvs161*Δ and *rvs167*Δ show strong negative interactions with actin cytoskeleton genes, such as *BBC1*, *JSN1* and *BZZ1,* (Figure 2a, insert 2a;(Breton and Aigle, 1998; Brizzio et al., 1998; Friesen et al., 2006; Tong et al., 2004)). In addition, we found several previously unrecognized relationships, including negative

interactions between the *RVS*-genes and *ire1*Δ and *hac1*Δ, two mediators of the unfolded protein response (UPR) control system for endoplasmic reticulum function. Possibly, cells react to Rvs-deficiency by altering lipid synthesis or transport, which in turn activates the UPR. Cells lacking the UPR in addition to the Rvs-proteins could have decreased fitness. Consistent with this notion, a recent genome-wide study found the UPR activated in *rvs*Δ cells (Jonikas et al., 2009).

We also detected many genetic interactions and highly correlated profiles between genes encoding actin patch components. For example, *sla1*Δ and *ede1*Δ that function in endocytosis are highly correlated (correlation = 0.64, Figure 2a, insert 1) and display a negative genetic interaction (interaction score = -7.7). Surprisingly, given its function in exocytosis rather than endocytosis, we also found *chs6*Δ highly correlated with *sla1*Δ and *ede1*Δ (correlations *ede1*Δ -*chs6*Δ =0.53; *sla1*Δ -*chs6*Δ = 0.43, Figure 2a, insert 1). Furthermore, these three genes all result in strong negative genetic interactions when any two of them are combined. Collectively, this indicates that Chs6 might function in coordinating exo- and endocytosis, perhaps by delivering a subset of cargos to the plasma membrane (Valdivia et al., 2002). In this scenario, *chs6*Δ would lead to depletion of an endocytic factor from the plasma membrane and as a consequence decrease in endocytosis efficiency. Combination with mutants defective in this process would further decrease the fitness of the resulting strains.

We also observed many strong genetic interactions between trafficking complexes. Genes encoding the retromer complex (*VPS17, VPS29, VPS35, PEP8)*, the COG complex (*COG5, COG6, COG7, COG8)* or the AP3 complex (*APM3, APL5*) all

formed highly correlated clusters in the plasma membrane E-MAP (Figure 2a, insert 4). In addition, potential new connections between these complexes and so far poorly characterized components of the endocytic machinery are apparent in these clusters. As an example, the retromer complex co-clusters with deletion of *MON2* (correlation = 0.48), a gene encoding an evolutionary conserved scaffolding protein functioning in endosome to Golgi trafficking (Efe et al., 2005). Our data suggest that Mon2 acts together with the retromer in this process.

Many genes encoding members of signaling cascades showed strong genetic relationships. For example, two kinases of the cell integrity MAP kinase signaling module, Slt2 (the MAP kinase) and Bck1 (the MAPKKK) (Levin, 2005), showed one of the highest correlations (0.75). Similarly, genes encoding components of retrograde signaling (*RTG1, RTG2, RTG3* and *MKS1*) all cluster together (correlation = 0.44) indicating that all pairs have high correlation coefficients (e.g. *MKS1*/*RTG1* CC = 0.59; Figure 2b).

Functional links involving eisosomes

Although the eisosome has been linked to endocytosis regulation, details regarding its biological roles remain unresolved. To understand eisosome function *in vivo*, we genetically analyzed its core components *PIL1* and *LSP1*. Since the encoded proteins are more than 70% identical and are stoichiometric components of the eisosome, we expected very similar genetic profiles for them. Surprisingly however, *PIL1* and *LSP1* showed very different genetic interactions and, accordingly, cluster in different regions of the E-MAP (correlation = 0.038, Figure 2a, arrows).

This parallels the cell biological observation that deletion of *PIL1,* but not of *LSP1,* results in strong effects on plasma membrane organization and protein turnover.

To gain further insight into eisosome function, we analyzed the triplet genetic motifs (TGMs) that *pil1*Δ is participating in (Fiedler et al., 2009). TGMs are the simplest motifs, apart from binary interactions, and can exist in four forms: (Type I), all three genes displaying positive genetic interactions; (Type II), two positive and one negative; (Type III), two negative and one positive; and (Type IV), three negative interactions (Figure 3a). We have previously shown that genes displaying all positive genetic interactions (Type I TGM) are enriched for functioning in the same pathway (Fiedler et al., 2009). We therefore assembled a complete map of Type I TGMs found in the plasma membrane E-MAP (Supplementary Figure 4). Since Pil1 plays a more prominent role than Lsp1 in eisosome and plasma membrane function, we extracted out all Type I TGMs involving *pil1*Δ (Figure 3b). In this representation, we highlighted genes that are important for eisosome localization or are closely related to such genes (*YMR031c* and *EMP70*, respectively (Frohlich et al., 2009); green nodes in Figure 3b) and characterized them further.

EIS1/YMR031c encodes a novel eisosome component

Since *ymr031c*Δ and *pil1*Δ have a positive genetic interaction and a correlated interaction profile (Figure 3b), we tested whether the corresponding proteins physically associate. To this end, we fused the sequence encoding the green fluorescent protein (GFP) tag to *PIL1* at its endogenous location in the yeast genome

and immune-purified the expressed Pil1-GFP from a yeast culture that was metabolically labeled with heavy, non-radioactive lysine (SILAC) (Ong et al., 2002). In parallel, we performed a mock purification from control, light labeled wild type cells. We identified 533 proteins present over a 10000-fold dynamic range in the mixed eluates from both purifications. As expected, we found Pil1 and Lsp1 as well as the recently identified eisosomes binding protein Mrp8 as significant outliers with a high ratio of labeled to non-labeled protein, indicating that they are specific interactors ($P < 0.0001$; Figure 4a (Walther et al., 2006; Wang et al., 2009). In addition, we found a number of other specific interactors, including Ymr031c, which is consistent with a recent report (Deng et al., 2009). To independently confirm this observation, we performed immunoprecipitations of TAP-tagged Ymr031c and, as a control, Lsp1 and found that both specifically precipitated Pil1 (Figure 4b). To test whether Ymr031c colocalizes with Pil1, we fluorescently tagged both proteins. The signal from Pil1 and Ymr031c perfectly overlapped at eisosomes (Figure 4c, upper panel; Pearson correlation = 0.81 ± 0.06). Consistent with these data, Ymr031c was recently detected at MCCs (Grossmann et al., 2008). One prediction for a genuine eisosome component is that it re-localizes to eisosome remnants in a *PIL1* deletion strain (Walther et al., 2006). We therefore investigated Ymr031c-GFP localization in *pil1*Δ cells and found that both Ymr031c and the eisosome component Lsp1 localized to one or a few eisosome remnants in the cell periphery (Figure 4d). To investigate whether *YMR031c* plays a role in eisosome architecture or assembly, we deleted it and analyzed localization of eisosome core components in the resulting strain. For both Pil1 and Lsp1-GFP, we observed substantially increased cytosolic

fluorescence in *ymr031c*Δ cells (Figure 4e and 4f). Collectively, these data demonstrate that Ymr031c is physically associated with eisosomes and required for their normal formation. We have therefore named this gene *EIS1*.

EMP70 is an early endosomal and vacuolar protein

In the genetic network of the plasma membrane E-MAP, *EMP70* is the strongest candidate for a functional relationship with *PIL1* since: i) the two genes have highly correlated genetic profiles (Correlation of *PIL1* and *EMP70* = 0.37, *EMP70* has the most similar profile to *PIL1* of all the EMAP genes*;* Figure 5a); ii) the two genes participate in two Type I TGMs (Figure 3b); and iii) the Emp70 homologue Tmn2 is required for normal Pil1-GFP localization (Frohlich et al., 2009). In addition, our modular analysis identified *EMP70* and *PIL1* as part of the same six gene module (Supplementary Figure 3; *S-score:* between *PIL1* and *EMP70* = 1.78; Supplementary Table 5).

These genetic links prompted us to investigate *EMP70* in more detail. We fluorescently tagged Emp70 with GFP and found it localizes in a complex pattern, consisting of a central ring reminiscent of vacuoles and several bright foci in the cytoplasm that often seem connected to the vacuole (Figure 5b and Supplementary Movie 1). Emp70 was previously found in an endosomal membrane fraction (Schimmoller et al., 1998). We therefore tested whether cytosolic Emp70 foci represent endosomes. We used a number of endosomal markers and found Emp70-GFP foci to colocalize with Kex2 marking the early endosome, which in yeast is functionally continuous with the trans-Golgi network. In contrast, Emp70

localization did not overlap with the late endosomal/ pre-vacuolar marker Vps5 (Figure 5b and Supplementary Figure 5a).

To test whether the Emp70-labeled compartments are part of the endocytic route, we used the endocytosis tracer FM4-64. This lipid dye is incorporated in the plasma membrane, taken-up by endocytosis and trafficked through the endosomal system to the vacuole (Vida and Emr, 1995). We found in pulse chase experiments that early FM4-64 intermediates co-localize with Emp70 foci (Figure 5c, "0 min"). As the dye migrated through the endocytic system, it also colocalized with a subset of Emp70 positive foci towards the end of the reaction, but markedly less at intermediate time points (Figure 5c, "30 min"). At the final time point, FM4-64 clearly labeled the vacuole delimiting membrane where it colocalized with the Emp70-GFP ring staining. Trafficking from early endosomes can be blocked by incubation of cells at 16˚C, which leads to accumulation of FM4-64 (Zheng et al., 1998). Emp70-GFP almost perfectly colocalized with FM4-64 when the latter was accumulated in such a "16˚C compartment", further arguing that Emp70 localizes to early endosomes (Supplementary Figure 5b). Strains harboring a deleted or C-terminally tagged *SNF7* (an ESCRT–III gene) show a "Class E" vacuolar protein sorting defect. This is characterized by collapse of endosomes to one or a few large "Class E" compartments (Babst et al., 1998; Teis et al., 2008). Under these conditions, Emp70-GFP formed fewer, very large clusters that colocalize with Snf7-RFPmars marked "Class E" compartments and displayed reduced vacuolar membrane staining (Figure 5d). From these data, we conclude that Emp70 localizes to early endosomes and the vacuole. To better characterize the localization of

Emp70 in these two pools, we quantitated the relative amount of Emp70 colocalizing with markers for each organelle and found 48% of Emp70 to localize in the TGN/endosomal compartment and 41% at the vacuolar membrane (Figure 5g).

During our localization studies, we often observed early endosomal foci marked by FM4-64 dynamically associating with the plasma membrane. To test whether the genetic link of *EMP70* with *PIL1* was reflected in recruitment of Emp70-GFP foci to eisosomes, we investigated the Emp70-GFP localization in respect to fluorescently tagged eisosomes. Strikingly, we found many spots of Emp70-GFP at eisosomes (Figure 5e). In any given cell, 4% of the total Emp70-GFP signal colocalized with an eisosome marker (Figure 5g and Movie 2). To test whether this association has functional relevance, we investigated the Emp70 localization in *pil1D* cells and found a marked reduction of foci close to the plasma membrane (Figure 5f).

The EMP70 family of proteins is required for normal endosomal sorting

To test whether Emp70 is important for early endosome to vacuole trafficking, we analyzed Kex2-GFP localization in an *emp70D* strain and found a dramatic Kex2 re-localization from early endosomes to the vacuole (Figure 6a). Kex2 steady-state localization depends on signals that send it to early endosomes, which subsequently mature into late endosomes, from where it is actively retrieved (Brickner and Fuller, 1997; Sipos et al., 2004). It is possible that vacuolar mis-localization of Kex2 in *emp70*Δ cells results from a defect in retrieval from the late endosome, or a complex trafficking problem affecting early endosome function. If

retrieval is compromised, normally vacuolar sorted carboxypeptidase Y (CPY) is secreted. We tested this and found that in contrast to the control *vps1*Δ, *emp70*Δ alone does not lead to CPY secretion (Robinson et al., 1988; Rothman et al., 1989). *EMP70* has two homologues in the genome, *TMN2* and *TMN3*. To address whether they could compensate for Emp70 function in its absence, we tested CPY secretion in strains with different combinations of the family members deleted. *TMN2* deletion alone has no, and *TMN3* deletion alone only a weak effect on CPY sorting (Figure 6b). In contrast, combining *emp70*Δ with either *tmn2*Δ or *tmn2*Δ*tmn3*Δ results in CPY secretion, showing Emp70 is functionally redundant with Tmn2 in vacuolar protein sorting and that the Emp70 protein family is required for normal endosomal function.

Sphingolipid metabolism and its regulation

The plasma membrane E-MAP interrogates relationships within metabolic networks important for plasma membrane function, including sphingolipid metabolism (Figure 2b; insert 3a and 3b and Figure 7a). Consistent with their common function, many of the sphingolipid pathway genes showed high correlation (> 0.2). Figure 7b shows the distance of the action of enzymes in the pathway plotted against the correlation coefficient of the corresponding genes. The linear best fit on all data points revealed that genes encoding enzymes catalyzing subsequent steps are more highly correlated than genes further away in the metabolic network. Moreover, whereas most mutations in genes encoding enzymes catalyzing early steps of sphingolipid synthesis have negative genetic interactions

amongst each other (Figure 2b; insert 3b and Figure 7a), they show positive genetic interactions when combined with mutations in genes acting late in complex sphingolipid formation (Figure 2b insert 3a and Figure 7a). This might indicate that deficiency in late acting enzymes leads to build-up of toxic intermediates, which can be suppressed by deleting genes encoding upstream acting enzymes. Precedence for this includes inhibition of Aur1, which converts ceramide to inositolphosphoceramide by aureobasidin A, leading to complex sphingolipids depletion and concomitant accumulation of ceramide, which both contribute to toxicity (Cerantola et al., 2009).

The plasma membrane E-MAP also revealed that *ROM2*, encoding a Rho1 GTPase exchange factor, has strong genetic connections to sphingolipid synthesis genes. For example, *ROM2* has correlated genetic profiles with *FEN1*, *SUR2*, *LCB3* and *SUR4,* all acting early, but shows negative correlation with *CSG2* acting late in sphingolipid synthesis (Figure 8a). In addition, a *ROM2* deletion mutation results in a strong synthetic sick phenotype with *lcb3*Δ*, sur2*Δ *dpl1*Δ and *ysr3*Δ*,* all genes encoding enzymes catalyzing different steps of sphingolipid metabolism (Interaction score < -2, see Figure 2b; insert 3b and data not shown). Together, this suggests that Rom2 is an activator of sphingolipid metabolism. To test this model, we profiled the lipidome of *rom2*Δ and several other mutants in the sphingolipid pathway by "shotgun" lipidomics (Ejsing et al., 2009). *ROM2* deletion resulted in a lipid phenotype similar to that of *sur2*Δ or *sur4*Δ cells (Figure 8b and Supplementary Table 7). Particularly, *ROM2* deletion led primarily to accumulation of long chain bases and a small decrease of ceramides. This argues that Rom2 activates

sphingolipid synthesis by regulating the conversion of long chain bases to ceramides.

Discussion

The plasma membrane E-MAP quantitatively describes interactions between genes involved in plasma membrane processes. Together with previous studies, it shows that the E-MAP technology can be used to detect protein interactions and signaling pathways as well as uncover complex biological connections. Here, we highlighted several examples of novel insights into plasma membrane function derived from the E-MAP, focusing on its spatial organization and homeostasis. As an example of a physical interaction revealed from the E-MAP data, we investigated Eis1/Ymr031c and defined it as an eisosome component. Based on its much lower abundance compared to the eisosome core components, it might play a special architectural or regulatory role there. This is also a case where we combined data from the plasma membrane E-MAP with our visual screen for genes affecting Pil1-GFP localization(Frohlich et al., 2009), which provides an example how the combination of different high-throughput datasets helps to uncover previously unrecognized relationships.

Mining of the plasma membrane E-MAP also yielded information on more functional interactions not reflected in physical associations. The transmembrane protein Emp70 has a fascinatingly complex localization and genetically interacts with eisosome components. Particularly intriguing is the Emp70 pool localized in endosomal structures that often appear connected with the vacuolar membrane.

This observation raises the interesting possibility that endosomes reach out to the plasma membrane and pick up their cargo. It also suggests that at least parts of the endosomal membrane system might be a tubular network connected to the vacuole, but further detailed cell biological studies will have to clarify this point.

We also used the plasma membrane E-MAP to interrogate metabolic networks and their regulation. The strong correlation profiles of sphingolipid synthesis genes argues that novel functionally related genes could be found by using the genetic profiles from the plasma membrane E-MAP. For example, genes that function in sphingolipid metabolism or are involved in its regulation would be expected to cluster with known sphingolipid synthesis genes. Using this logic, we identified Rom2 as a regulator of sphingolipid metabolism. Mechanistically, its activator function could occur either through ceramide synthesis activation by Rom2 or by negative regulation of ceramidase. Between these two hypotheses, we consider the first one more likely since *rom2*Δ clusters with genes encoding ceramide synthase (*lag1*Δ *lac1*Δ), but not *ydc1*Δ, encoding ceramidase (Figure 2b, insert 3b). This is consistent with previous findings that connect the Tor2 kinase pathway with Rho1-signalling via Rom2 and recent findings that Tor kinase complex 2 (TORC2) is required for ceramide biosynthesis (Aronova et al., 2008; Schmidt et al., 1997). This previous study implicated an alternative branch of TORC2 signaling through the Ypk2 kinase in regulation of ceramide biosynthesis, but did not rule out involvement of Rom2. The effect of *ROM2* deletion could either be directly on ceramide synthase or alternatively could block synthesis of its substrate, long chain fatty acid-CoA. In the latter model, depletion of long chain fatty acids would slow

ceramide synthesis and therefore lead to the accumulation of long chain bases, the second substrate of ceramide-synthase. In either scenario, Rom2 has a stimulatory function in sphingolipid synthesis at the step converting long chain bases to ceramides. Consistent with this notion, inhibition of sphingolipid synthesis by the antifungal drug myriocin leads to a re-localization of Rom2 from the plasma membrane (Kobayashi et al., 2005). Rom2 is recruited to the plasma membrane through binding of phosphoinositol-(4,5)-bisphosphate ($PI(4,5)P_2$) by its pleckstrin homolog (PH) domain and also reduction of $PI(4,5)P_2$ relocalizes Rom2 (Audhya and Emr, 2002). This raises the interesting possibility that Rom2 serves to connect phosphoinositide and sphingolipid signaling pathways. The details of this intriguing regulation of sphingolipid metabolism remain to be worked out, but it demonstrates how genetic interactions in the plasma membrane E-MAP yield novel insights on metabolic networks and their regulation.

We anticipate this dataset will fuel many more mechanistic studies. Particularly integration with other data from lipidomics, interaction proteomics or systematic visuals screens are likely to reveal novel insights into the regulation of plasma membrane processes. In addition, many antifungal drugs target functions connected to the plasma membrane, such as cell wall and ergosterol synthesis. Probing the set of genes on the E-MAP presented here with a battery of drugs and comparison of the resulting drug profiles to the mutant profiles is an effective way to identify putative drug targets. This would facilitate the identification of compounds impinging on these various processes and could potentially have therapeutic value.

Acknowledgements:

We thank members of the Walther and Krogan lab for critical reading and comments and Jason Brickner (Northwestern University, Chicago) for suggestions. Ulrike Laabs provided excellent technical assistance. Thanks to Ole Nørregaard Jensen (University of Southern Denmark) for providing access to the Nanomate Triversa used for the lipidomic experiments and to Patrick Kemmeren for establishing the database. This work was supported by the Max Planck Society (T.C.W.), the German Research Foundation (DFG; T.C.W.), the German-Israeli Foundation (T.C.W), NIH, the Searle, Sandler and Keck Foundations (N.J.K), the International Human Frontier Science Program (HFSP; T.C.W.), the Israel Science Foundation (grant no. 802/08; R.S.), the Edmond J. Safra Bioinformatics program at Tel Aviv University (I.U.), the Programa de Apoyo Sectorial a la Estrategia Nacional de Innovación (INNOVA URUGUAY, DCI-ALA/2007/19.040, P.S.A.) and the Agencia Nacional de Investigación e Innovación (ANII, A.O-C.).

Figures

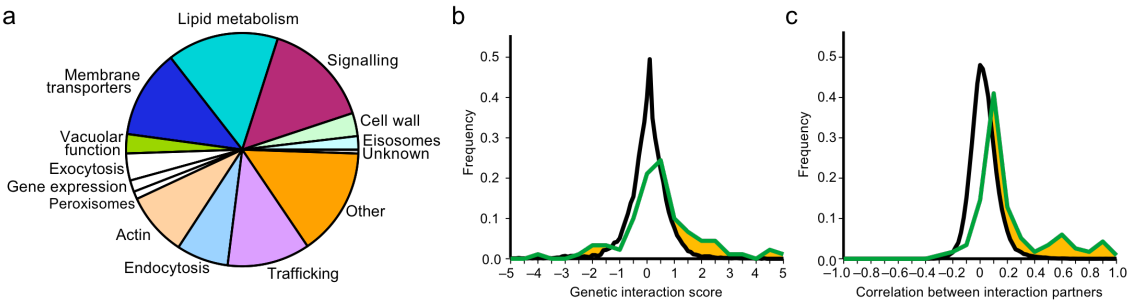


Figure 1: Composition of the plasma membrane E-MAP. **(a)** Genes selected for the PM E-MAP are classified according to their biological function. **(b, c)** Genes encoding proteins interacting with each other are more likely to show positive genetic interactions (b) and correlated genetic interaction profiles (c). Interaction and correlation scores of gene pairs known to encode interacting proteins are shown in green, the remainder of gene pairs black.

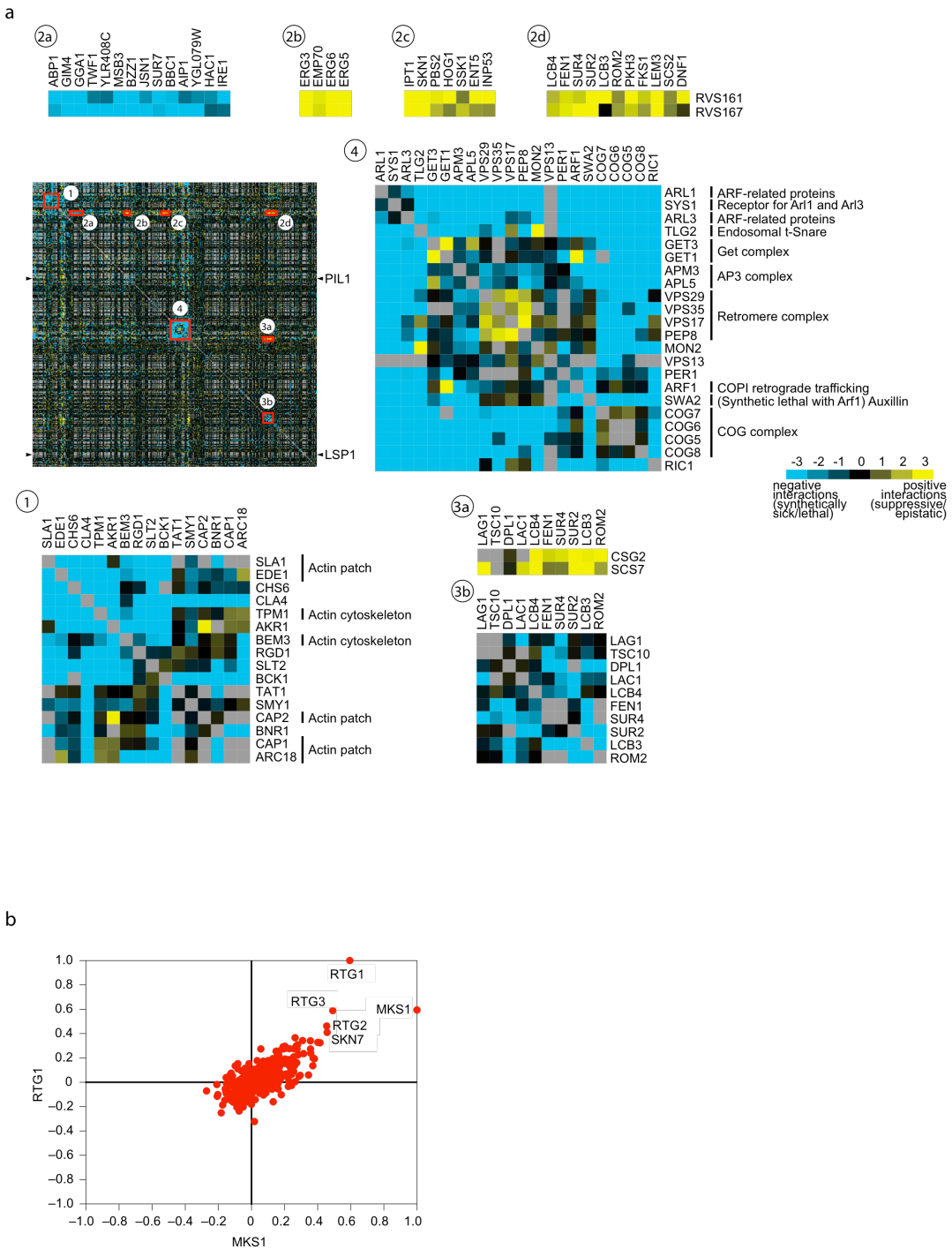


Figure 2: Overview of the clustergram of the plasma membrane E-MAP. **(a)** Selected areas are marked in the overview and highlighted as inserts 1-4. Positive genetic interactions are shown in yellow, negative ones in blue. **(b)** Genes with correlating

genetic profiles are shared between RTG1 and MKS1. Pairwise correlations between RTG1, and MKS1 and all other genes in the plasma membrane E-MAP were calculated and plotted against each other.

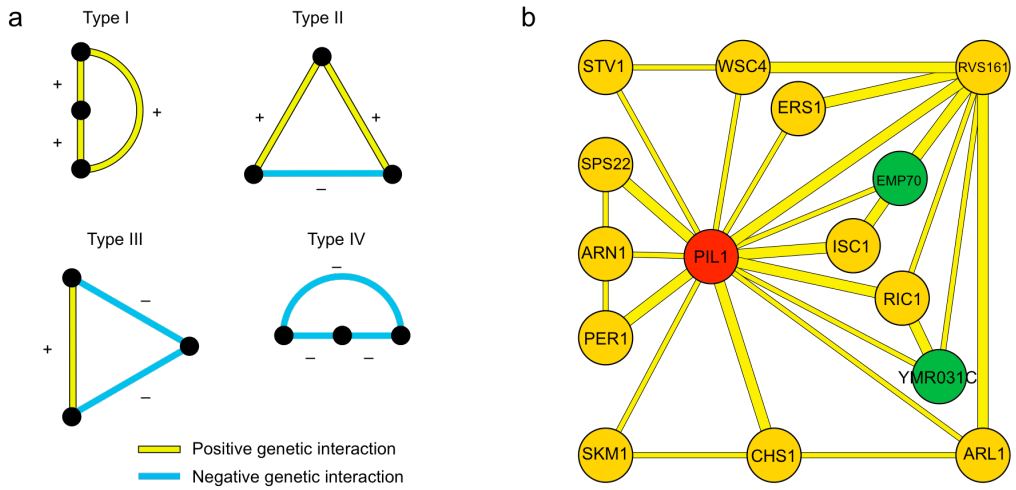


Figure 3: Triplet Genetic Motives (TGMs) of the plasma membrane E-MAP. **(a)** All four potential TGMs are shown. Nodes in vertical order represent involvement in the same pathway, horizontal orientation indicates possible parallel pathways. **(b)** Type I TGMs that have *PIL1* as a node. Nodes in green represent a gene important for Pil1-GFP localization (*YMR031C)* or a homologue of such a gene (*EMP70*) (Frohlich et al., 2009).

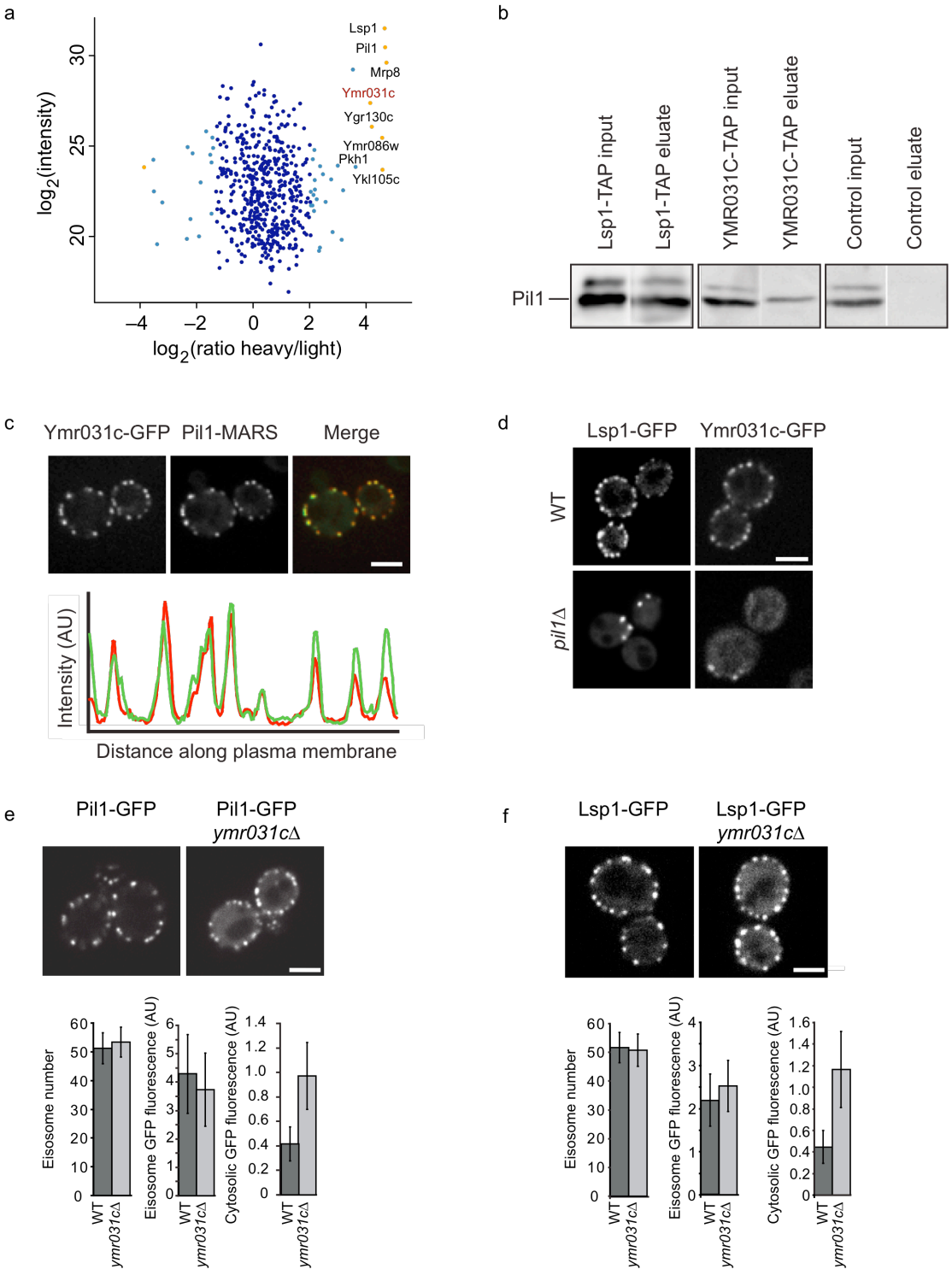


Figure 4: *YMR031C/EIS1* encodes an eisosome component. **(a)** Affinity purification and mass spectrometric analysis of heavy labeled cells expressing GFP-tagged Pil1

and untagged control cells. Averaged peptide intensities are plotted against Heavy/Light SILAC ratios. Significant outliers (P< 0.0001) are colored in orange or light blue (P< 0.05), other identified proteins are shown in dark blue **(b)** Pull-down purification from cells expressing tandem affinity tagged Lsp1, Ymr031c or untagged control cells. Inputs and eluates form the pull-down were blotted and probed with antibodies against Pil1. **(c)** Colocalization of GFP-tagged Ymr031C with RFPmars-tagged Pil1. Representative confocal mid-sections are shown. The graph shows the intensity profiles for both channels along the perimeter of the cell. **(d)** *PIL1* is required for normal localization of Ymr031c. Ymr031c-GFP or Lsp1-GFP was expressed and imaged either in wt or *pil1D* cells. Representative confocal mid sections are shown. **(e)** Ymr031c is required for normal eisosomes formation. Pil1-GFP (left panels) or Lsp1-GFP (right panels) was expressed in *ymr031cD* or control cells. Representative mid sections are shown. For each experiment, the number of eisosomes per cell, the GFP fluorescence per eisosomes and the cytosolic GFP fluorescence were quantified from at least 100 cells and is shown below the images. Bar = 2.5 mm.

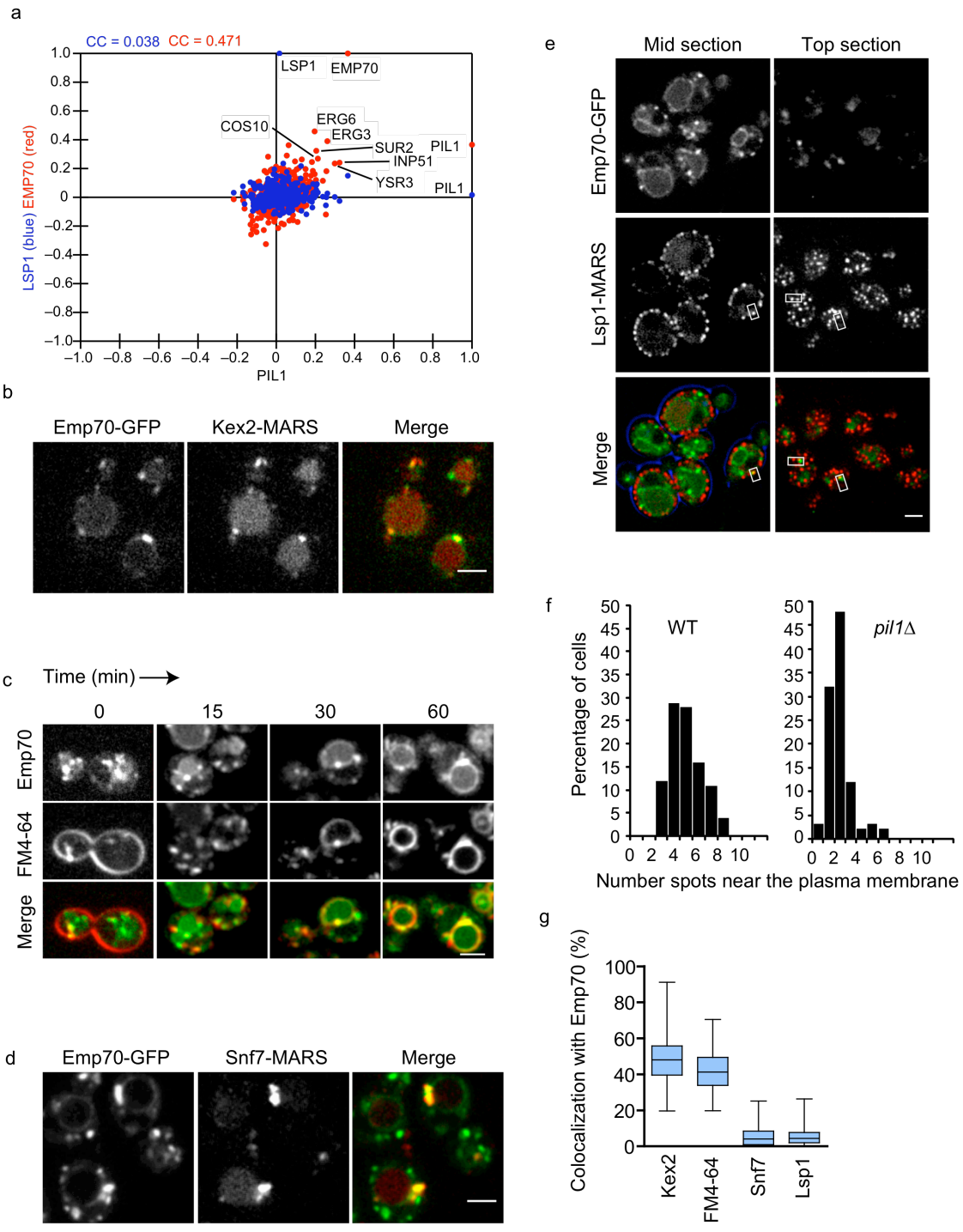


Figure 5: The eisosome-linked Emp70 is an early endosomal protein. **(a)** Genes with correlating genetic profiles are shared between *PIL1* and *EMP70*, but not *PIL1* and *LSP1*. Correlation coefficients between the genetic profile of *PIL1* and each of

the other 373 profiles in the E-MAP are plotted on the X-axis against on the Y-axis either the similar set of values for the *LSP1* profile with all other profiles (blue), or those for *EMP70* with all other profiles (red). Labeled points indicate some genes with profiles that are positively correlated with both the profile of *PIL1* and that of *EMP70*. CC values in blue and red indicate the correlation coefficients for the full set of blue or red points plotted. **(b)** Emp70 colocalizes with Kex2. Emp70-GFP and Kex2-RFPmars were co-expressed and imaged. Representative confocal midsections are shown. **(c)** Emp70 localizes to a FM4-64 marked endocytic compartment. Cells expressing Emp70-GFP (green) were pulse-labeled with FM4-64 (red) and imaged for 1h. Images of mid sections of cells at selected time are shown as indicated. **(d)** Emp70 localizes to the "Class E" compartment in *SNF7* mutants. GFP tagged Emp70 was expressed in cells harboring non-functional Snf7-RFPmars, resulting in the clustering of endosomal proteins in the class E compartment. Representative confocal mid sections are shown. **(e)** Emp70-GFP foci localize to the cell periphery. Emp70-GFP (green) was expressed in cells harboring the fluorescent eisosomes marker Lsp1-MARS. Representative mid (left panel) and top sections (right panel) are shown. Boxes highlight selected areas of colocalization. **(f)** *PIL1* is required for normal Emp70 localization to the cell periphery. Emp70-GFP was expressed in cells expressing the plasma membrane marker Ylr413w-RFPmars and foci overlaying this marker were counted in more than 100 wt and *pil1D* cells. Results are shown as a histogram of number of spots opposed to the plasma membrane in each cell. **(g)** Quantitation of the organelle distribution of Emp70. Emp70-GFP was imaged in live cells and analyzed for colocalization with Kex2-RFP Mars (n = 100), vacuolar FM4-

64 (n = 91), Snf7-RFP Mars (n = 93, diploid strain expressing one tagged Snf7 allele) and Lsp1-Cherry (n=107). The relative area of overlap between signals was quantified as a percentage of total area occupied by Emp70 signal. Box plots representing maxima, 75th percentile, median, 25th percentile and minima are shown for the colocalization with each marker.

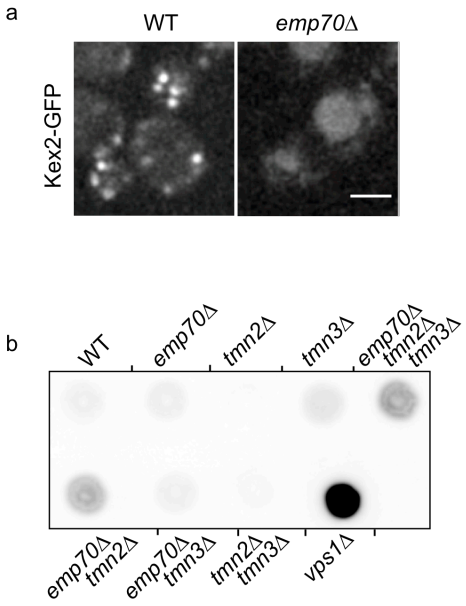


Figure 6: Emp70 is required for normal endosome function. **(a)** *EMP70* is required for normal localization of Kex2-GFP. Kex2-GFP was expressed either in WT or *emp70D* cells and representative confocal mid sections are shown. **(b)** Emp70 family members are required for late endosomal protein retrieval. Mutants of *EMP70, TMN2 or TMN3* were tested alone or in combination for CPY secretion. A representative colony blot is shown. Bars = 2.5 mm.

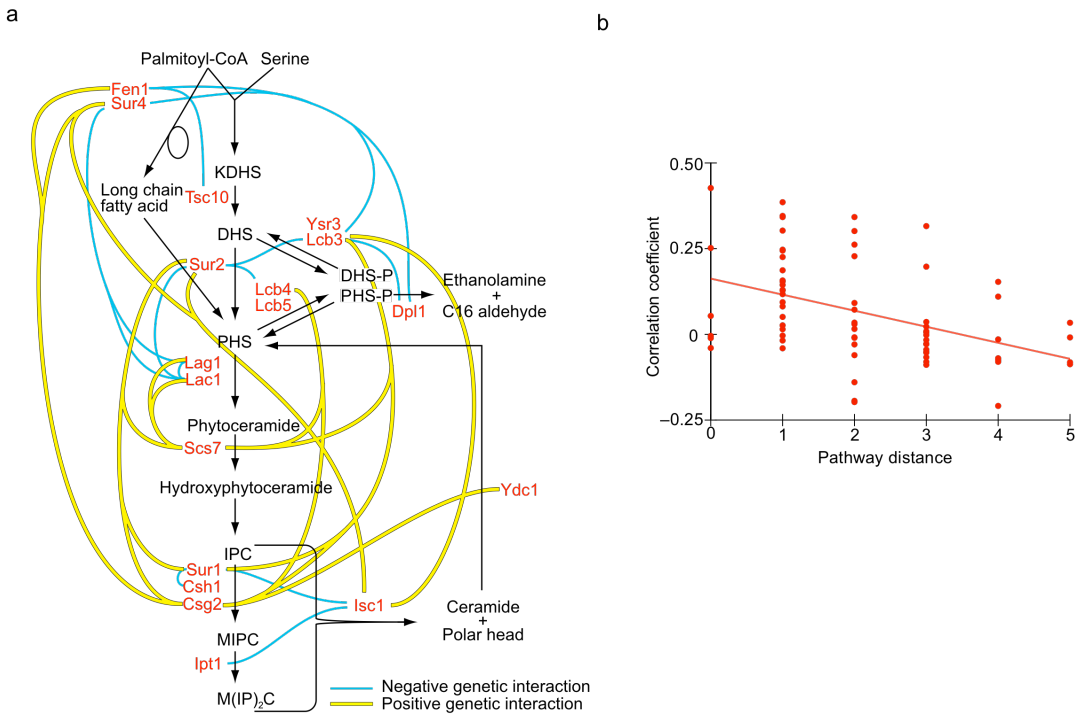


Figure 7: Genetic interactions of sphingolipid metabolism. **(a)** Graphic representation of the sphingolipid synthesis pathway with negative genetic interactions shown in blue, positive interactions shown in yellow. **(b)** Genes encoding enzymes acting in succession in sphingolipid synthesis show higher correlation than genes further apart in the metabolic network. For each gene pair in sphingolipid synthesis, the pathway distance of genes, that is the number of metabolic intermediates between the catalyzed reactions is plotted against the correlation coefficient of the gene pairs. The red line is a best fit linear regression line fitted for all the data points on the graph.

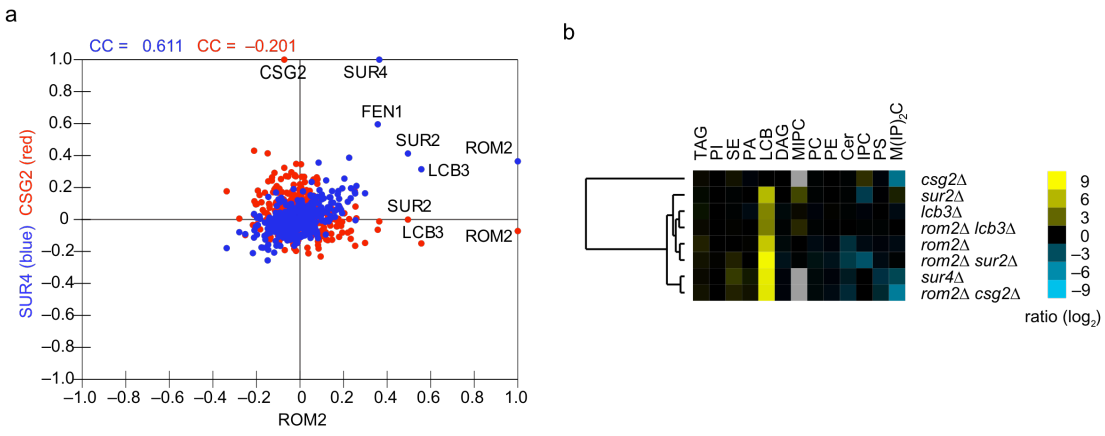
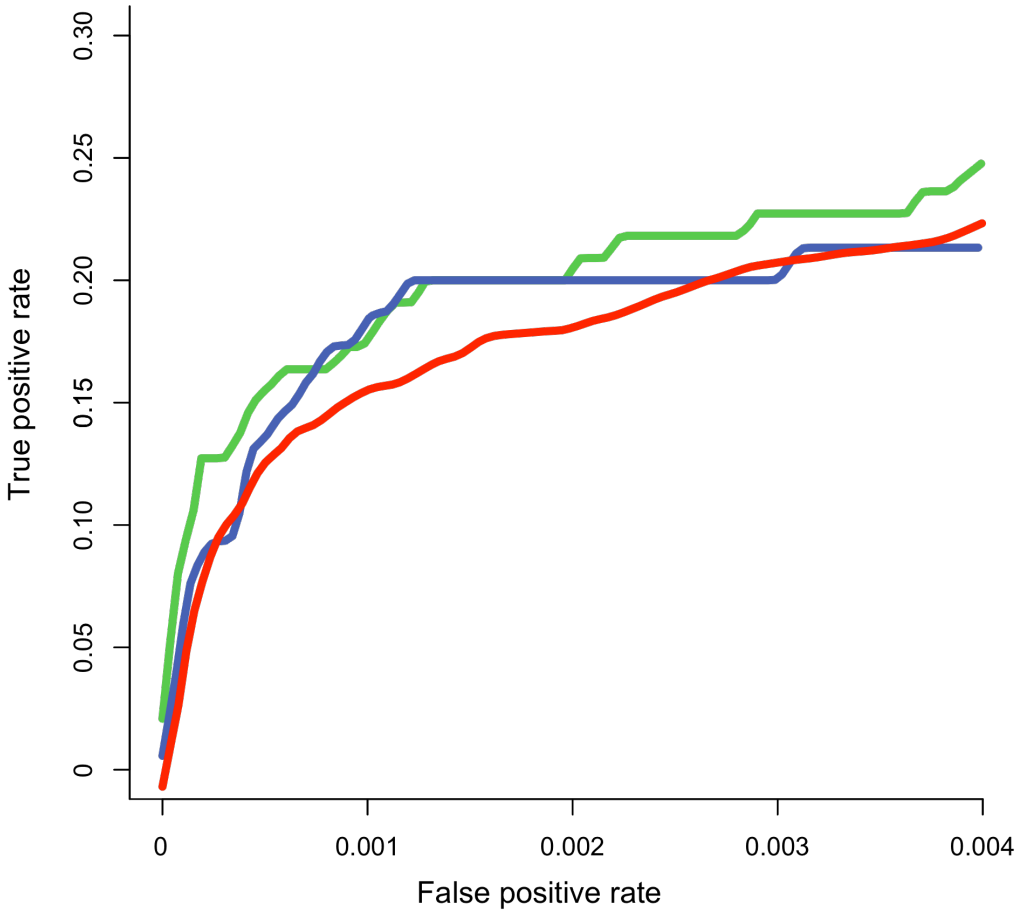
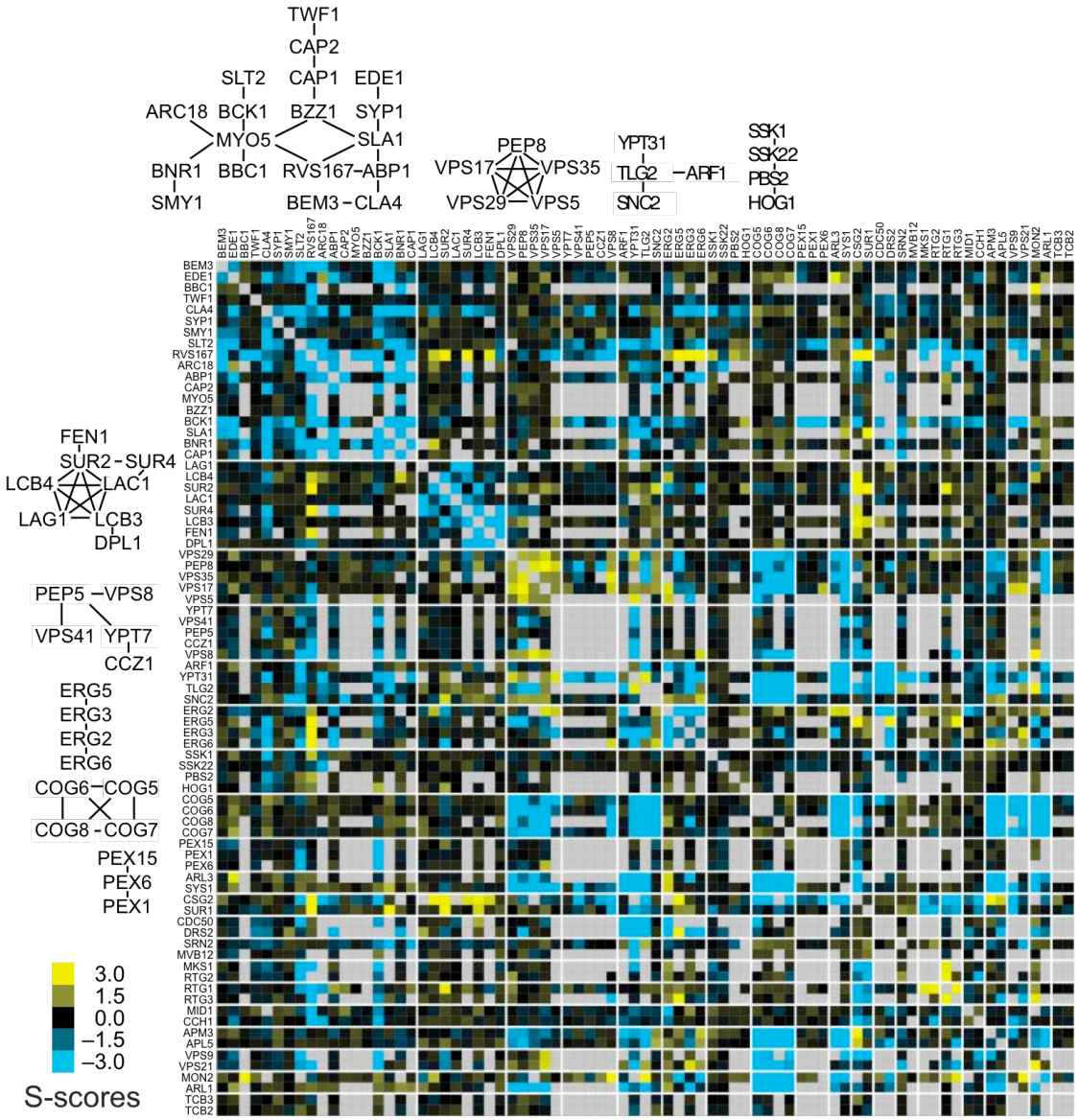


Figure 8: Rom2 Interacts with sphingolipid metabolism. **(a)** Genes with correlating genetic profiles are shared between *SUR4* and *ROM2*, but not between *CSG2* and *ROM2*. Correlation coefficients between the genetic profile of *ROM2* and each of the other 373 profiles in the E-MAP are plotted on the X-axis against on the Y-axis either the similar set of values for the *SUR4* profile with all other profiles (blue), or those for *CSG2* with all other profiles (red). Labeled points indicate genes with profiles that are positively correlated with the profile of *ROM2*. CC values in blue and red indicate the correlation coefficients for the full set of blue or red points plotted. **(b)** Lipidome profiling of *rom2D* and selected sphingolipid metabolism mutants Lipid class abundances were normalized to wild type leves. Sterol esters (SE), phosphatidic acid (PA), triacylglycerol (TAG), long chain base (LCB) mannosylinositol phosphoceramide (MIPC), phosphatidylethanolamine (PE), diacylglycerol (DAG), phoshphatidylcholine (PC), phoshphatidyinositol (PI), ceramide (Cer), phosphatidylserine (PS) mannosylinositol-2-phosphoceramide ($M(IP)_2C$) and inositol phosphoceramide (IPC) levels are shown.

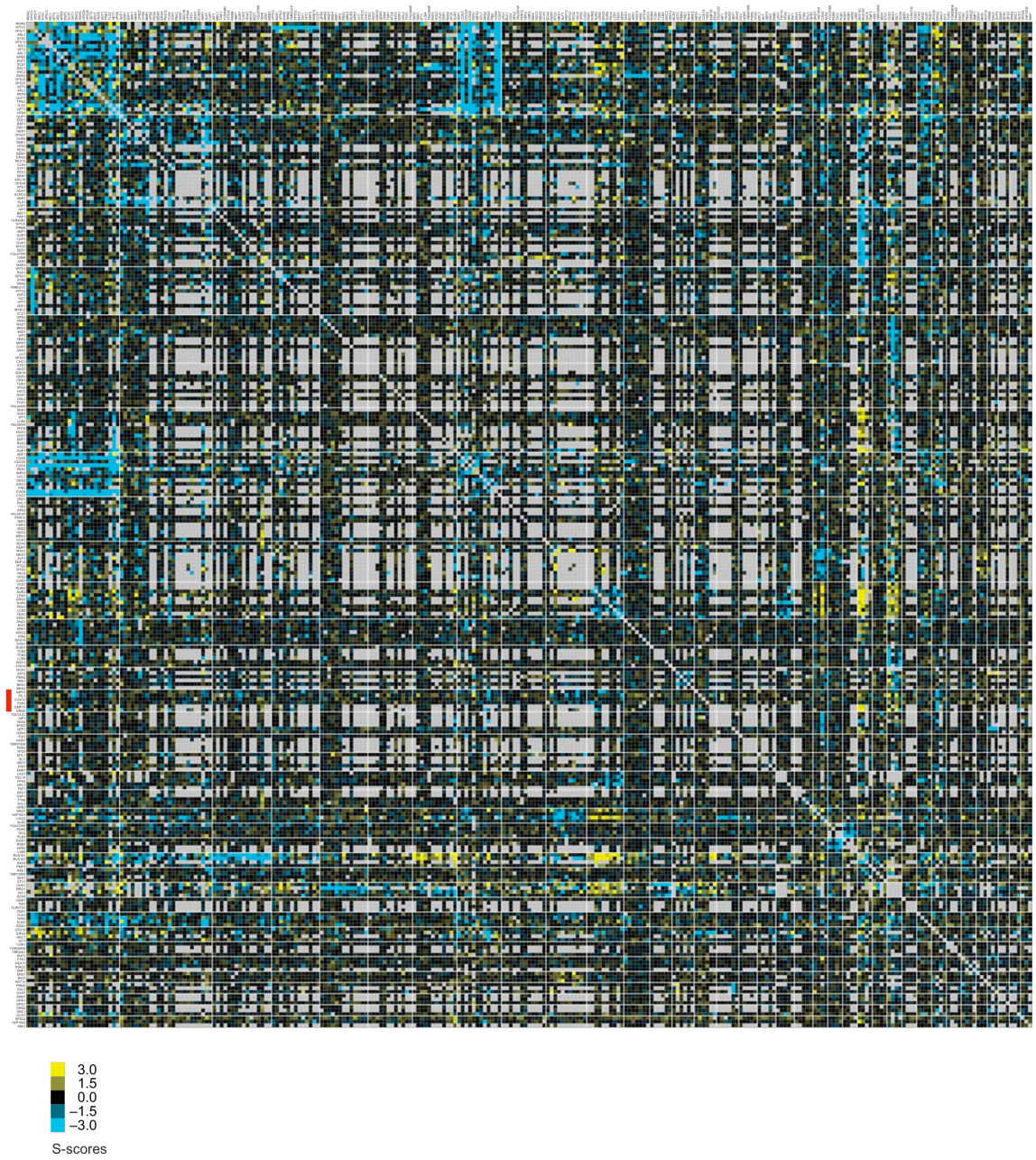
Supplementary Figures



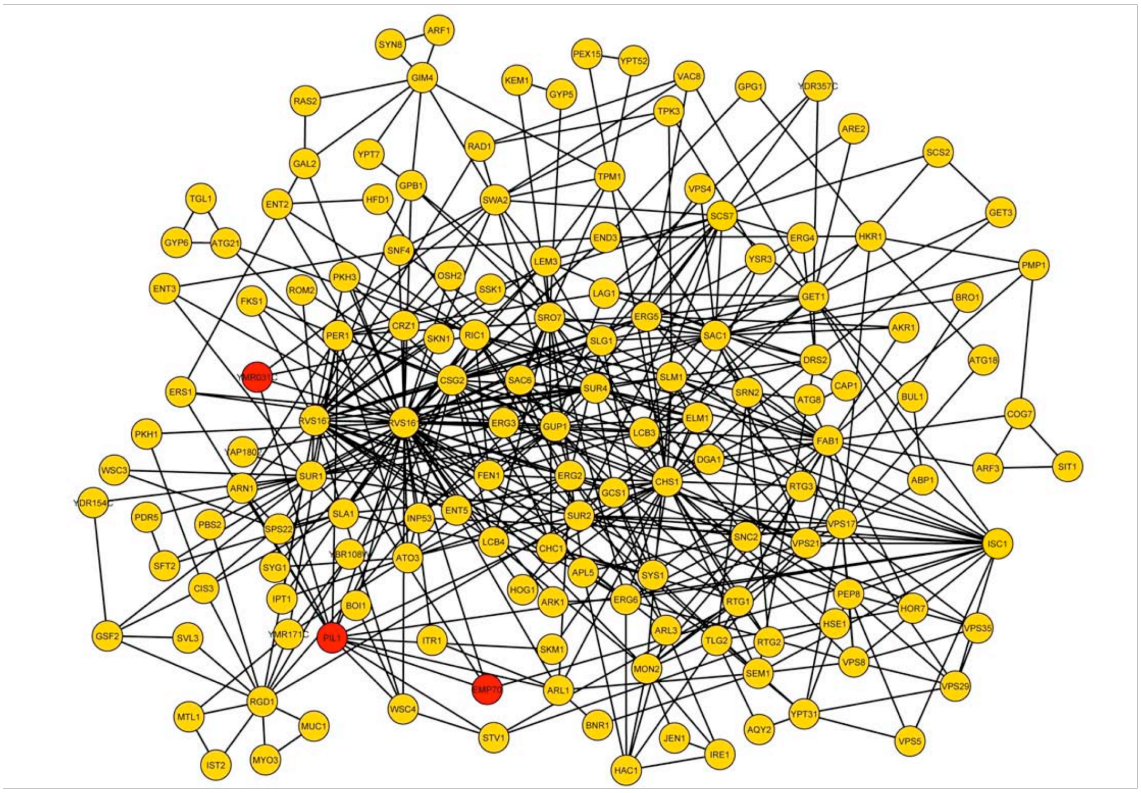
Supplementary Figure 1: Receiver operating characteristic (ROC) curves gauging the power of pairwise genetic profile correlations to predict physical interactions between pairs of proteins. The predictive power of the plasma membrane E-MAP (blue) is comparable to that of previously published E-MAPs, focused on chromosome biology ((Collins et al., 2007b), red) and the early secretory pathway ((Schuldiner et al., 2006), green). The slope of the initial portion of the curve serves as a measure of the predictor's discriminatory power at its minimal false positive rate. Physical interaction definitions were collected from (Collins et al., 2007a), imposing a threshold of PE score > 2.



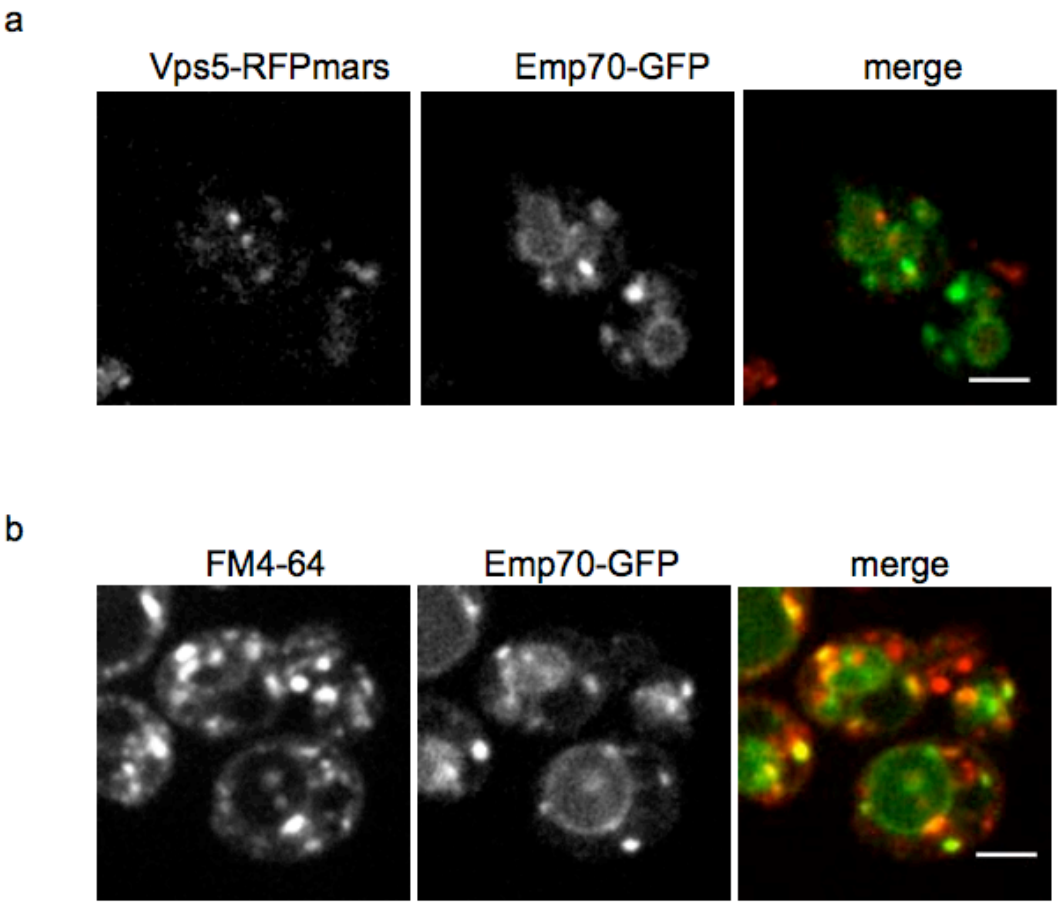
Supplementary Figure 2: Functional modules identified using the plasma membrane E-MAP data and protein-protein interactions. Modules were identified as described in (Ulitsky et al., 2008). White lines separate modules, which are ordered by size. Missing values are in grey. The protein-protein interaction sub-network is shown next to each module with at least three genes.



Supplementary Figure 3: Functional modules identified using the plasma membrane E-MAP data for pairs without PPIs. Modules were identified as described in, but without requiring PPI network connectivity. White lines separate modules, which are ordered by size. Missing values are in grey. The *PIL1-EMP70* module is marked by a red bar.



Supplementary Figure 4: Map of type I TGMs in the plasma membrane E-MAP. Overview of all Type I TGMs found in the plasma membrane E-MAP. *PIL1, YMR031c* and *EMP70* are highlighted in red.



Supplementary Figure 5: Emp70 localization in respect to endosomes. (a) Strains expressing Emp70-GFP and Vps5-RFPmars from the endogenous locus were imaged by confocal microscopy. Representative images are shown. (b) FM4-64 uptake experiment was performed at 16°C and representative confocal midsections are shown. Bar = 2.5 μm.

Supplementary Tables, Data and Movies

Available online at:

http://www.nature.com/nsmb/journal/v17/n7/full/nsmb.1829.html#supplementary-information

Supplementary Methods

E-MAP analysis

Generation and analysis of the quantitative genetic interaction data was carried out as previously described (Collins et al., 2006; Schuldiner et al., 2005; Schuldiner et al., 2006).

Detection of functional modules

Modules were identified using an algorithm described previously (Ulitsky et al., 2008), which was applied to the S-scores in the plasma membrane E-MAP and a PPI network containing 49,010 interactions between 5,815 yeast genes, compiled from several databases (Cherry et al., 1998; Kanehisa and Goto, 2000; Reguly et al., 2006). The algorithm identifies a collection of modules and a set of module pairs by optimizing a probabilistic scoring function. This function takes into account both similarities of the S-score profiles and the raw S-scores, by preferring modules of genes having high correlation coefficients (CC-scores) between members of the same complex and strongly negative S-scores among members of a module pair. Significance of the CC-scores and S-scores was assessed by fitting Gaussian distributions using an EM algorithm.

Yeast strains and culture

All yeast strains used are listed in the Table S1. C-terminal fusions and gene deletions were generated by homologous recombination of PCR-generated DNA fragments (Janke et al., 2004). All single, double or triple deletion strains were

confirmed by PCR. Strains expressing tagged proteins were confirmed by PCR and fluorescence microscopy or Western blot. Yeast cells were grown according to standard procedure. For SILAC labeling, Pil1-TEV-GFP expressing and WT yeast cells and were grown in the presence of the presence of 20 mg l^{-1} L-Lysine-U-$^{13}C_6$,$^{15}N_2$ 20 mg l^{-1} or normal L-Lysine, respectively with at least ten doublings to an OD_{600nm} = 0.7.

Microscopy

Cells were grown to an OD_{600nm}= 0.6 in synthetic medium at 30º C unless indicated. Cells were mounted in synthetic media onto concanavalin A coated cover slips and imaged with an ANDOR/TiLL iMIC CSU22 spinning disk confocal microscope, using an ANDOR iXonEM 897 EM CCD camera and an Olympus 100x 1.4 NA oil immersion objective. 16-bit images were collected using Andor Image iQ 1.9 in the linear range of the camera. For presentation, images were filtered with a smoothening filter averaging 2 pixels, converted to 8-bit images and cropped using ImageJ (www.rsb.info.nih.gov/ij/). For quantitation of colocalization, we collected stacks and extracted 4D images for individual cells. The area of overlap was quantified dividing the total area of the Emp70 signal by the area of overlap determined using the RG2B colocalization ImageJ plugin.

FM4-64 uptake assay

Cells exponentially growing at an OD_{600nm}= 0.7 (1ml), were harvested, resuspended in 50mL of media and chilled on ice for 5 min. FM4-64 was added to a

final concentration of 10mM and incubated for another 10 min. Cells were washed with ice-cold media, resuspended and incubated for different time points at which cells were killed by 10mM NaN_3 and 10mM NaF and immediately analyzed by microscopy.

CPY secretion assay

The CPY secretion colony blot assay was performed as described using anti-CPY antibodies (Invitrogen-A6428) (Mullins and Bonifacino, 2001).

Proteomics

Protein extracts from 70 ODs of "light" and "heavy" labeled cells were obtained as described (Frohlich et al., 2009). For immunopurification, equivalent amounts of proteins were incubated with anti-GFP AB-conjugated magnetic nanobeads (Miltenyi Biotech) for 5min at 4°C and loaded on µMacs columns (Miltenyi Biotech) in a magnetic µMacs Separator (Miltenyi Biotech) washed 3 times with 1 ml lysis buffer with 1% Triton-X100, 3 times with 1ml of lysis buffer without Triton-X100 and eluted by TEV cleavage. Eluates were mixed, diluted 5x in 8M urea, reduced for 20 min at RT in 1 mM dithiothreitol and then alkylated for 30 minutes by 5.5 mM iodoacetamide in the dark. Then the eluates were digested, desalted and concentrated as described (Frohlich et al., 2009). Peptides were separated on-line using an Easy nLC™ system (Proxeon Biosystems, Odense, Denmark). Samples (5 µl) were loaded as described (Frohlich et al., 2009). Peptides were eluted with a segmented gradient of 2-60% solvent B over 102 min with a constant flow of 250 nl

min^{-1}. The HPLC system was coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) via a nanoscale LC interface (Proxeon Biosystems). Spray voltage was 2.2 kV and temperature of the heated capillary was 180°C. Survey full scan spectra (m/z=300-1600) were acquired in positive ion mode with a resolution of 60,000 at m/z = 400 after accumulation of 1,000,000 ions. Up to ten most intense ions were sequenced by collision induced dissociation in the LTQ. Precursor ion charge state screening was enabled and all unassigned charge states as well as singly charged peptides were rejected. The dynamic exclusion list was restricted to a maximum of 500 entries with a maximum retention period of 90 sec and a relative mass window of 10 ppm. Orbitrap measurements were performed enabling the lock mass option for survey scans to improve mass accuracy (Olsen et al., 2005). Data were acquired using the Xcalibur software (version 2.1.0) and MaxQuant, version 1.0.1 (Cox and Mann, 2008). The data was searched against the yeast database concatenated with reversed copies of all sequences (Moore et al., 2002; Peng et al., 2003) and supplemented with frequent contaminants using Mascot (version 2.2.0, Matrix Science (Perkins et al., 1999)). Carbamidomethylated cysteins were set as fixed, oxidation of methionine and N-terminal acetylation as variable modifications. Maximum allowed mass deviation for MS/MS peaks and missed cleavages were 0.5 Da and 3, respectively. Maximum false discovery rates (FDR) were 0.01 both on peptide and protein levels. Minimum required peptide length was 6 amino acids. Proteins with at least two peptides were considered identified.

Lipidomics

Cell were harvested from 20 ml cell logarithmically growing cultures in synthetic media at 30°C and washed in 4°C water. Cell pellets were frozen immediately in liquid nitrogen, thawed later by adding 155 mM NH_4HCO_3 (pH 8), followed by cell disruption using zirconia beads (0.5 mm; BioSpec Products, USA). Lipids were extracted from lysates as previously described (Ejsing et al., 2009) and profiled analysis using a LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Germany) equipped with a robotic nanoflow ion source TriVersa (Advion BioSciences Ltd, USA). MS survey scans were acquired in positive and negative ion mode using the Orbitrap analyzer with target mass resolution of 100,000 and automatic gain control set at $1e^5$ as the target value (Ejsing et al., 2009).

References

Aronova, S., Wedaman, K., Aronov, P.A., Fontes, K., Ramos, K., Hammock, B.D., and Powers, T. (2008). Regulation of ceramide biosynthesis by TOR complex 2. Cell Metab *7*, 148-158.

Audhya, A., and Emr, S.D. (2002). Stt4 PI 4-kinase localizes to the plasma membrane and functions in the Pkc1-mediated MAP kinase cascade. Dev Cell *2*, 593-605.

Babst, M., Wendland, B., Estepa, E.J., and Emr, S.D. (1998). The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. EMBO J *17*, 2982-2993.

Bandyopadhyay, S., Kelley, R., Krogan, N.J., and Ideker, T. (2008). Functional maps of protein complexes from quantitative genetic interaction data. PLoS Comput Biol *4*, e1000065.

Breton, A.M., and Aigle, M. (1998). Genetic and functional relationship between Rvsp, myosin and actin in Saccharomyces cerevisiae. Curr Genet *34*, 280-286.

Brickner, J.H., and Fuller, R.S. (1997). SOI1 encodes a novel, conserved protein that promotes TGN-endosomal cycling of Kex2p and other membrane proteins by modulating the function of two TGN localization signals. J Cell Biol *139*, 23-36.

Brizzio, V., Gammie, A.E., and Rose, M.D. (1998). Rvs161p interacts with Fus2p to promote cell fusion in Saccharomyces cerevisiae. J Cell Biol *141*, 567-584.

Cerantola, V., Guillas, I., Roubaty, C., Vionnet, C., Uldry, D., Knudsen, J., and Conzelmann, A. (2009). Aureobasidin A arrests growth of yeast cells through both ceramide intoxication and deprivation of essential inositolphosphorylceramides. Mol Microbiol *71*, 1523-1537.

Cherry, J.M., Adler, C., Ball, C., Chervitz, S.A., Dwight, S.S., Hester, E.T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., *et al.* (1998). SGD: Saccharomyces Genome Database. Nucleic Acids Res *26*, 73-79.

Collins, S.R., Kemmeren, P., Zhao, X.C., Greenblatt, J.F., Spencer, F., Holstege, F.C., Weissman, J.S., and Krogan, N.J. (2007a). Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics *6*, 439-450.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol *7*, R63.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology *26*, 1367-1372.

deHart, A.K., Schnell, J.D., Allen, D.A., and Hicke, L. (2002). The conserved Pkh-Ypk kinase cascade is required for endocytosis in yeast. J Cell Biol *156*, 241-248.

Deng, C., Xiong, X., and Krutchinsky, A.N. (2009). Unifying fluorescence microscopy and mass spectrometry for studying protein complexes in cells. Mol Cell Proteomics *8*, 1413-1423.

Efe, J.A., Plattner, F., Hulo, N., Kressler, D., Emr, S.D., and Deloche, O. (2005). Yeast Mon2p is a highly conserved protein that functions in the cytoplasm-to-vacuole transport pathway and is required for Golgi homeostasis. J Cell Sci *118*, 4751-4764.

Ejsing, C.S., Sampaio, J.L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R.W., Simons, K., and Shevchenko, A. (2009). Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. Proc Natl Acad Sci U S A *106*, 2136-2141.

Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S.*, et al.* (2009). Functional organization of the S. cerevisiae phosphorylation network. Cell *136*, 952-963.

Friant, S., Lombardi, R., Schmelzle, T., Hall, M.N., and Riezman, H. (2001). Sphingoid base signaling via Pkh kinases is required for endocytosis in yeast. EMBO J *20*, 6783-6792.

Friesen, H., Humphries, C., Ho, Y., Schub, O., Colwill, K., and Andrews, B. (2006). Characterization of the yeast amphiphysins Rvs161p and Rvs167p reveals roles for the Rvs heterodimer in vivo. Mol Biol Cell *17*, 1306-1321.

Frohlich, F., Moreira, K., Aguilar, P.S., Hubner, N.C., Mann, M., Walter, P., and Walther, T.C. (2009). A genome-wide screen for genes affecting eisosomes reveals Nce102 function in sphingolipid signaling. J Cell Biol *185*, 1227-1242.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B.*, et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Grossmann, G., Malinsky, J., Stahlschmidt, W., Loibl, M., Weig-Meckl, I., Frommer, W.B., Opekarova, M., and Tanner, W. (2008). Plasma membrane microdomains regulate turnover of transport proteins in yeast. J Cell Biol *183*, 1075-1088.

Grossmann, G., Opekarova, M., Malinsky, J., Weig-Meckl, I., and Tanner, W. (2007). Membrane potential governs lateral segregation of plasma membrane proteins and lipids in yeast. EMBO J *26*, 1-8.

Guan, X.L., Souza, C.M., Pichler, H., Dewhurst, G., Schaad, O., Kajiwara, K., Wakabayashi, H., Ivanova, T., Castillon, G.A., Piccolis, M.*, et al.* (2009). Functional interactions between sphingolipids and sterols in biological membranes regulating cell physiology. Mol Biol Cell *20*, 2083-2095.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., *et al.* (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947-962.

Jonikas, M.C., Collins, S.R., Denic, V., Oh, E., Quan, E.M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J.S., *et al.* (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science *323*, 1693-1697.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res *28*, 27-30.

Kobayashi, T., Takematsu, H., Yamaji, T., Hiramoto, S., and Kozutsumi, Y. (2005). Disturbance of sphingolipid biosynthesis abrogates the signaling of Mss4, phosphatidylinositol-4-phosphate 5-kinase, in yeast. J Biol Chem *280*, 18087-18094.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., *et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Levin, D.E. (2005). Cell wall integrity signaling in Saccharomyces cerevisiae. Microbiol Mol Biol Rev *69*, 262-291.

Moore, R.E., Young, M.K., and Lee, T.D. (2002). Qscore: an algorithm for evaluating SEQUEST database search results. J Am Soc Mass Spectrom *13*, 378-386.

Mullins, C., and Bonifacino, J.S. (2001). Structural requirements for function of yeast GGAs in vacuolar protein sorting, alpha-factor maturation, and interactions with clathrin. Mol Cell Biol *21*, 7981-7994.

Olsen, J.V., de Godoy, L.M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol Cell Proteomics *4*, 2010-2021.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res *2*, 43-50.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551-3567.

Reguly, T., Breitkreutz, A., Boucher, L., Breitkreutz, B.J., Hon, G.C., Myers, C.L., Parsons, A., Friesen, H., Oughtred, R., Tong, A.*, et al.* (2006). Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae. J Biol *5*, 11.

Ren, G., Vajjhala, P., Lee, J.S., Winsor, B., and Munn, A.L. (2006). The BAR domain proteins: molding membranes in fission, fusion, and phagy. Microbiol Mol Biol Rev *70*, 37-120.

Revardel, E., Bonneau, M., Durrens, P., and Aigle, M. (1995). Characterization of a new gene family developing pleiotropic phenotypes upon mutation in Saccharomyces cerevisiae. Biochim Biophys Acta *1263*, 261-265.

Robinson, J.S., Klionsky, D.J., Banta, L.M., and Emr, S.D. (1988). Protein sorting in Saccharomyces cerevisiae: isolation of mutants defective in the delivery and processing of multiple vacuolar hydrolases. Mol Cell Biol *8*, 4936-4948.

Rothman, J.H., Howald, I., and Stevens, T.H. (1989). Characterization of genes required for protein sorting and vacuolar function in the yeast Saccharomyces cerevisiae. EMBO J *8*, 2057-2065.

Schimmoller, F., Diaz, E., Muhlbauer, B., and Pfeffer, S.R. (1998). Characterization of a 76 kDa endosomal, multispanning membrane protein that is highly conserved throughout evolution. Gene *216*, 311-318.

Schmidt, A., Bickle, M., Beck, T., and Hall, M.N. (1997). The yeast phosphatidylinositol kinase homolog TOR2 activates RHO1 and RHO2 via the exchange factor ROM2. Cell *88*, 531-542.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F*., et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Schuldiner, M., Collins, S.R., Weissman, J.S., and Krogan, N.J. (2006). Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344-352.

Sipos, G., Brickner, J.H., Brace, E.J., Chen, L., Rambourg, A., Kepes, F., and Fuller, R.S. (2004). Soi3p/Rav1p functions at the early endosome to regulate endocytic trafficking to the vacuole and localization of trans-Golgi network transmembrane proteins. Mol Biol Cell *15*, 3196-3209.

Tabuchi, M., Audhya, A., Parsons, A.B., Boone, C., and Emr, S.D. (2006). The phosphatidylinositol 4,5-biphosphate and TORC2 binding proteins Slm1 and Slm2 function in sphingolipid regulation. Mol Cell Biol *26*, 5861-5875.

Teis, D., Saksena, S., and Emr, S.D. (2008). Ordered assembly of the ESCRT-III complex on endosomes is required to sequester cargo during MVB formation. Dev Cell *15*, 578-589.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M*., et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Ulitsky, I., Shlomi, T., Kupiec, M., and Shamir, R. (2008). From E-MAPs to module maps: dissecting quantitative genetic interactions using physical interactions. Mol Syst Biol *4*, 209.

Valdivia, R.H., Baggott, D., Chuang, J.S., and Schekman, R.W. (2002). The yeast clathrin adaptor protein complex 1 is required for the efficient retention of a subset of late Golgi membrane proteins. Dev Cell *2*, 283-294.

Vida, T.A., and Emr, S.D. (1995). A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast. J Cell Biol *128*, 779-792.

Walther, T.C., Aguilar, P.S., Frohlich, F., Chu, F., Moreira, K., Burlingame, A.L., and Walter, P. (2007). Pkh-kinases control eisosome assembly and organization. EMBO J *26*, 4946-4955.

Walther, T.C., Brickner, J.H., Aguilar, P.S., Bernales, S., Pantoja, C., and Walter, P. (2006). Eisosomes mark static sites of endocytosis. Nature *439*, 998-1003.

Wang, H., Kakaradov, B., Collins, S.R., Karotki, L., Fiedler, D., Shales, M., Shokat, K.M., Walther, T.C., Krogan, N.J., and Koller, D. (2009). A complex-based reconstruction of the Saccharomyces cerevisiae interactome. Mol Cell Proteomics *8*, 1361-1381.

Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S.*, et al.* (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell *32*, 735-746.

Zheng, B., Wu, J.N., Schober, W., Lewis, D.E., and Vida, T. (1998). Isolation of yeast mutants defective for localization of vacuolar vital dyes. Proc Natl Acad Sci U S A *95*, 11721-11726.

Chapter 7

Systematic Triple Mutant Analysis Uncovers Functional Connectivity Between Pathways Involved in Chromosome Regulation

Systematic Triple Mutant Analysis Uncovers Functional Connectivity Between Pathways Involved in Chromosome Regulation

James E. Haber[1,*], Hannes Braberg[2,3], Qiuqin Wu[1], Richard Alexander[2,3], Julian Haase[4], Colm Ryan[2,3], Zach Lipkin-Moore[1], Kathleen E. Franks-Skiba[2,3], Tasha Johnson[2,3,5], Michael Shales[2,3], Tineke L. Lenstra[6], Frank C. P. Holstege[6], Jeffrey R. Johnson[2,3,5], Kerry Bloom[4] and Nevan J. Krogan[2,3,5,*]

1 Department of Biology and Rosenstiel Basic Medical Sciences Research Center, Waltham, MA, 02454, USA;

2 Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, 94158, USA;

3 California Institute for Quantitative Biosciences, QB3, San Francisco, CA, 94158, USA;

4 Department of Biology, University of North Carolina, Chapel Hill, NC, 27599, USA;

5 J. David Gladstone Institutes, San Francisco, CA, 94158, USA;

6 Molecular Cancer Research, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands;

*To whom correspondence should be addressed: James E. Haber (haber@brandeis.edu) and Nevan J. Krogan (nevan.krogan@ucsf.edu)

Abstract

Genetic interactions reveal the functional relationships between pairs of genes. In this study, we describe a method for the systematic generation and quantitation of triple mutants, termed Triple Mutant Analysis (TMA). We have used this approach to interrogate partially redundant pairs of genes in *S. cerevisiae*, including *ASF1* and *CAC1*, two histone chaperones. After subjecting *asf1*Δ *cac1*Δ to TMA, we found that the Swi/Snf Rdh54 protein, compensates for the absence of Asf1 and Cac1. Rdh54 more strongly associates with the chromatin apparatus and the pericentromeric region in the double mutant. Moreover, Asf1 is responsible for the synthetic lethality observed in *cac1*Δ strains lacking the HIRA-like proteins. A similar TMA was carried out after deleting both *CLB5* and *CLB6*, cyclins that regulate DNA replication, revealing a strong functional connection to chromosome segregation. This approach can reveal functional redundancies that cannot be uncovered using traditional double mutant analyses.

Introduction

The systematic study of genetic interactions by high-throughput methods in numerous organisms has provided insight into a variety of complicated biological processes (Beltrao et al., 2010). For example, in *S. cerevisiae* double mutants obtained by crossing a given gene knockout with a large number of other viable single gene ablations or modifications can reveal either synthetic reductions in growth or they can exhibit epistasis (Collins et al., 2010; Schuldiner et al., 2005; Tong et al., 2004). Synthetic lethality or synthetic sickness occurs when two genes have overlapping, important functions so that the absence of both genes results in a severe growth defect that is only seen when both genes are absent. In the same fashion, double mutant combinations can reveal sensitivity to environmental conditions or to antibiotics. In contrast, the combination of two knockouts that remove elements of the same pathway shows no additional phenotype (epistasis). Epistatic interactions often correspond to gene products that participate in protein-protein interactions and/or function in the same biological pathway (Beltrao et al., 2010; Collins et al., 2007).

More recently, this type of systematic genetic analysis has been extended to other organisms, including *S. pombe* (Roguev et al., 2008; Roguev et al., 2007; Ryan et al., 2012), *C. elegans* (Lehner et al., 2006), *D. melanogaster* (Horn et al., 2011) and mammalian cells (Bassik et al., 2013; Lin et al., 2012; Roguev et al., 2013) and has also been carried out in the presence of specific exogenous stresses (Ideker and Krogan, 2012) and using specific point mutants of multifunctional genes (unpublished data). To date, almost all genetic interaction data collected in any

organism has been generated using systems that create mutants in a pair-wise fashion, even though great mechanistic insight could be uncovered through higher perturbation studies. In this study, we describe an approach, termed Triple Mutant Analysis (TMA), which facilitates higher-level genetic interaction analysis by allowing for the generation and quantitative analysis of triple mutants in budding yeast.

In budding yeast, Asf1 and the CAF-1 complex, comprising Cac1, Cac2 and Cac3, are the two known histone H3-H4 chaperones (De Koning et al., 2007). Deletion of *ASF1* results in slower growth, with an accumulation of cells in G2, as well as a marked increase in gross chromosomal rearrangements (GCR) (Kats et al., 2006; Ramey et al., 2004). Asf1, interacting with histones H3-H4, also recruits Rtt109, an enzyme responsible for histone H3-K56 acetylation (Collins et al., 2007; Driscoll et al., 2007; Han et al., 2007b). Furthermore, Asf1 plays a role in the monitoring of replication and chromatin assembly, through interaction with the checkpoint kinase Rad53. Similarly, deletion of any of the three subunits of CAF-1 (Cac1, Cac2 or Cac3) results in elevated GCR (Myung et al., 2003) and mutation of *CAC1*, also known as *RLF2*, affects the structure of telomeric chromatin and the transcriptional silencing of *HML* and *HMR* (Enomoto and Berman, 1998). As expected, the double mutant *asf1*Δ *cac1*Δ is more seriously compromised with cells having a doubling time about twice that of the wild type cells, with a long delay in the G2/M phase of the cell cycle (Kats et al., 2006). Moreover, Asf1 and Cac1 have also been shown to have overlapping roles in deactivating the DNA damage checkpoint after cells have repaired a double-strand break (Kim and Haber, 2009).

Despite this, *asf1*Δ *cac1*Δ strains are surprisingly robust (Kats et al., 2006; Ramey et al., 2004). Therefore, we hypothesized that there could be additional chromatin remodelers acting as functional "back-ups" to Asf1 and CAF-1, and hence we subjected the double mutant to TMA.

Similar to Asf1 and CAF-1, two of yeast six cyclin B homologs, Clb5 and Clb6, act redundantly in regulating the timing of DNA replication in mitosis and meiosis (Decesare and Stuart, 2012; Donaldson et al., 1998); however a *clb5*Δ *clb6*Δ double mutant is not particularly sick, suggesting an additional pathway exists for this process. In both of these instances, we show novel genetic outcomes derived from the TMA approach that provide mechanistic insight into the pathways being interrogated that could not have been gleaned from the analysis of single mutants, illustrating the utility of this approach. We propose that the experimental and computational framework described here could be used to interrogate higher order relationships in various biological processes in many organisms, including in mammalian cells.

Results and Discussion

In some instances, eliminating redundant genes results in cell death (e.g. deleting the three cyclin A homologs (*CLN1*, *CLN2* and *CLN3*) or deleting 2 cyclin B homologs (*CLB1* and *CLB2*) are lethal (Schneider et al., 1996; Surana et al., 1991)). Similarly, deleting components of parallel pathways can result in cell death (e.g. ablating pairs of genes involved in chromosome transmission, such as *CTF18* and *CTF4*, is synthetically lethal (Xu et al., 2007)). However, in other situations,

removing two parallel processes leaves the double mutant slow-growing, but still viable. One such case involves the absence of two known histone H3-H4 chaperone complexes – Asf1 and the CAF-1 complex. If Asf1 and CAF-1 are the only histone H3-H4 chaperones, how do double mutant cells remain viable? In mammalian cells, the HIRA complex also acts as an independent histone chaperone (Ray-Gallet et al., 2002); in budding yeast, the components of this complex (Hir1, Hir2, Hir3 and Hpc2) appear to be involved in regulating histone transcription, but are also involved in histone deposition working in association with Rtt106 (Green et al., 2005; Krawitz et al., 2002; Sharp et al., 2002; Silva et al., 2012). To address this question, we have used the SGA technique (Tong et al., 2004) and the quantitative scoring system associated with the E-MAP approach (Schuldiner et al., 2005) to develop a quantitative strategy to analyze triple mutants, which we term Triple Mutant Analysis (TMA). Normally, double mutant analysis in *S. cerveisiae* involves the crossing of two differently marked mutant strains to generate a double mutant strain. TMA involves the use of a double mutant strain that is crossed to a panel of mutant strains, allowing for the systematic creation of triple mutants. In this study, a haploid *MAT*α *asf1*Δ::*HPH cac1*Δ::*URA3* strain was crossed against *MAT***a** haploids with knockouts of non-essential genes or hypomorphic alleles of essential genes, marked by *KAN*.

We anticipated three scenarios that might emerge from the selection of Ura^+ HPH^R KAN^R triple mutants. First, the triple mutant might exhibit an even more severe growth defect, revealing a third gene whose product might act as an additional histone H3-H4 chaperone. Second, mutations that caused more severe

growth defects might be uncovered if, for example, the absence of both chaperones exacerbates the phenotype of another mutant not involved in chromatin remodeling (e.g. by altering the expression of other genes). A third scenario, which would represent a positive interaction, could correspond to a suppression of the intrinsically poor growth of *asf1*Δ *cac1*Δ via removal of a gene whose mis-regulated product was responsible for the poor growth associated with the double mutant.

TMA of asf1Δ cac1Δ

In the genetic cross, we can select for two different sets of double mutants (*asf1*Δ *geneX*Δ; *cac1*Δ *geneX*Δ) as well as for the triple mutants (*asf1*Δ *cac1*Δ *geneX*Δ) using the appropriate antibiotic resistant markers (**Figure 1A; Supp. Data 1**). Analysis of this data using previously described analytical tools can provide genetic scores (or S-scores) (Collins et al., 2010) for all three sets of mutants that range from negative (i.e. synthetic sickness) (blue) to positive (suppression) (yellow) (**Figure 1B**). However, inspection of the triple mutant scores alone can be deceiving if not compared to the scores derived from each corresponding double mutant. For example, there are interactions of a gene knockout with both *asf1*Δ and *cac1*Δ (i.e. the triple mutant) that were not significantly different from a deletion of the gene with either *asf1*Δ or *cac1*Δ. Indeed, the S-score from the *asf1*Δ *slx8*Δ double mutant (-6.7) is similar to what is observed with the *asf1*Δ *cac1*Δ *slx8*Δ triple mutant (-6.9) (**Figure 1B, middle**). Other examples include *asf1*Δ *doa1*Δ and *cac1*Δ *nup60*Δ (**Figure 1B**), suggesting that only one of the two starting mutants is contributing to

the phenotype observed in the triple mutant. To separate these cases from situations where the triple mutant score is strong when each double mutant has little effect, we performed a minimum difference comparison (MinDC) between the score derived from the triple mutant compared to the more significant score of the two double mutant combinations. The resulting MinDC scores, which also range from negative (dark blue) to positive (red) (**Figure 1B; Supp. Data 2**), reveal that the most significant negative interactions found with both *ASF1* and *CAC1* absent are with *swm1*Δ (-10.9), *hsl1*Δ (-10.7), *clb2*Δ (-9.8), *rad27*Δ (-9.3), and *rpn4*Δ (-8.3). Hsl1 is a Nim1-like kinase that acts on the cell cycle regulator, Swe1 (Booher et al., 1993). Clb2, a B-type cyclin (Surana et al., 1991) and Swm1, a component of the Anaphase Promoting Complex (APC) (Hall et al., 2003) are implicated in cell cycle progression, as is Rpn4, a component of the 19S proteasome (Xie and Varshavsky, 2001). Rad27 is a 5' flap endonuclease involved in DNA repair (Reagan et al., 1995). Among the genes that appeared to act redundantly with Asf1 and Cac1 is Radh54, also known as Tid1, a Swi/SNF homolog that has been implicated in chromatin remodeling (Kwon et al., 2008; Prasad et al., 2007). Deletion of *RDH54* resulted in the 14th lowest MinDC score (-6.7) (**Figure 1B**).

To confirm these results, we carried out tetrad analysis after sporulating diploids heterozygous for *asf1*Δ *cac1*Δ and either *clb2*Δ or *rdh54*Δ (**Figure 1C**), as well as *hsl1*Δ (data not shown). In each case, *asf1*Δ *cac1*Δ segregants are slow-growing, but the phenotype of the triple mutants are far more severe. Hsl1 was originally identified as a synthetic lethal mutation (histone synthetic lethality) in cells lacking the N-terminal tail of histone H3 (Ma et al., 1996) and it appears that

*asf1*Δ *cac1*Δ phenocopies the H3 tail mutation. Hsl1 is required for degradation of the Swe1 kinase and therefore in a *hsl1*Δ mutant background, Swe1 is constitutively active and cells are prevented from completing the cell cycle (Sia et al., 1996). Consistent with this interpretation, we showed that deleting *SWE1* in the *asf1*Δ *cac1*Δ *hsl1*Δ background suppressed its lethality (data not shown). Interestingly, the cyclin Clb2 is also genetically linked to the Swe1-Hsl1 network (Simpson-Lavy et al., 2009). We also confirmed the TMA negative interaction with Rdh54 (**Figure 1C**). Radh54 plays important roles in meiotic recombination, where it partners with the meiosis-specific recombination protein Dmc1, but has relatively minor roles in mitotic recombination (Dresser et al., 1997; Klein, 1997). In mitotic cells, its role in recombination is less evident, but *rdh54D* cells have defects in DNA damage checkpoint adaptation after induction of a single, unrepaired DSB (Lee et al., 2001). Rdh54 also appears to act redundantly with Rad54 and Uls1, two other Swi/Snf homologs that work to free Rad51 from dsDNA to enable it to repair DSBs (Chi et al., 2011). However, neither *rad54*D nor *uls1*D exhibit any significant genetic interaction with *asf1*D *cac1*D (**Supp. Data 1, 2**).

On the positive end of the MinDC scores, we again see cases where one of the double mutant combinations is similar to what is seen in the triple mutant; for example the S-score for *asf1*Δ *cac1*Δ *rtt109*Δ (+5.3) is similar to that observed for *rtt109*Δ *asf1*Δ (+4.5) (**Figure 1B**). This is an example of classic epistasis, where the Rtt109 histone acetyltransferase is known to be dependent on Asf1 for its activity (Collins et al., 2007; Driscoll et al., 2007; Han et al., 2007a). However, when the MinDC analysis was applied, we found all five members of the HIRA-like complex

(*HIR1*, *HIR2*, *HIR3*, *HPC2* and *RTT106*) to be the genes with largest MinDC scores (**Figure 1B**). Double mutants lacking *cac1*Δ and one of the HIR complex genes display severe growth defects (Sharp et al., 2002), a finding that is recapitulated in our study. However, this severe defect is dramatically suppressed by deleting *ASF1*. The entire dataset are displayed in a more comprehensive fashion in **Figure 2**. The S-scores for the *asf1*Δ and *cac1*Δ double mutants are displayed along two axes. For each gene, the MinDC score for the triple mutant is displayed by node color. Red indicates cases where the triple mutant grew better than the single mutants (MinDC scores >+5); dark blue shows where the triple mutant was more severely impaired (MinDC scores <-7)). Overall, the genetic profile obtained from the double mutant is more similar to that seen with *asf1*Δ than *cac1*Δ (**Figure 2B**), suggesting that defects in *asf1*Δ *cac1*Δ are largely a reflection of those that arise when Asf1 is absent.

In summary, our data suggest that Rdh54 is potentially acting in a parallel pathway to Asf1 and CAF-1, one that may only function when the other pathways are disabled (**Figure 2C**) - a conclusion supported by additional observations presented later. Moreover, that deleting *ASF1* apparently suppresses double deletions of HIRA and CAF-1 is surprising because one might have anticipated that components of the HIRA complex would be the functionally substitute for Asf1 and CAF-1 and hence have strongly negative MinDC scores. The utility of the TMA scoring system is that it reveals very strong suppression even though the triple mutants are not generally much better growing than average triple mutant strains and hence would not provide significant S-scores. These suppressive relationships are consistent with the observation that HIRA/Rtt106 proteins normally act to

regulate Asf1. When HIRA/Rtt106 is absent, Asf1 apparently acts in an abnormal fashion, causing a severe defect in growth when CAF-1 is also absent (**Figure 2C**). Thus deleting *ASF1* rescues the extreme defect of *cac1*Δ *hir/rtt106/hpc2*Δ mutants. For example, in the absence of Cac1 and HIRA/Rtt106, the level of histones may fall too low, but *asf1*Δ may ensure adequate histone abundance.

Functional comparison of Asf1 and Rtt109 using triple mutant analysis

One major role of Asf1 is to promote the acetylation of histone H3-K56 by the histone acetyltransferase, Rtt109 (Collins et al., 2007; Driscoll et al., 2007; Han et al., 2007a). To determine what degree the defects associated with *asf1*Δ are related to Rtt109 function, we carried out a similar TMA using a strain deleted for both *RTT109* and *CAC1* (**Supp. Data 3**). Overall, there is a strong correlation of the double mutant scores from *asf1*Δ *cac1*Δ and *rtt109*Δ *cac1*Δ (r=0.60) (**Figure 3A**) (for scatter plots using MinDC scores, see **Supp. Figure 1**). For example, deletion of *RAD27*, *SWM1*, *HSL1* or *RDH54*, in combination with *rtt109*Δ *cac1*Δ, all provided negative S-scores, as was also observed with *asf1*Δ *cac1*Δ. Furthermore, negative interactions were also seen with both sets of triple mutants when Rrm3, a DNA helicase (Ivessa et al., 2002), or Tsa1, a thioredoxin peroxidase involved in DNA replication stress (Inoue et al., 1999), were removed (**Figure 3A,B**). Similarly, positive interactions were observed with both pairs of double mutants combined with deletions of *MMS22*, *MMS1*, *CSM3* and *MRC1*, genes involved in monitoring DNA replication progression (Mohanty et al., 2006; Vaisica et al., 2011). These data

support the idea that Asf1 and Rtt109 primarily work together in regulating chromatin function.

Nonetheless, several interesting differences between *asf1*Δ *cac1*Δ and *rtt109*Δ *cac1*Δ hint at functional differences between Asf1 and Rtt109. For example, the negative interactions when *cac1*Δ is combined with *hir1*Δ, *hir2*Δ, *hir3*Δ or *hpc2*Δ are not suppressed by *rtt109*Δ (**Figure 3A,B**), unlike what is seen with *asf1*Δ. This suggests that Asf1 has at least some roles related to CAF-1 and HIRA functions that are independent of Rtt109. Similar results were observed with deletions of one of the two genes encoding for histones H3 and H4 (*HHT1*, *HHT2* or *HHF1*) (**Figure 3A,B**), consistent with the notion that the suppressive activity is related to histone levels. Furthermore, only *asf1*Δ *cac1*Δ, but not *rtt109*Δ *cac1*Δ, displayed strong negative interactions when combined with deletions of components of the SWR-C complex (*SWR1*, *VPS71*, *VPS72*, *SWC3* and *HTZ1*), which replaces histone H2A with the variant Htz1 in chromatin (Kobor et al., 2004; Krogan et al., 2003; Mizuguchi et al., 2004). Again, these data imply that the Asf1 functions associated with these genetic interactions are not related to Rtt109.

Finally, we observed a striking different genetic interaction pattern with *rtt109*Δ *cac1*Δ versus *asf1*Δ *cac1*Δ double and triple mutants with deletions of *MUS81* and *MMS4*, encoding an endonuclease complex that promotes crossovers during homologous recombination (de los Santos et al., 2003; Ho et al., 2010). Only *mus81*Δ, and not *mms4*Δ, displays a strong negative interaction with *rtt109*Δ *cac1*Δ (**Figure 3B**), as confirmed by tetrad analysis (**Figure 3C**), even though Mus81 and Mms4 are thought to function exclusively together. These data suggest that Mus81

works independently of Mms4 in certain instances. It is interesting to note that in mammals, Mus81 has been shown to pair with two distinct partners, Eme1 and Eme2 (Shin et al., 2012). Although no other partner as been found in budding yeast, our results raise the possibility of a second partner or that in some circumstances, Mus81 can act alone. This result is an example where Rtt109 seemingly functions outside of its role with Asf1 since *asf1*Δ does not provide the same pattern of interactions (**Figure 3B**). We suggest that the Rtt109 acetyltransferase has substrates other than histone H3-K56 that are unrelated to Asf1. Collectively, these data demonstrate the insight that can be gained from higher-level genetic analysis that could not be gleaned from information derived from double mutants, as it teases out important differences between factors that were thought to work in the same functional pathway.

Rad54 plays a role in chromatin regulation in the absence of Asf1 and CAF-1

The fact that deleting *RDH54* results in a marked growth defect in an *asf1*Δ *cac1*Δ or *rtt109*Δ *cac1*Δ background suggests that Rdh54 plays a more major role in chromatin regulation in the absence of the two chaperones. To test this idea, we TAP-tagged and purified Rdh54 from both wild type and *asf1*Δ *cac1*Δ strains (Krogan et al., 2006). As shown in **Figure 4A**, there was a marked increased in the number and quantity of co-purifying proteins when the purification was carried out in a strain lacking Asf1 and Cac1. This increase did not reflect a difference in the amount of Rdh54-TAP that was *in vivo* or purified from the two strains (**Figure 4A**). Mass spectrometry analysis revealed that several of the additional factors seen

specifically co-purifying with Rad54-TAP in the double deletion background are components of chromatin modifying enzymes, including Ino80 and RSC, as well as the core histones H2A, H2B and H4 (**Figure 4B**, **Supp. Data 4**). We also recovered one member of the SWR-C complex, Arp6, as well as several proteins that are part of both Ino80 and SWR-C. Interestingly, we had identified the catalytic subunit of the SWR-C, Swr1, another member of the SNF2 family of ATPases, as having a MinDC score of -6.4 (**Supp. Data 2**), suggesting that SWR-C also functions in a redundant pathway to Rdh54. Furthermore, we identified 11/15 subunits of the general transcription factor, TFIID, 6 proteins of RNA polymerase I and Rfc2 and Rfc3, components of multiple RFC complexes involved in DNA repair and replication. We also recovered several components of the Anaphase Promoting Complex (APC), again consistent with a strong negative TMA score with deletion of *SWM1*, a component of the APC, in the *asf1*Δ *cac1*Δ background. One explanation for the increased association of these proteins with Rdh54 is that they could be more highly expressed in the double mutant background. To test this, we globally assessed the gene expression changes in *asf1*Δ *cac1*Δ and found that the majority of the genes corresponding to these proteins were actually down-regulated (**Supp. Figure 2**; **Supp. Data 5**). That there are so many complexes involved in chromatin regulation found specifically associated with Rdh54-TAP *asf1*Δ *cac1*Δ strongly suggests that Rdh54 plays a more central role in overall chromatin regulation in the absence of the well-characterized histone chaperones.

Interestingly, we also found that components of cohesion (Mcd1/Scc1, Smc1 and Smc3) preferentially associated with Rdh54-TAP in the *asf1*Δ *cac1*Δ mutant

background (**Figure 4**), suggesting that Rdh54 plays a role in aspects of chromosome segregation. To further explore this connection, we carried out a cytological analysis of Rdh54-GFP in the presence and absence of Asf1 and Cac1. In wild-type cells, Rdh54 is frequently localized to the metaphase spindle in mitosis (64% of cells). Specifically, Rdh54 is concentrated at centromeres and flanking pericentric chromatin and exists exclusively as "spots" (**Figure 5A**). Unlike kinetochore proteins, including Mtw1, Rdh54 does not always appear as two spots of equal intensity representing sister kinetochores. Instead, one or two foci are seen (e.g. Rdh54-GFP, Metaphase: first and second panel, respectively). In the absence of Asf1 and Cac1 in metaphase, there is 4.8-fold increase in the concentration of Rdh54-GFP along the pericentric chromatin and this localization is apparent in all cells; 63% have foci spots whereas 37% now have elongated "bars", or linear extensions between the spindle poles (>2 microns) (**Figure 5A**). Upon anaphase onset, only 20% of the Rad54-GFP wild-type cells display fluorescence, with an equal amount showing focal spots and bars (**Figure 5B**). In *asf1*Δ *cac1*Δ cells, there is a substantial increase of overall fluorescence (7.7 fold) and it is seen in all cells, 89% as bars. In general, the strong prevalence of this pattern is reminiscent of condensin, which is concentrated in the pericentric chromatin and lies proximal to the spindle axis in metaphase (Stephens et al., 2011). This suggests that Rad54 plays a role in chromatin remodeling and/or chromosome condensation.

The pericentromere surrounding the point centromere is functionally distinct from the bulk chromosome arms during mitosis (Verdaasdonk et al., 2012). Several chromatin remodeling complexes, including RSC and ISW2, are required for

histone stability and chromatin packaging within the pericentromere as this region experiences spindle tension force. The enrichment of Rdh54 in the absence of the Asf1 and CAF-1 histone chaperones suggests it plays a role in the maintenance of chromatin under tension. Its localization along the spindle axis also places Rdh54 in a position to regulate chromatin loops formed by the action of condensin (Stephens et al., 2011). Condensin and cohesin compact the pericentromere into axial loops that constitute the chromatin spring. The combined function of pericentric compaction, together with chaperones to regulate histone dynamics, reflect a critical function in building a chromatin spring that resists microtubule-based pulling forces throughout mitosis. The recruitment of Rdh54 to this region in the absence of histone chaperones reflects the ability of the cell to respond to changes in the stoichiometry of components that regulate chromatin stiffness. Collectively, this localization data is consistent with the protein-protein interaction data (**Figure 4**) that suggests Rdh54 is more functionally relevant in the absence of the two chaperones.

TMA of CLB5 and CLB6 reveals a connection of the cyclins to chromosome segregation

In a similar fashion to Asf1 and CAF-1, we analyzed another situation of apparent redundancy. The two cyclin B proteins, Clb5 and Clb6, function in the regulation of the timing of DNA replication, although they are not essential individually or when both are absent. Among B-type cyclins, Clb5 and Clb6 play distinct roles in regulating the initiation of S phase. Either Clb5 or Clb6 can activate

early-firing replication origins, but late origin firing is dependent only on Clb5 (Donaldson, 2000). However, neither *CLB5* nor *CLB6* is essential and even the *clb5*Δ *clb6*Δ double mutant is viable. Moreover, early expression of Clb2 during S phase does not restore normal replication timing in the absence of Clb5 and Clb6 (Donaldson, 2000). Intriguingly, the *clb5*Δ *clb6*Δ double mutant is unable to promote DNA replication in meiosis, most likely by the absence of *CLB2* expression (Mai and Breeden, 2000). Clb5 has also been shown to play an important role in mitotic spindle positioning, a function that is not replaced by overexpressing Clb2 (Segal et al., 1998). Clb5 and Clb6 has also been implicated in the establishment of sister-chromatid cohesion in cells under replication stress; cells lacking Clb5, Clb6 and the securin, Pds1 are inviable (Hsu et al., 2011).

To establish if there were novel gene functions that maintained the viability of the *clb5*Δ *clb6*Δ double mutant, we performed TMA as described above, crossing *clb5*Δ::*URA3 clb6*Δ::HPH against KAN-marked gene knockouts (**Figure 6A; Supp. Data 6 and 7**). Again, there were a number of synthetically sick interactions where there was little or no defect for either single mutant. Unlike the situation with *asf1*Δ and *cac1*Δ, there were very few instances where the synthetic lethality of *clb5*Δ *clb6*Δ was also seen in one or the other single mutant (**Figure 6B**). However, *mtc1*Δ is synthetically sick with *clb5*Δ *clb6*Δ (-16.4) and *clb5*Δ (-10.5) but not *clb6*Δ (0.5); conversely *vps53*Δ is synthetically sick with *clb5*Δ *clb6*Δ (-10.3) and *clb6*Δ (-11.8) but not at all with *clb5*Δ (1.2). Furthermore, there were only very weak correlations between the pattern of genetic interactions between *clb5*Δ *clb6*Δ and either *clb5*Δ (r = 0.26) or *clb6*Δ (r = 0.18) (**Figure 6B**). In many cases, the basis of this lethality is

not evident. Mbf1 is required for gene expression in the G1 to S transition and it is possible that some other early S phase functions are essential without Clb5 and Clb6. One striking set of genes with common genetic interactions was the HIR complex where *hir1*Δ (-14.6), *hir2*Δ (-11.8), *hir3*Δ (-12.8), and *hpc2*Δ (-14.6) all show a common strongly negative interaction in *clb5*Δ *clb6*Δ but not with either single mutant (-0.3, -0.5, -2.5, -0.5 for *clb5*Δ and -0.9, 0.3, -0.3, -0.1 for *clb6*Δ, respectively).

To characterize more generally the phenotype of *clb5*Δ *clb6*Δ mutants we compared interactions across all tested gene knockouts with the hierarchically clustered quantitative genetic interaction data from previous studies (Ryan et al., 2012). We found that *clb5*Δ *clb6*Δ most strongly correlated with mutations affecting the establishment of sister-chromatid cohesion, including point mutants in cohesin, *smc1-259*, *smc3-1*, *smc3-42* and in the cohesin loading factor Scc2 (*scc2-4*) as well as with deletions in *CTF4* and *CTF18,* (Stirling et al., 2012; Yuen et al., 2007) (**Figure 6C; Supp. Data 8**). *clb5*Δ *clb6*Δ showed especially strong negative interactions with *ctf3*Δ (-12.1), *ctf19*Δ (-13.3), *irc15*Δ (-13.9), *mcm21*Δ (-13.8) and *chl1*Δ (-15.1) (**Figure 6a**). The set of strongly negative genes is strongly reminiscent of genes that were synthetically lethal with hypomorphic mutations of three cohesion proteins (McLellan et al., 2012). We confirmed that *clb5*Δ *clb6*Δ, but not the single mutants, exhibit a chromosome loss phenotype, consistent with a defect in cohesion (**Figure 6D**).

The idea that Clb5 and Clb6 might play an important role in chromosome segregation was further strengthened by our finding that the double mutant

exhibited strong synthetic sickness with the spindle assembly checkpoint mutants *mad1*Δ (S-score = -10.7) and *mad2*Δ (S-score = -13.3), however there was little negative interaction with *mad3*Δ (S-score = -1.4). To verify these interactions, we carried out tetrad analysis, scoring spore colony size. We confirmed that both *mad1*Δ and *mad2*Δ, but not *mad3*Δ showed strong synthetic lethality (**Supp. Table 1**). In addition, tetrad analysis showed that *bub1*Δ, but not *bub2*Δ, also showed strong negative interactions. Despite the fact that Mad2 and Mad3 often appear to function together, there are other observations that suggest they have separate functions; for example the phosphorylation of Mad1 when the mitotic checkpoint is activated depends on Mad2 and Bub1 but not on Bub2 and Mad3 (Hardwick and Murray, 1995). Similarly, accurate meiotic chromosome segregation depends on Mad2 but not Mad3 (Tsuchiya et al., 2011). Moreover, a screen of synthetic genetic interactions with various chromosome transmission mutations found a number of instances where *mad1*Δ and *mad2*Δ, but not *mad3*Δ, exhibited synthetic lethality (Daniel et al., 2006). We suggest that the absence of both Clb5 and Clb6 establishes a delay in kinetochore assembly and/or spindle assembly that requires Mad1, Mad2 and Bub1 to delay mitosis, but not Mad3 or Bub2.

We summarize these findings in the scheme shown in **Figure 6E**. Especially around centromeres, DNA replication is coupled to the establishment of kinetochore attachment to microtubules and to extensive sister-chromatid cohesion in the pericentromeric region. In the absence of Clb5 and Clb6, replication is delayed and these critical steps may not be completed by the time mitosis is normally initiated. Consequently in *clb5*Δ *clb6*Δ, cells become dependent on the action of Mad1, Mad2

and Bub1 in establishing a spindle-assembly checkpoint and cannot tolerate defects in the establishment of cohesion. It is also possible that the chromosome loss phenotype is a reflection of a more general consequence of replication stress similar to the increased genome instability seen in mammalian cells with activated oncogenes (Halazonetis et al., 2008). However, *clb5*Δ *clb6*Δ does not show strong synthetic interactions with deletions either of DNA repair genes or components of the DNA damage checkpoint. Hence, we believe the effects are more directly linked to kinetochore assembly and the recruitment of cohesins in pericentromeric regions.

Perspective

Great insight into a variety of different biological processes has been extracted from genetic interaction maps in a variety of different organisms, in both budding and fission yeast (Collins et al., 2007; Fiedler et al., 2009; Roguev et al., 2008; Schuldiner et al., 2005; Wilmes et al., 2008) and bacteria (Butland et al., 2008; Typas et al., 2008), as well as more complex organisms such as *D. melanogaster* (Horn et al., 2011) and *C. elegans* (Lehner et al., 2006). More recently, quantitative genetic interaction mapping have been developed for mammalian cell lines (Bassik et al., 2013; Lin et al., 2012; Roguev et al., 2013). However, to date, essentially all of these data have been generated in a systematic, pair-wise fashion even though a deeper understanding of functional pathways could be gleaned from analyzing more than two genetic perturbations at a time. There have been genetic studies involving triple mutants in budding yeast (Tong et al., 2004; Zou et al., 2009), however these

were qualitative in nature. Here we describe a quantitative approach, termed TMA (triple mutant analysis), that allows for higher-order interactions. We interrogated two pairs of genes known to act redundantly in budding yeast: histone chaperones Asf1 and Cac1 and the cyclins, Clb5 and Clb6. This has led to several discoveries. First, Rdh54/Tid1 functionally "backs-up" Asf1 and CAF-1 and interacts much more strongly with chromatin complexes when the two major chaperones are absent. Second, loss of Asf1 suppresses the defect seen when HIR-A proteins are removed in combination with *cac1*Δ. Such a complex relationship could not have been gleaned simply from double mutant analysis. Third, we found Clb5 and Clb6 play an important role in the regulation chromosome segregation. Finally, comparison of the TMA profiles of *asf1*Δ *cac1*Δ and *rtt109*Δ *cac1*Δ revealed functional differences between Asf1 and Rtt109 that, again, could not have been uncovered using more traditional genetic analysis. We suggest this approach will be a powerful way to uncover subtle functional differences between factors that were thought to exclusively function in the same pathway.

Combined with genetic interaction mapping under different conditions (Ideker and Krogan, 2012) and analysis that allows for genetic data to be collected at sub-protein resolution (unpublished data), quantitative TMA allows for the probing into a previously unexplored interactome space. As many genes in budding yeast do not have strong genetic interaction profiles (e.g. *RDH54*), more systematic TMA will be crucial in assigning their functions. Furthermore, extending this analysis into other organisms that have tools for genetic interaction mapping, will be important to further understand various processes in these organisms and will

facilitate evolutionarily analysis of higher order genetic interactions. Finally, as platforms are developed in mammalian cells that are amenable for higher order genetic interactions using combinatorial RNAi approaches (Roguev et al., 2013), TMA will be crucial to identify pathway organization in normal conditions and in specific disease states. Ultimately, we argue work of this nature will be key for identifying targets for polypharmacy therapeutic intervention.

Methods

Detailed methods are presented in **Supplemental Information**. E-MAP analysis was conducted as described in (Schuldiner et al., 2006). TAP-tagged proteins were purified as previously described (Krogan et al., 2006).

Acknowledgements

We thank members of the Krogan and Haber groups for helpful discussion. This work was supported by grants from NIH (GM084448, GM084279, GM081879 and GM098101 to NJK; GM61766 and GM20056 to JEH and R37 GM32238 to KB). NJK is a Searle Scholar and a Keck Young Investigator.

Figures

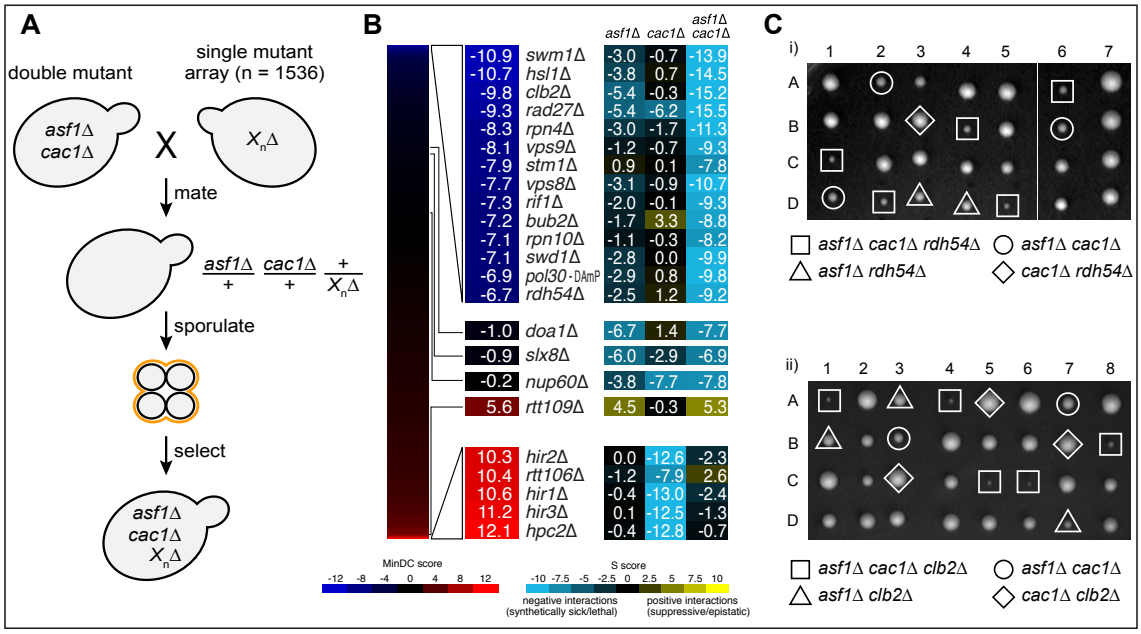


Figure 1

Figure 1. Triple mutant analysis of *asf1Δ cac1Δ*. **A**) TMA using a strain deleted for both *ASF1* and *CAC1* crossed to a library of 1536 different mutants. The mutants represent all major biological processes, with a particular emphasis on chromatin biology (Ryan et al., 2012). Following mating, the diploid cells are sporulated and the triple mutant haploid strains are selected. **B**) Double and triple mutant S-scores range from positive (yellow) to negative (blue). A minimum difference comparison was obtained by subtracting the triple mutant S-score from the S-score of the more severe of the two double mutant combinations. MinDC scores range from positive (red) to negative (dark blue). **C**) Meiotic tetrad dissection yields the triple mutants *asf1Δ cac1Δ radh54Δ* and *asf1Δ cac1Δ clb2Δ*, as well as all corresponding double mutants.

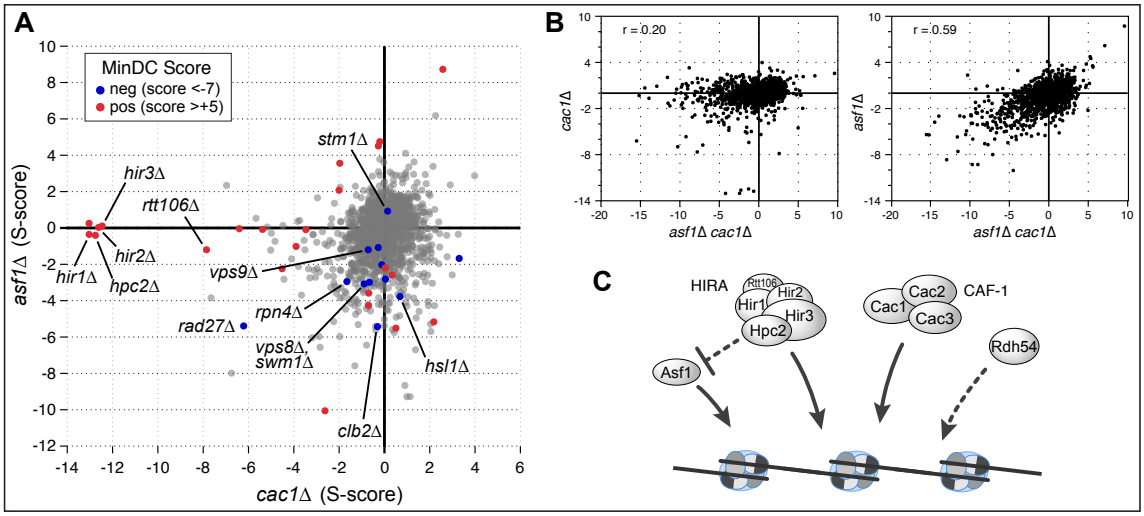


Figure 2

Figure 2. Comparison of the S- and MinDC scores from the *asf1*Δ *cac1*Δ TMA. **A**) Scatter plot of the S-scores derived from *asf1*Δ and *cac1*Δ double mutants with the corresponding MinDC scores highlighted in red (positive: >+5) and dark blue (negative: <-7). **B**) Scatter plot of S-scores from *asf1*Δ and *cac1*Δ double mutants compared to *asf1*Δ *cac1*Δ triple mutants. **C**) Model of how chromatin regulators, Asf1, HIR-C, CAF-1 and Rdh54, functionally interact to ensure efficient chromatin regulation (see text for details).

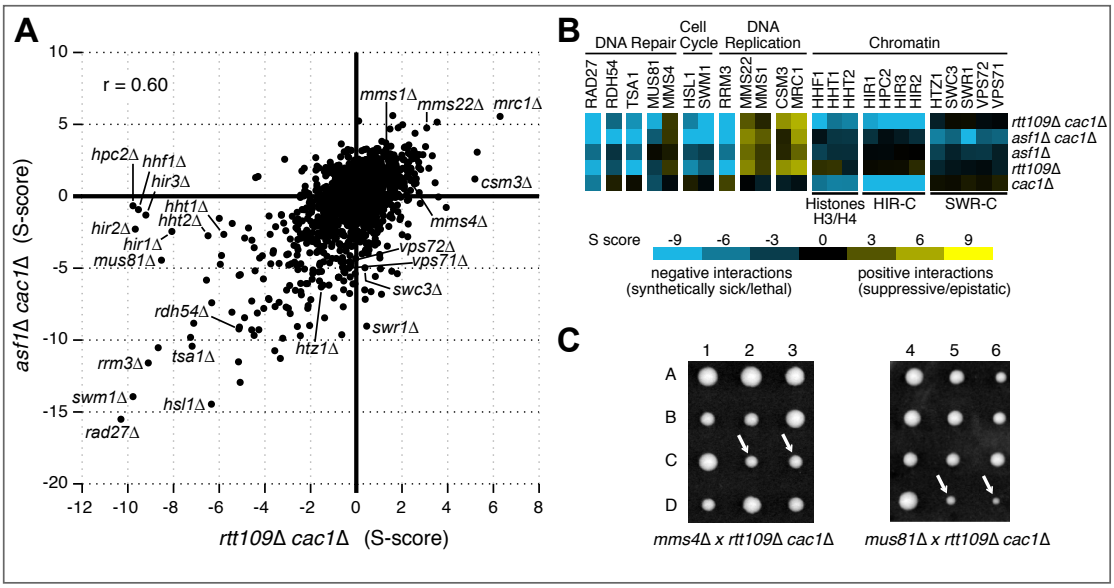


Figure 3

Figure 3. Comparison of the S-scores from *asf1*Δ *cac1*Δ and *rtt109*Δ *cac1*Δ. A) Scatter plot of the triple mutant S-scores from *asf1*Δ *cac1*Δ versus *rtt109*Δ *cac1*Δ double mutants. **B**) A selection of genetic interactions (S-scores) derived from the triple and double mutants analyses. Yellow and blue correspond to positive and negative genetic interactions, respectively. Note the *cac1*Δ double mutant scores were averaged from data obtained from both *asf1*Δ *cac1*Δ and *rtt109*Δ *cac1*Δ starter strains. **C**) Tetrad analysis shows a difference in the viability of *mms4*Δ *rtt109*Δ *cac1*Δ and *mus81*Δ *rtt109*Δ *cac1*Δ segregants, marked by white arrows.

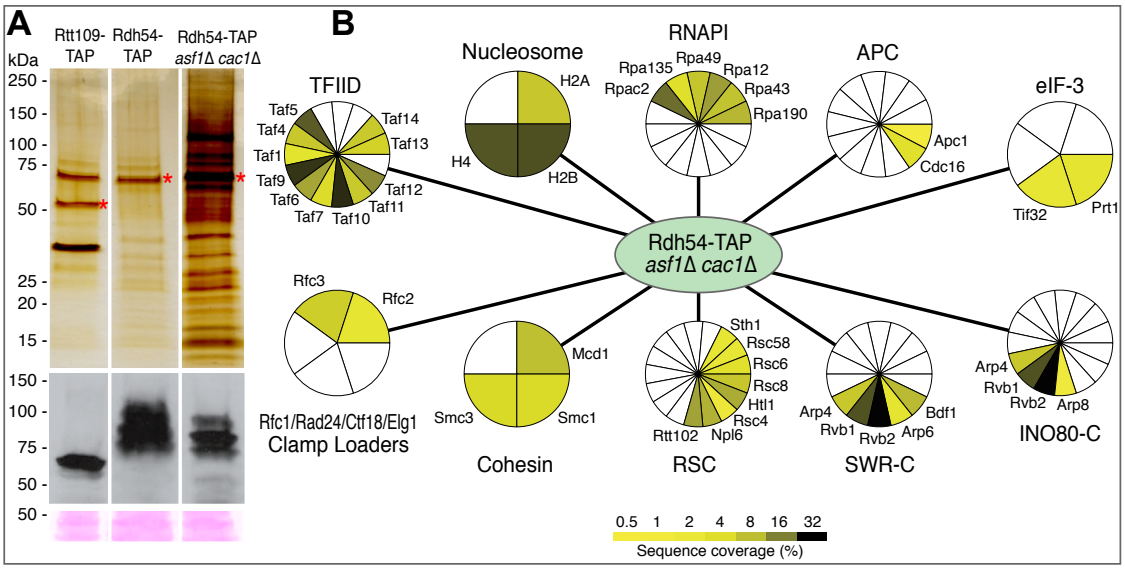


Figure 4

Figure 4. Affinity tag purifications of Rdh54-TAP and Rad54-TAP *asf1*Δ *cac1*Δ.

A) Rad54-TAP was purified from strains in the presence and absence of Asf1 and Cac1. The purified material was subjected to SDS-PAGE and stained with silver (Top). As a control, Rtt109-TAP was purified in the same manner. Red asterisks mark the tagged proteins. (Bottom) Western blot analysis was carried out using extracts prior to purification. Bands from a Ponceau stain serve as a loading control. **B)** Network of interactions derived from Rdh54-TAP *asf1*Δ *cac1*Δ using affinity tag purification-mass spectrometry. Proteins found to interact with either Rdh54-TAP or Rtt109-TAP were removed. Complexes or pathways are labeled and the sequence coverage (%) from the mass spectrometry analysis is shown using a yellow/green color scheme.

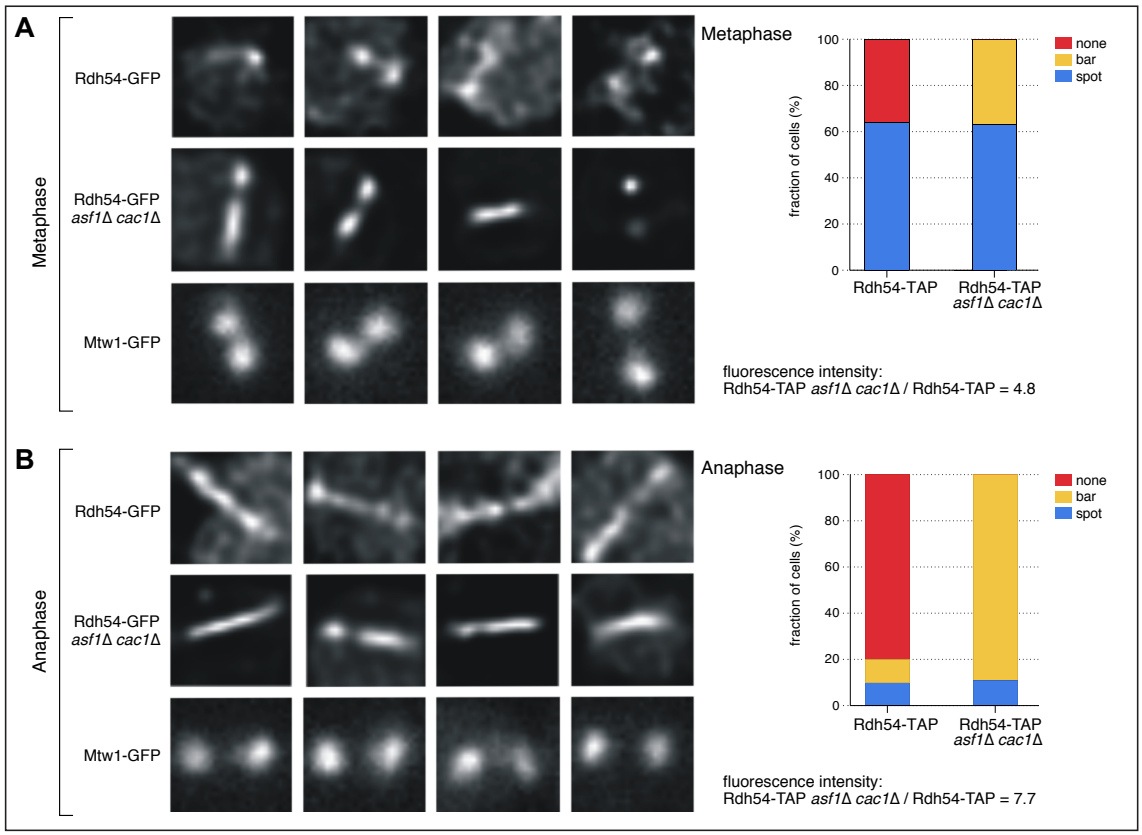


Figure 5

Figure 5. Rdh54-GFP is redistributed to pericentric chromatin in *asf1Δ cac1Δ* mutants during mitosis. Representative images of Rdh54-GFP (C-terminal fusion) in wild-type (WT) and *asf1*Δ *cac1*Δ cells. **A)** Metaphase: Rdh54-GFP localization is compared to a core component of the yeast kinetochore (Mtw1-GFP). The concentration of Rdh54-GFP along the spindle is ~ 5-fold greater in the mutant vs. WT. **B)** Anaphase: In the absence of Asf1 and Cac1, Rdh54 appears as a bar along the metaphase spindle in nearly 90% of cells. The concentration of Rdh54-GFP along the spindle is approximately 7.7-fold greater in the mutant vs. WT. Foci are defined as diffraction-limited (or slightly larger) fluorescent spots. For the kinetochore protein Mtw1 (bottom row) spots are approximately spherical, with sister kinetochores exhibiting similar shape and intensity (Haase et al., 2012). Bars are

defined as linear extensions of fluorescence (ex. bars in 1st and 3rd panel of Rdh54-GFP *asf1*Δ *cac1*Δ Metaphase). In instances where a bar and a foci are observed (ex. 2nd panel Rdh54-GFP, *asf1*Δ *cac1*Δ Anaphase) the cell is scored as containing a bar.

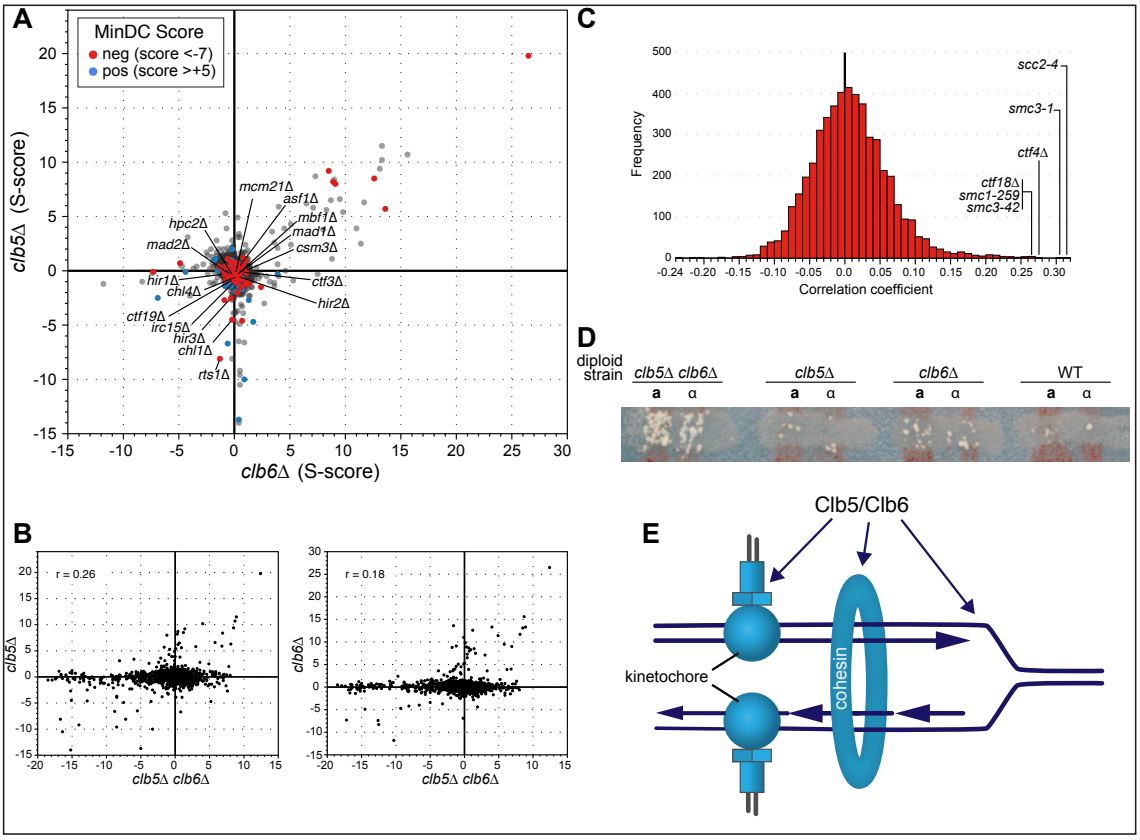


Figure 6

Figure 6. Triple mutant analysis of *clb5*Δ *clb6*Δ. **A**) Scatter plot of the S-scores derived from *clb5*Δ and *clb5*Δ double mutants with the corresponding MinDC scores highlighted in dark blue (positive: >+5) and red (negative: <-7). **B**) Scatter plot of S-scores from *clb5*Δ or *clb6*Δ double mutants compared to *clb5*Δ *clb6*Δ triple mutants. **C**) Comparison of genetic profile generated from *clb5*Δ *clb6*Δ to profiles from large genetic interaction dataset (Data Set S4 from (Ryan et al., 2012)). The most highly

correlated profiles belong to strains harboring mutations in *SCC2*, *SMC3*, *CTF4*, *CTF18* and *SMC1*. **D**) Diploids homozygous for the indicated genotype and for *ade2* were cross-streaked with haploid *MAT***a** and *MAT*α *ade5* cells. Chromosome loss (or mitotic crossing-over) was scored as papillae that grow when the matings are replica-plated to minimal medium (see **Methods**). **E**) Model of how Clb5 and Clb6 function to ensure efficient chromosome segregation (see text for details).

Supplementary Figures

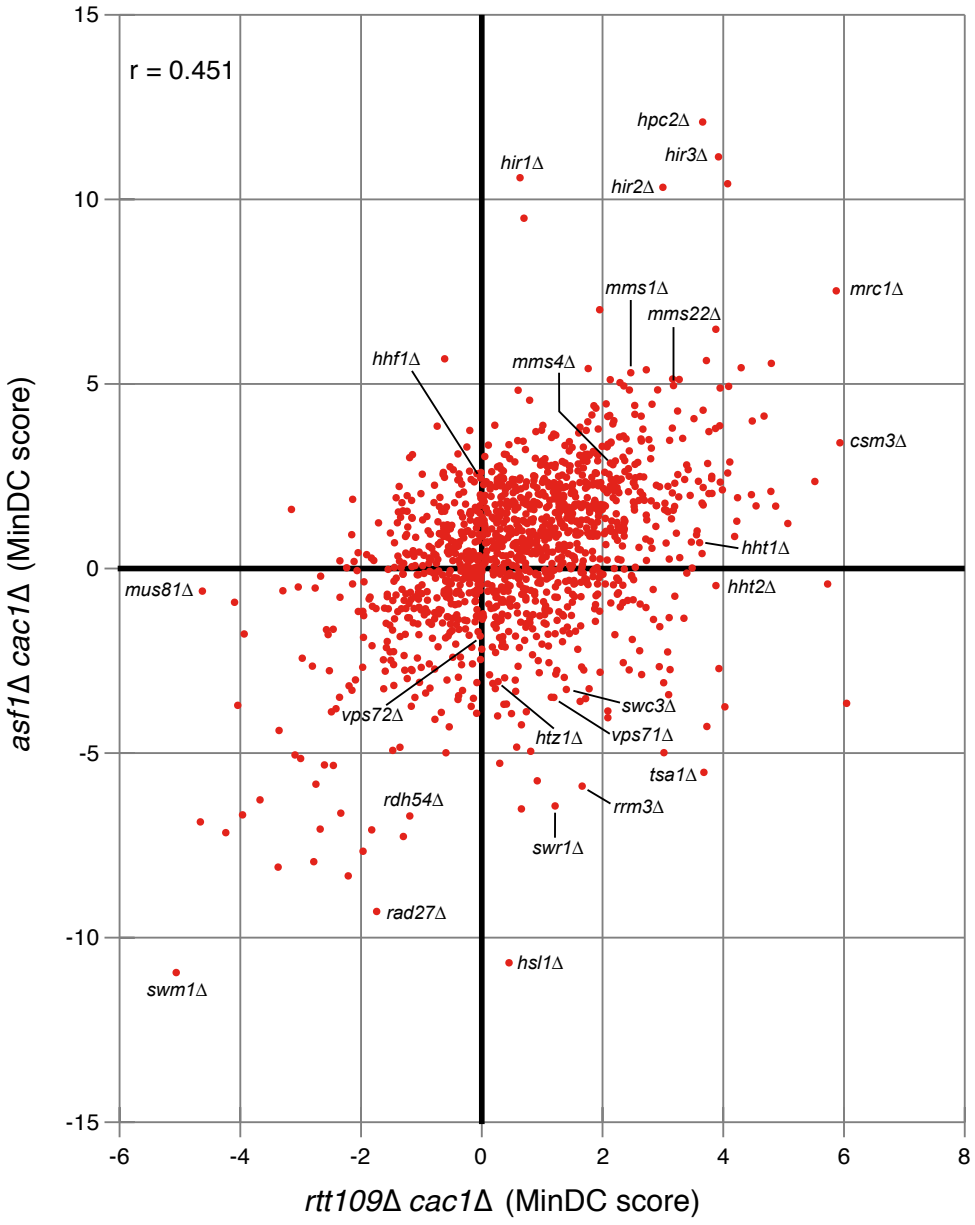


Figure S1. Scatter plot of the MinDC scores derived from *asf1Δ cac1Δ* and *rtt109Δ cac1Δ* double mutants. Related to Figure 3. The genes labeled are the same ones labeled in the S-score plot in Figure 3A.

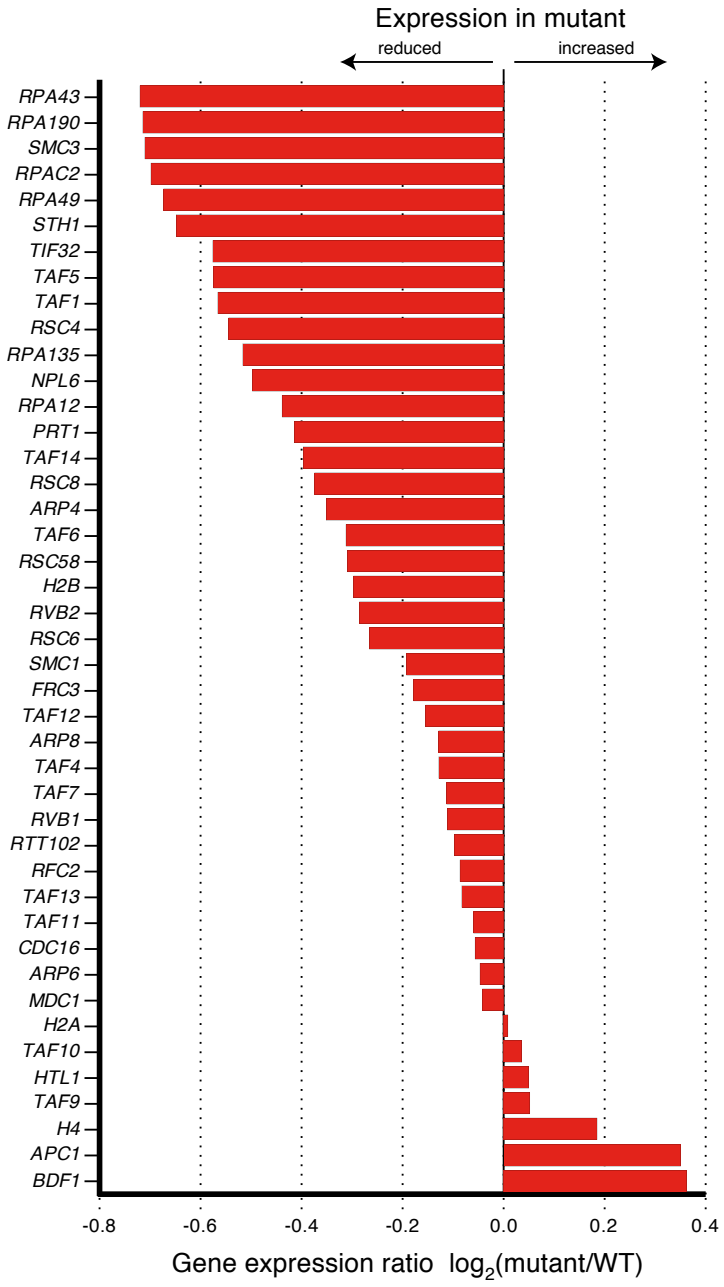


Figure S2. In *asf1Δ cac1Δ* cells, the change in expression of the genes that code for proteins more physically associated with Rdh54 after removal of Asf1 and Cac1. Related to Figure 4. The entire gene expression dataset is contained within Supp. Data 5.

Supplemental Tables and Data

Available online at:

http://www.sciencedirect.com/science/article/pii/S2211124713002246

Extended Experimental Procedures

E-MAP analysis

Single and double mutant deletion strains amenable for E-MAP analysis were generated by insertion of *URA3* (for *ASF1* and *CLB5*) and/or HPH^r (for *CAC1*, *RTT109* and *CLB6*) cassettes into the relevant loci of an E-MAP amenable *MAT*α starter strain described in (Schuldiner et al., 2006). High-throughput generation of double/triple mutant strains was then carried out as described in section 4 of (Schuldiner et al., 2006) with the modifications below. For strains marked with HPH^r cassettes, HPH (200mg/l) was substituted for NAT in all steps. For strains marked with URA3 cassettes, the following media was used for diploid selection: Agar (20g/l) + dextrose (20g/l) + YNB w/o amino acids and w/o ammonium sulfate (BD 233520) (1.7g/l) + URA "drop out mix" (CSM-URA Sunrise Science 1004-100) (2g/l) + monosodium glutamic acid (Sigma G1626_500G) (1g/l) + G418 (100mg/l) and + HPH (200mg/l) if HPH^r cassette. In addition, the following changes were carried out in the last two selection steps ("SM" and "DM" in (Schuldiner et al., 2006)): Library mutations were not selected for in "SM", and G418 was thus left out from this step. Instead, the query mutation(s) were selected for in "SM", by adding HPH (200mg/l) (for HPH^r cassette), excluding URA from the "DROP OUT" mix (for

URA3 cassette), or both (for HPH^r and URA3 cassettes). Final selection media was same as "SM" above but with G418 (100mg/l) added.

Tetrad analysis

Isogenic strains used in the E-MAP analysis have been previously described (Collins et al., 2007). A *MAT***a** strain carrying *asf1*Δ::*URA3 cac1*Δ::HPH or *clb5*Δ::*URA3 clb6*Δ::HPH were crossed to several *MAT*α strains carrying different gene deletions marked by KAN. Diploids were sporulated and tetrads microdissected on YEPD plates. Colony size was scored visually on a scale from 0 (inviable) to 4. Segregants were picked onto a YEPD plate and replica-plated on various nutritional or antibiotic containing plates to score for the various gene deletions. The mean colony size for different genotypes was calculated.

Minimum Difference Comparison (MinDC)

To identify the instances in which the S-scores for a triple mutant were distinctively different from either double mutant (i.e. when *asf1*Δ *cac1*Δ *gene x*Δ was different from *asf1*Δ gene *x*Δ or *cac1*Δ *x*Δ) a minimum difference comparison was performed. From the S-score of the triple mutant was subtracted the most negative of the two S-scores of the two relevant double mutants.

Protein purification and mass spectrometry analysis

TAP-tagged proteins were purified as previously described (Krogan et al., 2006). Purified proteins eluates were digested with trypsin for LC-MS/MS analysis.

Samples were denatured and reduced in 2M urea, 10 mM NH4HCO3, 2 mM DTT for 30 min at 60C, then alkylated with 2 mM iodoacetamide for 45 min at room temperature. Trypsin (Promega) was added at a 1:100 enzyme:substrate ratio and digested overnight at 37C. Following digestion, samples were concentrated using C18 ZipTips (Millipore) according to the manufacturer's specifications. Digested peptide mixtures were analyzed by LC-MS/MS on a Thermo Scientific Velos Pro ion trap mass spectrometry system equipped with a Proxeon Easy nLC 1000 ultra high pressure liquid chromatography and autosampler system. Sample were injected onto a pre-column (2 cm x 100 um I.D. packed with ReproSil Pur C18 AQ 5um particles) in 0.1% formic acid and then separated with a two-hour gradient from 5% to 30% ACN in 0.1% formic acid on an analytical column (10 cm x 75 um I.D. packed with ReproSil Pur C18 AQ 3 um particles). The mass spectrometer collected data in a data-dependent fashion, collecting one full scan followed by 20 collision-induced dissociation MS/MS scans of the 20 most intense peaks from the full scan. Dynamic exclusion was enabled for 30 seconds with a repeat count of 1. Each sample was analyzed in technical duplicate. Blanks were run between different samples to reduce carryover effects. The results raw data was matched to protein sequences by the Protein Prospector algorithm. Data were searched against the SwissProt Human protein sequence database (downloaded March 6, 2012) concatenated to a decoy database where each sequences was randomized in order to estimate the false positive rate. The searches considered a precursor mass tolerance of 1 Da and fragment ion tolerances of 0.8 Da, and considered variable modifications for protein N-terminal acetylation, protein N-terminal acetylation and oxidation, glutamine to

pyroglutamate conversion for peptide N-terminal glutamine residues, protein N-terminal methionine loss, protein N-terminal acetylation and methionine loss, and methionine oxidation, and constant modification for carbamidomethyl cysteine. Prospector data was filtered using a maximum protein expectation value of 0.01 and a maximum peptide expectation value of 0.05.

Image Deconvolution

Deconvolution was performed with Huygens Essential software (Huygens compute engine 4.2.1p4 32b Huygens Essential for Win32 Built on Jun 19, 2012 13:00:09, Scientific Volume Imaging) using the classic maximum-likelihood estimation and predicted point-spread function.

Chromosome stability assay

A chromosome loss assay was carried out as described previously (Haber, 1974). A patch of non-mating diploids homozygous for *ade2* were mated with *MAT***a** and *MAT*α tester strains carrying a different nutritional requirement, *ade5*, grown overnight on rich medium (YEPD) and replica-plated to minimal medium so that only matings between diploids that had lost chromosome 3 carrying *MAT****a*** or *MAT*α can mate. Chromosome loss of mitotic recombination is measured by the frequency of wild-type papillae that appear with both mating testers.

Microarray analysis

Gene expression analysis was by two channel DNA microarray analysis with all procedures including statistical analyses as described previously in detail (Lenstra et al., 2011). Each strain was profiled four times from two biological replicates in balanced dye-swap replicate using a single batch of wild type RNA as common reference. Average fold-changes and p-values were calculated by Limma using a large set of wild type strains analyzed in parallel.

References

Bassik, M.C., Kampmann, M., Lebbink, R.J., Wang, S., Hein, M.Y., Poser, I., Weibezahn, J., Horlbeck, M.A., Chen, S., Mann, M.*, et al.* (2013). A systematic Mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell *152*, 909-922.

Beltrao, P., Cagney, G., and Krogan, N.J. (2010). Quantitative genetic interactions reveal biological modularity. Cell *141*, 739-745.

Booher, R.N., Deshaies, R.J., and Kirschner, M.W. (1993). Properties of Saccharomyces cerevisiae wee1 and its differential regulation of p34CDC28 in response to G1 and G2 cyclins. EMBO J *12*, 3417-3426.

Butland, G., Babu, M., Diaz-Mejia, J.J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A.G., Pogoutse, O.*, et al.* (2008). eSGA: E. coli synthetic genetic array analysis. Nat Methods *5*, 789-795.

Chi, P., Kwon, Y., Visnapuu, M.L., Lam, I., Santa Maria, S.R., Zheng, X., Epshtein, A., Greene, E.C., Sung, P., and Klein, H.L. (2011). Analyses of the yeast Rad51 recombinase A265V mutant reveal different in vivo roles of Swi2-like factors. Nucleic Acids Res *39*, 6511-6522.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S.R., Roguev, A., and Krogan, N.J. (2010). Quantitative genetic interaction mapping using the E-MAP approach. Methods Enzymol *470*, 205-231.

Daniel, J.A., Keyes, B.E., Ng, Y.P., Freeman, C.O., and Burke, D.J. (2006). Diverse functions of spindle assembly checkpoint genes in Saccharomyces cerevisiae. Genetics *172*, 53-65.

De Koning, L., Corpet, A., Haber, J.E., and Almouzni, G. (2007). Histone chaperones: an escort network regulating histone traffic. Nat Struct Mol Biol *14*, 997-1007.

de los Santos, T., Hunter, N., Lee, C., Larkin, B., Loidl, J., and Hollingsworth, N.M. (2003). The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. Genetics *164*, 81-94.

Decesare, J.M., and Stuart, D.T. (2012). Among B-Type Cyclins Only CLB5 and CLB6 Promote Premeiotic S Phase in Saccharomyces cerevisiae. Genetics *190*, 1001-1016.

Donaldson, A.D. (2000). The yeast mitotic cyclin Clb2 cannot substitute for S phase cyclins in replication origin firing. EMBO Rep *1*, 507-512.

Donaldson, A.D., Raghuraman, M.K., Friedman, K.L., Cross, F.R., Brewer, B.J., and Fangman, W.L. (1998). CLB5-dependent activation of late replication origins in S. cerevisiae. Mol Cell *2*, 173-182.

Dresser, M.E., Ewing, D.J., Conrad, M.N., Dominguez, A.M., Barstead, R., Jiang, H., and Kodadek, T. (1997). *DMC1* functions in a *Saccharomyces cerevisiae* meiotic pathway that is largely independent of the *RAD51* pathway. Genetics *147*, 533-544.

Driscoll, R., Hudson, A., and Jackson, S.P. (2007). Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science *315*, 649-652.

Enomoto, S., and Berman, J. (1998). Chromatin assembly factor I contributes to the maintenance, but not the re-establishment, of silencing at the yeast silent mating loci. Genes Dev *12*, 219-232.

Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S.*, et al.* (2009). Functional organization of the S. cerevisiae phosphorylation network. Cell *136*, 952-963.

Green, E.M., Antczak, A.J., Bailey, A.O., Franco, A.A., Wu, K.J., Yates, J.R., 3rd, and Kaufman, P.D. (2005). Replication-independent histone deposition by the HIR complex and Asf1. Curr Biol *15*, 2044-2049.

Haase, J., Stephens, A., Verdaasdonk, J., Yeh, E., and Bloom, K. (2012). Bub1 kinase and Sgo1 modulate pericentric chromatin in response to altered microtubule dynamics. Curr Biol *22*, 471-481.

Haber, J.E. (1974). Bisexual mating behavior in a diploid of Saccharomyces cerevisiae: evidence for genetically controlled non-random chromosome loss during vegetative growth. Genetics *78*, 843-858.

Halazonetis, T.D., Gorgoulis, V.G., and Bartek, J. (2008). An oncogene-induced DNA damage model for cancer development. Science *319*, 1352-1355.

Hall, M.C., Torres, M.P., Schroeder, G.K., and Borchers, C.H. (2003). Mnd2 and Swm1 are core subunits of the Saccharomyces cerevisiae anaphase-promoting complex. J Biol Chem *278*, 16698-16705.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.M., and Zhang, Z. (2007a). Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science *315*, 653-655.

Han, J., Zhou, H., Li, Z., Xu, R.M., and Zhang, Z. (2007b). Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity. J Biol Chem *282*, 28587-28596.

Hardwick, K.G., and Murray, A.W. (1995). Mad1p, a phosphoprotein component of the spindle assembly checkpoint in budding yeast. J Cell Biol *131*, 709-720.

Ho, C.K., Mazon, G., Lam, A.F., and Symington, L.S. (2010). Mus81 and Yen1 promote reciprocal exchange during mitotic recombination to maintain genome integrity in budding yeast. Mol Cell *40*, 988-1000.

Horn, T., Sandmann, T., Fischer, B., Axelsson, E., Huber, W., and Boutros, M. (2011). Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. Nat Methods *8*, 341-346.

Hsu, W.S., Erickson, S.L., Tsai, H.J., Andrews, C.A., Vas, A.C., and Clarke, D.J. (2011). S-phase cyclin-dependent kinases promote sister chromatid cohesion in budding yeast. Mol Cell Biol *31*, 2470-2483.

Ideker, T., and Krogan, N.J. (2012). Differential network biology. Mol Syst Biol *8*, 565.

Inoue, Y., Matsuda, T., Sugiyama, K., Izawa, S., and Kimura, A. (1999). Genetic analysis of glutathione peroxidase in oxidative stress response of Saccharomyces cerevisiae. J Biol Chem *274*, 27002-27009.

Ivessa, A.S., Zhou, J.Q., Schulz, V.P., Monson, E.K., and Zakian, V.A. (2002). Saccharomyces Rrm3p, a 5' to 3' DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. Genes Dev *16*, 1383-1396.

Kats, E.S., Albuquerque, C.P., Zhou, H., and Kolodner, R.D. (2006). Checkpoint functions are required for normal S-phase progression in Saccharomyces cerevisiae RCAF- and CAF-I-defective mutants. Proc Natl Acad Sci U S A *103*, 3710-3715.

Kim, J.A., and Haber, J.E. (2009). Chromatin assembly factors Asf1 and CAF-1 have overlapping roles in deactivating the DNA damage checkpoint when DNA repair is complete. Proc Natl Acad Sci U S A *106*, 1151-1156.

Klein, H.L. (1997). *RDH54*, a *RAD54* homologue in *Saccharomyces cerevisiae*, is required for mitotic diploid-specific recombination and repair and for meiosis. Genetics *147*, 1533-1543.

Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D., Gin, J.W., Jennings, J.L., Link, A.J., Madhani, H.D., and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol *2*, E131.

Krawitz, D.C., Kama, T., and Kaufman, P.D. (2002). Chromatin assembly factor I mutants defective for PCNA binding require Asf1/Hir proteins for silencing. Mol Cell Biol *22*, 614-625.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P.*, et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V.*, et al.* (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell *12*, 1565-1576.

Kwon, Y., Seong, C., Chi, P., Greene, E.C., Klein, H., and Sung, P. (2008). ATP-dependent chromatin remodeling by the Saccharomyces cerevisiae homologous recombination factor Rdh54. J Biol Chem *283*, 10445-10452.

Lee, S.E., Pellicioli, A., Malkova, A., Foiani, M., and Haber, J.E. (2001). The Saccharomyces recombination protein Tid1p is required for adaptation from G2/M arrest induced by a double-strand break. Curr Biol *11*, 1053-1057.

Lehner, B., Crombie, C., Tischler, J., Fortunato, A., and Fraser, A.G. (2006). Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. Nat Genet *38*, 896-903.

Lenstra, T.L., Benschop, J.J., Kim, T., Schulze, J.M., Brabers, N.A., Margaritis, T., van de Pasch, L.A., van Heesch, S.A., Brok, M.O., Groot Koerkamp, M.J.*, et al.* (2011). The specificity and topology of chromatin interaction pathways in yeast. Mol Cell *42*, 536-549.

Lin, Y.Y., Kiihl, S., Suhail, Y., Liu, S.Y., Chou, Y.H., Kuang, Z., Lu, J.Y., Khor, C.N., Lin, C.L., Bader, J.S.*, et al.* (2012). Functional dissection of lysine deacetylases reveals that HDAC1 and p300 regulate AMPK. Nature *482*, 251-255.

Ma, X.J., Lu, Q., and Grunstein, M. (1996). A search for proteins that interact genetically with histone H3 and H4 amino termini uncovers novel regulators of the Swe1 kinase in Saccharomyces cerevisiae. Genes Dev *10*, 1327-1340.

Mai, B., and Breeden, L. (2000). CLN1 and its repression by Xbp1 are important for efficient sporulation in budding yeast. Mol Cell Biol *20*, 478-487.

McLellan, J.L., O'Neil, N.J., Barrett, I., Ferree, E., van Pel, D.M., Ushey, K., Sipahimalani, P., Bryan, J., Rose, A.M., and Hieter, P. (2012). Synthetic Lethality of Cohesins with PARPs and Replication Fork Mediators. PLoS Genetics *8*, e1002574.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Mohanty, B.K., Bairwa, N.K., and Bastia, D. (2006). The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *103*, 897-902.

Myung, K., Pennaneach, V., Kats, E.S., and Kolodner, R.D. (2003). Saccharomyces cerevisiae chromatin-assembly factors that act during DNA replication function in the maintenance of genome stability. Proc Natl Acad Sci U S A *100*, 6640-6645.

Prasad, T.K., Robertson, R.B., Visnapuu, M.L., Chi, P., Sung, P., and Greene, E.C. (2007). A DNA-translocating Snf2 molecular motor: Saccharomyces cerevisiae Rdh54 displays processive translocation and extrudes DNA loops. J Mol Biol *369*, 940-953.

Ramey, C.J., Howar, S., Adkins, M., Linger, J., Spicer, J., and Tyler, J.K. (2004). Activation of the DNA damage checkpoint in yeast lacking the histone chaperone anti-silencing function 1. Mol Cell Biol *24*, 10313-10327.

Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Mol Cell *9*, 1091-1100.

Reagan, M.S., Pittenger, C., Siede, W., and Friedberg, E.C. (1995). Characterization of a mutant strain of Saccharomyces cerevisiae with a deletion of the RAD27 gene, a structural homolog of the RAD2 nucleotide excision repair gene. J Bacteriol *177*, 364-371.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J.*, et al.* (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science *322*, 405-410.

Roguev, A., Talbot, D., Negri, G.L., Shales, M., Cagney, G., Bandyopadhyay, S., Panning, B., and Krogan, N.J. (2013). Quantitative genetic-interaction mapping in mammalian cells. Nat Methods *advance online publication 2013/02/15*.

Roguev, A., Wiren, M., Weissman, J.S., and Krogan, N.J. (2007). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat Methods *4*, 861-866.

Ryan, C.J., Roguev, A., Patrick, K., Xu, J., Jahari, H., Tong, Z., Beltrao, P., Shales, M., Qu, H., Collins, S.R., *et al.* (2012). Hierarchical Modularity and the Evolution of Genetic Interactomes across Species. Mol Cell *46*, 691-704.

Schneider, B.L., Yang, Q.H., and Futcher, A.B. (1996). Linkage of replication to start by the Cdk inhibitor Sic1. Science *272*, 560-562.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F., *et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Schuldiner, M., Collins, S.R., Weissman, J.S., and Krogan, N.J. (2006). Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344-352.

Segal, M., Clarke, D.J., and Reed, S.I. (1998). Clb5-associated kinase activity is required early in the spindle pathway for correct preanaphase nuclear positioning in Saccharomyces cerevisiae. J Cell Biol *143*, 135-145.

Sharp, J.A., Franco, A.A., Osley, M.A., and Kaufman, P.D. (2002). Chromatin assembly factor I and Hir proteins contribute to building functional kinetochores in S. cerevisiae. Genes Dev *16*, 85-100.

Shin, Y.K., Amangyeld, T., Nguyen, T.A., Munashingha, P.R., and Seo, Y.S. (2012). Human MUS81 complexes stimulate flap endonuclease 1. FEBS J *279*, 2412-2430.

Sia, R.A., Herald, H.A., and Lew, D.J. (1996). Cdc28 tyrosine phosphorylation and the morphogenesis checkpoint in budding yeast. Mol Biol Cell *7*, 1657-1666.

Silva, A.C., Xu, X., Kim, H.S., Fillingham, J., Kislinger, T., Mennella, T.A., and Keogh, M.C. (2012). The replication-independent histone H3-H4 chaperones HIR, ASF1, and RTT106 co-operate to maintain promoter fidelity. J Biol Chem *287*, 1709-1718.

Simpson-Lavy, K.J., Sajman, J., Zenvirth, D., and Brandeis, M. (2009). APC/CCdh1 specific degradation of Hsl1 and Clb2 is required for proper stress responses of S. cerevisiae. Cell Cycle *8*, 3003-3009.

Stephens, A.D., Haase, J., Vicci, L., Taylor, R.M., 2nd, and Bloom, K. (2011). Cohesin, condensin, and the intramolecular centromere loop together generate the mitotic chromatin spring. J Cell Biol *193*, 1167-1180.

Stirling, P.C., Crisp, M.J., Basrai, M.A., Tucker, C.M., Dunham, M.J., Spencer, F.A., and Hieter, P. (2012). Mutability and mutational spectrum of chromosome transmission fidelity genes. Chromosoma *121*, 263-275.

Surana, U., Robitsch, H., Price, C., Schuster, T., Fitch, I., Futcher, A.B., and Nasmyth, K. (1991). The role of CDC28 and cyclins during mitosis in the budding yeast S. cerevisiae. Cell *65*, 145-161.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M.*, et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Tsuchiya, D., Gonzalez, C., and Lacefield, S. (2011). The spindle checkpoint protein Mad2 regulates APC/C activity during prometaphase and metaphase of meiosis I in Saccharomyces cerevisiae. Mol Biol Cell *22*, 2848-2861.

Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L.*, et al.* (2008). High-throughput, quantitative analyses of genetic interactions in E. coli. Nat Methods *5*, 781-787.

Vaisica, J.A., Baryshnikova, A., Costanzo, M., Boone, C., and Brown, G.W. (2011). Mms1 and Mms22 stabilize the replisome during replication stress. Mol Biol Cell *22*, 2396-2408.

Verdaasdonk, J.S., Gardner, R., Stephens, A.D., Yeh, E., and Bloom, K. (2012). Tension-dependent nucleosome remodeling at the pericentromere in yeast. Mol Biol Cell *23*, 2560-2570.

Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S*., et al.* (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell *32*, 735-746.

Xie, Y., and Varshavsky, A. (2001). RPN4 is a ligand, substrate, and transcriptional regulator of the 26S proteasome: a negative feedback circuit. Proc Natl Acad Sci U S A *98*, 3056-3061.

Xu, H., Boone, C., and Brown, G.W. (2007). Genetic dissection of parallel sister-chromatid cohesion pathways. Genetics *176*, 1417-1429.

Yuen, K.W., Warren, C.D., Chen, O., Kwok, T., Hieter, P., and Spencer, F.A. (2007). Systematic genome instability screens in yeast and their potential relevance to cancer. Proc Natl Acad Sci U S A *104*, 3925-3930.

Zou, J., Friesen, H., Larson, J., Huang, D., Cox, M., Tatchell, K., and Andrews, B. (2009). Regulation of cell polarity through phosphorylation of Bni4 by Pho85 G1 cyclin-dependent kinases in Saccharomyces cerevisiae. Mol Biol Cell *20*, 3239-3250.

Chapter 8

Individual Lysine Acetylations on the N Terminus of *Saccharomyces cerevisiae* H2A.Z are Highly but Not Differentially Regulated

Individual Lysine Acetylations on the N Terminus of *Saccharomyces cerevisiae* H2A.Z are Highly but Not Differentially Regulated

Monika Mehta[‡], Hannes Braberg[§¶], Shuyi Wang[§¶], Anita Lozsa[‖], Michael Shales[§¶], Alejandra Solache[‖], Nevan J. Krogan[§¶], Michael-Christopher Keogh[‡,1]

From the [‡]Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461,

the [§]Department of Cellular and Molecular Pharmacology and

the [¶]California Institute for Quantitative Biomedical Research, University of California, San Francisco, California 94158, and

the [‖]Millipore Corporation, Temecula, California 92590

[1] To whom correspondence should be addressed: Chanin 415A, Dept. of Cell Biology, 1300 Morris Park Ave., Bronx, NY 10461., Tel.: Phone: 718-430-8796; Fax: 718-430-8574; E-mail: michael.keogh@einstein.yu.edu

Abstract

The multi-functional histone variant Htz1 (*Saccharomyces cerevisiae* H2A.Z) is acetylated on up to four N-terminal lysines at positions 3, 8, 10, and 14. It has thus been posited that specific acetylated forms of the histone could regulate distinct roles. Antibodies against Htz1-K8^{Ac}, -K10^{Ac}, and -K14^{Ac} show that all three modifications are added by Esa1 acetyltransferase and removed by Hda1 deacetylase. Completely unacetylatable *htz1* alleles exhibit widespread interactions in genome scale genetic screening. However, singly mutated (*e.g. htz1-K8R*) or singly acetylable (*e.g.* the triple mutant *htz1-K3R/K10R/K14R*) alleles show no significant defects in these analyses. This suggests that the N-terminal acetylations on Htz1 are internally redundant. Further supporting this proposal, each acetylation decays with similar kinetics when Htz1 transcription is repressed, and proteomic screening did not find a single condition in which one Htz1^{Ac} was differentially regulated. However, whereas the individual acetylations on Htz1 may be redundant, they are not dispensable. Completely unacetylatable *htz1* alleles display genetic interactions and phenotypes in common with and distinct from *htz1*Δ. In addition, each Htz1 N-terminal lysine is deacetylated by Hda1 in response to benomyl and reacetylated when this agent is removed. Such active regulation suggests that acetylation plays a significant role in Htz1 function.

Introduction

Chromatin is the highly ordered assembly that packages eukaryotic DNA and regulates nuclear processes such as replication, repair, and transcription. The basic

repeating unit of chromatin is the nucleosome core particle: ~146 bp of DNA wrapped around a core histone octamer composed of a $(H3\text{-}H4)_2$ tetramer and two (H2A-H2B) dimers (Luger, 2003). The major histones (H2A, H2B, H3, and H4) are each encoded by multi-copy genes, highly expressed during the S phase, and deposited in chromatin during DNA replication. Histone variants are nonallelic isoforms that are usually (although not exclusively (Dryhurst et al., 2009)) encoded by single copy genes. In many cases variants can substitute for the major histones in specific nucleosomes through the action of dedicated deposition machineries (Henikoff et al., 2004). Both major and variant histones are subject to extensive post-translational modification by the addition of small chemical moieties, including phosphorylation, acetylation, methylation, sumoylation, and ubiquitylation (Latham and Dent, 2007). These groups are thought to regulate access to the DNA in the modified nucleosome by various means, including directly modulating the charge on the nucleosome surface or generating sites for the recruitment of regulatory proteins (Taverna et al., 2007; Caterino and Hayes, 2007).

Histone H2A has one of the largest variant families and includes H2A.Z, a protein that is highly conserved across eukaryotes but differs considerably from the major H2A in each species (Mehta et al., 2010). H2A.Z has been ascribed a large number of roles (Raisner and Madhani, 2006; Zlatanova and Thakar, 2008), including most recently suppressing antisense RNAs (Zofall et al., 2009) and stabilizing the association of condensin with mitotic chromosomes (Kim et al., 2009). We still have a poor understanding of precisely how the variant mediates any specific function, although differential enrichment at specific locations and distinct

post-translational modifications are likely to contribute. In all species examined, their respective H2A.Zs are subject to multiple N-terminal tail acetylations (Thambirajah et al., 2009; Ishibashi et al., 2009). This modification is integral to fission yeast H2A.Z function, with completely unacetylatable alleles phenocopying complete deletion of the histone in genome scale transcriptome and genetic analyses (Kim et al., 2009). Multiple regulatory roles have been assigned to acetylated H2A.Z in *Saccharomyces cerevisiae* (*Sc*),[2] including heterochromatin restriction (Babiarz et al., 2006), transcription (Millar et al., 2006), and chromosome stability (Keogh et al., 2006a). However, many of these ascribed functions were derived from completely unacetylatable alleles, so the importance of any individual acetylation is unclear. Finally, it is also unknown whether the effects of any of these modifications are direct (*e.g.* steric hindrance or charge modulation) or indirect (*e.g.* via the recruitment of regulatory proteins).

In this work we examine the regulation and function of four individual acetylations on the *Sc* H2A.Z (gene name *HTZ1*) N terminus: $K3^{Ac}$, $K8^{Ac}$, $K10^{Ac}$, and $K14^{Ac}$. We have raised antibodies to the latter three acetyl forms and show that each is chromatin-associated and primarily regulated by the NuA4 acetyltransferase and Hda1 deacetylase complexes. We have identified a range of novel $Htz1^{Ac}$ regulators, all of which have equivalent effects on $K8^{Ac}$, $K10^{Ac}$, and $K14^{Ac}$. This suggests that each acetylation is redundant, a proposal supported by epistasis mapping analyses of a comprehensive panel of unacetylatable *htz1* alleles. However, N-terminal acetylation is important for Htz1 function, and completely unacetylatable *htz1* alleles show a large number of genetic interactions in common with and distinct

from *htz1*Δ. Furthermore, each Htz1^{Ac} is actively regulated in response to benomyl, a microtubule destabilizing agent. Thus our results indicate that N-terminal acetylation is important for the Htz1 function, but the cell does not distinguish between individual acetylated lysines on this histone.

Experimental Procedures

Materials

The antibodies and strains used in this study are listed in supplemental Tables S1 and S2.

Antibody Generation

Rabbit polyclonal affinity-purified anti-*Sc* Htz1-K8^{Ac}, -K10^{Ac}, and -K14^{Ac} were from Millipore, with αK14^{Ac} described previously (Keogh et al., 2006a). For each new target (αK3^{Ac}, αK8^{Ac}, and αK10^{Ac}) five rabbits (prescreened for minimal reactivity to *Sc* proteins) were immunized (peptides as in Fig. 1*A* coupled to keyhole limpet hemocyanin), and test bleeds/boosts were performed monthly. Initial screening of each bleed was by immunoblotting against whole cell extracts (WCEs) from WT and *htz1*Δ cells. Attempts to raise αK3^{Ac} failed at this stage because all rabbits failed to raise a detectable immune response (not shown). Higher titer sera (usable at >1/200 dilution) from each rabbit were then tested for specificity against WCEs from unacetylatable *htz1* point mutant cells (*e.g. htz1-K8R*; see Fig. 1*B*). Specificity was further characterized by Luminex bead assay against a comprehensive peptide panel, including the relative immunogen, its acetylated

form, and all other acetyl peptides from the immunization series (supplemental Fig. S1). These analyses confirmed the absence of cross-reaction with an inappropriate Htz1^{Ac}species. To derive preparations for this study, positive bleeds from optimal responders were pooled, affinity-purified on the immunizing acetyl peptide, and subtracted with the relevant unmodified peptide.

Cell Fractionation

Sc cell fractionations were as described (Keogh et al., 2006a). The total, cytoplasmic, nuclear, and chromatin fractions were analyzed by SDS-PAGE and immunoblotting.

Creation of htz1 Mutant Strains

Various unacetylatable point mutants at *HTZ1* (*4KR*, *K3R*, *K3Ac**, etc.) were created by a modified *Delitto Perfetto* approach (Storici and Resnick, 2006) in a "magic marker" strain compatible with the synthetic genetic array protocol (Tong and Boone, 2006) (*KFY1069*; see supplemental Table S2). In brief, *HTZ1* was first replaced by a [Kan^R/*URA3*] cassette. The *htz1* ORF was then PCR-amplified from *Sc* genomic DNA (with the desired mutation inserted by megapriming (Keogh et al., 2002)), and the product was transformed into exponentially growing cells. Colonies that replaced [Kan^R/*URA3*] with *htz1* by homologous recombination were identified on 5-fluoroorotic acid (which counterselects *URA3*) and confirmed by sequencing. To construct *htz1-N*Δ, the *loxP-KanMX-loxP* cassette from pOM10 (Gauss et al., 2005) was used to replace the N terminus of *htz1* (residues 3–14, including all

four acetylatable lysines; see Fig. 1*A*) *in situ. Cre*-induced recombination resulted in cassette removal with retention of the 24-residue *loxP* element. In the final step a Nourseothricin resistance cassette (*NAT*) was incorporated immediately downstream of each *htz1* to facilitate locus selection during synthetic genetic array (Tong and Boone, 2006). An *HTZ1*::*NAT* locus-tracking strain was also created as a WT control.

Growth Curve Analysis

Growth curves were monitored with a Bioscreen C (*Oy Growth Curves*). Seed cultures were grown to mid-log in nonselective medium and diluted to $A_{600} \leq 0.1$ in medium with the appropriate agents, *e.g.* synthetic complete ± 6-azauracil (6AU) or mycophenolic acid (MPA), YPD ± benomyl, TBZ, or camptothecin. All of the analyses were performed in triplicate, and A_{600} curves (30 °C, constant agitation, 15-min time points) were monitored for >48 h.

Htz1 Stability Analyses

Translational shut-off: Cycloheximide (final concentration, 35 μg/ml (Fillingham et al., 2008)) was added to exponentially growing cultures ($A_{600} = \sim 0.5$), and 5-ml aliquots were taken at the indicated intervals for WCE isolation (TCA method; see below). Target abundance was estimated by immunoblotting at each time point (see Fig. 3*A*).

Transcriptional Shut-off

To create *GAL1$_p$.HTZ1.HA$_3$*, the *GAL1* promoter was integrated by homologous recombination to replace the endogenous promoter upstream of *HTZ1.HA$_3$*. Immunoblotting (αHA) confirmed that comparable Htz1.HA$_3$ levels are derived from each promoter in YPGR (2% galactose, 1% raffinose) medium (not shown). An exponentially growing *GAL1$_p$.HTZ1.HA$_3$* culture (A_{600} = ~0.5 in YPGR) was collected by centrifugation, washed in double distilled H_2O, and resuspended in YPD (containing 2% glucose to repress transcription from *GAL1$_p$*). Aliquots were taken at the indicated intervals for WCE isolation. Target abundance was estimated by immunoblotting at each time point (see Fig. 3, *B* and *C*).

Synthetic Genetic Array Screening

Genetic interactions were determined by the partially automated synthetic genetic array method (Tong and Boone, 2006). *htz1* alleles in the magic marker background (supplemental Table S2) were mated in quadruplicate to either a mutant library of 1,286 factors involved in chromatin metabolism (Collins et al., 2007) or two libraries that together cover >98% of yeast genes (*i.e.* ~4800 nonessential genes individually deleted with a KanMX cassette (Winzeler et al., 1999) or hypomorphic alleles of ~842 essential genes with KanMX disrupting their polyadenylation site (Breslow et al., 2008)). All of the libraries were arrayed at 1536 colony density/12.5 × 8.5-cm plate and replica plated with a Singer RoToR. The growth of all double-mutant haploid daughters was compared with the respective single-mutant parents to identify and quantify positive or negative genetic

interactions (Collins et al., 2006; Schuldiner et al., 2005). For epistasis mapping (Collins et al., 2007), the genetic profile of each *htz1* allele was compared within a set of 2,255 profiles to calculate pair-wise Pearson correlation coefficients.

TCA Cell Extracts for Western Blotting

WCEs were isolated by the TCA method, which efficiently extracts chromatin and preserves labile modifications (Keogh et al., 2006b). In brief, ~10-ml cultures were grown to mid-log (A_{600} = ~1.0), and cells were collected by centrifugation and washed with 20% TCA. All further steps were performed on ice with prechilled solutions. Cell pellets were resuspended in 250 µl of 20% TCA and subjected to glass bead lysis. The suspension minus the glass beads was collected, 1 ml of 5% TCA was added, and the precipitate was collected by centrifugation. The pellets were washed with 750 µl of 100% ethanol, and the proteins were solubilized in 50 µl of 1 M Tris, pH 8.0, 100 µl of 2× SDS-PAGE loading buffer (60 mM Tris, pH 6.8, 2% SDS, 10% glycerol, 100 mM DTT, 0.2% bromphenol blue). After 5 min at 95 °C, insoluble material was removed by centrifugation, and the supernatant was analyzed further by immunoblotting.

Results

Htz1-K8^{Ac}, K10^{Ac}, and K14^{Ac} Are Chromatin-associated and Regulated by Esa1 and Hda1

Sc H2A.Z (Htz1) has four acetylatable lysines on its N terminus at positions 3, 8, 10, and 14 (Fig. 1*A*) (Babiarz et al., 2006; Keogh et al., 2006a; Millar and

Grunstein, 2006). Antibodies to each modification would greatly aid in their analysis. We have previously described anti-Htz1-K14^{Ac} (αK14^{Ac}) in detail (Keogh et al., 2006a) and have now raised αK8^{Ac} and αK10^{Ac} (Fig. 1*B* and supplemental Fig. S1). Attempts to raise αK3^{Ac} were unsuccessful (see "Experimental Procedures"). Each reagent (αK8^{Ac}, αK10^{Ac}, and αK14^{Ac}) shows strong specificity for its respective target in immunoblotting, losing recognition if the appropriate lysine is mutated to arginine (*e.g. K10R*; Fig. 1*B*). These antibodies also indicate that individual Htz1 N-terminal acetylations show no interdependence (*e.g.* K10^{Ac} is not impacted by*K3R*, *K8R*, or *K14R*) or compensation (*e.g.* K10^{Ac} levels do not increase relative to WT if K10 is the sole acetylatable residue on the Htz1 N terminus) (Fig. 1*B* andsupplemental Fig. S2).

The abundance of Htz1 K8^{Ac}, K10^{Ac}, and K14^{Ac} are strongly reduced on deletion of *SWR1*, the eponymous ATPase subunit of the SWR complex (SWR-C; Fig. 1*C*). Because Swr1 is required for the insertion of Htz1 into nucleosomes (Krogan et al., 2003; Mizuguchi et al., 2004), this indicates that all three lysines are acetylated after chromatin assembly. Cell fractionation confirmed that each Htz1^{Ac} is chromatin-associated (Fig. 1*C*) and further that its acetylation status does not regulate entry to this cellular compartment (supplemental Fig. S3).

The Esa1 and Gcn5 acetyltransferases or Hda1 deacetylase are each reported to regulate the acetylation of Htz1 (Millar et al., 2006; Keogh et al., 2006a; Lin et al., 2008). In direct analyses K8^{Ac}, K10^{Ac}, and K14^{Ac} are abolished in cells containing mutants of NuA4, including a temperature-sensitive allele of Esa1 (*esa1-L254P*) or deletion of various complex subunits (Fig. 1, *D* and *E*). In contrast *gcn5*Δ has no

effect on K8^{Ac}, K10^{Ac}, or K14^{Ac} levels (Fig. 1*D*), although we note that Gcn5 could still target Htz1-K3^{Ac}; this could not be determined without an antibody to the latter modification. Regarding the deacetylation of Htz1, Hda1 has been shown to target Htz1-K14^{Ac} (Lin et al., 2008). We have confirmed this and determined that Hda1 also deacetylates K8^{Ac} and K10^{Ac} (Fig. 1*F*). Thus three distinct Htz1^{Ac} species are chromatin-associated and metabolized by the same enzymes in rapidly growing cells.

Genome Scale Genetic Analyses Indicate That Individual Htz1 Acetylations Are Redundant

To determine whether each Htz1 acetylation could have a specific role, we tested the genetic interactions of a comprehensive range of unacetylatable alleles. A number of synthetic genetic array compatible strains were created, including complete deletion (*htz1*Δ), N-terminal deletion (*N*Δ), all four lysines mutated to arginine (*4KR*) or glutamine (*4KQ*), each individual lysine mutated to arginine (*e.g. K3R*) or glutamine (*e.g. K3Q*), or a single remaining acetylatable residue (*e.g. K8R/K10R/K14R* = *K3Ac**). A locus-marked *HTZ1*::*NAT* strain was also created as a WT control. In immunoblotting each mutant was expressed at similar levels to WT with the exception of *htz1-N*Δ, which is weakly hypomorphic (not shown). Notably this hypomorphism is also seen with a comparable allele of *Schizosaccharomyces pombe* H2A.Z (*pht1-N*Δ) (Kim et al., 2009).

Each strain was individually mated to a mutant library of 1,286 factors involved in chromatin metabolism (Collins et al., 2007), and double-mutant haploid

daughters were isolated. We then derived scores covering each negative (*i.e.* synthetic sick/lethal (SS/SL)) or positive (*i.e.* epistasis or suppression) genetic interaction using colony size as a quantitative readout (Schuldiner et al., 2005; Schuldiner et al., 2006). The genetic interaction profile of a particular mutant can be used as a high resolution phenotype, with functionally related factors displaying similar profiles (Collins et al., 2007; Schuldiner et al., 2005). Thus deletions of individual members of the SWR-C are highly correlated with *htz1*Δ and *N*Δ (Fig. 2*A*, *green nodes*). This is expected because SWR-C is required for the insertion of Htz1 into chromatin (Fig. 1*C*) (Krogan et al., 2003; Mizuguchi et al., 2004). The profiles for *htz1*Δ, *N*Δ, *4KR*, and *4KQ* are also highly correlated (Fig. 2*A*), suggesting that significant functionality of Htz1 resides in its N terminus and furthermore that the primary role of this region is to harbor the acetylations. Indeed the *4KR* and *4KQ* profiles were most similar to *N*Δ within those 2,255 profiles collated to date (Fig. 2*B*) (Collins et al., 2007; Schuldiner et al., 2005; Fiedler et al., 2009). Furthermore the *4KR* and *4KQ* profiles were themselves very highly correlated (Pearson correlation coefficient (CC) 0.876; Fig. 2*C*), suggesting that the charge on these residues does not regulate their function (see "Discussion").

In contrast to the widespread genetic interactions of completely unacetylatable *htz1* alleles, single point (*e.g. K3R*) or singly acetylatable (*e.g. K3Ac**) mutants gave few consistent interactions and clustered near the wild-type control (Fig. 2*A*, *white node*). Thus the loss of any single Htz1 acetylation does not confer a genotype strong enough to distinguish the allele, suggesting no significant loss-of-

function in any case. Because the retention of any single acetylatable lysine has a similar effect, this suggests that the individual Htz1 acetylations are redundant.

Closer inspection of the individual genetic interactions for *htz1*Δ, *N*Δ, *4KR*, and *4KQ* revealed many that are common between these four mutants. This includes SS/SL interactions with components of the PAF complex (*RTF1* and *CDC73*), the replication checkpoint complex (*CSM3* and *MRC1*), and the chromatin regulators*BRE1*, *SET2*, and *SET3* (Fig. 2*D*). However, and as expected given their imperfect CCs, *htz1*Δ also shows many interactions distinct from the completely unacetylatable alleles. This may be related to the loss of the Htz1 core domain or C terminus in the complete deletion (see "Discussion"). For example, we observed SS/SL interactions between *htz1*Δ and components of the Isw1, Ino80, and Elongator complexes, as well as factors involved in kinetochore (*MCM16*, *MCM21*, and*MCM22*) and spindle (*MAD2*, *MAD3*, and *RAD61*) function. Finally, we also observed SS/SL interactions common to *htz1*Δ and *N*Δ but not shared by *4KR* or *4KQ*, such as with components of the Rpd3L deacetylase complex (Fig. 2*D*). These last may be related to the hypomorphic nature of the *N*Δ allele or indicate that the tail contains functional residues other than the acetylatable lysines.

To more comprehensively investigate the various pathways in which $Htz1^{Ac}$ may play a role, we further mated the *N*Δ or *4KR* strains against two libraries that together cover >95% of the entire *Sc* genome: ~4800 nonessential deletions (Winzeler et al., 1999) and ~900 essential DAmP hypomorphs (Breslow et al., 2008). As above, *N*Δ and *4KR* shared many genetic interactors, with ~30% of these conserved with *htz1*Δ (supplemental Fig. S4). This was a surprise because we had

originally predicted that completely unacetylatable *htz1* mutants would simply display a subset of the genetic interactions seen in *htz1*Δ (see "Discussion"). However, gene ontology analysis supports the validity of each group SS/SL with *htz1-N*Δ, *i.e.* those in common with *htz1*Δ or unique to the tail-less mutant. For example, the "telomere maintenance" gene ontology category is over-represented in each group, and this function is known to be mediated by the Htz1 N terminus (Babiarz et al., 2006).

Htz1 Is Recycled into Chromatin after Nucleosome Displacement

To further investigate any potential distinction between Htz1-K8^{Ac}, -K10^{Ac}, and -K14^{Ac}, we measured the half-life (t½) of each modification. Htz1 is enriched in "hot" nucleosomes, defined as those with a high turnover rate (Dion et al., 2007). However, it is not known whether the histone variant is degraded or reused after nucleosome eviction, although the former would likely result in a protein with a relatively short half-life. Furthermore if each Htz1^{Ac} species is enriched at distinct genomic locations, this could lead to a range of decay kinetics.

Two approaches were used: general inhibition of the translational machinery by cycloheximide (Fig. 3*A*) or the specific transcriptional repression of *GAL1p-HTZ1.HA$_3$* by glucose (Fig. 3*B*). After each shut-off, the abundance of Htz1 and Htz1-K8^{Ac}, -K10^{Ac}, and -K14^{Ac} were monitored for up to 6 h. In each case total Htz1 had a t½ greater than 3 h or approximately two complete cell cycles, strongly suggesting reuse after nucleosome eviction. Each Htz1^{Ac} decayed significantly faster than total Htz1 on cycloheximide (Fig. 3*A*). However, this is likely due to the translational

inhibition of an acetylation regulator (*e.g.* a subunit of NuA4; Fig. 1,*D* and *E*), because each Htz1^{Ac} tracked with total Htz1 when transcription of the histone was repressed (Fig. 3*B*). Furthermore the observation that Htz1-K8^{Ac}, -K10^{Ac}, and -K14^{Ac} decayed at similar rates on transcriptional shut-off also argues against their differential usage at distinct locations.

To examine whether its acetylation could regulate Htz1 turnover, we compared *GAL1p-HTZ1.HA$_3$* to two additional strains: *GAL1p-htz1–4KR.HA$_3$* and [*GAL1p-HTZ1.HA$_3$*/*hda1*Δ]. The former contains an unacetylatable form of Htz1, the latter is increased for all acetylated species of the histone (Fig. 1*F*). On transcriptional repression, total Htz1 levels decayed with similar rates in all three backgrounds (Fig. 3*C*), indicating that its acetylation status has no impact on the turnover of this histone variant.

Each Htz1^{Ac} Is Similarly Regulated in a Range of Mutant Backgrounds

To further investigate the possibility of their differential regulation, we used a proteomic screening approach (Schneider et al., 2004) to identify modifiers of each Htz1^{Ac}. Viable mutants of ~50 genes, primarily factors related to the gene ontology terms "histone acetylation" or "histone deacetylation," were isolated from the *Sc* haploid deletion collection and immunoblotted for Htz1-K8^{Ac}, -K10^{Ac}, and -K14^{Ac} (see supplemental Table S3). This confirmed that all three Htz1^{Ac} forms are dependent on individual subunits of NuA4 (*YAF9*, *YNG2*, and *EAF1*) but independent of SAGA components (see also Fig. 1, *D* and *E*). The approach identified many additional regulators of Htz1^{Ac} (*e.g.* Asf1, Rpd3L, and Elongator) but no differential

regulators of any individual acetylation, *i.e.* K8^{Ac}, K10^{Ac}, and K14^{Ac} tracked together in every background (supplemental Table S3 and data not shown). This includes Bud14 and Clb2, whose deletions were previously identified as reducing K14^{Ac} levels (Fiedler et al., 2009); these deletions impact K8^{Ac} and K10^{Ac} to a similar degree (not shown). In many of these deletion backgrounds, the effect on Htz1^{Ac} may be indirect, almost certainly in Rpd3L complex mutants where reduced K8^{Ac}, K10^{Ac}, and K14^{Ac} is an unlikely outcome when inhibiting a deacetylase.

We chose Asf1 as a novel Htz1^{Ac} regulator to examine further. Deletion of this H3-H4 chaperone reduces the level of each Htz1^{Ac} without affecting the abundance of the histone (Fig. 4*A*). Each acetylation takes place after Htz1 is assembled into chromatin (Fig. 1*C*), so we tested whether *asf1*Δ impacts the deposition and/or acetylation step(s). In fractionation analyses, the chromatin of *asf1*Δ cells shows WT levels of total Htz1 but a reduction in each Htz1^{Ac} (Fig. 4*B*). Thus Asf1 appears to regulate the Htz1 acetylation step. Asf1 is multi-functional and among other roles regulates various acetylations on histones H3 and H4 (H3-K9^{Ac}, K27^{Ac}, K56^{Ac}; H4-K5^{Ac}, K8^{Ac}, K12^{Ac}; see Fig. 3*C* (Ransom et al., 2010)). We thus tested whether Htz1^{Ac} was dependent on the prior modification of any of these residues using a range of mutant strains with unacetylatable alanine at each position (*i.e.* H3-K9A, K27A, K56A; H4-K5A, K8A, K12A (Nakanishi et al., 2008)). In each mutant the abundance of each Htz1^{Ac} was comparable with WT (supplemental Fig. S5).

Given the multi-functionality of Asf1, it could prove difficult to identify whether a specific role is required for efficient Htz1^{Ac}. We thus took an approach where Asf1 was rapidly removed from cells, and each Htz1^{Ac} was monitored over a

succeeding time course. We used two shut-off strains, namely *GAL1$_p$.ASF1.HA$_3$* and*GAL1$_p$.ASF1.HA$_3$.PEST* as the only source of the factor, with *ASF1* transcription in each repressed by glucose addition (Zabaronick and Tyler , 2005). The Cln2 C-terminal PEST domain confers a rapid, cell cycle-independent turnover on heterologous proteins (Mateus and Avery, 2000), and its addition to Asf1.HA$_3$ significantly shortens protein half-life (*t*½ Asf1.HA$_3$ > 60 min; *t*½ Asf1.HA$_3$.PEST < 15 min; Fig. 4*D*) (Zabaronick and Tyler , 2005). However, the abundance of each Htz1^{Ac} (K8^{Ac}, K10^{Ac}, or K14^{Ac}) remained unchanged over each 6-h time course (*i.e.* more than three complete cell cycles under experimental conditions) rather than reducing as expected (Fig. 4*D*). This is not a strain background issue because their prolonged (> 24 h) incubation in glucose-containing medium recapitulates the original observation of reduced Htz1^{Ac} (Fig. 4*E*). Thus the reduced Htz1^{Ac} in *asf1*Δ cells may not be due to a direct regulation by Asf1 of Htz1 acetylation but rather may involve some adaptation to the chronic absence of the histone chaperone. This also raises a cautionary note about overinterpreting the relationship of those additional Htz1 regulators identified by the proteome screening approach (supplemental Table S3).

Htz1 Acetylation Regulates the Response to Microtubule Depolymerizing Agents

The deletion of H2A.Z in *S. cerevisiae* or *S. pombe* leads to genome instability, synthetic genetic interactions with components of the kinetochore and spindle checkpoint machineries, and sensitivity to benomyl, a microtubule-depolymerizing agent (Figs. 2*D* and and55*A*) (Kim et al., 2009; Keogh et al., 2006a; Wu et al., 2001).

H2A.Z acetylation contributes to the genome stability role in each species, with strains containing unacetylatable alleles showing increased rates of chromosome loss (Kim et al., 2009; Keogh et al., 2006a).

To further examine how individual acetylations on Htz1 may regulate benomyl resistance, we tested how the full range of unacetylatable alleles (*htz1*Δ, *N*Δ, *4KR*,*K3R*, *K3Ac**, etc.) responded to this agent. Rather than spotting onto plates, growth curves were monitored in liquid culture because the latter approach can more easily distinguish the reduced fitness of *htz1*Δ relative to WT on rich media (YPD; Fig. 5, compare *A* and *B*). In this manner completely unacetylatable *htz1* alleles (*N*Δ, *4KR* or *4KQ*) show a dose-dependent sensitivity to benomyl, although not to the same degree as *htz1*Δ (Fig. 5*B*). We (and others) had previously reported an*htz1-K14R* strain as sensitive to benomyl (Keogh et al., 2006a; Lin et al., 2008). However, we have been unable to consistently repeat this phenotype in current analyses using any single unacetylatable (*e.g. K14R*) or acetylatable (*e.g. K14Ac**) *htz1* allele (not shown). This may be related to strain differences or some other as yet uncharacterized factor.

Because Htz1^{Ac} contributes to benomyl resistance, we next asked whether the acetylations are themselves regulated by the agent. In parallel analyses, cells were treated with thiabendazole (TBZ, another microtubule depolymerizer) or camptothecin (a topoisomerase I inhibitor). These were chosen for comparison because*htz1*Δ (and to a lesser degree *N*Δ, *4KR* and *4KQ*) strains are sensitive to thiabendazole, whereas *htz1*Δ shows sensitivity to camptothecin under some conditions (*i.e.*on plates but not in liquid culture: Fig. 5*A* and supplemental Fig. S6).

The cells were exposed to each agent for 2 h, and the abundance of Htz1 and each acetylated form (K8^{Ac}, K10^{Ac}, K14^{Ac}) was determined. Under these conditions each Htz1^{Ac} was reduced in a dose-dependent manner by benomyl but unchanged in response to thiabendazole or camptothecin (Fig. 5*C*). The reduction in Htz1^{Ac} by benomyl is post-translational, because it is not accompanied by a change in Htz1 levels (Fig. 5*C*). It is also reversible, largely returning to normal 2 h after removal of the agent (Fig. 5*D*).

Benomyl could modulate Htz1^{Ac} levels by various means, including impacting Htz1 deposition (by Swr1; Fig. 1*C*), acetylation (by Esa1; Fig. 1*D*), or deacetylation (by Hda1; Fig. 1*F*). We thus exposed WT or *hda1*Δ cells to benomyl as above and determined that those lacking the deacetylase (or its accessory subunits Hda2 and Hda3) show no reduction in any Htz1^{Ac} (Fig. 6*A*, supplemental Fig. S7*A*, and data not shown). Benomyl has no effect on the abundance of any members of the HDA complex (Hda1, Hda2, or Hda3; supplemental Fig. S7, *B* and *C*), so the agent is likely to regulate either its activity or localization. In addition, the response is rapid, with Htz1^{Ac} levels visibly reduced within 10 min of exposure to benomyl and bottoming out by 30 min (Fig. 6*B*).

Finally we determined whether the benomyl-induced deacetylation of Htz1 is the sole chromatin change in response to this agent. Of four acetylations tested (H3-K9^{Ac}, H3-K27^{Ac}, H4-K5^{Ac}, and H4-K12^{Ac}), all are reduced in response to benomyl (Fig. 6*C*). However, each is independent of Hda1 (Fig. 6*C*), even although H3-K9^{Ac} and K27^{Ac} are substrates of this deacetylase at Tup1-repressed promoters (Wu et al., 2001). Thus benomyl induces widespread changes in histone acetylation and likely

via multiple mediators. The benomyl sensitivity of strains containing unacetylatable *htz1* alleles suggests that deacetylation of this histone variant is one component of the effective cellular response to this agent.

Discussion

Histones are subject to a bewildering array of covalent modifications (Lall, 2007). In this work we set out to test whether each acetylation on a single histone tail could mediate a specific function. We examined the acetylations of H2A.Z (*S. cerevisiae* Htz1) rather than of a major histone because the variant provides a number of advantages. It is encoded by a single, nonessential gene, and the protein has a dedicated deposition machinery (the SWR-C) and limited genomic distribution (*e.g.*highly enriched in the nucleosomes flanking the nucleosome-free region over promoters (Mehta et al., 2010; Zlatanova and Thakar, 2008). In addition Htz1 is multiply acetylated and multi-functional. It has thus been posited that specific acetylated forms could regulate distinct roles (Babiarz et al., 2006; Millar et al., 2006; Keogh et al., 2006a). We find that this is not the case: rather the modifications are redundant, with any single lysine on the Htz1 tail sufficient to mediate function. However, we also show that the acetylation of Htz1 is highly regulated and required for full functionality of this histone variant.

N-terminal Acetylation Is a Widespread Modification of H2A.Z Orthologs

The H2A.Z N terminus is acetylated in *S. cerevisiae* (Babiarz et al., 2006; Millar et al., 2006; Keogh et al., 2006a), *S. pombe* (Kim et al., 2009), *Tetrahymena*

thermophila (Ren and Gorovsky, 2001; Ren and Gorovsky, 2003), and metazoans (Ishibashi et al., 2009; Tanabe et al., 2008; Bruce et al., 2005). In each organism multiple acetyl-lysines are possible: up to four in budding and fission yeasts, five in metazoans, and six in *Tetrahymena*. However, these acetylations are unequally distributed across all potential lysine substrates in each species. For example, in *Sc* $K3^{Ac}$, $K8^{Ac}$, $K10^{Ac}$, and $K14^{Ac}$ are ~3, 7, 14, and 38% of total Htz1, respectively (Millar et al., 2006). In contrast the histone is ~90% unmodified in chicken erythrocytes, with the remainder primarily mono-acetylated on the more N-terminal lysines ($K4^{Ac}$ ~ $K7^{Ac}$ > $K11^{Ac}$ > $K13^{Ac}$ ⋙ $K15^{Ac}$) (Dryhurst et al., 2009; Ishibashi et al., 2009). Thus the yeast and chicken H2A.Z N-terminal tails differ in both the abundance and precise location of their acetylations. This, coupled with the fact that the N terminus is one of the most variable regions across H2A.Z orthologs, would tend to suggest differences in how these modifications are regulated. Despite this, the respective *Kat5* family acetyltransferase appears to be the primary H2A.Z modifier in all tested species, *i.e. S. pombe* Mst1 (Kim et al., 2009), *Drosophila* Tip60 (Kusch et al., 2004), or *Caenorhabditis elegans* MYS-1 (Updike and Mango, 2006). This is certainly the case in*Sc*, with Htz1-$K8^{Ac}$, -$K10^{Ac}$, and -$K14^{Ac}$ each dependent on Esa1 (the budding yeast *Kat5* homolog) and multiple members of the Esa1-containing NuA4 complex (Fig. 1, *D* and *E*) (Babiarz et al., 2006; Keogh et al., 2006a). Esa1 thus shows great promiscuity in its ability to acetylate each lysine on the Htz1 N terminus (MSGKAHGGKGKSGAKD). However, the acetyltransferase is not infinitely capable, with an N-terminal HA_3-tagged form of Htz1 being very poorly acetylated (not shown). N-terminal GFP tagging

of*Drosophila* H2A.Z also disrupts its acetylation and compromises the ability of the variant to regulate the heat shock response (Tanabe et al., 2008). Alternatively, because the Htz1 N terminus (residues 1–53) regulates nuclear localization (Straube et al., 2010), N-terminal tagging of the histone may simply interfere with nuclear import and thus availability for nucleosome assembly and subsequent acetylation.

Coregulation and Genetics Indicate Minimal Specialization of Each Htz1^{Ac}

Disparate analyses demonstrate that the individual N-terminal acetylations on Htz1 are coregulated. Thus K8^{Ac}, K10^{Ac}, and K14^{Ac} are added by NuA4 acetyltransferase after Htz1 is inserted into nucleosomes by the SWR-C (Fig. 1, *C*–*E*) and subsequently removed by the Hda1 deacetylase complex (Fig. 1*F* and supplemental Fig. S7). All three acetylations decay with similar kinetics after Htz1 transcription is repressed (Fig. 3). All three are similarly affected in every mutant background tested (Fig. 4 and supplemental Table S3). Finally, all three acetylations are similarly effected by a diverse range of agents, including 6-azauracil, benomyl, camptothecin, mycophenolic acid, and TBZ (Fig. 5*C* and supplemental Fig. S8).

Their coregulation strongly suggests that individual Htz1 N-terminal acetylations are redundant, a proposal supported by comparing the genetic profiles of a comprehensive range of *htz1* alleles (Fig. 2). If the cell utilizes each individual acetylation for a specific role, we would expect to see distinct defects (and thus unique genetic interactions) associated with the mutation of these residues. However, completely unacetylatable mutants (*N*Δ, *4KR*, and *4KQ*) are highly

correlated with*htz1*Δ, whereas singly mutated (*e.g. K3R*) or singly acetylatable (*e.g. K3Ac**) alleles cluster closer to *HTZ1* (Fig. 2*A*). This suggests that a significant degree of the function of Htz1 resides in its N terminus or more specifically in the acetylatable lysines of this region. Furthermore it is strong evidence against an individual function for each N-terminal acetylation.

How Does Acetylation Regulate Htz1 Function?

Previous analyses have identified some role for Htz1 acetylation in the maintenance of euchromatin-heterochromatin boundaries (Babiarz et al., 2006), transcription (Millar et al., 2006), and chromosome transmission (Keogh et al., 2006a). Gene ontology categories covering these functions are among the most highly represented when the genetic interactions of completely unacetylatable *htz1* alleles are examined (supplemental Fig. S4). In addition, and consistent with a role in chromosome stability, unacetylatable *htz1* alleles are sensitive to the microtubule destabilizing agents benomyl and TBZ (Fig. 5, *A* and *B*, and supplemental Fig. S6, *D* and *E*). One possibility is that Htz1^{Ac} somehow regulates the response to spindle stress; this is supported by the observation that each acetylation is regulated by the HDA complex on exposure to benomyl (Fig. 5*C*). This could in turn suggest how this deacetylase complex regulates normal chromosome segregation (Kanta et al., 2006).

Approximately 30% of the genetic interactions of *htz1*Δ are conserved with *N*Δ or *4KR* (Fig. 2 and supplemental Fig. S4). However, the list of interactors with unacetylatable *htz1* alleles are not simply a subset of those with the complete

deletion. The *htz1*Δ-specific group may be due to a loss of the Htz1 core domain or C terminus. Alternatively it may be caused by the inappropriate activity of SWR-C in the absence of Htz1 (Morillo-Huesca et al., 2010; Halley et al., 2010). However, the genetic interactors unique to unacetylatable *htz1* alleles are more difficult to explain. H2A.Z acetylation does not regulate the incorporation of the histone into chromatin in either *Sc*(supplemental Fig. S3*A*) or *S. pombe* (Kim et al., 2009). In addition the acetylation status of Htz1 does not impact nucleosome turnover (Fig. 3*C*). It is possible that nucleosomes containing unacetylatable *htz1* are unable to use an $H2A^{Ac}$-dependent compensation pathway. This might indicate that cells differentiate certain histone tails as one of two states, acetylated or not, and make little distinction between the particular modified residue. This would appear to be a retrograde step in our understanding of chromatin function. However, we note that NuA4 acetylates Htz1 (Fig. 1, *D* and *E*), H2A, and H4 (Suka et al., 2001), and mutant combinations that reduce the total number of NuA4-dependent acetylations (*e.g.* [*htz1–4KR*, *eaf1*Δ] or [*htz1–4KR*, *hhf2-K5/8/12R*]) are invariably synthetic (Babiarz et al., 2006). In addition the Esa1-targeted H4 N terminus also shows significant redundancy, such that the loss of all four acetylatable lysines leads to slow growth and camptothecin sensitivity, but the retention of any single lysine complements both phenotypes (Bird et al., 2002). Furthermore an ectopic lysine in an otherwise unacetylatable H4 N terminus is acetylated (likely by Esa1) and also repairs the slow growth and camptothecin sensitivity (Bird et al., 2002).

As noted above, the simultaneous loss of all N-terminal acetylation on Htz1 has widespread effects. Thus the modifications regulate the functions of this histone

variant. However, is this by charge modulation or effector recruitment? Arginine and glutamine are respectively considered charge mimics for unacetylated and acetylated lysine (Ishibashi et al., 2009), although there is no evidence that they mimic the unmodified/modified state as it relates to recognition by acetyl-binding proteins. We note that the *htz1–4KR* and *4KQ* mutants are indistinguishable throughout this work, an observation also made with comparable mutants in fission yeast (Kim et al., 2009). This strongly argues against the acetylation of Htz1 acting via charge modulation. However, no chromatin effector has yet been identified that is so selectively promiscuous in its binding (*i.e.* only to K^{Ac}s on a H2A.Z tail). One possibility is that the Htz1 tail functions by dynamic charge modulation; *i.e.* the acetylations must be added and then removed (constitutively one or the other will not work). If this is indeed the case, our current belief that the simple presence of many histone marks is enough may need some revision.

Acknowledgments

We thank Steve Buratowski (Harvard Medical School), Jeffrey Fillingham (Ryerson University), Dan Finley (Harvard Medical School), Jack Greenblatt (University of Toronto), Jessica Tyler (University of Colorado, Denver), Fred Winston (Harvard Medical School), and Ali Shilatifard (Stowers Institute) for the generous supply of antibodies and yeast strains (detailed in supplemental Tables S1 and S2). We also thank Brian Strahl (University of North Carolina, Chapel Hill), Joseph Wade (Wadsworth Center), Jon Warner (AECOM), and members of the Krogan and Keogh labs for advice and comments.

*This work was supported, in whole or in part, by the National Institutes of Health. This work was also supported by the Searle and Keck Foundations, by National Cancer Institute Cancer Centre Grant 2P30CA013330 to Albert Einstein College of Medicine, and by the Speaker's Fund for Biomedical Research: Toward the Science of Patient Care awarded by the City of New York.

Figures

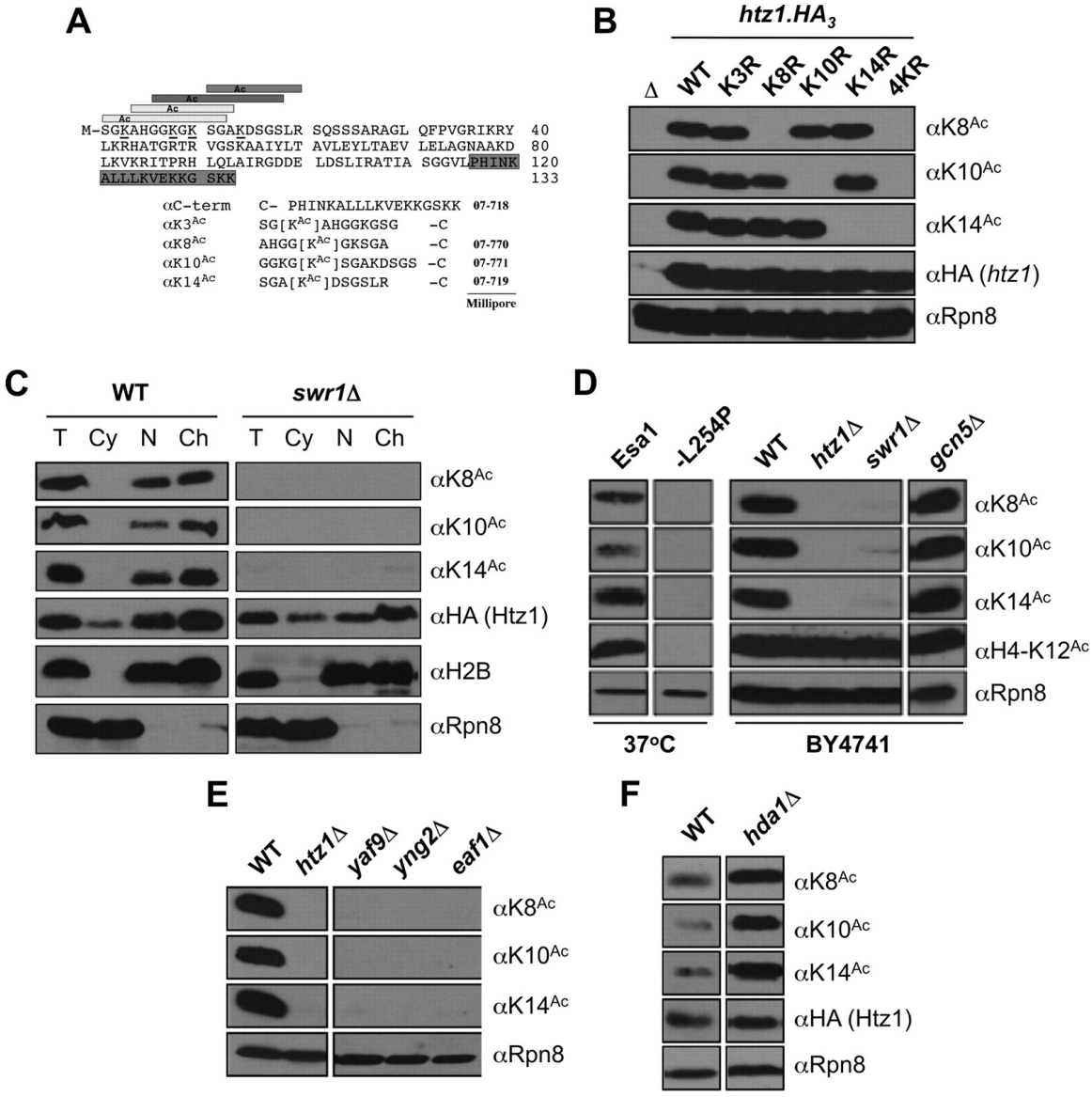


Figure 1. Htz1 lysines 8, 10, and 14 are acetylated by NuA4 as a component of chromatin. *A*, peptides used to raise αHtz1^{Ac} antibodies (see "Experimental Procedures"), with the four acetylatable lysines (at positions 3, 8, 10, and 14) on the Htz1 N terminus *underlined*. The terminal cysteine (–*C*) on each peptide was used to couple to KLH for immunization or a sulfolink resin for affinity purification. *B*, αK8^{Ac}, αK10^{Ac}, and αK14^{Ac} are highly specific. C-terminally HA$_3$-tagged forms of the

indicated *htz1* alleles were individually expressed as the sole source of the histone and WCEs analyzed by immunoblotting. Specific mutation of the appropriate lysine to unacetylatable arginine (*e.g. K8R*) abolishes recognition by the relevant affinity-purified antibody. Δ is an *htz1*Δ strain. αHA demonstrates equivalent expression of each *htz1.HA*$_3$. αRpn8 (19 S proteasome subunit) serves as a loading control. *C*, each acetylated form of Htz1 is chromatin-associated. WT and *swr1*Δ cells were spheroblasted (total, *T*); fractionated into cytoplasm (*Cy*), nucleus (*N*), and chromatin (*Ch*); and immunoblotted as indicated. αHA monitors the localization of Htz1.HA$_3$. Appropriate segregation of histone H2B and Rpn8 indicates efficient fractionation; the former is primarily localized in insoluble chromatin, and the latter is primarily in soluble cytoplasm. *D*, all of the tested forms of Htz1^{Ac} are Esa1-dependent and Gcn5-independent. *Left panel*, *ESA1* or *esa1-L254P* strains were incubated at a nonpermissive temperature (37 °C, 2 h), WCEs were harvested and immunoblotted as indicated. Esa1-dependent H4-K12^{Ac} is a positive control. Rpn8 is a loading control. *Right panel*, BY4741. The indicated WCEs were harvested at 30 °C. *E*, all of the tested forms of Htz1^{Ac} are dependent on indicated members of the NuA4 acetyltransferase complex. These deletions also reduce H4^{Ac} levels but have no impact on the abundance of Htz1 (Keogh et al., 2006a). *F*, each Htz1^{Ac} is increased on deletion of the Hda1 deacetylase.

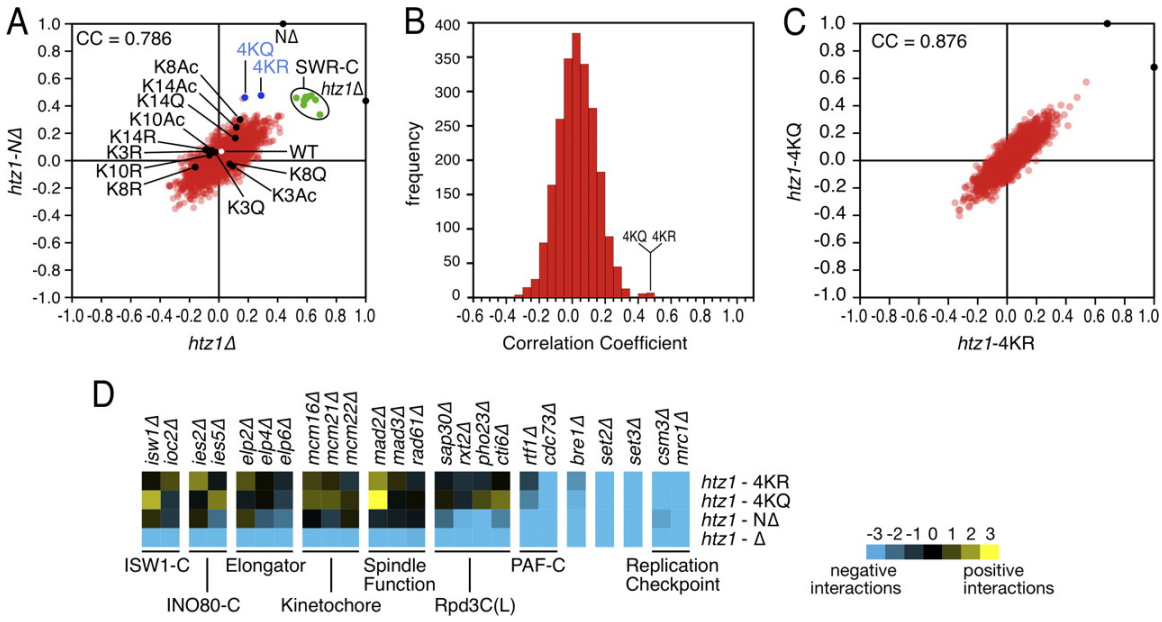


Figure 2. The primary function of the Htz1 N terminus is to harbor lysine acetylations. *A*, Pearson CC plot compares the genetic interaction profiles of *htz1*Δ or *htz1-N*Δ (*X*- or *Y*-axes) with 2,255 profiles performed to date. *Blue* and *black spots*, as labeled; *green spots*, SWR-C components (*swr1*Δ, *yaf9*Δ, *vps71*Δ, *vps72*Δ, *swc3*Δ, *swc5*Δ, and *arp6*Δ); *white spot*, WT control; *red spots*, all other mutants on array. *B*, distribution plot of the CCs of *N*Δ against 2,245 mutant profiles. The most highly correlated profiles include those derived from *4KR* and *4KQ. C*, a correlation plot of CCs for 2,245 individual genetic screens relative to *4KR* or *4KQ. D*, as expected from their strong but imperfect pairwise CCs *htz1*Δ, *N*Δ, *4KR*, and *4KQ* show both common and unique genetic interactions. *Blue* and *yellow* correspond to negative or positive genetic interactions.

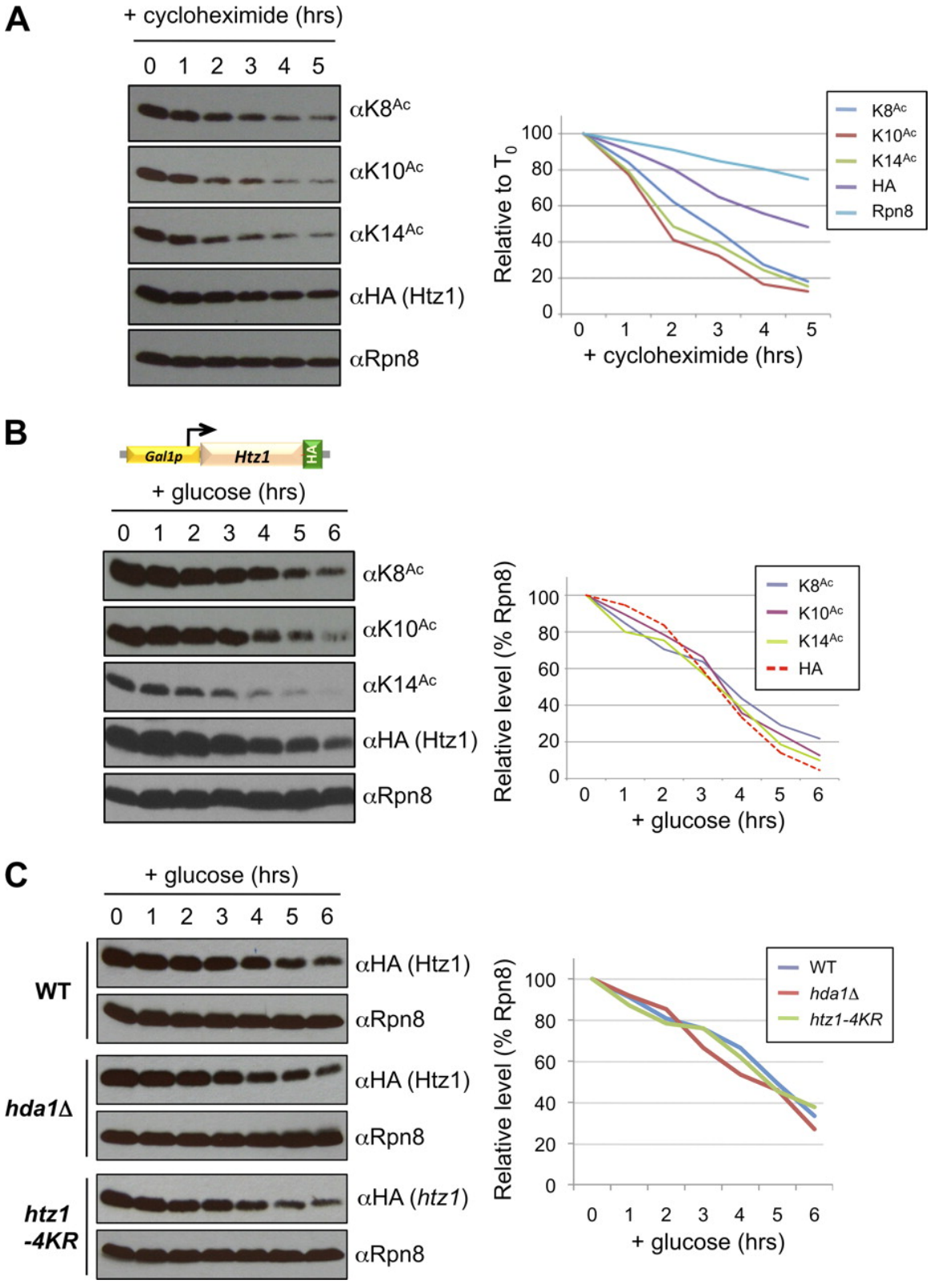

A
+ cycloheximide (hrs)
0 1 2 3 4 5
αK8Ac
αK10Ac
αK14Ac
αHA (Htz1)
αRpn8
Relative to T0
100
80
60
40
20
0
0 1 2 3 4 5
+ cycloheximide (hrs)
K8Ac
K10Ac
K14Ac
HA
Rpn8
B
Gal1p
Htz1
HA
+ glucose (hrs)
0 1 2 3 4 5 6
αK8Ac
αK10Ac
αK14Ac
αHA (Htz1)
αRpn8
Relative level (% Rpn8)
100
80
60
40
20
0
0 1 2 3 4 5 6
+ glucose (hrs)
K8Ac
K10Ac
K14Ac
HA
C
+ glucose (hrs)
0 1 2 3 4 5 6
WT
αHA (Htz1)
αRpn8
hda1Δ
αHA (Htz1)
αRpn8
htz1 -4KR
αHA (htz1)
αRpn8
Relative level (% Rpn8)
100
80
60
40
20
0
0 1 2 3 4 5 6
+ glucose (hrs)
WT
hda1Δ
htz1-4KR

Figure 3. Htz1 acetylation does not regulate turnover of the histone variant. *A*, cycloheximide was used to repress translation in a *HTZ1.HA*$_3$ strain, with samples taken for WCE isolation and immunoblotting at times indicated. Quantitation measures each immunoblotted species relative to their abundance at T_0 (precycloheximide). *B*, glucose (final concentration, 2%) was added to repress *GAL1* promoter-driven transcription in a *GAL1*$_p$*.HTZ1.HA*$_3$ strain, with samples taken for WCE isolation and immunoblotting at times indicated. Quantitation measures each immunoblotted species relative to Rpn8 at the same time point. *C*, glucose was added to repress transcription from *GAL1*$_p$*.HTZ1.HA*$_3$ (WT or *hda1*Δ backgrounds) or *GAL1*$_p$*.htz1–4KR.HA*$_3$, with samples taken for WCE isolation and immunoblotting at times indicated. Quantitation measures each *htz1.HA*$_3$ relative to Rpn8 at the same time point in the same background.

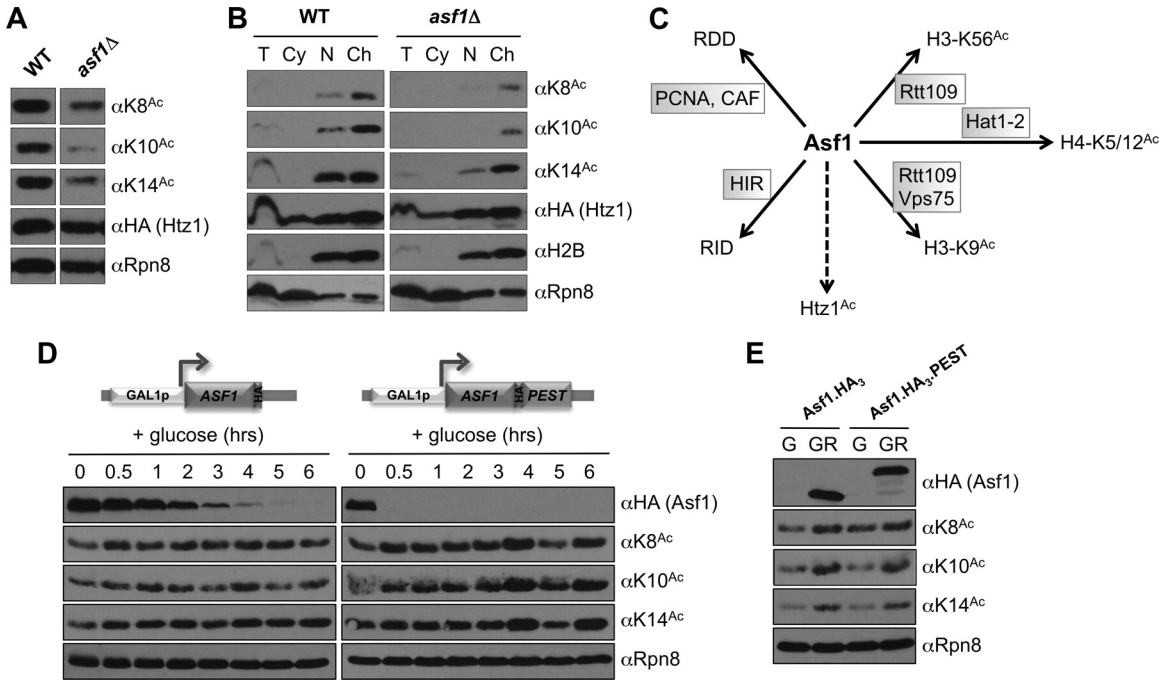


Figure 4. Htz1^{Ac} is regulated in *asf1*Δ cells by an indirect mechanism. *A*, each Htz1^{Ac} (but not total Htz1.HA$_3$) is decreased in *asf1*Δ cells. Rpn8 serves as a loading control. *B*, *asf1*Δ reduces each Htz1^{Ac} but has no effect on the amount of total Htz1 associated with chromatin. WT and *asf1*Δ cells were spheroblasted (total, *T*); fractionated into cytoplasm (*Cy*), nucleus (*N*), and chromatin (*Ch*); and immunoblotted as indicated. Appropriate localization of H2B and Rpn8 indicates efficient fractionation (as in Fig. 1*C*). *C*, schematic depicts the various roles of Asf1, a histone H3-H4 donor to various acetyltransferase complexes (substrates indicated), CAF/proliferating cell nuclear antigen (PCNA) at replication forks (for replication-dependent deposition (*RDD*)), and HIR in actively transcribed regions (for replication-independent deposition (*RID*)). *D*, glucose (final concentration, 2%) was added to repress *GAL1$_p$*-driven transcription in *GAL1$_p$.ASF1.HA$_3$* or *GAL1$_p$.ASF1.HA$_3$.PEST* strains, with samples taken for WCE isolation and immunoblotting at times indicated. Rpn8 serves as a loading control. *E*,

GAL1$_p$.ASF1.HA$_3$ or *GAL1$_p$.ASF1.HA$_3$.PEST* strains were maintained in glucose (*G*, 2%) or galactose/raffinose (*GR*, 2%/1%) containing medium for >24 h before the samples were collected for WCE isolation and immunoblotting. Rpn8 serves as a loading control.

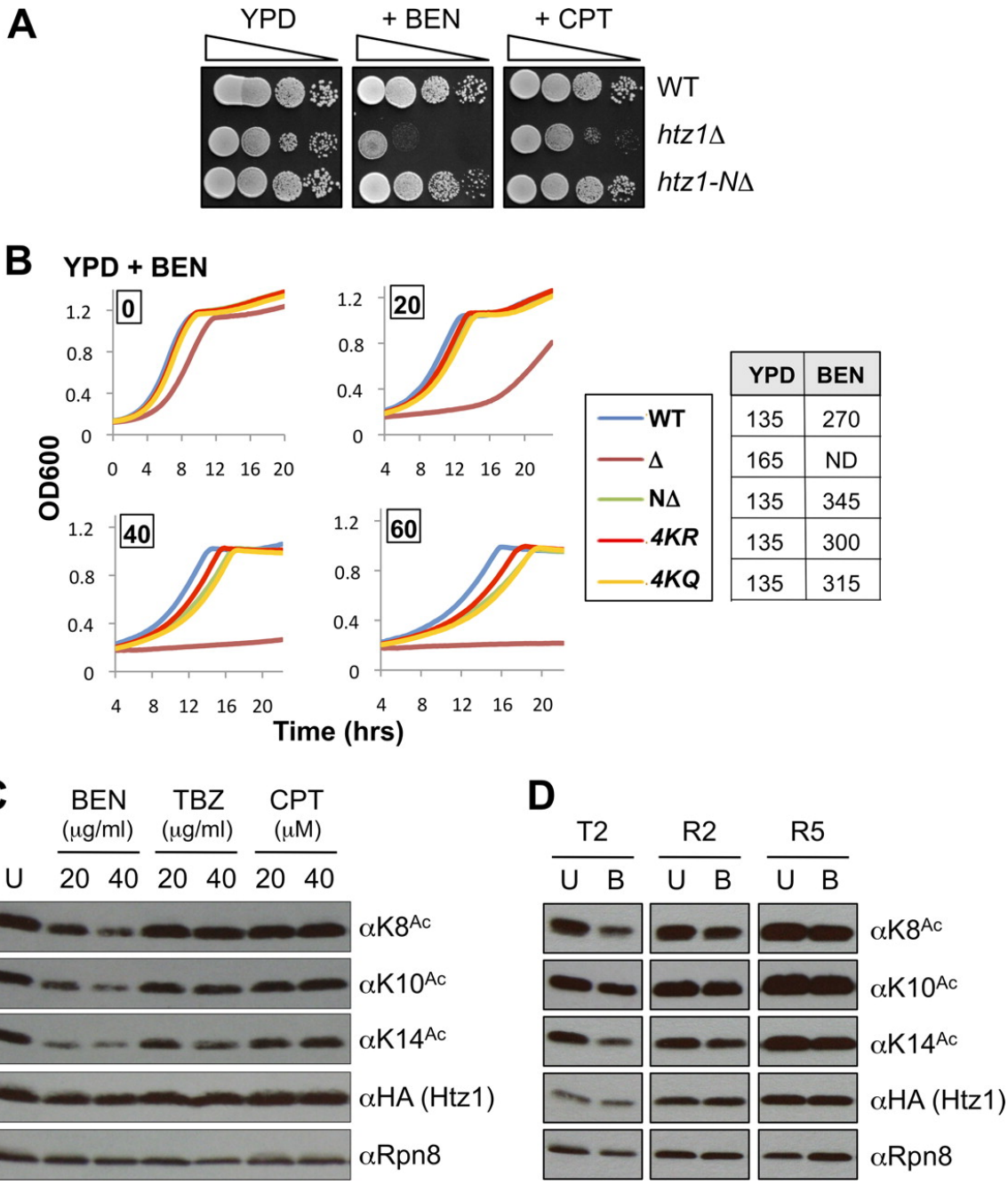


Figure 5. Htz1^{Ac} is reduced in response to benomyl. *A*, *htz1*Δ (and to a lesser degree *N*Δ) cells are sensitive to benomyl (*BEN*, 15 μg/ml). In contrast *htz1*Δ but not *N*Δ cells are sensitive to camptothecin (*CPT*, 20 μM). Spots are 10-fold serial dilutions on YPD plates ± each agent after 2 days at 30 °C. *B*, unacetylatable *htz1* strains (*N*Δ, *4KR*, or *4KQ*) are weakly sensitive to benomyl. Growth curves were

analyzed on a Bioscreen (see “Experimental Procedures”) during culture in YPD + benomyl (0–60 μg/ml as indicated). The table indicates the time taken (in minutes) for each strain to traverse A_{600} 0.45 to 0.9 (*i.e.* exponential part of growth curve) in YPD ± benomyl (60 μg/ml). *ND*, not determined. *C*, each Htz1^{Ac} is reduced on benomyl treatment. WT cells were grown in YPD to A_{600} ~0.5 and benomyl, TBZ, or camptothecin was added (concentrations indicated) before growth for another 2 h. WCEs were harvested and immunoblotted as indicated. Rpn8 is a loading control. *U*, untreated. *D*, the decrease of each Htz1^{Ac} in response to benomyl is reversible. WT cells were grown in YPD to A_{600} ~0.5, and benomyl was added (40 μg/ml) for 2 h (T2). The cells were then washed, resuspended in YPD, and allowed to recover for 2 (*R2*) or 5 (*R5*) h. WCEs were harvested and immunoblotted as indicated. Rpn8 is a loading control. *U*, untreated; *B*, benomyl.

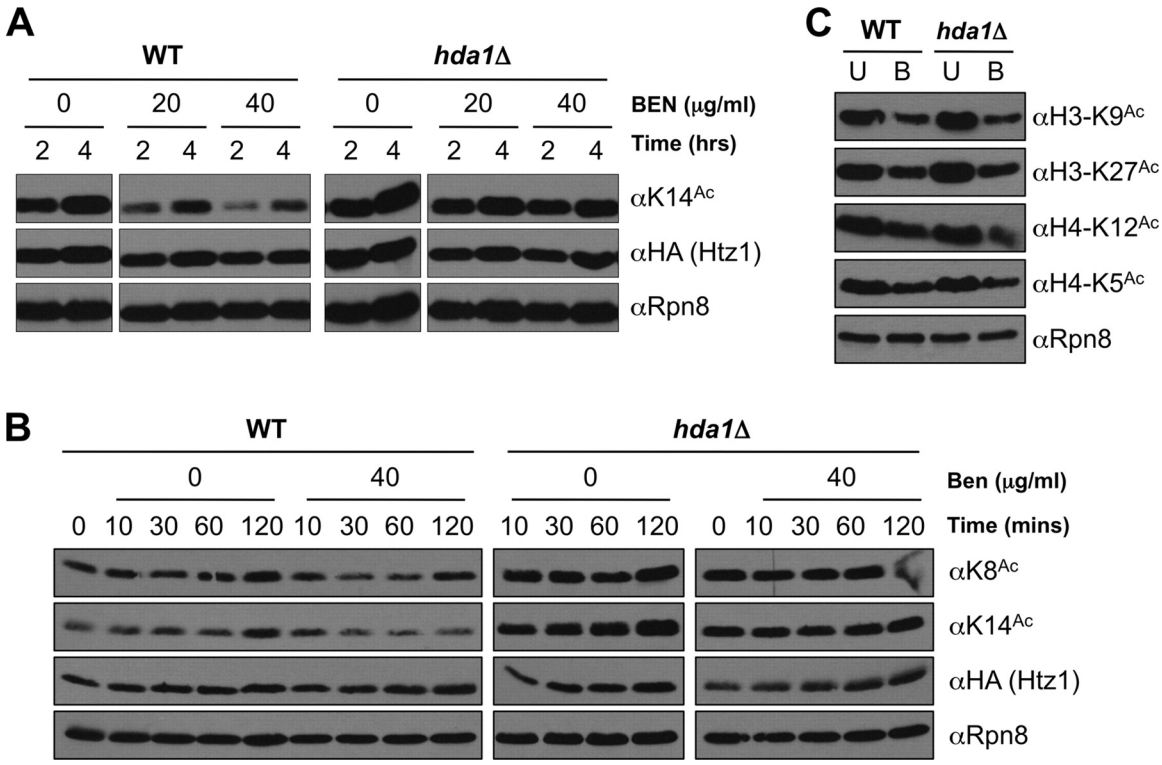


Figure 6. The benomyl-induced decrease of each Htz1^{Ac} is Hda1-dependent. *A*, cultures were grown in YPD to A_{600} ~0.5, benomyl was added (40 μg/ml) and incubated as indicated before WCEs were harvested and immunoblotted. *B*, Htz1^{Ac} is reduced within 30 min of benomyl treatment. The cultures were grown in YPD to A_{600} ~0.5 before benomyl addition (40 μg/ml). WCEs were harvested and immunoblotted as indicated. *C*, acetylations on histones H3 (K9^{Ac} and K27^{Ac}) and H4 (K5^{Ac} and K12^{Ac}) are reduced by benomyl but independently of Hda1. The cultures were grown in YPD to A_{600} ~0.5 before benomyl addition (40 μg/ml, 2 h). WCEs were harvested and immunoblotted as indicated. *U*, untreated; *B*, benomyl.

Supplementary Figures

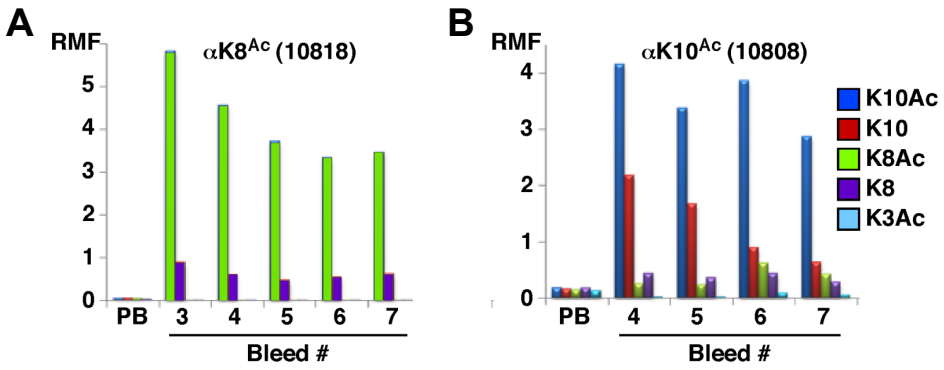


Figure S1. *The generation of Htz1Ac-specific antibodies* **(A)** αHtz1-K8Ac sera from rabbit 10818, or **(B)** αHtz1-K10Ac sera from rabbit 10808, were tested against a comprehensive peptide panel (sample shown) in Luminex assays. PB, Pre-Bleed; RMF, Relative Mean Fluorescence.

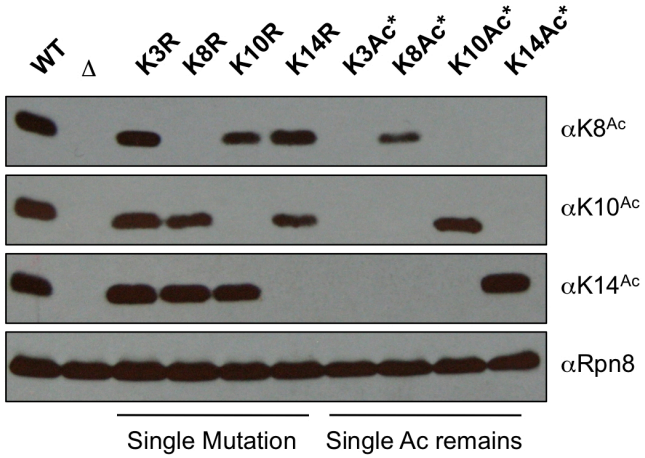


Figure S2. *The acetylation of individual Htz1 N-terminal lysines is not interdependent* Whole cell extracts from the indicated strains were immunoblotted as indicated. As an example *K8R* is the specific mutation of lysine 8 to arginine, while *K8Ac** is the

mutation of all lysines other than K8 (i.e. *K3R, K10R, K14R*). WT, wild type; Δ is *htz1*Δ. Rpn8 is a loading control.

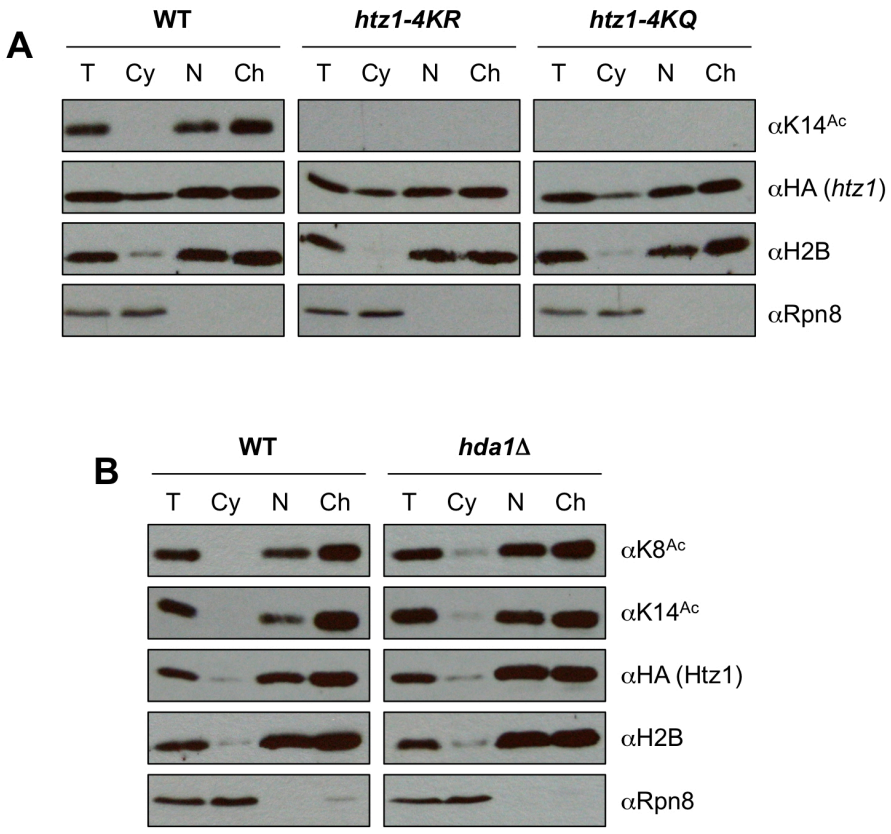
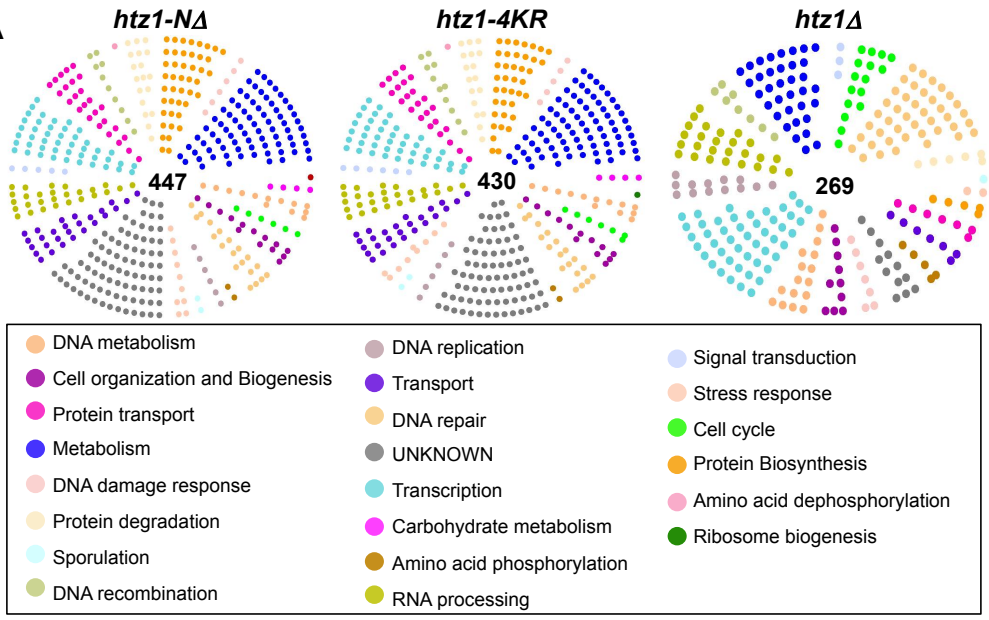


Figure S3. *The acetylation status of Htz1 does not regulate association of the variant with chromatin* Htz1 acetylation is abolished in *4KR* or *4KQ* and increased in *hda1*Δ cells. **(A - B)** Indicated strains were spheroblasted (Total), fractionated into Cytoplasm, Nucleus and Chromatin, and immunoblotted as indicated. αHA monitors the localization of *htz1.HA3*. Appropriate segregation of histone H2B and Rpn8 indicate efficient fractionation: the former is primarily localized in insoluble chromatin, the latter in soluble cytoplasm.

A



B

htz1	SS/SL	GO term	Category	P-value
Δ	272	GO:0006325	Establishment &/or maintenance of chromatin architecture	0
		GO:0016568	chromatin modification	0
		GO:0051276	chromosome organization & biogenesis	0
		GO:0006323	DNA packaging	0
		GO:0005667	transcription factor complex	1.11E-77
		GO:0006351	transcription, DNA-dependent	2.21E-62
		GO:0016570	histone modification	8.84E-54
		GO:0006974	response to DNA damage stimulus	2.00E-47
		GO:0032200	telomere organization & biogenesis	1.90E-45
		GO:0006366	transcription from RNA polymerase II promoter	3.83E-43
		GO:0006338	chromatin remodeling	5.79E-35
		GO:0045449	regulation of transcription	6.12E-35
4KR	386	GO:0032200	telomere organization and biogenesis	2.39E-45
		GO:0051276	chromosome organization and biogenesis	1.43E-43
		GO:0016570	histone modification	1.43E-22
		GO:0006351	transcription, DNA dependent	3.86E-20
		GO:0016568	chromatin modification	8.85E-20
		GO:0006323	DNA packaging	1.10E-17
		GO:0006366	transcription from RNA polymerase II promoter	7.26E-17
		GO:0007035	vacuolar acidification	5.46E-09
		GO:0045449	regulation of transcription	1.10E-08
		GO:0005667	transcription factor complex	1.71E-08
		GO:0006338	chromatin remodelling	7.89E-08
		GO:0006885	regulation of cellular pH	8.61E-08
NΔ	398	GO:0032200	telomere organization and biogenesis	1.85E-46
		GO:0051276	chromosome organization and biogenesis	1.67E-44
		GO:0016570	histone modification	1.08E-21
		GO:0016568	chromatin modification	6.67E-20
		GO:0006351	transcription, DNA dep	1.12E-19
		GO:0006323	DNA packaging	9.94E-18
		GO:0006366	transcription from RNA polymerase II promoter	8.22E-16
		GO:0007034	vacuolar transport	5.24E-09
		GO:0005667	transcription factor complex	6.03E-09
		GO:0007035	vacuolar acidification	7.50E-09
		GO:0007033	vacuole organization and biogenesis	7.55E-09
		GO:0045449	regulation of transcription	1.31E-08

Figure S4. *Genetic interactions show that [htz1-NΔ ≈ htz1-4KR] ≠ htz1Δ* **(A)** Gene Ontology (GO) categories for the synthetic sick / synthetic lethal (SS/SL) interactors of *htz1*Δ, *N*Δ or *4KR.* This highlights the relative similarity between these mutants: i.e. the latter two are most closely related. Numbers in the center of each group indicate their respective SS/SL interactors identified. Note that *N*Δ and *4KR* were mated to the non-essential deletion and DAmP collections, while *htz1*Δ was only mated to the former (hence the reduced number of interactors identified). Each map was generated with the OSPREY network visualization system (1). **(B)** The top hits for over-represented GO functional categories within the SS/SL interactors of *htz1*Δ, *4KR,* or *N*Δ. Each list was generated with GOstat, which considers and ranks the GO terms contained in the data (2). The precise SS/SL interrogation group differs from the Osprey list (e.g. *htz1*Δ: 269 or 272 interactors) as the gene space used by each program varies slightly.

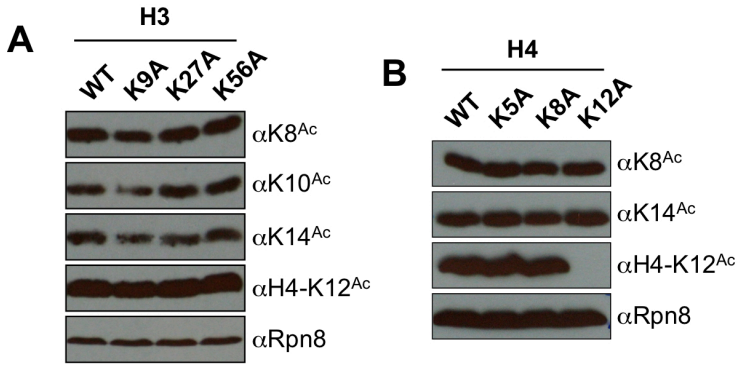


Figure S5. *Efficient Htz1Ac does not require individual Asf1-dependent acetylations on histones H3 or H4* **(A - B)** WCEs from the indicated unacetylatable histone H3 or H4 point mutants (from the SHIMA collection (3)) were immunoblotted as indicated. H4-K12Ac and Rpn8 are controls.

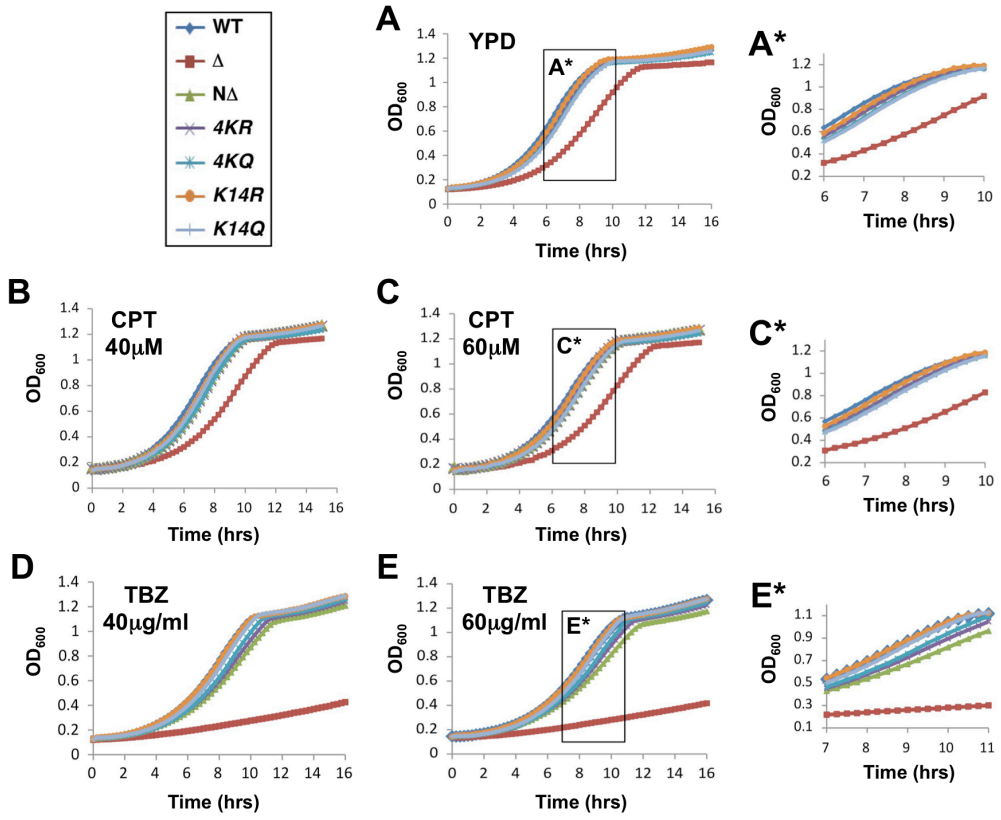


Figure S6. *Htz1Ac is required for resistance to TBZ but not to Camptothecin* **(A)** Growth curves on rich non-selective media (YPD) were analyzed on a Bioscreen (see Experimental Procedures). **A***: Boxed area (from **A**) is re-plotted to facilitate comparison across strains. **(B - C)** *htz1*Δ shows a mild sensitivity to the topoisomerase inhibitor camptothecin (CPT) on plates (Fig. 5A) but not in liquid culture. Growth curves on YPD containing CPT (μM as indicated) were analyzed on a Bioscreen. **C***: Boxed area (from **C**) is re-plotted to facilitate comparison across strains. **(D - E)** *htz1*Δ, and to a lesser degree *N*Δ, *4KR* or *4KQ*, are sensitive to the microtubule inhibitor TBZ. Growth curves on YPD containing TBZ (μg/ml as indicated) were analyzed on a Bioscreen. **E***: Boxed area (from **E**) is re-plotted to facilitate comparison across strains.

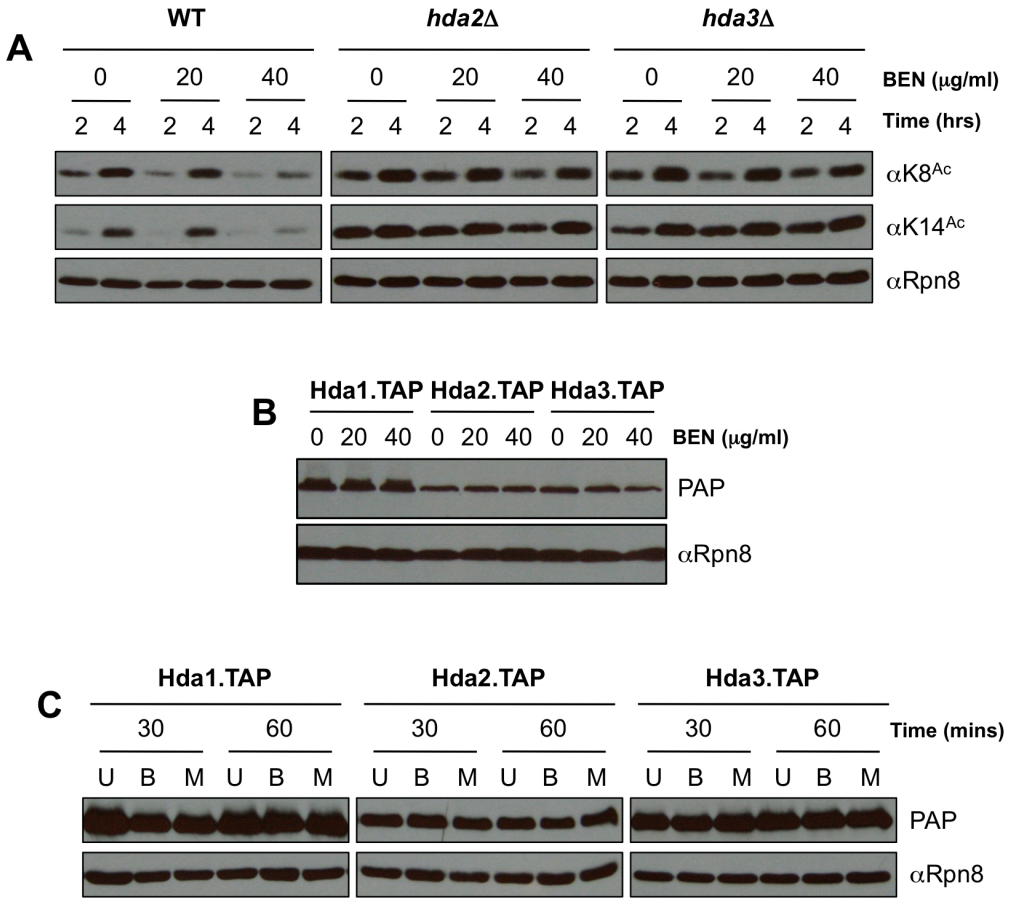


Figure S7. *Htz1Ac is regulated by benomyl via the Hda1 deacetylase complex* **(A)** The reduction of Htz1Ac on benomyl treatment is dependent on all members of the Hda1-complex. Strains were grown in YPD to OD600 ~ 0.5 and benomyl (BEN; μg/ml) added. WCEs were harvested and immunoblotted as indicated. Rpn8 is a loading control. WT, Wild type. **(B)** Benomyl has no impact on the abundance of Hda1, Hda2 or Hda3. Strains containing C-terminally TAP-tagged forms were grown in YPD to OD600 ~ 0.5 and benomyl (BEN; μg/ml) added for two hours. WCEs were immunoblotted and probed as indicated. PAP is Peroxidase anti- peroxidase to detect the immunoglobulin component of the TAP tag. Rpn8 is a loading control. **(C)** Hda1-complex subunits show no mobility shift (which could indicate post-

translational modifications such as ubiquitylation or phosphorylation) after treatment of cells with benomyl (B; 40μg/ml) or MMS (M; 0.05%). Gels in the PAP panel were run for significantly longer than those in **B** to increase resolution. Rpn8 is a loading control. U, untreated.

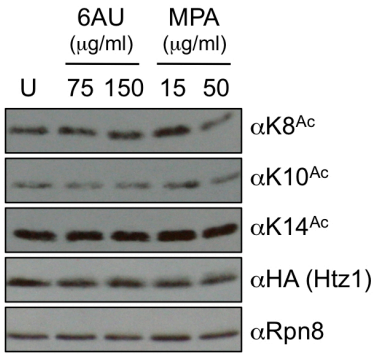


Figure S8. *Htz1Ac levels are unaffected by 6AU or MPA treatment* URA+ wild type cells were grown in synthetic complete medium (lacking uracil) to OD600 ~ 0.5 and 6-azauracil (6AU) or mycophenolic acid (MPA) added at concentration indicated before growth for another two hours. WCEs were harvested and immunoblotted as indicated. Rpn8 is a loading control. U, Untreated.

Supplementary Tables

Available online at:

http://www.jbc.org/content/early/2010/10/14/jbc.M110.185967/suppl/DC1

References

Babiarz J. E., Halley J. E., Rine J. (2006) Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae. Genes Dev. *20*, 700–710.

Bird A. W., Yu D. Y., Pray-Grant M. G., Qiu Q., Harmon K. E., Megee P. C., Grant P. A., Smith M. M., Christman M. F. (2002) Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. Nature *419*, 411–415.

Breslow D. K., Cameron D. M., Collins S. R., Schuldiner M., Stewart-Ornstein J., Newman H. W., Braun S., Madhani H. D., Krogan N. J., Weissman J. S. (2008) A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. Nat. Methods *5*, 711–718.

Bruce K., Myers F. A., Mantouvalou E., Lefevre P., Greaves I., Bonifer C., Tremethick D. J., Thorne A. W., Crane-Robinson C. (2005) The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. Nucleic Acids Res. *33*, 5633–5639.

Caterino T. L., Hayes J. J. (2007) Chromatin structure depends on what's in the nucleosome's pocket. Nat. Struct. Mol. Biol. *14*, 1056–1058.

Collins S. R., Schuldiner M., Krogan N. J., Weissman J. S. (2006) A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol. *7*, R63.

Collins S. R., Miller K. M., Maas N. L., Roguev A., Fillingham J., Chu C. S., Schuldiner M., Gebbia M., Recht J., Shales M., Ding H., Xu H., Han J., Ingvarsdottir K., Cheng B., Andrews B., Boone C., Berger S. L., Hieter P., Zhang Z., Brown G. W., Ingles C. J., Emili A., Allis C. D., Toczyski D. P., Weissman J. S., Greenblatt J. F., Krogan N. J. (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806–810.

Dion M. F., Kaplan T., Kim M., Buratowski S., Friedman N., Rando O. J. (2007) Dynamics of replication-independent histone turnover in budding yeast. Science *315*, 1405–1408.

Dryhurst D., Ishibashi T., Rose K. L., Eirín-López J. M., McDonald D., Silva-Moreno B., Veldhoen N., Helbing C. C., Hendzel M. J., Shabanowitz J., Hunt D. F., Ausió J. (2009) Characterization of the histone H2A.Z-1 and H2A.Z-2 isoforms in vertebrates. BMC Biol. *7*, 86.

Fiedler D., Braberg H., Mehta M., Chechik G., Cagney G., Mukherjee P., Silva A. C., Shales M., Collins S. R., van Wageningen S., Kemmeren F., Holstege F. C., Weissman J. S., Keogh M. C., Koller D., Shokat K. M., Krogan N. J. (2009) Functional organization of the S. cerevisiae phosphorylation network. Cell *136*, 952–963.

Fillingham J., Recht J., Silva A. C., Suter B., Emili A., Stagljar I., Krogan N. J., Allis C. D., Keogh M. C., Greenblatt J. F. (2008) Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Mol. Cell. Biol. *28*, 4342–4353.

Gauss R., Trautwein M., Sommer T., Spang A. (2005) New modules for the repeated internal and N-terminal epitope tagging of genes in Saccharomyces cerevisiae. Yeast *22*, 1–12.

Halley J. E., Kaplan T., Wang A. Y., Kobor M. S., Rine J. (2010) Roles for H2A.Z and its acetylation in GAL1 transcription and gene induction, but not GAL1-transcriptional memory. PLoS Biol. *8*, e1000401.

Henikoff S., Furuyama T., Ahmad K. (2004) Histone variants, nucleosome assembly and epigenetic inheritance. Trends Genet *20*, 320–326.

Ishibashi T., Dryhurst D., Rose K. L., Shabanowitz J., Hunt D. F., Ausió J. (2009) Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome. Biochemistry *48*, 5007–5017.

Kanta H., Laprade L., Almutairi A., Pinto I. (2006) Suppressor analysis of a histone defect identifies a new function for the hda1 complex in chromosome segregation. Genetics *173*, 435–450.

Keogh M. C., Cho E. J., Podolny V., Buratowski S. (2002) Kin28 is found within TFIIH and a Kin28-Ccl1-Tfb3 trimer complex with differential sensitivities to T-loop phosphorylation. Mol. Cell. Biol. *22*, 1288–1297.

Keogh M. C., Mennella T. A., Sawa C., Berthelet S., Krogan N. J., Wolek A., Podolny V., Carpenter L. R., Greenblatt J. F., Baetz K., Buratowski S. (2006a) The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes Dev. *20*, 660–665.

Keogh M. C., Kim J. A., Downey M., Fillingham J., Chowdhury D., Harrison J. C., Onishi M., Datta N., Galicia S., Emili A., Lieberman J., Shen X., Buratowski S., Haber J. E., Durocher D., Greenblatt J. F., Krogan N. J. (2006b) A phosphatase complex that dephosphorylates gammaH2AX regulates DNA damage checkpoint recovery. Nature *439*, 497–501.

Kim H. S., Vanoosthuyse V., Fillingham J., Roguev A., Watt S., Kislinger T., Treyer A., Carpenter L. R., Bennett C. S., Emili A., Greenblatt J. F., Hardwick K. G., Krogan N. J., Bähler J., Keogh M. C. (2009) An acetylated form of histone H2A.Z regulates chromosome architecture in Schizosaccharomyces pombe. Nat. Struct. Mol. Biol. *16*, 1286–1293.

Krogan N. J., Keogh M. C., Datta N., Sawa C., Ryan O. W., Ding H., Haw R. A., Pootoolal J., Tong A., Canadien V., Richards D. P., Wu X., Emili A., Hughes T. R., Buratowski S., Greenblatt J. F. (2003) A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol. Cell *12*, 1565–1576.

Kusch T., Florens L., Macdonald W. H., Swanson S. K., Glaser R. L., Yates J. R., 3rd, Abmayr S. M., Washburn M. P., Workman J. L. (2004) Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. Science *306*, 2084–2087.

Lall S. (2007) Primers on chromatin. Nat. Struct. Mol. Biol. *14*, 1110–1115.

Latham J. A., Dent S. Y. (2007) Cross-regulation of histone modifications. Nat. Struct. Mol. Biol. *14*, 1017–1024.

Lin Y. Y., Qi Y., Lu J. Y., Pan X., Yuan D. S., Zhao Y., Bader J. S., Boeke J. D. (2008) A comprehensive synthetic genetic interaction network governing yeast histone acetylation and deacetylation. Genes Dev. *22*, 2062–2074.

Luger K. (2003) Structure and dynamic behavior of nucleosomes. Curr. Opin. Genet. Dev. *13*, 127–135.

Mateus C., Avery S. V. (2000) Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. Yeast *16*, 1313–1323.

Mehta M., Kim H. S., Keogh M. C. (2010) Sometimes one just isn't enough: do vertebrates contain an H2A.Z hyper-variant? J. Biol. *9*, 3.

Millar C. B., Xu F., Zhang K., Grunstein M. (2006) Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. Genes Dev. *20*, 711–722.

Millar C. B., Grunstein M. (2006) Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. Nat. Rev. Mol. Cell Biol. *7*, 657–666.

Mizuguchi G., Shen X., Landry J., Wu W. H., Sen S., Wu C. (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343–348.

Morillo-Huesca M., Clemente-Ruiz M., Andújar E., Prado F. (2010) The SWR1 histone replacement complex causes genetic instability and genome-wide transcription misregulation in the absence of H2A.Z. PLoS One *5*, e12143.

Nakanishi S., Sanderson B. W., Delventhal K. M., Bradford W. D., Staehling-Hampton K., Shilatifard A. (2008) A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. Nat. Struct. Mol. Biol. *15*, 881–888.

Raisner R. M., Madhani H. D. (2006) Patterning chromatin: form and function for H2A.Z variant nucleosomes. Curr. Opin. Gen Dev *16*, 119–124.

Ransom M., Dennehey B. K., Tyler J. K. (2010) Chaperoning histones during DNA replication and repair. Cell *140*, 183–195.

Ren Q., Gorovsky M. A. (2001) Histone H2A.Z acetylation modulates an essential charge patch. Mol. Cell *7*, 1329–1335.

Ren Q., Gorovsky M. A. (2003) The nonessential H2A N-terminal tail can function as an essential charge patch on the H2A.Z variant N-terminal tail. Mol. Cell. Biol. *23*, 2778–2789.

Schneider J., Dover J., Johnston M., Shilatifard A. (2004) Global proteomic analysis of S. cerevisiae (GPS) to identify proteins required for histone modifications. Methods Enzymol. *377*, 227–234.

Schuldiner M., Collins S. R., Thompson N. J., Denic V., Bhamidipati A., Punna T., Ihmels J., Andrews B., Boone C., Greenblatt J. F., Weissman J. S., Krogan N. J. (2005) Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507–519.

Schuldiner M., Collins S. R., Weissman J. S., Krogan N. J. (2006) Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344–352.

Storici F., Resnick M. A. (2006) The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods Enzymol. *409*, 329–345.

Straube K., Blackwell J. S., Jr., Pemberton L. F. (2010) Nap1 and Chz1 have separate Htz1 nuclear import and assembly functions. Traffic *11*, 185–197.

Suka N., Suka Y., Carmen A. A., Wu J., Grunstein M. (2001) Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. Mol. Cell *8*, 473–479.

Tanabe M., Kouzmenko A. P., Ito S., Sawatsubashi S., Suzuki E., Fujiyama S., Yamagata K., Zhao Y., Kimura S., Ueda T., Murata T., Matsukawa H., Takeyama K., Kato S. (2008) Activation of facultatively silenced Drosophila loci associates with increased acetylation of histone H2AvD. Genes Cells *13*, 1279–1288.

Taverna S. D., Li H., Ruthenburg A. J., Allis C. D., Patel D. J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol. *14*, 1025–1040.

Thambirajah A. A., Li A., Ishibashi T., Ausió J. (2009) New developments in post-translational modifications and functions of histone H2A variants. Biochem. Cell Biol. *87*, 7–17.

Tong A. H., Boone C. (2006) Synthetic genetic array analysis in Saccharomyces cerevisiae. Methods Mol. Biol. *313*, 171–192.

Updike D. L., Mango S. E. (2006) Temporal regulation of foregut development by HTZ-1/H2A.Z and PHA-4/FoxA. PLoS Genetics *2*, e161.

Winzeler E. A., Shoemaker D. D., Astromoff A., Liang H., Anderson K., Andre B., Bangham R., Benito R., Boeke J. D., Bussey H., Chu A. M., Connelly C., Davis K., Dietrich F., Dow S. W., El Bakkoury M., Foury F., Friend S. H., Gentalen E., Giaever G., Hegemann J. H., Jones T., Laub M., Liao H., Liebundguth N., Lockhart D. J., Lucau-Danila A., Lussier M., M'Rabet N., Menard P., Mittmann M., Pai C., Rebischung C., Revuelta J. L., Riles L., Roberts C. J., Ross-MacDonald P., Scherens B., Snyder M., Sookhai-Mahadeo S., Storms R. K., Véronneau S., Voet M., Volckaert G., Ward T. R., Wysocki R., Yen G. S., Yu K., Zimmermann K., Philippsen P., Johnston M., Davis

R. W. (1999) Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901–906.

Wu J., Suka N., Carlson M., Grunstein M. (2001) TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast. Mol. Cell *7*, 117–126.

Zabaronick S. R., Tyler J. K. (2005) The histone chaperone anti-silencing function 1 is a global regulator of transcription independent of passage through S phase. Mol. Cell. Biol. *25*, 652–660.

Zlatanova J., Thakar A. (2008) H2A.Z: view from the top. Structure *16*, 166–179.

Zofall M., Fischer T., Zhang K., Zhou M., Cui B., Veenstra T. D., Grewal S. I. (2009) Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. Nature *461*, 419–422.

Chapter 9

RNA Polymerase II Carboxyl-terminal Domain Phosphorylation Regulates Protein Stability of the Set2 Methyltransferase and Histone H3 Di- and Trimethylation at Lysine 36

RNA Polymerase II Carboxyl-terminal Domain Phosphorylation Regulates Protein Stability of the Set2 Methyltransferase and Histone H3 Di- and Trimethylation at Lysine 36

Stephen M. Fuchs[‡§], Kelby O. Kizer[‡,1], Hannes Braberg[¶], Nevan J. Krogan[¶], Brian D. Strahl[‡§,2]

From the [‡]Department of Biochemistry and Biophysics and

[§]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599 and

the [¶]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158

[2] To whom correspondence should be addressed: brian_strahl@med.unc.edu.

[1]Present address: Physical Sciences Directorate, US Army Research Office, Research Triangle Park, NC 27709.

Abstract

Methylation of lysine 36 on histone H3 (H3K36) is catalyzed by the Set2 methyltransferase and is linked to transcriptional regulation. Previous studies have shown that trimethylation of H3K36 by Set2 is directed through its association with the phosphorylated repeats of the RNA polymerase C-terminal domain (RNAPII CTD). Here, we show that disruption of this interaction through the use of yeast mutants defective in CTD phosphorylation at serine 2 results in a destabilization of Set2 protein levels and H3K36 methylation. Consistent with this, we find that Set2 has a short half-life and is co-regulated, with RNAPII CTD phosphorylation levels, during logarithmic growth in yeast. To probe the functional consequence of uncoupling Set2-RNAPII regulation, we expressed a truncated and more stable form of Set2 that is capable of dimethylation but not trimethylation *in vivo*. Results of high throughput synthetic genetic analyses show that this Set2 variant has distinct genetics from either *SET2* or *set2*Δ and is synthetically sick or lethal with a number of transcription elongation mutants. Collectively, these results provide molecular insight into the regulation of Set2 protein levels that influence H3K36 methylation states.

Introduction

The N-terminal tails of histones are richly decorated with post-translational modifications that regulate a variety of DNA-templated processes through the recruitment of additional factors to chromatin. These factors interact with modifications such as acetylation, methylation, and phosphorylation through specific protein domains such as bromo-, chromo-, PHD, and TUDOR domains (Yun et al., 2011). Furthermore, these domains are usually found in the context of multiprotein complexes that perform diverse functions in chromatin remodeling, replication, transcription, and DNA repair, among others (Gardner et al., 2011; Ruthenburg et al., 2007; Taverna et al., 2007). Thus, there is great interest in both identifying biological important post-translational modifications on histones and understanding how these modifications function within chromatin.

One such post-translational modification is methylation of lysine 36 on histone H3. In yeast, H3K36[3] is modified by the methyltransferase Set2, which can add one, two, or three methyl groups to the ϵ-amino group of the lysine side chain using *S*-adenosylmethionine as the methyl donor (Strahl et al., 2002). Genome-wide chromatin immunoprecipitation studies have demonstrated that H3K36 methylation is generally associated with actively transcribed genes; however, dimethylation (H3K36me2) and trimethylation (H3K36me3) show different patterns of localization (Pokholok et al., 2005; Rando, 2007; Rao et al., 2005). H3K36me2 appears uniformly distributed across transcribed genes and is not correlated with transcription frequency, whereas H3K36me3 seems to be most prevalent at the 3′ end of coding regions and is positively correlated with the rate of

transcription. These data suggest that the deposition of different levels of H3K36 methylation is controlled within the cell and that the different degrees of this modification likely direct distinct biological functions.

H3K36 methylation has been reported to be recognized by the Rpd3(S) deacetylase complex in *Saccharomyces cerevisiae* through the chromodomain of Eaf3 (a member of both the Rpd3(S) deacetylase and NuA4 histone acetyltransferase complexes) (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). Subsequent studies revealed that H3K36me2 is the preferred substrate for Rpd3(S) binding (Li et al., 2007a; Li et al, 2009; Youdell et al., 2008) and that a primary function for H3K36 methylation is to recruit the repressive activity of Rpd3(S) to genes to "reset" or reestablish chromatin structure between multiple rounds of transcription (Cheung et al., 2008; Li et al, 2007b; Lickwar et al., 2009). Such resetting, or at least maintaining a more compact chromatin structure within gene bodies, also protects genes from spurious transcription from improper initiation sites (cryptic transcription) (Cheung et al., 2008; Li et al, 2007b; Lickwar et al., 2009; Kaplan et al., 2003). To date, no separate function has been reported for H3K36me3. However, H3K36me3 can be recognized by both the chromodomain of Eaf3 and *in vitro* by the PHD domain of Nto1, a member of the NuA3 histone acetyltransferase complex, suggesting that H3K36me3 might have a positive role in transcription through the recruitment of histone acetyltransferase complexes (Bua et al., 2009). Consistent with this, NuA3 association to chromatin has been shown to be partially dependent on Set2 methylation (Martin et al., 2006).

Set2 interacts with the repeating C-terminal domain (CTD) of RNA polymerase II (RNAPII) (Li et al., 2007a; Kizer et al., 2005; Krogan et al., 2003; Li et al., 2003; Schaft et al., 2003; Xiao et al., 2003). In yeast, the CTD of RNAPII consists of 26 repeats of a heptapeptide repeat sequence (YSPTSPS) in which all three serine residues can be phosphorylated (Chapman et al., 2007; Egloff et al., 2007; Fuchs et al., 2009). Set2 preferentially recognizes the Ser^2/Ser^5-modified form of the CTD through a region at the C terminus of Set2 termed the SRI (Set2-Rpb1 interacting) domain (Kizer et al., 2005; Vojnic et l., 2006). Consistent with this, H3K36me3 is well correlated with Ser^2- and Ser^2/Ser^5-modified forms of the CTD that occur within the bodies and 3′ ends of genes (Pokholok et al., 2005; Buratowski, 2003; Phatnani and Greenleaf , 2006; Tietjen et al., 2010).

We recently reported that deletion of the kinase responsible for Ser^2 phosphorylation of the CTD, Ctk1, results not only in loss of H3K36me but also decreased Set2 protein levels (Youdell et al., 2008). This finding suggested a novel mechanism for regulating histone methyltransferase activity. Here, we present a molecular basis for the regulation of Set2 protein levels. We find that events that alter or perturb Ser^2 CTD phosphorylation levels result in reduced levels of Set2 protein and global loss of H3K36me3 without much change in H3K36me2. Furthermore, we demonstrate that expression of a stable truncation of Set2, incapable of H3K36me3, greatly compromises the growth of several transcription elongation mutants. Lastly, high throughput epistatic (E-MAP) analysis reveals different interactions for strains capable of catalyzing different levels of H3K36 methylation *in vivo*. Collectively, our data delineate a novel mechanism for Set2

regulation and support the model that different degrees of H3K36 methylation likely harbor distinct functions.

Experimental Procedures

Yeast Strains and Plasmids

Strains used in this study are detailed in supplemental Table S1. Strains created for this study were constructed using heterologous gene replacement (Gelbart et al., 2001; Janke et al., 2004) or through classical yeast mating in combination with tetrad dissection or high throughput strain construction techniques (Tong and Boone, 2007). Set2 expression plasmids (Youdell et al., 2008; Kizer et al., 2005) were transformed using the lithium chloride technique (Gietz and Schiestl, 2007), and shuffle strains (where important genes were genomically deleted but carried on a uracil-selectable plasmid) were maintained on selective media until the time of the appropriate experiment, and plasmid loss was instigated by plating on media containing 5-fluoroorotic acid (5-FOA) (Boeke et al., 1987).

Western Blot Analysis

Yeast cultures were grown in either rich medium (YPD) or synthetic defined medium from a starting A_{600} of between 0.1 and 0.2 and grown to mid-log phase (A_{600} 0.6–1.0) with the exception of cultures started from stationary phase. In this instance, 200-ml cultures were started from saturated cultures of yeast grown in synthetic defined medium for >7 days and inoculated at an initial A_{600} of 0.5. In all

cases, five optical units of cells were harvested by centrifugation, and extracts were prepared using a protocol adapted from Ref. 38. Briefly, cell pellets were resuspended in 75 µl of SUMEB buffer (1% SDS, 8 M urea, 10 mM MOPS, pH 6.8, 10 mM EDTA, 0.01% bromphenol blue), mixed with 200 µl of glass beads, and vortexed for 3 min. To the beads was added 175 µl of additional buffer, and the total supernatant was collected and clarified to give a total volume of 250 µl of cell extract. Extracts were heated at 95 °C for 5 min and frozen at −80 °C until use. For Western blot analysis of histones, 5 µl of extract was separated by 17% SDS-PAGE. For Western analysis of Set2 and RNAPII, proteins were separated by 8% SDS-PAGE. Gels were transferred to PVDF for 90 min at 45 mA and dried in methanol to block according to the manufacturer's instructions (Millipore-Immobilon-P). Dried membranes were rehydrated briefly in methanol and incubated with primary antibodies overnight at 4 °C in PBST (phosphate-buffered saline including 0.05% Tween 20) containing 5% blocking (Pierce). Antibodies were used at the following dilutions: anti-H3K36me2 (Active Motif), 1:1,000; anti-H3K36me3 (Abcam ab9050), 1:2,500; anti-H3 1:5,000; anti-Set2 (1:8,000), anti-G6PDH (1:100,000), anti-Ser(P)^2 CTD (gift of D. Eick), 1:200. Western blots were visualized using HRP-conjugated secondary antibodies and ECL Plus chemiluminescence (GE Healthcare).

Protein Turnover Experiments

Turnover experiments were adapted from a protocol from Ref. 39. Briefly, 100-ml cultures of yeast in YPD were grown to an A_{600} of ~0.6, and cycloheximide was added to a final concentration of 200 µg/ml. Five optical units of cells were

taken at various time points, and 1 ml of 10 mM sodium azide and 1 ml of dimethyl sulfoxide were added. Cells were then immediately frozen in liquid nitrogen. Frozen samples were thawed simultaneously, and extracts were prepared using the SUMEB method described above.

Yeast Growth Assays

For phenotypic growth assays, yeast were grown overnight in appropriate media (YPD or synthetic complete medium with necessary additives). Saturated cultures were used to start fresh cultures in the same medium at an A_{600} of ~0.2. Cells were allowed to double at least two times, and cells were harvested and resuspended to an A_{600} of 1.0 in sterile water in a 96-well plate. Cells were 5-fold serially diluted and then spotted onto plates containing necessary selections using a 48-pin replicating tool. Plates were incubated at indicated temperatures and imaged daily.

Epistatic Miniarray Analysis

E-MAP analysis was carried out as described (Schuldiner et al., 2006; Tong et al., 2001). Growth rates were assessed by automated image analysis of colony size (Collins et al., 2007).

Results

Set2 and H3K36 Methylation Are Regulated Differentially during Cell Growth

It has been well established that methylation by Set2 is correlated with hyperphosphorylation of the CTD of RNAPII. More than a decade ago, West and Corden showed that CTD phosphorylation was highly correlated with log phase growth and the diauxic shift in yeast (West and Corden, 1995). We therefore measured Set2 protein and histone modifications in different stages of yeast growth. Starting from stationary phase cells, a 200-ml culture of yeast in YPD was inoculated at an initial A_{600} of 0.5, and time points/A_{600} measurements were taken at regular intervals (Fig. 1*A*). Extracts from each time point were separated by SDS-PAGE and probed for Set2 as well as RNAPII phosphorylation and histone modifications by Western blot analysis. As shown in Fig. 1*B*, and consistent with previous findings (West and Corden, 1995), RNAPII and CTD phosphorylation levels peaked at late log phase and during diauxic shift (t = 6–8 h). Significantly, we found that Set2 protein levels also changed under these same conditions and correlated strongly with the occurrence of RNAPII CTD phosphorylation (Fig. 1*B*). In examination of H3K36 methylation levels, we observed that H3K36me2 was most abundant at t = 0 and 2 h but decreased during mid- to late log phase (Fig. 1*C*). In contrast, H3K36me3 levels were lowest at t = 0–4 h but increased during mid- to late log phase at the same time where Set2 levels and CTD phosphorylation levels increased (Fig. 1*C*). Although Set2 and CTD phosphorylation levels decreased during late log phase, H3K36me3 levels remained constant. These data uncover that Set2 and H3K36me2 and H3K36me3

are dynamically regulated and that Set2 proteins levels and the occurrence of H3K36me3 are tightly coupled to CTD phosphorylation.

Set2 Protein Levels Are Regulated in Vivo

We previously noted that Set2 protein was lost in *CTK1* delete cells (*ctk1*Δ), and this was not the result of changes in *SET2* gene expression (Youdell et al., 2008; Xiao et al., 2003). These data, combined with our observation that Set2 protein levels correlate with RNAPII CTD phosphorylation levels, prompted us to explore the regulation of the Set2 protein further. Accordingly, we examined the abundance of Set2 protein upon treatment with the translational inhibitor cycloheximide (Fig. 2*A*). Results showed in a time course experiment that Set2 exhibited a rapid turnover with a half-life of ~9 min (Fig. 2, *A* and *B*). In contrast, performing the same experiment in a temperature-sensitive proteasome-deficient strain (*cim3-1*) at the nonpermissive temperature showed increased stability of Set2 (Fig. 2, *A* and*B*).

We next sought to identify the region responsible for regulating Set2 stability. Using strains harboring truncation mutants of Set2 (Fig. 2*C*), we were able to map the region responsible for Set2 instability to several hundred amino acids of unknown function between the Set2 domain and the SRI domain responsible for interaction with RNAPII (Fig. 2*D*). Interestingly, analysis of Set2 and histone methylation levels in these strains revealed unexpected behavior of the different Set2 truncations. Deleting solely the SRI domain of Set2 (Set2(1–618)) resulted in no significant change in the abundance of Set2 protein but also no observable

H3K36me2 or H3K36me3 (Fig. 2*D*). However, expression of just the SET domain (1–261) of Set2 resulted in a marked increase in Set2 protein and H3K36me2 but no detectable H3K36me3 (Fig. 2*D*).

Set2 Protein Levels Are Decreased in Mutants That Alter RNAPII CTD Phosphorylation

Set2 is known to bind to the Ser^2/Ser^5-phosphorylated form of RNAPII. Because Ctk1 is the major kinase responsible for Ser^2 phosphorylation, we next wanted to determine whether the lower Set2 protein levels found in *ctk1*Δ strains (Fig. 3*A*) were related to the ability of Set2 to bind to the CTD of RNAPII. We therefore examined the levels of Set2 protein in a number of yeast mutants known to alter the phosphorylation state of the CTD. Set2 protein levels were noticeably decreased in mutants of the PAF complex, which is known to play an important role in transcription elongation and affect CTD Ser^2 phosphorylation (Fig. 3*B*) (Nordick et al., 2008). Significantly, Set2 protein levels were only decreased in PAF mutants that also exhibited changes in CTD Ser^2 phosphorylation, *cdc73*Δ, *paf1*Δ, and *ctr9*Δ. Decreased levels of Set2 correlated with decreased levels of H3K36me3 but not H3K36me2, consistent with the model that CTD association is necessary for H3K36me3 by Set2. We also note that overall histone H3 levels are reduced in both the*ctr9*Δ and *paf1*Δ strains, likely also resulting in the lower levels of even the cross-reacting band in these two strains (Fig. 3*B*). Furthermore, Set2 levels were decreased in a temperature-sensitive allele of Kin28, as well as mutant alleles of the phosphatases Ssu72 (*ssu72-2*), Fcp1 (*fcp1-1*), and the cyclin-like protein Bur2

(supplemental Fig. S1). Because many of the mutants we tested were enzymes themselves, we wanted to confirm that the phenotype of decreased Set2 protein was directly due to a loss changes in its ability to recognize the CTD rather than other potential activities of these enzymes. Accordingly, we looked at Set2 protein levels in cells expressing mutant forms of the RNAPII CTD itself (Fig. 3*C*). Each strain carries an RNAPII gene with a CTD comprising wild-type repeats (WT) or mutations at either Ser^2 or Ser^5 within the given number of repeats. Again, we observed a decrease in Set2 protein levels when a number of Ser^2 or Ser^5 residues within the CTD repeat were mutated to alanine. However, we saw a strong reduction in H3K36me3, which was accompanied by either no change or a slight increase in the levels of H3K36me2 in these cells.

Overexpression of More Stable Form of Set2(1–261) Is Lethal in Combination with Various Transcription Mutants

A well studied phenotype for *set2*Δ is the ability to bypass deletion of essential transcription factors such as members of the Bur1 complex and Spt16 (Keogh et al., 2005; Biswas et al., 2006; Chu et al., 2006). We therefore hypothesized that Set2 degradation might alter the levels of H3K36me2/me3 to facilitate transcription elongation during periods of transcriptional stress, when, for example, chromatin compaction by Rpd3(S) would inhibit rapid gene activation of stress response genes (Lenstra et al., 2011; Shivaswamy and Iyer, 2008). To examine this possibility, by way of using mutant forms of Set2 that have different behaviors in degradation and H3K36me, we expressed the different truncated forms of Set2

(see Fig. 2*C*) on plasmids in a *bur2*Δ *set2*Δ double deletion strain harboring a wild-type copy of *BUR2* and then analyzed the growth of these mutants after selection for the loss of the *BUR2* plasmid on 5-FOA (Fig. 4*A*). As expected, deletion of *SET2* bypasses the poor growth of *bur2*Δ strain. However, reintroduction of Set2(1–261) resulted in worse growth than either the *SET2* control or vector-only controls. These results suggest that the inability to degrade Set2 and/or reduce H3K36me2 levels results in impaired growth due to inappropriate transcriptional repression through maintained recruitment of Rpd3(S). Accordingly, we next deleted *RCO1* (*rco1*Δ), a unique subunit of the Rpd3(S) complex, to see whether Rpd3(S) activity was necessary for this lethality. In fact, *rco1*Δ was able to rescue the lethality caused by expression of Set2(1–261) in the *bur2* background (Fig. 4*A*).

High Throughput Genetic Analysis of Set2 Truncation Mutants

The truncated form of Set2(1–261) offered a unique tool to assess the biological consequence of having a stable form of catalytically active Set2, which is not able to catalyze H3K36me3. We therefore performed epistatic miniarray analysis against 1,534 yeast mutant strains (Fig. 4*B* and supplemental Table S1). Our analysis revealed several classes of interactions. For example Set2(1–261) and *set2*Δ exhibit similar negative (synthetically sick) interactions with members of the SWR1 and PAF complexes, whereas only Set2(1–261) results in synthetic sickness with members of the COMPASS complex (Fig. 4*B*). We were able to confirm several interactions by classical genetic analysis, most notably those with *SWR1* (Fig. 4*C*). Five representative tetrads from crosses between *set2* mutant constructs and

a *swr1*Δ strain are shown in Fig. 4*C*. We found that Set2(1–261), but neither Set2-FLAG nor *set2*Δ was lethal in combination with *swr1*Δ (Fig. 4*C* and Table 1). Because the SWR1 complex is primarily responsible for depositing the histone variant *HTZ1* into nucleosomes, we also tested genetic interactions between our *set2* mutant constructs and *HTZ1* (Table 1). In this instance, both deletion of *SET2* and expression of the Set2(1–261) construct resulted in synthetic lethality (supplemental Table S2).

Discussion

Using a combination of genetics and biochemistry, we have uncovered a significant mechanism for the regulation of Set2 methyltransferase activity in yeast. The SRI domain is known to bind to the CTD of RNAPII when it is phosphorylated at both the Ser^2 and Ser^5 positions within the CTD repeats. It was shown previously that disruption of this interaction, by deleting the CTD kinase Ctk1, resulted in a loss of H3K36 methylation without affecting Set2 mRNA levels (Xiao et al., 2003). Here, we show that loss of H3K36me3 results from changes in the stability of the Set2 enzyme itself, which we demonstrate is mediated via its interaction with the phosphorylated RNAPII CTD.

In support of a role for Set2-RNAPII interaction in governing Set2 stability, we found that mutations in enzymes that either act on the CTD (Ctk1, Kin28, Bur1, Ssu72, Fcp1), or mutations to the CTD sequence itself resulted in lower levels of Set2 protein and H3K36me3 (Fig. 3 and supplemental Fig. S1). Interestingly, we also found that deletion of *CTK1* altered the protein levels of other factors that associate

with the CTD, suggesting this mechanism may not be restricted to just Set2 (supplemental Fig. S2).

Set2 levels correlated strongly with changes in CTD phosphorylation as wild-type cells pass through log phase and into diauxic shift ($t \sim 8–10$ h) (Fig. 1, *A* and *B*). Interestingly, we observed H3K36me2 in whole cell extracts at time points where Set2 was barely detectible but H3K36me3 was not observed until log phase where Set2 and CTD Ser(P)^2 levels peaked (Fig. 1). These data indicate that the low levels of transcription found in stationary phase are accompanied by low levels of Set2 and CTD phosphorylation, yet H3K36me2 levels are maintained. However, during rapid log phase growth where robust transcription is needed, an increase in CTD phosphorylation accompanied by increased stability of Set2 and H3K36me3 levels is observed. Not only are these results consistent with previous findings that the interaction of Set2 with the CTD is required for H3K36me3 (Youdell et al., 2008), but they provide important insight into the dynamic regulation of H3K36me2 *versus* H3K36me3 and suggest that these two methyl marks have distinct functions in the cell.

Consistent with SRI regulating H3K36me3, several other transcription complexes or kinases have been shown to regulate H3K36 trimethylation, notably the PAF and BUR kinase complexes (Chu et al., 2006; Chu et al., 2007). Prelich and co-workers demonstrated several years ago that deletion of several PAF complex members, *cdc73*Δ, *ctr9*Δ, or*paf1*Δ, resulted in a decrease in H3K36me3 (Chu et al., 2007). Similarly, Jaehning and co-workers showed that CTD Ser^2 phosphorylation is reduced in these same three PAF mutants (Nordick et al., 2008). Our data show that

Set2 protein levels are also decreased in *cdc73*Δ, *ctr9*Δ, and *paf1*Δ strains (Fig. 3*B*). We take these data to conclude that decreased CTD phosphorylation in PAF mutants results in reduced Set2 protein levels which ultimately changes the balance of H3K36 methylation in the cell. Consistent with PAF effects of Ser^2 phosphorylation, we also see a decrease of Set2 levels in *BUR2* deletions (data not shown). Given that Bur1/Bur2 has recently been demonstrated to function on both Spt5 (DSIF) and the CTD (Zhou et al, 2009), it is plausible that Bur1 also affects H3K36me3/Set2 levels by directly influencing Ser^2phosphorylation and/or through DSIF recruitment of the PAF complex.

In our analysis of Set2 regulation, we noted that Set2 protein turnover was reduced in a temperature-sensitive allele of a proteasome subunit. We take these data to suggest that Set2 is degraded via the proteasome (Fig. 2). Consistent with this notion, our data are in agreement with a high throughput study of protein stability where it was predicted that Set2 is an unstable protein (Belle et al., 2006). However, perhaps due to the rapid turnover of Set2 or the relatively small amounts of enzyme normally found in cells, we have been unable to identify a ubiquitin ligase, ubiquitin protease, or a site of ubiquitin attachment to Set2 (data not shown).

A significant question to then ask is why would cells need to dynamically regulate Set2 levels? In addition to providing a rapid means of regulating H3K36 di- and trimethylation during logarithmic cell growth, the control of Set2 levels might have a role in the response to cellular stress as mentioned above, where the transient changes in H3K36 methylation may alter recruitment of the repressor Rpd3(S) to sites of transcription to alleviate chromatin repression during

transcription as proposed in Fig. 5. This would explain why deletion of *SET2* is able to bypass a number of transcription elongation mutations including mutations in Spt6, Spt16, and Bur1/2 (Keogh et al., 2005; Youdell et al., 2008; Biswas et al., 2006). In this scenario, removal of Set2 and hence both H3K36me2/me3, would result in more open chromatin structure and improved transcription of stress-response genes. In support of this idea, studies have shown that a wide number of stress-response genes are up-regulated in a *SET2* deletion (Lenstra et al., 2011).

As mentioned above, the ability to rapidly alter Set2 levels to control the balance of H3K36me2 and H3K36me3 supports the idea that H3K36me2 and H3K36me3 have different biological functions. To explore this possibility further, we examined the genetic interactions of a Set2 fragment that shows an inability to catalyze H3K36me3 (Set2(1–261)). Importantly, this Set2 fragment is properly recruited to gene coding regions in a transcription-dependent manner (Youdell et al., 2008). Our study compared the Set2(1–261) fragment against wild-type Set2, where all three methylation states at H3K36 are present, and a *SET2* deletion, where no H3K36 methylation is observed. We found that Set2(1–261) shows distinct genetics from either deletion of *SET2* or the wild-type case when crossed with mutations in other factors important for chromatin structure and/or transcription elongation. This is particularly significant because Set2(1–261) exhibits partial function. This could be attributed to an imbalance in the levels of H3K36 methylation and supports that Set2 has both positive and negative roles in regulating transcription. For example, a lack of H3K36me3 could lead to increased recruitment of repressive factors such as Rpd3(S) to H3K36me2, or it may result in

the inability to recruit effector proteins that recognize H3K36me3. Interestingly, Eaf3, the factor that recognizes H3K36me in the Rpd3(S) complex, is also an integral member of the NuA4 histone acetyltransferase complex. Furthermore, evidence shows that a member of the NuA3 histone acetyltransferase complex also interacts with H3K36me3 (and this complex is known to associate with chromatin in a Set2-dependent manner) (Bua et al., 2009; Martin et al., 2006). Perhaps then, H3K36me3 recruits activating complexes such as NuA3 and NuA4 to certain genes to facilitate transcription, and methylation at H3K36 can act as both an activating and repressive mark, as is seen with lysine methylation at H3K4 (Kim and Buratowski, 2009; Taverna et al, 2006). If this were the case, it is possible that, in response to changes in CTD phosphorylation caused by an environmental cue or transcriptional stress, degradation of Set2, rather than just disruption of the interaction between Set2 and the CTD is necessary to maintain some proper balance of H3K36me2/me3 in the cell. As Keogh *et al.* proposed several years ago, such a mechanism might also be used to attenuate regulation of certain genes or control transcription from normally silent cryptic promoters (Keogh et al., 2005).

Acknowledgments

We thank members of the Strahl laboratory for critical reading of this manuscript and Michael Keogh, Mike Hampsey, Bill Tansey, David Stillman, Danny Reinberg, Dirk Eick, Steve Hanes, and David Atencio for valuable technical suggestions, plasmids, yeast strains, and antibodies.

*This work was supported, in whole or in part, by National Institutes of Health Grant R01GM68088 (to B. D. S.). This work was also supported by National Research Service Award Postdoctoral Fellowship F32 GM80896 (to S. M. F.).

Figures

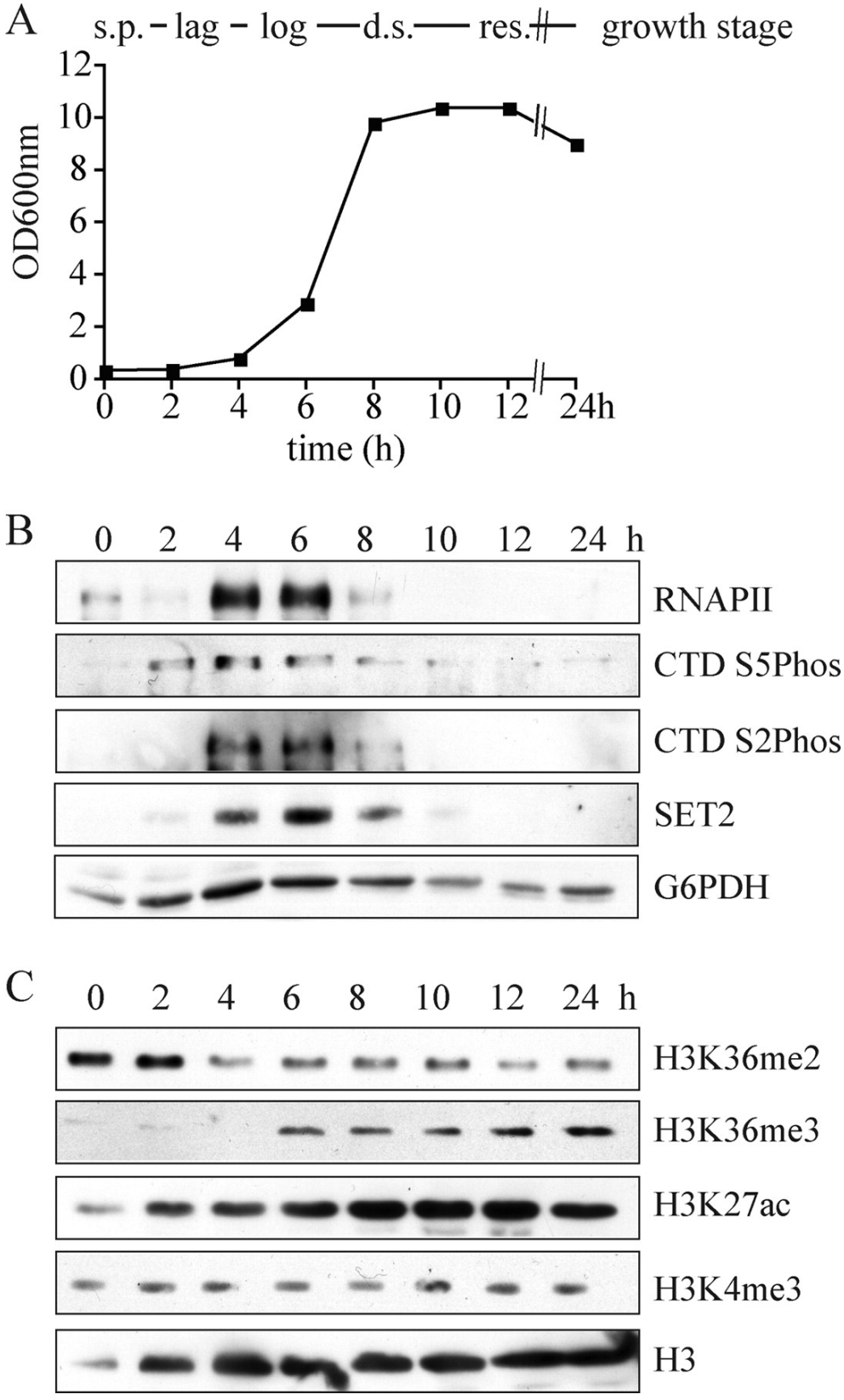

A
s.p.
lag
log
d.s.
res.
growth stage
OD600nm
12
10
8
6
4
2
0
0
2
4
6
8
10
12
24h
time (h)
B
0 2 4 6 8 10 12 24 h
RNAPII
CTD S5Phos
CTD S2Phos
SET2
G6PDH
C
0 2 4 6 8 10 12 24 h
H3K36me2
H3K36me3
H3K27ac
H3K4me3
H3

Figure 1. Set2, H3K36 methylation, and RNAPII phosphorylation levels are dynamically regulated during cell growth. *A*, growth curve of wild-type BY4741 cells in complete synthetic defined medium with approximate growth stages. *s.p.*, stationary phase; *lag*, lag phase; *log*, anaerobic fermentation of glucose, *d.s.*, diauxic shift; and *res.*, aerobic respiration of ethanol. *B*, Western blot analyses of extracts from the time course shown in *A*, indicating the presence of RNAPII as well as phosphorylated forms of the CTD and Set2. *C*, Western blot analyses from samples taken in *A* probing for the presence of various post-translational modifications (H3K36me2, H3K36me3, H3K4me3, and H3K27ac) on histone H3.

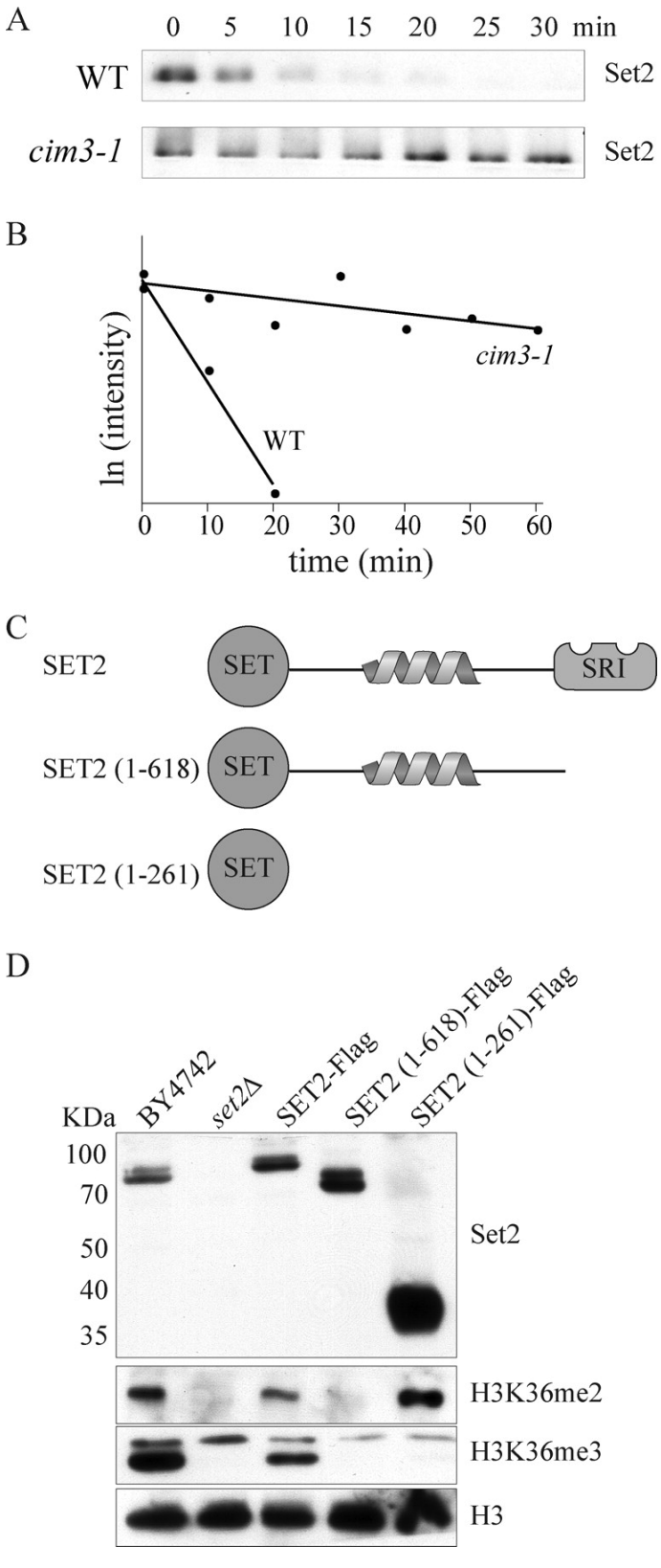


Figure 2. Set2 is a highly unstable protein. *A*, Western blot analysis of a time course experiment probing for Set2 protein levels in wild-type and proteasome-deficient (*cim3-1*) cells following addition of the translation inhibitor cycloheximide (200 µg/ml). *B*, quantification of Set2 protein levels as a function of time following cycloheximide treatment. Protein abundance was calculated by densitometry of Western blot analysis and plotted as the natural log of the intensity. *C*, schematic of Set2 domain structure and truncation mutants used in *D. D*, Western blot analyses showing the relative abundance of Set2 and the corresponding H3K36 methylation levels of several Set2 truncation mutants.

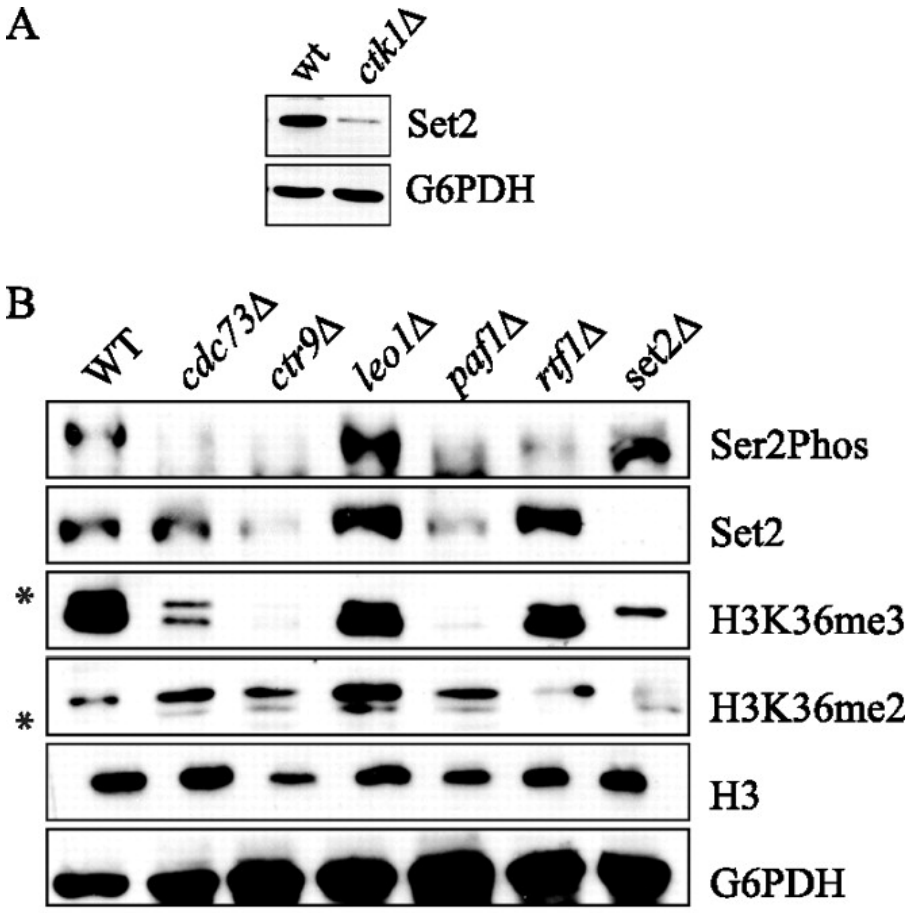

A
wt
ctk1Δ
Set2
G6PDH
B
WT
cdc73Δ
ctr9Δ
leo1Δ
paf1Δ
rtf1Δ
set2Δ
Ser2Phos
Set2
*
H3K36me3
*
H3K36me2
H3
G6PDH
C
26 repeats (WT)
10 CTD repeats
S2A(8) WT(7)
WT(7) S5A(5)
Ser2phos
Set2
H3K36me3
H3K36me2
H3
G6PDH

Figure 3. RNAPII CTD phosphorylation regulates Set2 protein levels. *A*, Set2 protein levels as measured by Western blotting of whole cell extract taken from WT and *ctk1*Δ strains. *B*, Western blots of extracts taken from PAF complex deletions strains as well as *set2*Δ cells to measure Set2 and H3K36 methylation. *Ser2phos* refers to phosphorylation at Ser^2 within the heptapeptide repeat of the RNAPII CTD. * denotes a cross-reactive band for the indicated antibody. *C*, Set2 and H3K36 methylation levels in RNAPII CTD mutants. *S2A* and *S5A* refer to mutations at either Ser^2 or Ser^5 within the CTD heptapeptide repeat, respectively. The *number* in *parentheses* refers to the number of tandem repeats in a given CTD sequence. The *black line* between the *third* and *fourth lanes* signifies that two unrelated lanes were removed during the editorial process. In all cases, all of the lanes originate from the same gel.

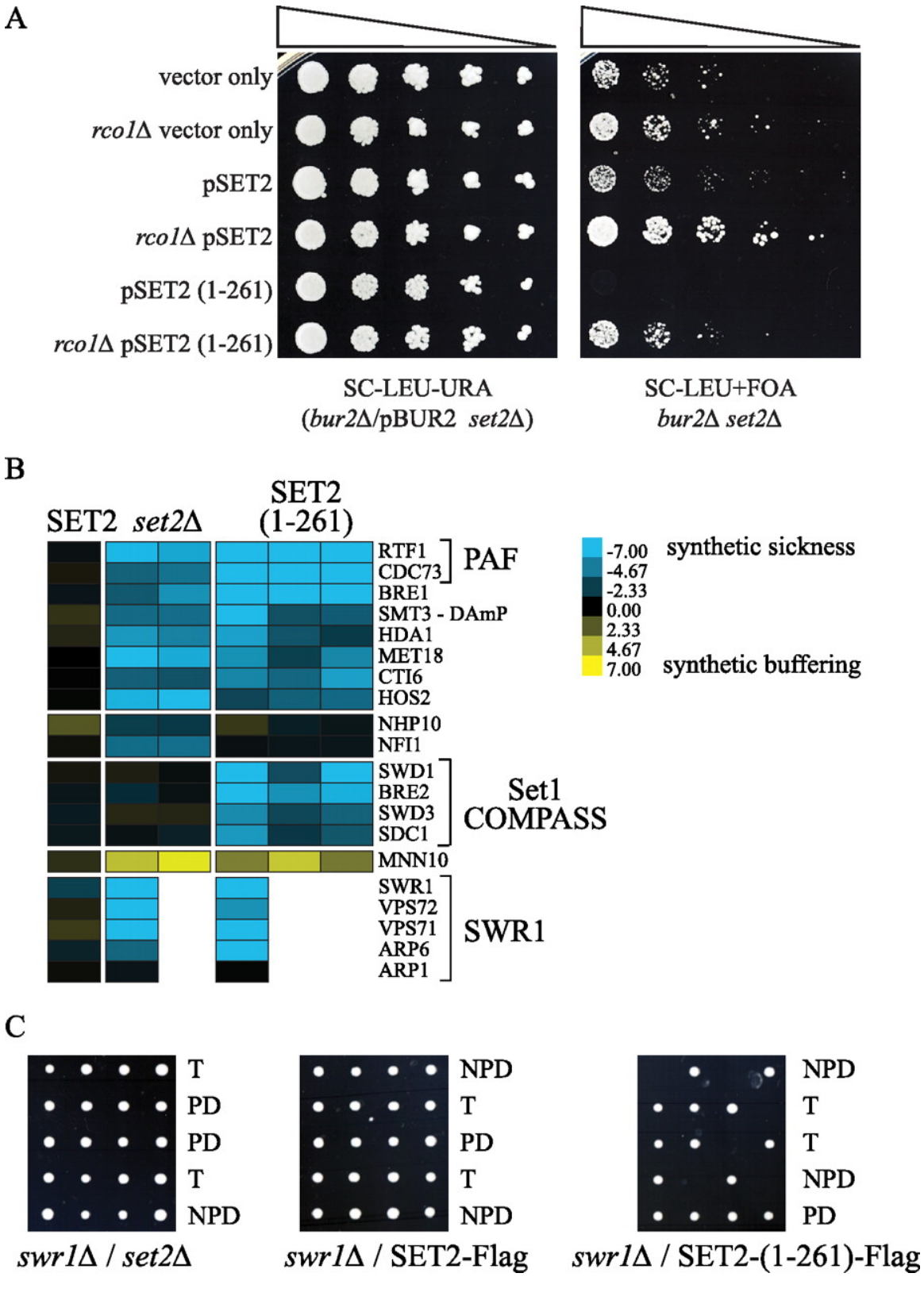

A
vector only
rco1Δ vector only
pSET2
rco1Δ pSET2
pSET2 (1-261)
rco1Δ pSET2 (1-261)
SC-LEU-URA
(bur2Δ/pBUR2 set2Δ)
SC-LEU+FOA
bur2Δ set2Δ
B
SET2
SET2 set2Δ
(1-261)
RTF1
CDC73
PAF
BRE1
SMT3 - DAmP
HDA1
MET18
CTI6
HOS2
NHP10
NFI1
SWD1
BRE2
SWD3
SDC1
Set1
COMPASS
MNN10
SWR1
VPS72
VPS71
ARP6
ARP1
SWR1
-7.00
-4.67
-2.33
0.00
2.33
4.67
7.00
synthetic sickness
synthetic buffering
C
T
PD
PD
T
NPD
swr1Δ / set2Δ
NPD
T
PD
T
NPD
swr1Δ / SET2-Flag
NPD
T
T
NPD
PD
swr1Δ / SET2-(1-261)-Flag

Figure 4. Set2(1–261) shows unique genetic interactions with transcription mutants. *A*, growth of *bur2*Δ*/set2*Δ or *bur2*Δ*/set2*Δ*/rco1*Δ strains harboring Set2 truncation plasmids plated on media to select for (−URA) or against (+FOA) a plasmid carrying a wild-type copy of *BUR2* (pBUR2). *B*, genetic interactions between *set2* mutants and a subset of the yeast deletion mutant collection. *Blue color* denotes combined sickness in the double mutant where yellow denotes better than expected growth (Schuldiner et al., 2006). *B*, representative spores grown on YPD from the dissection of genetic crosses between *swr1*Δ and *set2* mutant strains (*PD*, parental ditype; *NPD*, nonparental ditype; *T*, tetratype).

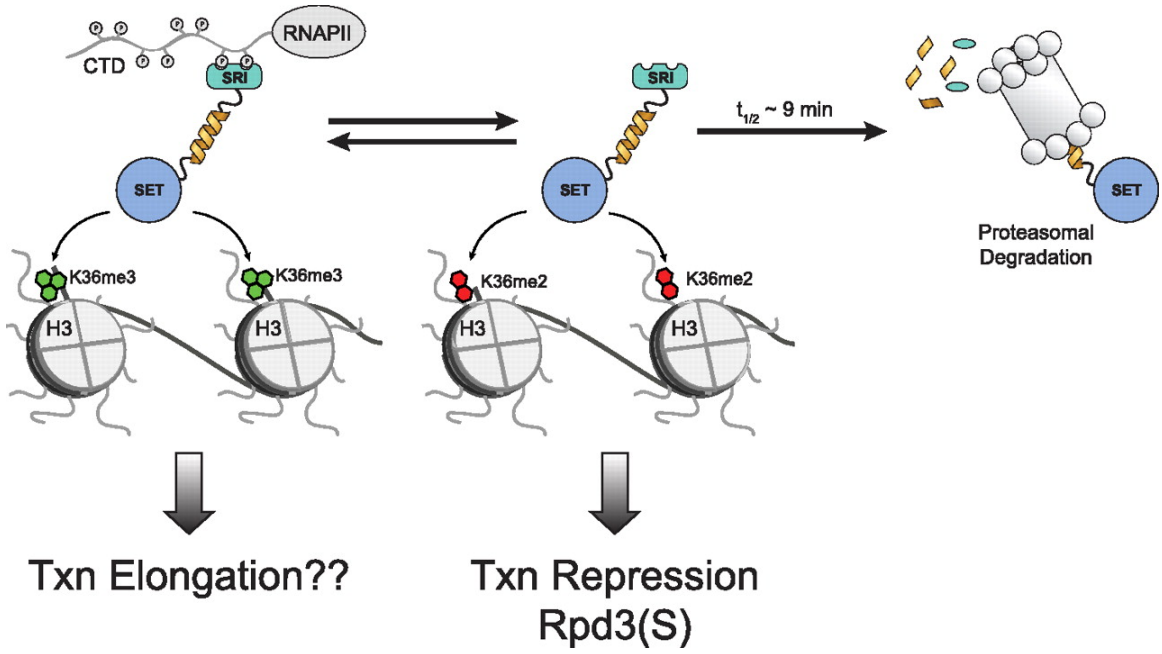


Figure 5. Proposed model of Set2 regulation. Set2 is in equilibrium between a free state and association with the hyperphosphorylated form of RNAPII. The free form of Set2 is degraded in a proteasome-dependent manner with a half-life in yeast of ~9 min. In its free form, Set2 is capable of dimethylation of H3K36, which is sufficient for the recruitment of the repressive Rpd3(S) complex. Binding to RNAPII is required for H3K36me3, which may perform a positive function in transcription elongation.

Table 1

Tetrad analysis of *set2* mutant strains with *htz1*Δ and *swr1*Δ

Mutant	Tetrads	Viable double mutant spores %
swr1Δx		
SET2	12	100 (12/12)
SET2 (1-261)	14	0 (0/15)
set2Δ	13	100 (15/15)
htz1Δx		
SET2	20	100 (15/15)
SET2 (1-261)	14	0 (0/13)
set2Δ	18	11 (2/19)

Supplementary Figures

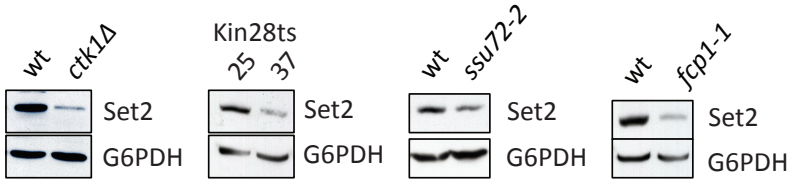


Figure S1. Set2 protein levels in mutants that disrupt RNAPII CTD phosphorylation

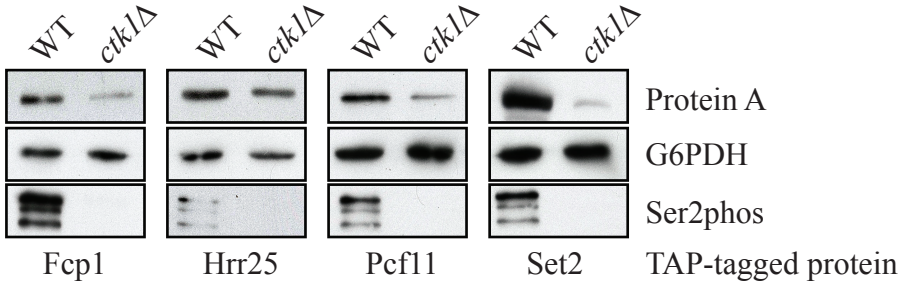


Figure S2. TAP-tagged strains abundance in *ctk1*Δ

Supplementary tables

Available online at: http://www.jbc.org/content/287/5/3249/suppl/DC1

References

Belle A., Tanay A., Bitincka L., Shamir R., O'Shea E. K. (2006) Quantification of protein half-lives in the budding yeast proteome. Proc. Natl. Acad. Sci. U.S.A.*103*, 13004–13009.

Biswas D., Dutta-Biswas R., Mitra D., Shibata Y., Strahl B. D., Formosa T., Stillman D. J. (2006) Opposing roles for Set2 and yFACT in regulating TBP binding at promoters. EMBO J. *25*, 4479–4489.

Boeke J. D., Trueheart J., Natsoulis G., Fink G. R. (1987) 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. *154*, 164–175.

Bua D. J., Kuo A. J., Cheung P., Liu C. L., Migliori V., Espejo A., Casadio F., Bassi C., Amati B., Bedford M. T., Guccione E., Gozani O. (2009) Epigenome microarray platform for proteome-wide dissection of chromatin-signaling networks. PLoS One *4*, e6789.

Buratowski S. (2003) The CTD code. Nat. Struct. Biol. *10*, 679–680.

Burke D., Dawson D., Stearns T. (2000) Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual, p. 174, Cold Spring Harbor Laboratory Press, New York.

Carrozza M. J., Li B., Florens L., Suganuma T., Swanson S. K., Lee K. K., Shia W. J., Anderson S., Yates J., Washburn M. P., Workman J. L. (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592.

Chapman R. D., Heidemann M., Albert T. K., Mailhammer R., Flatley A., Meisterernst M., Kremmer E., Eick D. (2007) Transcribing RNA polymerase II is phosphorylated at CTD residue serine 7. Science *318*, 1780–1782.

Cheung V., Chua G., Batada N. N., Landry C. R., Michnick S. W., Hughes T. R., Winston F. (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. PLoS Biol. *6*, e277.

Chu Y., Sutton A., Sternglanz R., Prelich G. (2006) The BUR1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by SET2. Mol. Cell. Biol. *26*, 3029–3038.

Chu Y., Simic R., Warner M. H., Arndt K. M., Prelich G. (2007) Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes.EMBO J. *26*, 4646–4656.

Collins S. R., Miller K. M., Maas N. L., Roguev A., Fillingham J., Chu C. S., Schuldiner M., Gebbia M., Recht J., Shales M., Ding H., Xu H., Han J., Ingvarsdottir K., Cheng B., Andrews B., Boone C., Berger S. L., Hieter P., Zhang Z., Brown G. W., Ingles C. J., Emili A., Allis C. D., Toczyski D. P., Weissman J. S., Greenblatt J. F., Krogan N. J. (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806–810.

Egloff S., O'Reilly D., Chapman R. D., Taylor A., Tanzhaus K., Pitts L., Eick D., Murphy S. (2007) Serine 7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. Science *318*, 1777–1779.

Fuchs S. M., Laribee R. N., Strahl B. D. (2009) Protein modifications in transcription elongation. Biochim. Biophys. Acta *1789*, 26–36.

Gardner K. E., Allis C. D., Strahl B. D. (2011) Operating on chromatin, a colorful language where context matters. J. Mol. Biol. *409*, 36–46.

Gelbart M. E., Rechsteiner T., Richmond T. J., Tsukiyama T. (2001) Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. Mol. Cell. Biol. *21*, 2098–2106.

Gietz R. D., Schiestl R. H. (2007) Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. *2*, 35–37.

Janke C., Magiera M. M., Rathfelder N., Taxis C., Reber S., Maekawa H., Moreno-Borchart A., Doenges G., Schwob E., Schiebel E., Knop M. (2004) A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947–962.

Joshi A. A., Struhl K. (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to pol II elongation. Mol. Cell *20*, 971–978.

Kaplan C. D., Laprade L., Winston F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096–1099.

Keogh M. C., Kurdistani S. K., Morris S. A., Ahn S. H., Podolny V., Collins S. R., Schuldiner M., Chin K., Punna T., Thompson N. J., Boone C., Emili A., Weissman J. S., Hughes T. R., Strahl B. D., Grunstein M., Greenblatt J. F., Buratowski S., Krogan N. J. (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593–605.

Kim T., Buratowski S. (2009) Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5′ transcribed regions. Cell *137*, 259–272.

Kizer K. O., Phatnani H. P., Shibata Y., Hall H., Greenleaf A. L., Strahl B. D. (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol. Cell. Biol. *25*, 3305–3316.

Krogan N. J., Kim M., Tong A., Golshani A., Cagney G., Canadien V., Richards D. P., Beattie B. K., Emili A., Boone C., Shilatifard A., Buratowski S., Greenblatt J. (2003) Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. Mol. Cell. Biol. *23*, 4207–4218.

Lenstra T. L., Benschop J. J., Kim T., Schulze J. M., Brabers N. A., Margaritis T., van de Pasch L. A., van Heesch S. A., Brok M. O., Groot Koerkamp M. J., Ko C. W., van Leenen D., Sameith K., van Hooff S. R., Lijnzaad P., Kemmeren P., Hentrich T., Kobor M. S., Buratowski S., Holstege F.

C. (2011) The specificity and topology of chromatin interaction pathways in yeast. Mol. Cell *42*, 536–549.

Li B., Howe L., Anderson S., Yates J. R., 3rd, Workman J. L. (2003) The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. *278*, 8897–8903.

Li B., Gogol M., Carey M., Lee D., Seidel C., Workman J. L. (2007a) Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. Science *316*, 1050–1054.

Li B., Gogol M., Carey M., Pattenden S. G., Seidel C., Workman J. L. (2007b) Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. Genes Dev. *21*, 1422–1430.

Li B., Jackson J., Simon M. D., Fleharty B., Gogol M., Seidel C., Workman J. L., Shilatifard A. (2009) Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. J. Biol. Chem. *284*, 7970–7976.

Lickwar C. R., Rao B., Shabalin A. A., Nobel A. B., Strahl B. D., Lieb J. D. (2009) The Set2/Rpd3S pathway suppresses cryptic transcription without regard to gene length or transcription frequency. PLoS One *4*, e4886.

Martin D. G., Grimes D. E., Baetz K., Howe L. (2006) Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin.Mol. Cell. Biol. *26*, 3018–3028.

Nordick K., Hoffman M. G., Betz J. L., Jaehning J. A. (2008) Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II. Eukaryot. Cell *7*, 1158–1167.

Pal B., Chan N. C., Helfenbaum L., Tan K., Tansey W. P., Gething M. J. (2007) SCFCdc4-mediated degradation of the Hac1p transcription factor regulates the unfolded protein response in *Saccharomyces cerevisiae*. Mol. Biol. Cell *18*, 426–440.

Phatnani H. P., Greenleaf A. L. (2006) Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. *20*, 2922–2936.

Pokholok D. K., Harbison C. T., Levine S., Cole M., Hannett N. M., Lee T. I., Bell G. W., Walker K., Rolfe P. A., Herbolsheimer E., Zeitlinger J., Lewitter F., Gifford D. K., Young R. A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517–527.

Rando O. J. (2007) Global patterns of histone modifications. Curr. Opin. Genet. Dev. *17*, 94–99.

Rao B., Shibata Y., Strahl B. D., Lieb J. D. (2005) Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide.Mol. Cell. Biol. *25*, 9447–9459.

Ruthenburg A. J., Li H., Patel D. J., Allis C. D. (2007) Multivalent engagement of chromatin modifications by linked binding modules. Nat. Rev. Mol. Cell Biol. *8*, 983–994.

Schaft D., Roguev A., Kotovic K. M., Shevchenko A., Sarov M., Shevchenko A., Neugebauer K. M., Stewart A. F. (2003) The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. Nucleic Acids Res. *31*, 2475–2482.

Schuldiner M., Collins S. R., Weissman J. S., Krogan N. J. (2006) Quantitative genetic analysis in *Saccharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344–352.

Shivaswamy S., Iyer V. R. (2008) Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response.Mol. Cell. Biol. *28*, 2221–2234.

Strahl B. D., Grant P. A., Briggs S. D., Sun Z. W., Bone J. R., Caldwell J. A., Mollah S., Cook R. G., Shabanowitz J., Hunt D. F., Allis C. D. (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol. Cell. Biol. *22*, 1298–1306.

Taverna S. D., Ilin S., Rogers R. S., Tanny J. C., Lavender H., Li H., Baker L., Boyle J., Blair L. P., Chait B. T., Patel D. J., Aitchison J. D., Tackett A. J., Allis C. D. (2006) Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Mol. Cell *24*, 785–796.

Taverna S. D., Li H., Ruthenburg A. J., Allis C. D., Patel D. J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol. *14*, 1025–1040.

Tietjen J. R., Zhang D. W., Rodríguez-Molina J. B., White B. E., Akhtar M. S., Heidemann M., Li X., Chapman R. D., Shokat K., Keles S., Eick D., Ansari A. Z. (2010) Chemical-genomic dissection of the CTD code. Nat. Struct. Mol. Biol. *17*, 1154–1161.

Tong A. H., Evangelista M., Parsons A. B., Xu H., Bader G. D., Pagé N., Robinson M., Raghibizadeh S., Hogue C. W., Bussey H., Andrews B., Tyers M., Boone C. (2001) Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364–2368.

Tong A. H. Y., Boone C. (2007) High-throughput strain construction and systematic synthetic lethal screening in *Saccharomyces cerevisiae*. Method Microbiol.*36*, 369–386.

Vojnic E., Simon B., Strahl B. D., Sattler M., Cramer P. (2006) Structure and carboxyl-terminal domain (CTD) binding of the Set2 SRI domain that couples histone H3 Lys-36 methylation to transcription. J. Biol. Chem. *281*, 13–15.

West M. L., Corden J. L. (1995) Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. Genetics *140*, 1223–1233.

Xiao T., Hall H., Kizer K. O., Shibata Y., Hall M. C., Borchers C. H., Strahl B. D. (2003) Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev. *17*, 654–663.

Youdell M. L., Kizer K. O., Kisseleva-Romanova E., Fuchs S. M., Duro E., Strahl B. D., Mellor J. (2008) Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. Mol. Cell. Biol. *28*, 4915–4926.

Yun M., Wu J., Workman J. L., Li B. (2011) Readers of histone modifications. Cell Res. *21*, 564–578.

Zhou K., Kuo W. H., Fillingham J., Greenblatt J. F. (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. Proc. Natl. Acad. Sci. U.S.A. *106*, 6956–6961.

Chapter 10

From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II

From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II

Hannes Braberg[1,2], Huiyan Jin[3,+], Erica A. Moehle[4,+], Yujia A. Chan[5], Shuyi Wang[1,2], Michael Shales[1,2], Joris J. Benschop[6], John H. Morris[7], James S. Fraser[1,2], Chenxi Qiu[3], Fuqu Hu[3], Leung K. Tang[3], Frank C.P. Holstege[6], Philip Hieter[5], Christine Guthrie[4], Craig D. Kaplan[3] and Nevan J. Krogan[1,2,8]

1 Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, 94158, USA;

2 California Institute for Quantitative Biosciences, QB3, San Francisco, CA, 94158, USA;

3 Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX, 77843, USA;

4 Department of Biochemistry and Biophysics, University of California, San Francisco, CA, 94158, USA;

5 Departments of Medical Genetics and Biochemistry, Michael Smith Laboratories, University of British Columbia, Vancouver, Canada;

6 Molecular Cancer Research, University Medical Centre Utrecht, The Netherlands;

7 Resource on Biocomputing, Visualization and Informatics, University of California, San Francisco, CA, 94158, USA;

8 J. David Gladstone Institutes, San Francisco, CA, 94158, USA;

[+]authors contributed equally;

To whom correspondence should be addressed: Christine Guthrie, christineguthrie@gmail.com; Craig D. Kaplan, cdkaplan@tamu.edu and Nevan J. Krogan, nevan.krogan@ucsf.edu

Summary

Dynamic control of gene expression is crucial for most aspects of cell physiology and at its core is RNA polymerase II (RNAPII). Using 53 RNAPII point mutants, we generated a **p**oint mutant **e**pistatic **mini**a**r**ray **p**rofile (pE-MAP) comprising ~60,000 quantitative genetic interactions in *Saccharomyces cerevisiae*. This enabled functional assignment of RNAPII sub-domains and uncovered connections to protein complexes. Using splicing microarrays and point mutants that alter elongation rates *in vitro*, we found an inverse relationship between RNAPII speed and *in vivo* splicing efficiency. Furthermore, the pE-MAP classified fast and slow mutants that favor upstream and downstream start site selection, respectively. Finally, the pE-MAP identified Sub1 as a positive transcription factor that regulates start site selection and influences splicing. The striking coordination of polymerization rate with transcription initiation and splicing suggests transcription rate is tuned to coordinate multiple gene expression steps. The pE-MAP approach provides a powerful strategy to understand other multi-functional machines at amino acid resolution.

Introduction

Alterations within a genome often cause specific as well as global phenotypic changes to a cell. Combining two alterations in the same cell allows for measurement of the genetic interaction between them: negative genetic interactions (synthetic sick/lethal) describe cases where two mutations in combination cause a stronger growth defect than expected from the two single mutations. This is often thought to be a result of the factors participating in redundant pathways or as non-essential components of the same essential complex. In contrast, positive interactions correspond to cases where the double mutant is either no sicker (epistatic) or healthier (suppressive) than the sickest single mutant (Beltrao et al., 2010) and may indicate that the factors are components of a non-essential complex and/or perform antagonizing roles in the cell. However, single genetic interactions are often difficult to interpret in isolation; an interaction pattern for a given mutation can be more informative as it reports on the phenotype in a large number of mutant backgrounds (Schuldiner et al., 2005; Tong et al., 2004). These genetic profiles provide highly specific readouts that can be used to identify genes that are functionally related (Beltrao et al., 2010).

Although genetic interaction datasets have been collected in a variety of different organisms, including *C. elegans* (Lehner et al., 2006), *D. melanogaster* (Horn et al., 2011) and mammalian cells (Bassik et al., 2013; Laufer et al., 2013; Lin et al., 2012; Roguev et al., 2013), one of the first organisms to be genetically interrogated on a genome-wide scale was *S. cerevisiae*, where non-quantitative genetic interaction data could be collected using the SGA (synthetic genetic array)

(Tong et al., 2001) or dSLAM (heterozygous diploid-based synthetic lethality analysis on microarrays) (Pan et al., 2004) approaches. We developed a technique, termed E-MAP (epistatic miniarray profile) (Collins et al., 2010; Schuldiner et al., 2005; Schuldiner et al., 2006), which utilizes the SGA methodology and allows for the quantitative collection of genetic interaction data on functionally related subsets of genes, including those involved in chromatin regulation (Collins et al., 2007b), RNA processing (Wilmes et al., 2008), signaling (Fiedler et al., 2009), plasma membrane (Aguilar et al., 2010) and mitochondrial function (Hoppins et al., 2011). Genetic approaches have been extended to *S. pombe* (Roguev et al., 2008; Roguev et al., 2007; Ryan et al., 2012) and *E. coli* (Butland et al., 2008; Typas et al., 2008) and recently, a variation of E-MAP, termed differential or dE-MAP, has been used where genetic interactions are measured in the presence of an exogenous stress (Bandyopadhyay et al., 2010; Ideker and Krogan, 2012). However, the vast majority of strains subjected to systematic genetic screening have had either deletions of non-essential genes or hypomorphic alleles (or "knock-downs") of essential genes. Since many genes, especially essential ones, are multi-functional, these methods would perturb all activities associated with a given protein.

Here, we describe an important advance of the E-MAP approach, which allows us to address higher levels of complexity by examining the genetic interaction space of point mutant alleles of multi-functional genes in a technique we term point mutant E-MAP (or pE-MAP). This method greatly increases the resolution achievable by gene function analysis, as it allows assignment of specific genetic relationships to individual residues and domains. We initially applied this

system to *S. cerevisiae* and chose to genetically dissect RNA polymerase II (RNAPII), due to its central role in the dynamic control of gene expression. It is an attractive but challenging target, as it is subject to multiple layers of regulation, including at the steps of initiation, elongation and termination of transcription, and the composition of the RNAPII holoenzyme itself is dynamic throughout the transcription cycle. In addition to querying transcription *per se*, we were hopeful that this approach would identify factors that regulate crosstalk between transcription and co-transcriptional processes, including mRNA splicing. Finally, application of our knowledge of this enzyme from *in vitro* biochemical analysis was expected to aid our interpretation of *in vivo* genetic interaction profiles. We used a set of traditional plate assay screens thought to report on specific transcription defects to isolate alleles in the 5 essential subunits unique to RNAPII with altered function *in vivo*, and further probed the effects of those mutations on *in vivo* function in a new pE-MAP.

In this study, we have used our pE-MAP to successfully dissect RNAPII function using viable, partial alteration-of-function alleles. We assign transcriptional activity and specific functions to different residues and regions of RNAPII. By examining the relationship between transcription rate and genetic interaction partners, we have revealed putative transcription rate-sensitive factors. Through the characterization of multiple stages of gene regulation, including start site selection, transcription elongation rate and mRNA splicing, our pE-MAP technique has provided both global and specific insight into structure-function

relationships of RNAPII. We argue that this strategy represents a paradigm for the high-resolution interrogation of any multi-functional protein.

Results

A set of alleles for the functional dissection of RNAPII

To identify residues important for transcriptional regulation *in vivo*, we isolated RNAPII alleles that conferred one or more of the following transcription-related phenotypes: suppression of sensitivity to galactose conferred by *gal10Δ56* (Gal^R) (Greger and Proudfoot, 1998; Kaplan et al., 2005), the Spt^- phenotype (Suppressor of Ty) (Winston and Sudarsanam, 1998) or mycophenolic acid (MPA) sensitivity (Shaw and Reines, 2000) (these phenotypes are described in detail in the legend of **Figure 1A**, **Methods**). Each of these phenotypes relates to a gene-specific transcription defect that can be monitored using simple plate growth assays (**Figure 1A**). Random mutagenesis by PCR was carried out on the entire coding regions of RNAPII subunit genes *RPB2*, *RPB3*, *RPB7* and *RPB11* and most of *RPO21/RPB1* in *S. cerevisiae* (**Methods**). These genes encode the essential subunits of RNAPII that are unique to this enzyme (Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12 are shared with RNAPI and RNAPIII, and Rpb4 and Rpb9 are non-essential subunits of RNAPII) (Archambault and Friesen, 1993). In total, 53 single point mutants that exhibit at least one of these phenotypes were identified (Kaplan et al., 2012) (**Figure 1B**, **Figure S1**, **Table S1**).

Analysis of the distribution of phenotypes relative to the RNAPII structure suggested that our alleles might be diverse in their functions. Gal^R and MPA

sensitivity mutations were broadly distributed, while those with Spt$^-$ phenotype were less common and more localized (**Figure 1B**, **Table S1**). Perhaps because these phenotypic screens are not highly specific for distinct altered functions of RNAPII, we found that mutations of individual residues were often identified in multiple screens. Our screens identified mutations in highly conserved residues and structural domains known to be important for RNAPII activity, including the Rpb1 trigger loop, Rpb1 nucleotide interacting residues, the Rpb1 bridge helix, RNA-DNA hybrid interaction surfaces, Rpb1 switch 1 and the Rpb2 lobe and protrusion (Cramer et al., 2001; Gnatt et al., 2001; Kaplan, 2013) (**Table S1**). We hypothesized that high-throughput genetic profiling quantitatively measuring the genetic interactions of specific residues might provide insight into the functions of these RNAPII regions, and may even be instructive in identifying protein-protein interaction interfaces.

An RNAPII point mutant epistatic miniarray profile (pE-MAP)

The 53 RNAPII point mutants (**Figure 1B**) were crossed against ~1200 mutants (deletions of non-essential genes and DAmP (Decreased Abundance by mRNA Perturbation) alleles of essential genes) (Schuldiner et al., 2005) (**Table S2**), a library representing all major biological processes. We thus created a quantitative, point mutant epistatic miniarray profile (or pE-MAP) comprising approximately 60,000 double mutants (**Supp. Data 1**). Two-dimensional hierarchical clustering of these data effectively grouped together genes from known complexes and pathways based on their individual interactions with the point mutant alleles (**Figure S2**,

Supp. Data 2). Our previous studies have demonstrated that genes encoding proteins that physically associate often have similar genetic interaction profiles when deleted or knocked down (Collins et al., 2007b; Roguev et al., 2008). We wondered whether similar observations would characterize our point mutant data, even though we have only examined 5 subunits of a single molecular machine. To test this, we generated a receiver operating characteristic (ROC) curve that measures how well the genetic profiles of the deletion and DAmP mutants in our pE-MAP predict known physical interactions between their encoded proteins (**Methods**). We found that the predictive power of the pE-MAP was similar to that of a previously published E-MAP (Collins et al., 2007b), indicating that the genetic interactions of the RNAPII point mutants report on connections among virtually all cellular processes (**Figure 2A**), including the DNA damage response pathway, peroxisome biogenesis, tRNA modification as well as transcriptional regulation (**Figure S2**, **Supp. Data 2**).

Next, to examine whether the spatial location of a mutated residue is a determinant of its function, we compared the similarity of pairs of RNAPII genetic profiles to the three-dimensional distance between the mutated residues (Wang et al., 2006). We observed a strong correlation ($r=-0.37$, $p<10^{-22}$) (**Table S3**). The trend is highly significant both for residue pairs residing in the same subunit ($r=-0.25$, $p<10^{-5}$) and for those in different subunits ($r=-0.28$, $p<10^{-6}$) (**Figure 2B**), suggesting that structural proximity correlates with functional similarity and that high resolution genetic interaction profiling could provide information for targets whose structures have not yet been determined.

Comparison of genetic and gene expression profiles derived from the RNAPII alleles

To address the possibility that any given genetic interaction might result from the point mutation directly affecting the expression of the corresponding gene, we subjected 26 of the RNAPII mutants to genome-wide gene expression analysis (**Supp. Data 3**). We found that there was no overall correlation (r=-0.003) between an RNAPII mutant's genetic interaction score with a gene deletion or DAmP and the gene expression change of that gene due to the RNAPII mutation (**Figure 2C**). Therefore, connections must be due to more complex relationships between the mutated residues and expression of the library genes. Nonetheless, these datasets allowed us to test whether the clustering of the RNAPII mutants in the pE-MAP (**Figure S2**, **Supp. Data 2**) can also be recapitulated using their gene expression profiles. We thus assessed pair-wise RNAPII mutant similarity based on genetic and gene expression profiles separately and found that the pair-wise similarities based on these two measures are highly correlated (r=0.71, **Figure 2D**). These results indicate that even though we observed virtually no overlap between individual genetic interaction partners and genes with altered expression, both datasets provide a common biological framework for functionally organizing the RNAPII mutants. More work will be necessary to mechanistically understand the intriguing link between these datasets.

Functional associations between RNAPII residues and protein complexes

Hierarchical clustering using the genetic interaction profiles allowed us to group together functionally related RNAPII point mutants. In an effort to link these individual mutations in RNAPII to specific cellular functions, the pE-MAP was compared to a previously published genetic interaction dataset that includes mutant profiles from >4000 genes (Costanzo et al., 2010). These genes were classified based on complex membership of their encoded proteins (Benschop et al., 2010) and Mann-Whitney U statistics were used to identify RNAPII mutants enriched in profile similarity to members of specific complexes (**Methods**, **Figure 3A**, **Table S4**). We uncovered a number of connections, including several point mutants having similar genetic profiles to mutants of components of the Mediator and the Rpd3C(L) histone deacetylase complexes (Carrozza et al., 2005; Keogh et al., 2005). Unexpectedly, we also observed that several RNAPII mutants are significantly correlated genetically to kinetochore mutants (**Figure 3A**). We carried out chromosome transmission fidelity (CTF) assays on 19 of our RNAPII mutants, selecting those either linked specifically to the kinetochore, to other unrelated complexes, or to no complex (**Methods**, **Table S5**) (Spencer et al., 1990). Only four of the tested mutants exhibited chromosome loss in more than 15% of their colonies, and all of these were genetically linked to the kinetochore in our analysis and had similar genetic interaction profiles in our pE-MAP (**Figure 3A**, **B**). Recent studies indicate that a certain level of transcription by RNAPII at the centromere is required for centromere function and high-fidelity chromosome segregation in budding yeast (Ohkuni and Kitagawa, 2011). Using specific constructs designed to ascertain centromere sensitivity to transcriptional

read-through, we did not observe any defects in kinetochore integrity in these RNAPII mutants (data not shown). Ultimately, further work will be required to understand the connection between these RNAPII point mutants and chromosome segregation. A full point mutant-module map from alleles of all subunits is presented in **Figure S3**.

pE-MAP identifies alleles involved in start site selection and can finely distinguish between different phenotypic categories

What transcription defects might underlie differences in the genetic interaction profiles of specific RNAPII alleles and distinguish between different allele subsets? Our recent work has shown that mutations in the Rpb1 trigger loop, a mobile element in the active site that couples RNA base recognition with catalysis, can alter transcription start site selection *in vivo* (Kaplan, 2013; Kaplan et al., 2012). For example, *rpb1* E1103G shifts transcription start site selection upstream at *ADH1* whereas *rpb1* H1085Y more often initiates transcription at downstream sites (Kaplan et al., 2012). The pE-MAP subcluster containing *rpb1* E1103G includes an additional 10 mutants in *RPB1*, *RPB2* and *RPB7* (**Figure 4A**, **Supp. Data 2**). We examined 8 of these for defects in start site selection at the *ADH1* gene by primer extension and found that, like *rpb1* E1103G, all 8 mutants had a preference for upstream start site selection to varying degrees (**Figure 4B**, **Figure S4A**, **Table S6**), consistent with their clustering with *rpb1* E1103G (**Figure 4A**). Not unexpectedly, four of these mutants fell within the trigger loop (*rpb1* F1084I, *rpb1* M1079R, *rpb1* A1076T, *rpb1* M1079V), however two were in other regions of *RPB1* (*rpb1* I1327V,

rpb1 S713P). Inspection of the structure of the RNAPII elongation complex (Kettenberger et al., 2004; Wang et al., 2006) reveals that these two mutated Rpb1 residues are in close proximity to the trigger loop (**Figure 4C**), suggesting they too may have a direct role in the regulation of the active site of RNAPII and its functions. Interestingly, some of the other mutants in this cluster are not in close proximity to the trigger loop (*rpb2* E328K and *rpb7* D166G), raising the possibility that these mutations may indirectly impinge on the active site via allosteric events or may function independent of the trigger loop by recruiting other factors to the transcription apparatus. Importantly, this is the first description of a mutation in *RPB7* that has an effect on transcriptional start site selection. We additionally examined start site selection at *ADH1* for an *rpb4Δ* strain and in a *rpb6* Q100R mutant background. Both mutants are expected to reduce the amount of Rpb4/Rpb7 associated with RNAPII (Edwards et al., 1991; Tan et al., 2003), yet neither mutant altered start site selection at *ADH1* (**Figure S4B),** indicating we have identified a specifically compromised function of Rpb7 in the D166G mutant (see **Discussion**). These data provide an example of how mechanistic information on structure-function relationships can be extracted from the pE-MAP.

In our screening process, we had also identified an *rpb2* allele mutated at two residues in close proximity (E437G/F442S) within the tip of the *rpb2* protrusion domain (Cramer et al., 2001; Gnatt et al., 2001), whose genetic profile also clusters with the set of mutants resulting in upstream start site selection (**Figure S4C, Supp. Data 4**). Rpb2 surfaces in the lobe and protrusion domains have been physically linked to TFIIF and TFIIE through extensive cross-linking studies (Chen et al., 2007;

Chen et al., 2010; Grunberg et al., 2012). We observed that *rpb2* E437G/F442S is sensitive to MPA (**Figure S4D**) and also has a preferential upstream start site selection (**Figure 4B**), similar to what has been reported for TFIIF alleles (Eichner et al., 2010; Ghazy et al., 2004) (H. Jin and C. Kaplan, unpublished data) (see **Discussion**). Additional *rpb2* alleles that confer MPA sensitivity also map near the protrusion (*rpb2* R120C) or to the lobe (*rpb2* D399H), however these do not alter *ADH1* start site selection and are genetically distinct from E1103G or other *rpb2* alleles, illustrating the fine resolving power of the pE-MAP even for mutations having similar locations and/or a subset of phenotypes (**Figure 4A, B**).

Finally, we found that the mutants displaying upstream start site selection in the larger cluster can be arranged into at least 5 subclusters based on varying degrees of Spt and MPA phenotypes. These include mutants that are Spt^- and show either weak (blue) or strong (green) sensitivity to MPA as well as Spt^+ mutants that display weak (purple), moderate (grey) or strong (orange) sensitivity to MPA (**Figure 4A**, **B**). However, many other mutants in the pE-MAP show similar MPA and Spt phenotypes but do not cluster in this group, including *rpb2* R120C (allelic to *rpb2-7*) (Martin and Young, 1989), which is Spt^+ and moderately sensitive to MPA (grey); the Spt^- mutant *rpb3* C86R (blue); and *rpb1* F1086S (purple), which is only weakly sensitive to MPA (**Figure 4A**). Consistent with the distinctive pE-MAP profiles of these three mutants, we found they have no upstream start site selection defect (**Figure 4B**, **Figure S4A**, **Table S6**). These data collectively suggest that the pE-MAP has the resolving power to identify the mutations causing upstream start site selection but within this group can further separate them into specific Spt and

MPA phenotypic categories and highlights the difficulty of interpreting phenotypes generated by these traditional plate assays.

In vitro biochemical activity correlates with pE-MAP profiles and gene expression

Next, we turned to a series of active-site mutants whose *in vitro* elongation rates range from <0.1 to >2-fold that of a wild-type polymerase (Kaplan et al., 2012; Kaplan et al., 2008; Kireeva et al., 2008; Malagon et al., 2006). These polymerase mutants allowed us to address questions regarding the *in vivo* consequences of altered elongation rates. We found that clustering these mutants based on their pE-MAP profiles yielded two distinct subsets that differ by their transcription rate (**Figure 5A**; **dendrogram**, **Supp. Data 5**). Specifically, one subset contains the catalytically fast mutants (*rpb1* E1103G, *rpb1* L1101S and *rpb1* F1084I) and the other contains those with a decreased rate of transcription relative to wild-type (*rpb1* F1086S, *rpb1* N1082S, *rpb1* N479S and *rpb1* H1085Q) (**Figure 5A**). The genome-wide gene expression profiles of these mutants also group them into the same two subsets. Thus, pairs of mutants with similar *in vitro* transcription rates show both highly similar genetic and gene expression profiles (**Figure 5B**). This observation highlights that increasing or decreasing the catalytic activity of RNA polymerase has a defined set of effects on genetic interaction patterns and genome-wide expression.

Genetic dissection of RNAPII Rpb1 trigger loop function has shown that combining two mutations that individually make polymerase fast and slow results in

a polymerase with near-wild-type rate of transcription *in vitro* (Kaplan, 2013; Kaplan et al., 2012). We wanted to determine if the genetic interactions derived from such double mutants would show similar suppression across the entire genetic interaction or gene expression profiles. Combination of the fast *rpb1* E1103G allele with *rpb1* F1086S, a moderately slow RNAPII mutant (**Figure 5C**), caused the double mutant to exhibit an essentially wild type genetic interaction profile, with very few high scoring genetic interactions compared to that of the single mutants (**Figure 5D**, **Supp. Data 6**), consistent with the overall growth defect suppression observed in the double mutant (**Figure S5A**). The same global suppressive trends were observed when the double mutant was subjected to gene expression profiling, specifically monitoring genes that were up- or down-regulated 1.7 fold or greater (**Figure 5D**, **Supp. Data 6**). Similar observations were seen when a *rpb1* E1103G (fast) *rpb1* N1082S (slow) double mutant was subjected to both genetic and gene expression profiling (data not shown). Taken together, these results are consistent with the notion that most of the genetic interactions and gene expression changes observed with the fast and slow mutants are related to the catalytic activity of RNAPII. However, the pE-MAP and gene expression resolutions were sufficiently high to also identify a double mutant, E1103G/H1085Q, which deviated from this general rule (**Figure 5C**, **D**, **Supp. Data 6**). Although this double mutant has a near-wild-type transcription rate *in vitro*, it exhibited a large number of genetic and gene expression defects that were also observed for E1103G alone, and its genetic interaction and gene expression profiles more highly correlate with E1103G than

with H1085Q (**Figure S5B**), suggesting a more complex genetic relationship with this double mutant.

Genome-wide and gene-specific effects of altering polymerase speed on in vivo splicing efficiency

Several steps in mRNA processing are coupled to the process of transcription (Bentley, 2005). Based primarily on experiments from metazoa, a kinetic coupling model has been invoked to explain how decreasing polymerase speed affects the usage of alternative splice sites, primarily favoring the recognition of otherwise weak splice sites (de la Mata et al., 2011). We sought to use the set of RNAPII mutants that differ in their *in vitro* elongation rates to determine whether this type of coupling exists in *S. cerevisiae*, where splice sites are of consensus or near-consensus sequence and the sole read-out is splicing efficiency. We used a splicing-sensitive microarray platform to measure the change in total (exon), pre-mRNA (intron), and mature mRNA (junction) for each intron-containing gene (**Figure 6A**; **top**) in response to specific, single point mutations in the RNAPII trigger loop. Many genes exhibit the reciprocal intron and junction ratios indicative of changes in splicing (**Figure 6A; side panels**), so to allow facile comparison across multiple mutants, we calculated pre/mature ratios (I/J) (Clark et al., 2002) for each gene (**Figure 6**, **Supp. Data 7**). Both gene-by-gene values (**Figure 6A**) and global trends (**Figure 6B**) show that *rpb1* mutants characterized as fast *in vitro* (E1103G and G1097D) lead to an increase in this ratio for many genes, which indicates a defect in the splicing of those transcripts. Conversely, mutants characterized *in vitro* as

having slow elongation (H1085Q and F1086S) lead to a decrease in this ratio for many genes, a phenotype consistent with more efficient splicing. Thus, the global trend we observe is an anti-correlation between polymerase rate and splicing efficiency ($p<10^{-5}$ for each mutant in **Figure 6A** compared to wild-type, **Table S7**). We confirmed this trend at several genes using RT-qPCR (**Figure S6**).

Given that polymerase rate and splicing efficiency are anti-correlated, it follows that the trigger loop double mutants (E1103G/F1086S and E1103G/H1085Q), which transcribe at near-wild-type rate, should also exhibit near-wild-type splicing. Consistent with this prediction, these strains had very few genes with splicing defects (**Figure 6B**, **Supp. Data 7**), demonstrating that suppressive relationships are observed in the double mutant with respect to not only growth, genetic and gene expression profiling, and in vitro transcription rate, but also mRNA splicing. Furthermore, slowing polymerase by chemical means should phenocopy a genetically slow polymerase mutant. We therefore evaluated splicing in wild-type cells treated with mycophenolic acid (MPA), which is known to impede transcriptional elongation (Howe et al., 2003; Mason and Struhl, 2005). The histogram of I/J ratios shows many genes with improved splicing (**Figure 6B**). It is, in fact, an even more striking phenotype ($p<10^{-15}$ compared to wild-type, **Table S7**) than mutation of RNAPII, possibly because MPA elicits an acute stress (10-minute treatment), as compared to the steady-state perturbation caused by a slow polymerase mutant. Taken together, these splicing phenotypes are consistent with a direct kinetic coupling between elongation rate and splicing *in vivo.*

Genetic interactions between RNAPII alleles and other mutants reveal relationships between transcription factors and RNAPII activity

Because we know that pairs of rate-altering RNAPII mutations of opposing effect suppress each other genetically (Kaplan et al., 2012) (**Figures 5C** and **D**), we hypothesized that some of the deletion/DAmP genetic interactors we detected in our pE-MAP might be direct regulators of or collaborators with RNAPII. We reasoned that disruption of a positively acting transcription factor would result in a positive genetic interaction (i.e. suppression) with fast RNAPII mutants but a negative interaction (i.e. synthetic sickness) with slow RNAPII mutants. Conversely, a negatively acting factor would show opposite RNAPII-activity genetic interactions. To identify factors that behaved in this manner among the pE-MAP deletion and DAmP mutants, we sorted these mutants based on the difference in average genetic interaction score with fast and slow RNAPII mutants (**Methods, Figure 7A**, **Supp. Data 8**). We focused our attention on genes that behaved as positively acting factors and observed that *sub1*Δ had the strongest pattern in this regard. Sub1 was first identified biochemically as a positive factor in *in vitro* transcription assays and genetically as a high copy suppressor of TFIIB mutants (Knaus et al., 1996), and *sub1Δ* was shown to be synthetically lethal with *sua7* (TFIIB) mutants (Wu et al., 1999). Sub1 is broadly recruited to RNAPII and RNAPIII promoters *in vivo* (Rosonina et al., 2009; Tavenet et al., 2009), and has recently been implicated as a member of the RNAPII preinitiation complex (PIC) (Sikorski et al., 2011). In addition to this evidence, suggestive of a role for Sub1 in initiation, other evidence implicates Sub1 in RNAPII elongation (Garcia et al., 2012) and 3' end processing

(reviewed in (Conesa and Acker, 2010)). Importantly, the genetic relationships observed in the pE-MAP study were confirmed using standard growth assays where *sub1Δ* exacerbated the growth defect caused by slow RNAPII alleles (**Figure 7B**) and partially suppressed fast RNAPII alleles (**Figure 7C**). We also saw exacerbation and suppression of specific RNAPII mutant phenotypes (MPA, Spt, Gal^R) with *sub1Δ* (**Figure S7A**), consistent with Sub1 being a factor whose deletion differentially modulates fast and slow mutants. Furthermore, gene expression analysis of *sub1*Δ, *rpb1* E1103G, and *sub1*Δ/*rpb1* E1103G showed an epistatic relationship between this fast RNAPII allele and deletion of *SUB1*, (**Figure 7D**), consistent with the observed positive interaction between the two and suggests there is bypass of the requirement for Sub1 in a catalytically fast mutant.

Our recent work has implicated changes in RNAPII activity with the alteration of start site selection *in vivo* (Kaplan et al., 2012). Because we know that our fast and slow RNAPII point mutants alter start site selection through their catalytic capabilities, we sought to determine whether Sub1 might modulate their start site choice. Using primer extension analysis at *ADH1*, we found that deleting *SUB1* led to a significant downstream shift in transcription starts (**Figure 7E, Figure S7B**). These data are similar to what was observed for a slow RNAPII trigger loop mutant *rpb1* F1086S (**Figure 4B**) and *sua7* alleles (Pinto et al., 1992), both of which *sub1Δ* exacerbates genetically. These data are consistent with the notion that Sub1 promotes transcription initiation. Double mutant analysis revealed that deletion of *SUB1* also exacerbated the downstream start site shift of the slow RNAPII trigger loop allele, *rpb1* F1086S, and slightly suppressed the *rpb1* E1103G allele (**Figure**

7E). Because *sub1Δ* has also been linked to another RNA processing step, namely 3' end processing, we examined the effect of *sub1*Δ on splicing and observed a statistically significant increase in splicing efficiency ($p<10^{-8}$, **Table S7**), again phenocopying the slow RNAPII mutants (**Figure 7F**, **Supp. Data 7**).

Based on these data, we propose a model where transcription start and splicing are intimately coupled with RNAPII elongation: fast RNAPII mutations result in upstream transcription start and diminished splicing, while slow mutations or *sub1*Δ give rise to downstream transcription start and enhanced splicing (**Figure 7G**). Given the possibility of direct coupling between start-site selection and downstream mRNA processing, we globally measured splicing defects in *sua7*-3 (TFIIB) (Pinto et al., 1994; Wu et al., 1999) and *tgf2Δ*261-273 (TFIIF) (Eichner et al., 2010), mutants that result in downstream and upstream start site selection, respectively (**Figure S8A**). If splicing were strictly coupled to start site choice, one would expect these mutants to have similar splicing defects to the slow and fast RNAPII mutants. However, such a correlation was not observed (**Figure S8B**), suggesting these processes are genetically separable. Taken together, these data support a model in which the catalytic rate of RNAPII has multiple, separable effects on start site selection and mRNA processing, and highlight the importance of wild-type elongation rate for multiple steps in gene expression.

Discussion

In this study, we have described an important extension of our E-MAP genetic interaction mapping strategy, the functional interrogation of individual protein residues. We have first used this approach to genetically dissect RNAPII. In the past, we have shown how quantitative genetic interaction maps can identify functionally distinct sub-modules within larger protein complexes, including the transcriptional regulator Mediator (Collins et al., 2007b). We now demonstrate how this platform can characterize functionally distinct individual amino acid residues within complex and multi-functional proteins. This analysis not only provided insight into global structure-function relationships within RNAPII but also specific details about how RNAPII regulates, and is regulated by, different factors and processes.

Insight into RNAPII Function Derived from the pE-MAP

Different aspects of the genetic interaction map were exploited in order to gain important insights into transcriptional regulation. First, by comparing the pE-MAP profiles with genetic interaction data from deletion/knockdown studies, function was assigned to individual residues based on similarity of profiles, an approach we used to generate a point mutant-protein complex connectivity map (**Figure 3**, **Figure S3**). Furthermore, generation of double mutants, both within RNAPII itself (**Figure 5C, D**) and between RNAPII and other genes (**Figure 7**), allowed for a better understanding of trigger loop function, as well as the identification of other factors that directly or indirectly impinge on the activity of RNAPII. As we link trigger loop-proximal residues to trigger loop mutations (*e.g.*

rpb1 E1103G) that are involved in altering translocation and catalytic properties of RNAPII (Kaplan et al., 2008; Kireeva et al., 2008; Larson et al., 2012), we surmise that these interactions might, in fact, meaningfully restrain the trigger loop during transcription. This study has also allowed for the identification of putative transcription factors that act negatively or positively, including Sub1, which we demonstrated plays a positive role in transcriptional initiation by influencing start site usage (**Figure 7**). Sub1 has been implicated as a component of the PIC (Sikorski et al., 2011), and more recent work has suggested that Sub1 may also function in RNAPII elongation either directly or through its relationships with Spt5, an elongation factor, and Bur1, a kinase with multiple roles in RNAPII elongation (Garcia et al., 2012). Our results demonstrating that epistatic relationships exist between fast RNAPII alleles and Sub1 suggest that the activity of Sub1 may be bypassed when RNAPII catalytic activity is increased. Collectively, our data indicate that Sub1 plays a direct role in transcriptional initiation, as well as influencing mRNA splicing, possibly via its effect on elongation (Garcia et al., 2012). Overall, the genetic interactions with trigger loop mutants should provide a rich dataset of potential positive and negative transcription factors for further study.

Finally, analysis of the 3-D structure of RNAPII revealed that there was a strong correlation between the similarity of genetic interaction profiles of the specific point mutants and the spatial distance between the residues, independent of the subunit that was mutated (**Figure 2B**). These data strongly suggest that structural information is ultimately contained within the p-EMAP and can be used to identify regions involved in mediating specific protein-protein interaction

interfaces. For example, the Rpb2 mutant (*rpb2* E437G/F442S), which results in upstream start site selection, is in a region which has been physically linked to TFIIF through contacts with its subunit, Tfg2 (Chen et al., 2007; Eichner et al., 2010). Interestingly, mutations in this region of TFIIF also show upstream start site shifts (Eichner et al., 2010). Our connection of an Rpb7 residue to start site selection is also intriguing as Rpb4/Rpb7 subunits are required for promoter-directed initiation *in vitro* (Edwards et al., 1991), but their exact roles *in vivo* are not known. Using a strain deleted for *RPB4*, we find that reduction of Rpb4/Rpb7 association with RNAPII does not alter start site selection, consistent with *rpb7* D166G altering an essential function of Rpb7. Interestingly, this region has been linked to TFIIH contacts in a recent cryo-EM structure of the RNAPII pre-initiation complex (He et al., 2013). Alleles in the Ssl2 subunit of TFIIH have recently been identified as altering start site selection (Goel et al., 2012), further suggesting that a Rpb7-TFIIH connection is involved in the initiation process. Therefore, in the future, as this analysis is extended to other molecular machines, the resolving power of the genetic profiles can be harnessed, with other biophysical techniques such as cross-linking or electron microscopy, to help identify the specific regions of proteins that physically interact.

Coordination of Transcriptional Rate with Start Site Selection and mRNA Splicing

The importance of maintaining wild-type rates of transcription is evidenced by the growth defects observed in the fast and slow mutants (**Figure 5**) (Kaplan et al., 2012), as well as the striking mutual suppression of these defects seen in double mutants that combine both fast and slow mutations in the same polymerase. What

parameters might account for the observed correlation between growth and transcription rates?

Interestingly, the pE-MAP identified two groups of RNAPII mutants: one that enables upstream start site selection and exhibits an increased rate of transcription, and one that preferentially starts transcription downstream and exhibits decreased transcriptional rate (**Figure 7**) (Kaplan et al., 2012). We propose that both these phenotypes are direct consequences of the efficiency of nucleotide selection and incorporation, since the addition of the first nucleotides at initiation is biochemically similar to adding nucleotides during elongation. Two models have been proposed to account for this apparent "scanning" (Giardina and Lis, 1993; Kaplan et al., 2012; Kuehner and Brow, 2006). The first assumes that RNAPII catalysis is the driver of scanning and that the ability of RNAPII to initiate productive transcription at an upstream site is determined by its increased ability to avoid abortive initiation. This model predicts that the fast, upstream mutants would exhibit less abortive initiation than the slow, downstream mutants. The second model requires scanning, perhaps driven by TFIIH, in the *absence* of nucleotide incorporation (Kaplan, 2013). In this scenario, the likelihood of initial phosphodiester bond formation at upstream sites is higher in the RNAPII mutants with increased catalytic activity, and lower in the slow mutants; notably, this model does not predict the creation of abortive transcripts. More experiments will be required to differentiate between these two models. We note that because changes in start-site selection alter the 5'UTR length and composition, this chromatin-associated process may have downstream effects on gene expression such as changes in RNA stability or translational efficiency of the

mRNA (Arribere and Gilbert, 2013; Hampsey et al., 1991; Rojas-Duran and Gilbert, 2012).

The pE-MAP has allowed important insights into the co-transcriptional process of mRNA splicing. It is now clear that most introns are removed while RNA polymerase is still associated with the DNA template. In metazoans, alternative splicing decisions can be influenced by factors impinging on transcription, including promoter identity and polymerase speed (reviewed in (Perales and Bentley, 2009)). Specifically, slowing the rate of elongation by mutation of RNA polymerase or chemical means has been shown to improve the recognition of splice sites that deviate from consensus signals (de la Mata et al., 2003; Ip et al., 2011). Because the spliceosome undergoes stepwise assembly on each intron, slowing transcription can afford more time for formation of the catalytically active machine before transcription of a downstream, stronger site. While budding yeast lack alternative splicing, it nonetheless follows that the efficiency of co-transcriptional splicing would be favored by allowing sufficient time for spliceosome assembly; indeed, recent work suggests that RNA polymerase may slow down to favor co- vs. post-transcriptional splicing (Aitken et al., 2011; Alexander et al., 2010; Carrillo Oesterreich et al., 2010). Our microarray analyses show a clear trend in which splicing efficiency is generally anti-correlated with transcription rate (**Figure 6**). This is the first genome-wide analysis to interrogate kinetic coupling in *S. cerevisiae*. Moreover, this is the first instance in which faster and slower polymerases have been directly compared; thus, this work satisfies the previously untested prediction that rate and splicing efficiency would be generally anti-correlated.

More challenging to understand is the apparent coordination between elongation rate, start-site selection and mRNA splicing efficiency. In fact, upstream transcriptional events are known to be able to influence downstream RNA transactions. For example, in other systems, promoter identity can influence alternative splicing outcome or mRNA stability (Cramer et al., 1997; Harel-Sharvit et al., 2010; Trcek et al., 2011). Furthermore, altered mRNA composition determined by start site selection in budding yeast has been shown to strongly alter translational efficiency, further suggesting that promoter proximal events can be propagated into downstream consequences (Rojas-Duran and Gilbert, 2012). However, when we measured splicing efficiency using mutants in the general transcription factors TFIIF and TFIIB that alter start site selection, we found not all initiation phenotypes are predictive of splicing efficiency (**Figure S8B**). This suggests that RNAPII catalytic rate has several, separable effects on gene expression, a claim supported by recent evidence showing kinetic coupling between RNAPII transcription and Sen1-dependent termination (Hazelbaker et al., 2013).

Taken together, our data highlight the important impact of transcription speed determined by the genetic status of RNAPII and trans-acting factors (e.g. Sub1) on start-site selection and mRNA splicing. We propose that RNA polymerase may have been evolutionarily tuned to maximize the coordination between multiple steps in gene expression and we predict that polymerase rate may influence multiple additional co-transcriptional steps in gene expression, including mRNP assembly, 3' end processing, and export.

Future Studies Using the pE-MAP Approach

The fact that such a wealth of information was gleaned from the genetic analysis of a series of point mutants in *S. cerevisiae* RNAPII suggests that more mutants should be analyzed in this complex as well as other key molecular machines, including those involved in translation (e.g. ribosome), protein quality control (e.g. proteasome, HSP70, HSP90), chromatin (histones) and DNA replication (DNA polymerases). This analysis can also be carried out in other organisms that are genetically tractable and where high-throughput genetic interaction mapping methodologies have been created, including *S. pombe* (Roguev et al., 2007) and *E. coli* (Butland et al., 2008; Typas et al., 2008). Furthermore, the platform could also be used fruitfully to gain structural insight into proteins and complexes that have not yet been structurally characterized. Finally, as genetic interaction mapping strategies become more prevalent in mammalian cells (Bassik et al., 2013; Laufer et al., 2013; Lin et al., 2012; Roguev et al., 2013) and with the development of genome editing using various double-strand break technologies (Gaj et al., 2013),similar work on characterizing the function of individual amino acids will have great impact on understanding how point mutations in specific genes result in different disease states.

Methods

Plate phenotyping

Plate phenotyping of RNAPII mutants was done as described (Kaplan et al., 2012). Briefly, *CEN LEU2* plasmids containing a mutant *rpb* RNAPII subunit gene

were transformed into an appropriate Leu$^-$ strain with corresponding endogenous RNAPII subunit gene knocked out but complemented with a *CEN URA3* WT *RPB* subunit gene. Leu$^+$ transformants were patched on solid medium lacking leucine and replica-plated to medium lacking leucine but containing 5-fluoroorotic acid (5-FOA, Gold Biotechnology) to select against cells maintaining *RPB* WT *URA3* plasmids. *sub1Δ* + RNAPII mutant strains for direct testing of double mutant phenotypes were constructed based on CKY283 and analyzed in the same fashion as RNAPII single mutants. Yeast media used in phenotyping were prepared as previously described (Amberg et al., 2005; Kaplan et al., 2012). YPD medium contained yeast extract (1% w/v final, BD), peptone (2% w/v final, BD), dextrose (2% final w/v), bacto agar (2%, BD) and supplemental adenine and tryptophan. Alternate carbon source YP media were YP raffinose (2% final w/v, USB) and YP raffinose (2% final w/v, USB) plus galactose (1% final w/v, Sigma-Aldrich), with plates also containing antimycin A (1µg/ml, Sigma-Aldrich). Synthetic complete medium lacking leucine (SC-Leu) was as described (Kaplan et al., 2012). Mycophenolic acid (MPA, Sigma-Aldrich) was added to SC-Leu at 20µg/ml final concentration from a 10mg/ml stock in ethanol.

Galactose sensitivity or resistance (Gal^S or Gal^R, determined based on strain growth on YP raffinose plus galactose) phenotypes are based on an allele of *GAL10*, *gal10Δ56*, that has a deletion of the major *GAL10* polyadenylation signal (Greger et al., 2000; Greger and Proudfoot, 1998; Kaplan et al., 2005). A consequence of this deletion is that *GAL10* transcription reads through into the downstream *GAL7* promoter, interfering with *GAL7* expression. Under conditions where other Gal genes are expressed, specifically Gal1p, Gal7 protein is required to prevent galactose

toxicity caused by Galactose catabolism intermediate. Thus defects in *GAL7* due to transcriptional interference from *gal10Δ56* cause the Gal^S phenotype, and resistance to this transcription defect leads to the Gal^R phenotype.

The Spt phenotype studied here relates to the *LYS2* allele, *lys2-128∂* (Simchen et al., 1984), and suppression of the lysine auxotrophy (Lys^- phenotype) conferred by a Ty *∂* element insertion into the 5' end of the *LYS2* gene present in this allele. The suppression of Lys^- phenotype (growth on medium lacking lysine) is referred to as the Spt^-.

RNAPII subunit mutagenesis and site-directed mutagenesis

PCR-based random mutagenesis was performed to introduce mutations into RNAPII subunits. For each individual RNAPII subunit, one or more screens were performed, mutagenizing different parts of longer subunit genes in different screens (*RPO21/RPB1* or *RPB2*), or the entire open reading frame (*RPB3*, *RPB7*, *RPB11*) for smaller subunit genes. Plasmids containing individual subunits and flanking DNA allowing for expression and termination from native elements were constructed in pRS-based *CEN*, low copy vectors (Sikorski and Hieter, 1989) for complementation of appropriate RNAPII subunit gene deletions. Mutagenesis conditions were under standard PCR conditions (1.5 mM MgCl2, 200 μM dNTPs) where 6 individual reactions for each mutagenized region were performed for 35 cycles of amplification followed by seeding of 6 additional reactions for 35 more cycles, followed by pooling of reactions for each mutagenized region. Mutagenized PCR products were transformed into yeast along with appropriate gapped plasmid

vector lacking most of the region to be mutagenized. Selection of vector marker (*LEU2*) after transformation allowed plasmids putatively created through gap repair recombination between gapped-vector and mutagenized PCR products to be obtained. These transformant populations represented putative mutant libraries that were screened for mutant phenotypes upon standard plasmid shuffling (Boeke et al., 1987) to remove the wild type *RPB* plasmid, allowing the transformed plasmid to be the only copy of the particular *RPB* subunit in the colony. Following replica-plating to 5FOA, plates were additionally replicated to a number of other media for phenotyping. Candidate mutants were streaked and rescreened for phenotypes followed by either dominance/recessive/plasmid linkage tests or direct plasmid isolation and retransformation to assay plasmid linkage/reproducibility of observed phenotypes. Mutant plasmids were isolated from yeast by standard procedures and portions of insert that had been subjected to PCR in initial mutagenesis were sequenced to identify alterations.

For *RPB3,* two independent replicate screens were performed, *Hin*dIII sites present in flanking genomic DNA were used to remove wild type sequence and gap-repair mutagenize the entirety of the *RPB3* ORF in an individual screen. For *RPB7*, vector *Xho*I and *Sac*I sites were used to remove the entire *RPB7* insert to allow mutagenesis of the entire *RPB7* ORF. For *RPB11*, a vector *Xho*I site and a *Hin*dIII site present in the 3' end of *RPB11* were used to remove wild type *RPB11* and mutagenize the majority if not all of the *RPB11* ORF. For *RPB2*, three individual screens were performed: the first utilized *Nde*I and *Hin*dIII sites to remove ~700 basepairs of RPB2 ORF 5' sequence for mutagenesis, the second utilized *Hin*dIII and

*Msc*I to remove a 1.6 kb internal ORF fragment for mutagenesis, and the third used an internal *Msc*I and vector *Sac*I site to remove a 1.6 kb *RPB2* ORF 3' sequence for mutagenesis. For *RPO21/RPB1*, three screens were performed. For the first, engineered silent *Bam*HI and *Pst*I restriction sites were utilized for removal of an internal *RPO21/RPB1* ORF fragment for mutagenesis. For the second, a 3' *RPO21/RPB1* ORF fragment up to the CTD was removed with digestion by *Xba*I and *Bsi*WI for mutagenesis. For the third, the entire *RPO21/RPB1* ORF 3' region was removed using an *Xba*I-*Sna*BI for mutagenesis. In general, 3000-5000 colonies were screened per independent screen and mutagenesis levels were approximated by number of 5FOA-sensitive colonies (representing lethal mutations and were between 2-7% in different screens).

For creation of *URA3*-marked *CEN* plasmids for pE-MAP screens, RNAPII alleles were moved from *LEU2*-marked *CEN* versions into pRS316 or pRS416 by standard cloning procedures. *URA3*-marked RNAPII mutant plasmids were then transformed into pE-MAP *rpbΔ* strains and screened for loss of wild-type *RPB LEU2* plasmids.

E-MAP analysis

E-MAP amenable strains for RNAPII mutants were created by transformation of MATa spore selectable marker (*mfa1p::HIS3*) into relevant MATa *rpbΔ* (complemented with WT *RPB CEN LEU2* plasmid) strains via recombination into the *CAN1* locus, generating *mfa1p::HIS3::can1Δ*. Mating type was switched to *MAT*a following standard procedures and mutant *rpb* RNAPII subunit genes were introduced on *CEN URA3* plasmids, followed by screening for loss of WT *RPB CEN*

LEU2 plasmid. E-MAP experiments and scoring of genetic interactions were then carried out as previously described (Collins et al., 2007b; Collins et al., 2006; Schuldiner et al., 2006).

ROC curves

All library deletion mutants that exist in both the RNAPII pE-MAP and the previously published deletion E-MAP (Collins et al., 2007b) were extracted, resulting in 404 mutants used for this analysis. First, pairs of proteins encoded by these genes were defined as physically interacting if listed with PE scores >2 and non-interacting if not (Collins et al., 2007a). Next, Pearson correlation coefficients were computed for all pairwise combinations of the 404 mutants, based on their pE-MAP profiles. An ROC curve was then generated to determine the power of these correlations to predict physical interactions between encoded proteins. Prior to generating the reference ROC curve, deletion E-MAP query strains with more missing data than the sparsest RNAPII mutant were removed, as were those that also exist in the library set. From this filtered E-MAP, 53 rows were randomly selected 1000 times, and an ROC curve was generated for each run. The median AROC and corresponding curve are reported in **Figure 2A**.

Genome-wide gene expression

Strains were streaked from -80°C stocks onto plates and grown for 3 days. Liquid cultures were inoculated with independent colonies and grown overnight in Synthetic Complete (SC) medium: 2g/l Drop out mix Complete and 6.71g/l Yeast

Nitrogen Base without AA, Carbohydrate & w/AS (YNB) from US Biological (Swampscott, USA) with 2% D-glucose. Overnight cultures were diluted to OD_{600}=0.15 in 1.5ml fresh medium and grown at 30°C in a 24 well plate in a Tecan Infinite F200 under continuous shaking. Growth curves were made for the mutant cultures (two cultures from two isolates) as well as for two wt inoculates, grown in parallel. Mutant and wt cells were harvested by centrifugation (6100 rpm, 3 min) at mid-log phase at OD_{600}=0.6, and pellets were immediately frozen in liquid nitrogen after removal of supernatant. Up to eleven mutant strains could be grown on a single day. wt cultures were grown parallel to the deletion mutants to assess day-to-day variance. RNA isolation and purification was performed (Lenstra et al., 2011). In short, total RNA was prepared by phenol extraction and cleaned up using a customized Sciclone ALH 3000 Workstation. For each sample, external control poly-A+ RNAs were added in equimolar amounts to the total RNA to enable monitoring of global changes in mutants (van de Peppel et al., 2003). For each micro-array analysis, 1.25μg labeled sample cRNA and 1.25μg reference cRNA was hybridized per slide. Each hybridization performed within this project was subjected to a number of quality controls. Some of these are based on the data from one single hybridization, while others are based on comparing data from one single hybridization against the wt grown in parallel (Lenstra et al., 2011). Two channel microarrays were used. RNA isolated from a large amount of wt yeast from a single culture was used as a common reference. This common reference was used in one of the channels for each hybridization and used in the statistical analysis to obtain an average expression-profile for each mutant relative to the wt. Two independent

cultures were hybridized on two separate microarrays. Each gene is represented twice on the microarray, resulting in four measurements per mutant.

Correlations

All correlations are Pearson correlation coefficients, unless otherwise noted. Genetic interaction correlations are based on the complete genetic interaction profiles, and gene expression correlations on the genome-wide expression profiles.

Identification of functional links between RNAPII mutants and protein complexes

Published genetic interaction data (Costanzo et al., 2010) were scaled to E-MAP format using a non-linear scaling method (Ryan et al., 2012), and processed to remove duplicate entries of genes represented by more than one allele (deletion, temperature sensitive or hypomorphic). If the deletion was present it was kept and all other alleles removed, and if no deletion existed the allele to keep was selected randomly. Pearson correlation coefficients were computed for each RNAPII mutant against all genes in this dataset, based on the genetic interaction profiles with the intersecting library mutants. The genes were then grouped by membership in physical complexes of their encoded proteins. Complex definitions were collected from (Benschop et al., 2010), and modified by first removing all binary complexes, and then excluding any gene assigned to more than one complex. To identify allele-complex pairs which were significantly correlated, we used a one-sided Mann-Whitney U test to compare the correlations between each RNAPII mutant and the

members of each complex to (i) the correlations between the same mutant and all genes not in that complex, and to (ii) the correlations between the same complex and all other mutants. The highest p-value of the comparison to (i) or (ii) was recorded. False discovery rates (FDR) were computed using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995), and are reported in **Table S4**.

Chromosome transmission fidelity (CTF) assay

RNA polymerase mutants containing a chromosome VII fragment and *ade2-101* were generated. Red colony color is caused by accumulation of pigment due to a block in adenine production caused by the *ade2-101* (ochre) mutation. The block is relieved in the presence of the *SUP11* gene on the telocentric arm of the chromosome fragment, encoding an ochre-suppressing $tRNA^{Tyr}$. Cells that lose the chromosome VII fragment develop red color so that colonies exhibiting unstable inheritance of the chromosome fragment develop red sectors. Mutant strains were plated at 100-400 cells a plate on synthetic complete medium with only 20% adenine and left to grow into colonies for 7 days at 30°C. Replicates were counted to obtain the percentage of sectored colonies over total colonies for each mutant strain (Spencer et al., 1990; Yuen et al., 2007).

Primer extension

Primer extension assays were performed as previously described (Ranish and Hahn, 1991) with the following modifications. Briefly, 30µg total yeast RNA purified

as described (Schmitt et al., 1990) was annealed with ^{32}P end-labeled oligo priming downstream of start sites in a 15µl reaction volume instead of 10µl, followed by reverse transcription reaction supplemented with RNase Inbibitor (Fermentas) by M-MLV Reverse Transcriptase (Fermentas), in a total volume of 45µl instead of 30µl. Products were separated on 8% polyacrylamide gels (19:1 acrylamide:bisacrylamide) containing 1XTBE and 7M urea. *tfg2Δ261-273* mutant plasmids were gifted from the Hahn lab and integrated into yeast strain CKY283 (**Table S1**). Reported values are means derived from at least three independent experiments.

Growth rate assay

Strains were cultured overnight in YPD, then resuspended to OD_{600} = 0.3 in the morning, and grown at 30°C until reaching OD_{600} = 0.7 - 1.0. These cultures were diluted in YPD to OD_{600} = 0.05, 100µl/well, in a Corning Costar 3631 clear bottom 96-well plate. Each culture was distributed into 3 wells as technical replicates. Plates were sealed with Corning 3930 clear polysterene lids, and edges taped to avoid evaporation. Growth curves were measured at 30°C in a Tecan Infinite 200 PRO microplate reader, controlled by Tecan Magellan software. OD_{600} measurements were performed every 15 minutes (top reading, 25 reads per well), with orbital shaking during incubation. Curves were fitted to the measured time series with a model-free spline method from the grofit package (Kahm et al., 2010), in the statistical computing environment R. Exponential growth rates were determined by grofit as the maximum slopes of the spline fits. For each set of

technical replicates, the $\log_2$ ratio of mutant to wt growth rate was computed. Reported averages and standard deviations were calculated over the $\log_2$ ratios in the three replicate sets.

Splicing microarray assay

Cultures were grown in rich medium according to standard techniques (Guthrie and Fink, 2002). Saturated cultures were diluted to OD_{600} = 0.1 in the morning and allowed to grow at 30°C until reaching mid-log phase (OD_{600}=0.5-0.7). Mutant strains and an isogenic wild-type were collected by centrifugation and snap frozen in liquid nitrogen. Where indicated, a wild-type log-phase culture was treated with mycophenolic acid (30μg/ml) for 10 minutes, harvested using vacuum filtration and snap frozen in liquid nitrogen. Total cellular RNA was isolated using a hot acid phenol extraction generally as outlined in (Schmitt et al., 1990), but with some specific modifications (Bergkessel et al., 2011). cDNA from each strain was synthesized, and labeled with Cy3 or Cy5 according to the low-throughput sample preparation method (Pleiss et al., 2007). The optimized oligos (Pleiss et al., 2007) were robotically arrayed onto poly-L-lysine coated glass slides (slides from ThermoScientific C40-5257-M20) and slides were processed according to (DeRisi et al., 1997; Pleiss et al., 2007). Microarrays were scanned using Axon Instruments GenePix 4000B at 635nm and 532nm wavelengths and image analysis was done using Axon Instruments GenePix Pro version 5.1. Spots were manually removed from analysis if they contained uncharacteristically high background or obvious defects; the ratio of the median intensity for 535nm and 625nm was calculated for

each remaining spot. Each biological replicate contains 6 technical replicates for each feature as well as dye-flipped replicates, which were combined and normalized (Pleiss et al., 2007). For all mutants, data from at least two biological replicates were used. To emulate a classic splicing measure that compares the levels of pre-mRNA and mature mRNA (Pikielny and Rosbash, 1985), we computed I/J ratios as $\log_2(I/J) = \log_2(Intron_{mutant}/Intron_{WT})-\log_2(Junction_{mutant}/Junction_{WT})$ for each gene. The reported p-values (**Table S7**) were calculated via one-sided Wilcoxon signed rank tests, comparing each I/J distribution to that resulting from direct hybridization of cDNA from two separately grown wild-type cultures.

qPCR assay for splicing efficiency

Splicing efficiency was measured by qPCR of cDNA as in (Pleiss et al., 2007) with modifications as described below. Ten µg of RNA were treated with 4U RQ1 DNaseI (Promega) in 16µL according to manufacturer instruction (20 min 37˚C). 4µL Stop Solution was added, and DNase was inactivated by incubation at 65˚C for 10 min. Samples were primed with random 9-mers in 40µL (50mM Tris-HCl pH8.3, 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 400ng dN9; 5 min 65˚C; 5 min on ice) and divided into equal 40µL reverse transcription reactions with M-MLV reverse transcriptase (+RT or -RT in 50mM Tris-HCl pH8.3, 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 250mM each dNTP; 2 hours 42˚C). qPCRs were performed in 25µL consisting of 1X Standard *Taq* Buffer and 1.25U *Taq* (New England Biolabs), 200mM each dNTP, 0.0013% SYBR Green I (Sigma), 600nM each primer, DMSO as needed (**Table S8**), and 25x – 2500x dilution of cDNA (-RT samples diluted equivalently to match

+RT). Amplification reactions were performed on a C1000 ThermoCycler (BioRad) (94˚C 3 min and 35 cycles of 94˚C 15 sec, 50˚C 15 sec, 72˚C 15 sec, plate read). Each qPCR run was finished with a melt curve to confirm homogeneity of the amplified product. Starting quantity was calculated from a genomic DNA standard curve for each primer set; standard curve reactions not falling in the linear range were removed manually. Two technical replicates were performed for 2-8 biological replicates (F1086S n=8, H1085Q n=5, sub1Δ n=2, MPA n=3, E1103G n=6, G1097D n=5). -RT control samples yielded negligible amplification (data not shown). Each gene was measured using primer sets specific for pre-mRNA and total mRNA (**Table S8**). To generate the graph in **Figure S6**, the pre-mRNA/total mRNA ratio for each mutant was normalized to a within-experiment wild-type before taking the mean of the biological replicates.

Sorting array mutants on their interactions with fast and slow mutants

Mean genetic interaction scores with the fast or slow RNAPII groups were calculated for all array mutants with at least two scores in each group. The 404 mutants with different signs of mean(fast) and mean(slow) were then sorted by the difference of the means.

Acknowledgements

The authors wish to thank members of the Krogan lab, Paivand Jalalian, Ada Cheng, Tanja Kortemme, David Agard, Andrej Sali, Keith Yamamoto and C. James Ingles for helpful discussion, Steve Buratowski, Olga Calvo and Eva Nogales for

sharing unpublished results, Colm Ryan for help with statistical analysis, and Jiewei Xu and Ricardo Almeida for assistance with growth assays. This work was supported by grants from QB3@UCSF, the NIH (GM084448, GM084279, GM081879 and GM098101 to NJK, GM36659 to Roger D. Kornberg for support of CDK for a portion of this work, GM097260 to CDK, GM21119 to CG and DP5OD009180 to JF). CG is an ACS Research Professor of Molecular Genetics. EAM was supported by a graduate fellowship from NSF. CDK acknowledges a Helen Hay Whitney Fellowship for the early stages of this work, whereas HB was supported by funding from the Biophysics Graduate Program at UCSF. NJK is a Searle Scholar and a Keck Young Investigator.

Figures

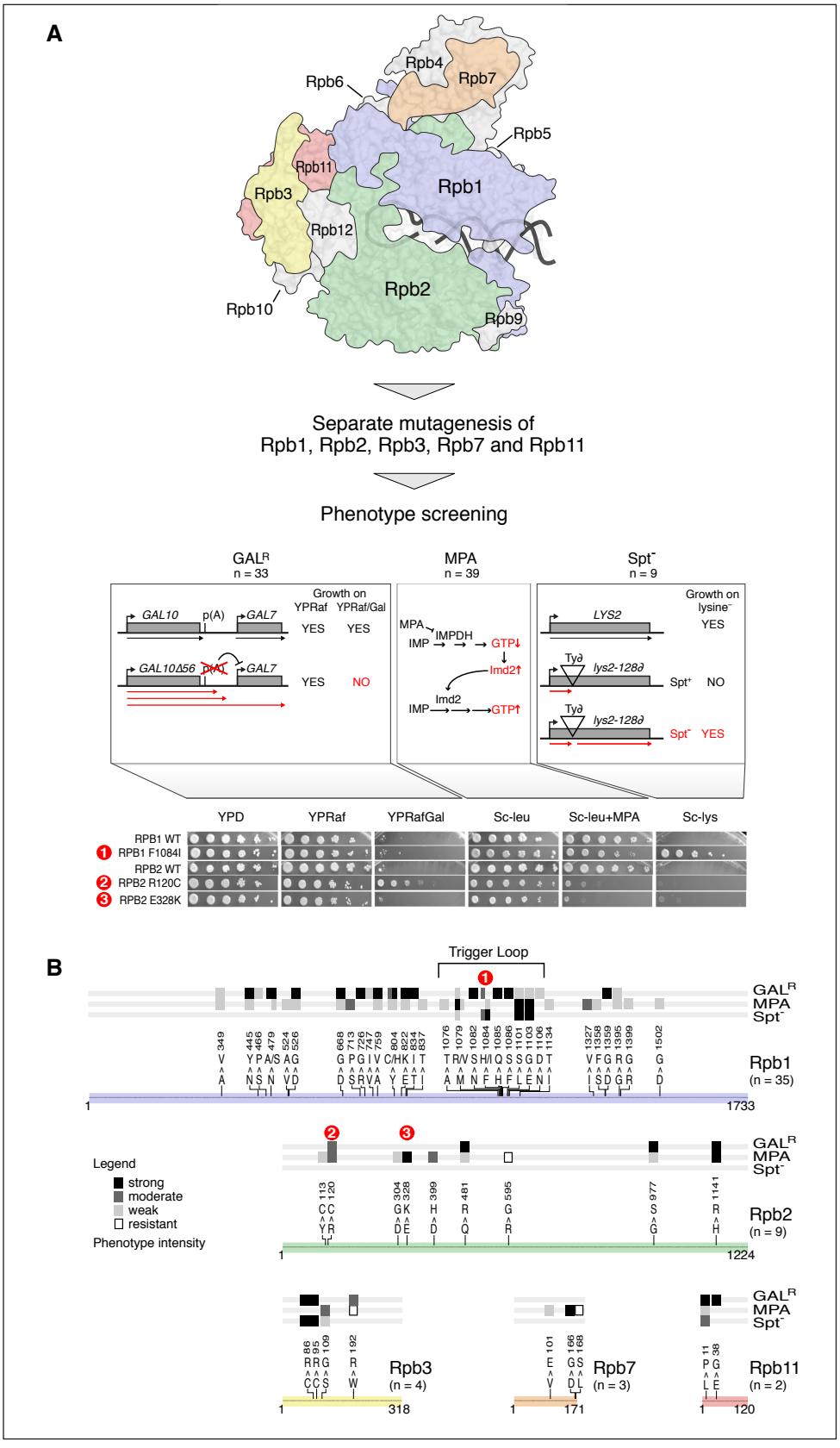

A
Rpb4
Rpb7
Rpb6
Rpb5
Rpb11
Rpb3
Rpb1
Rpb12
Rpb2
Rpb10
Rpb9
Separate mutagenesis of
Rpb1, Rpb2, Rpb3, Rpb7 and Rpb11
Phenotype screening
GALR
n = 33
MPA
n = 39
Spt-
n = 9
Growth on
YPRaf YPRaf/Gal
GAL10 p(A) GAL7
YES YES
GAL10Δ56 p(A) GAL7
YES NO
MPA
IMP IMPDH GTP↓
Imd2↑
Imd2
IMP GTP↑
Growth on
lysine-
LYS2
YES
Tyð lys2-128ð
Spt+ NO
Tyð lys2-128ð
Spt- YES
YPD
YPRaf
YPRafGal
Sc-leu
Sc-leu+MPA
Sc-lys
RPB1 WT
1 RPB1 F1084I
RPB2 WT
2 RPB2 R120C
3 RPB2 E328K
B
Trigger Loop
GALR
MPA
Spt-
Rpb1
(n = 35)
1
1733
Legend
strong
moderate
weak
resistant
Phenotype intensity
Rpb2
(n = 9)
1224
Rpb3
(n = 4)
318
Rpb7
(n = 3)
171
Rpb11
(n = 2)
120

Figure 1. Generation and selection of RNAPII point mutants. A) Random mutagenesis was carried out by PCR separately on each of the five essential subunits that are unique to RNAPII and the mutants were screened for three transcription-related phenotypes. Gal^R, left: Deletion of the major *GAL10* polyadenylation signal (*gal10*D56) results in RNAPII read-through and interference with initiation at *GAL7*, causing a Gal-sensitive phenotype. Certain mutants suppress Gal-sensitivity by increasing *GAL10* 3'-end formation and/or subsequent downstream termination, thereby increasing *GAL7* expression and resulting Gal resistance (Gal^R) (Kaplan et al., 2005). MPA sensitivity, middle: Mycophenolic acid (MPA) is a drug that inhibits IMP-dehydrogenase (IMPDH), which functions in GTP synthesis, but is counteracted by upregulation of *IMD2*, which encodes an MPA-resistant form of IMPDH. Transcriptional defects that are sensitive to lowered GTP pools, or that reduce *IMD2* expression, render cells sensitive to MPA. Spt^-, right: Insertion of a Ty-retrotransposon d element into *LYS2* (*lys2-128*d*)* results in a lysine auxotrophy due to transcription block. Certain mutants suppress *lys2-128*d and allow expression of *LYS2*, due to activation of an internal promoter, producing a shorter but functional *LYS2* mRNA. Spot tests used to identify each phenotype for three representative mutants are displayed in the lower panel. **B**) The positions and mutations of the 53 single point mutants that exhibit at least one phenotype and were selected for further analysis with the pE-MAP technology (**Figure S1**, **Table S1**). The colored lines correspond to the amino acid residue sequences of each subunit and mutations are denoted by their sequence position and single letter amino acid codes for wild-type and mutant. Counts under subunit labels show total number of mutants

identified. The trigger loop, highlighted, is a mobile element in the active site of Rpb1 that couples recognition of a correct NTP base with catalysis of phosphodiester bond formation.

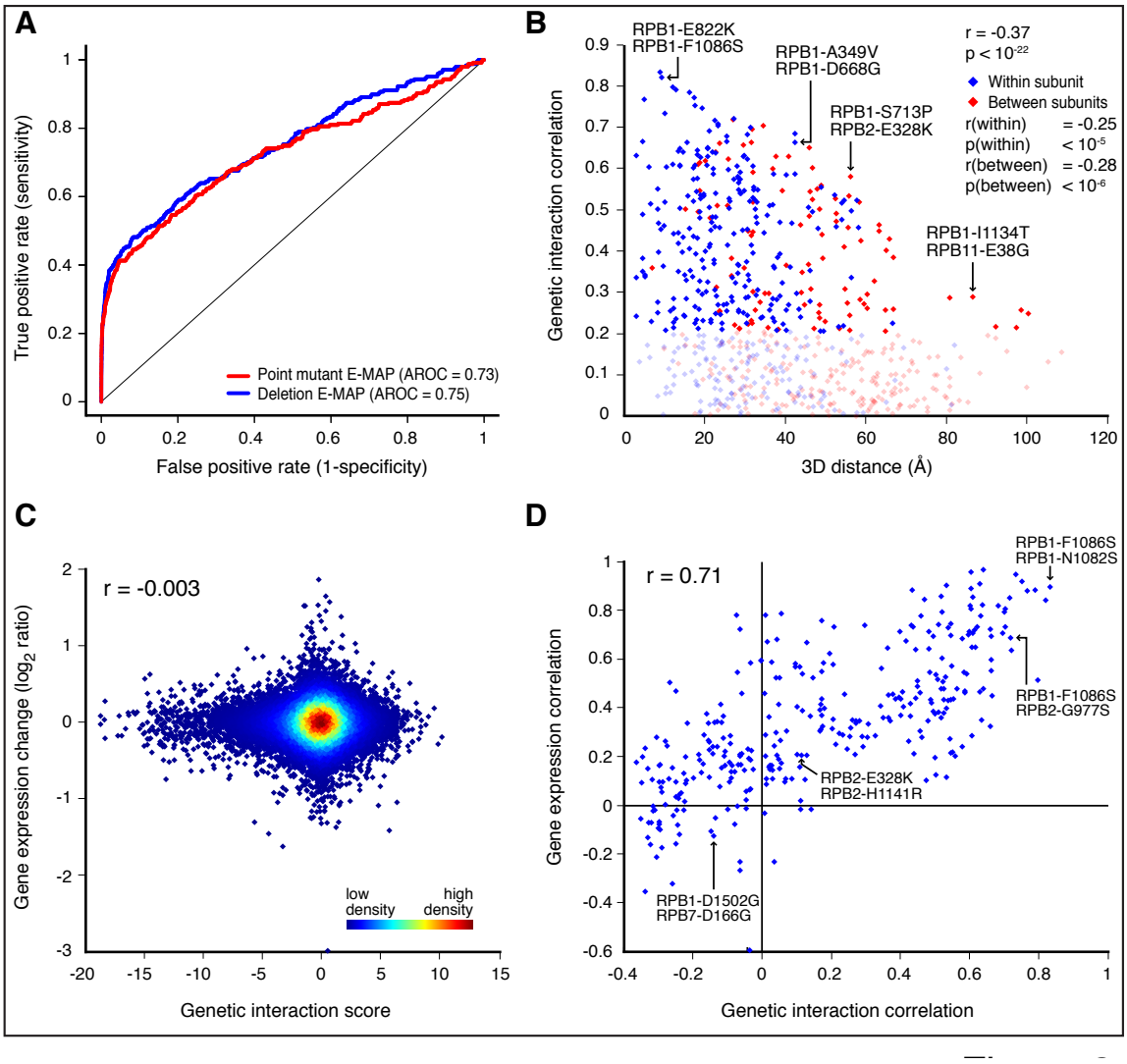


Figure 2. pE-MAP interactions span numerous biological processes, depend on spatial location of mutated residues and are not direct consequences of changes in gene expression. A) ROC curves comparing the power of genetic profile correlations from the pE-MAP (red) and an E-MAP focused on chromosome biology (Collins et al., 2007b) (blue) to predict physical interactions between pairs of proteins (Collins et al., 2007a) (**Supplementary Methods**). The area under the curve (AROC) is commonly used to determine predictive power. **B**) Genetic profile correlations between pairs of mutated residues compared to the three-dimensional

distance between their α-carbons (**Table S3**). Blue points correspond to residue pairs within the same subunit of RNAPII whereas red points represent pairs in different subunits. Negatively correlated residue pairs were excluded, as were four mutants of residues that are absent from the atomic coordinate file used (PDB ID 2E2H) (Rpb1 D1502, Rpb7 V101, Rpb7 D166, Rpb7 L168) (Wang et al., 2006). Correlations between 0 and 0.2 are dimmed to highlight the trends at higher correlations. **C**) Effect of RNAPII point mutations on gene expression, compared to the corresponding genetic interaction scores between RNAPII mutants and library gene deletion/DAmP alleles (**Supp. Data 3**). The scatter includes all combinations of the 26 RNAPII mutants and 1192 library genes/mutants that were examined via both pE-MAP and gene expression analyses. No global changes in gene expression were observed (measured by spike-in control RNA, **Supplementary Methods**). **D**) Comparison between pairwise RNAPII mutant correlations of genetic interaction profiles and gene expression profiles.

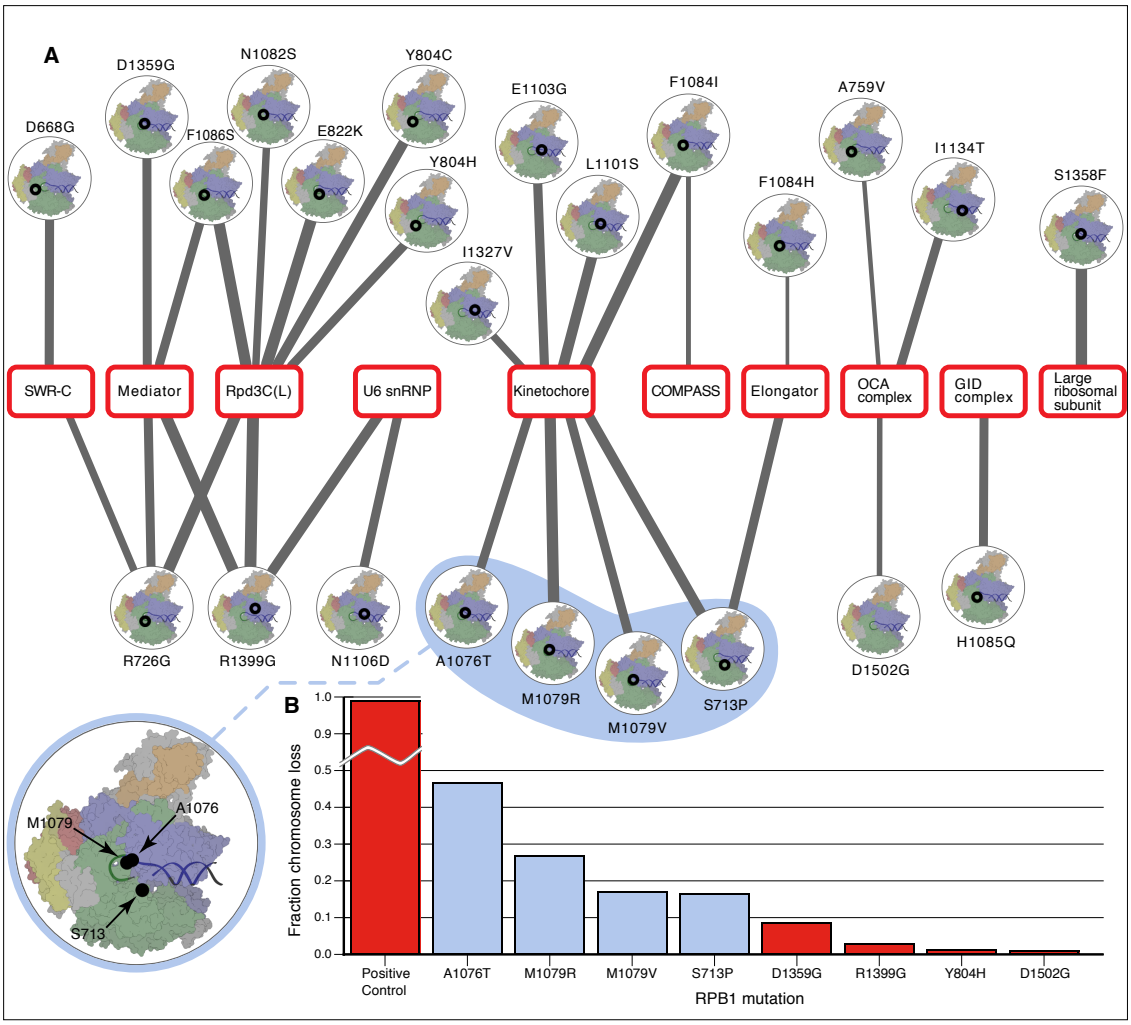


Figure 3. Comparison of the pE-MAP with previous collected genetic interaction data reveals functional associations between RNAPII residues and protein complexes. A) Correlations were computed between each RNAPII mutant genetic profile and previously published profiles for 4417 gene deletions and hypomorphic alleles (Costanzo et al., 2010). Genes were annotated by complex membership of encoded proteins and highly correlated intersections of RNAPII mutants and complexes were identified using Mann-Whitney U statistics. The edge widths correspond to the statistical significance of connections, and only *RPB1* edges with a false discovery rate (FDR) < 0.1 are displayed (**Table S4**). Four

mutated residues linked to the kinetochore are highlighted in blue and the blowup indicates their structural locations. A full point mutant-module genetic interaction map using mutations in all subunits is presented in **Figure S3**. **B**) 19 mutants were examined using a chromosome transmission fidelity (CTF) assay (**Table S5**). The four kinetochore-linked mutants highlighted in (**A**) exhibit chromosome loss in more than 15% of their colonies (blue bars), whereas unlinked mutants display no or weak phenotype (red bars, representative set).

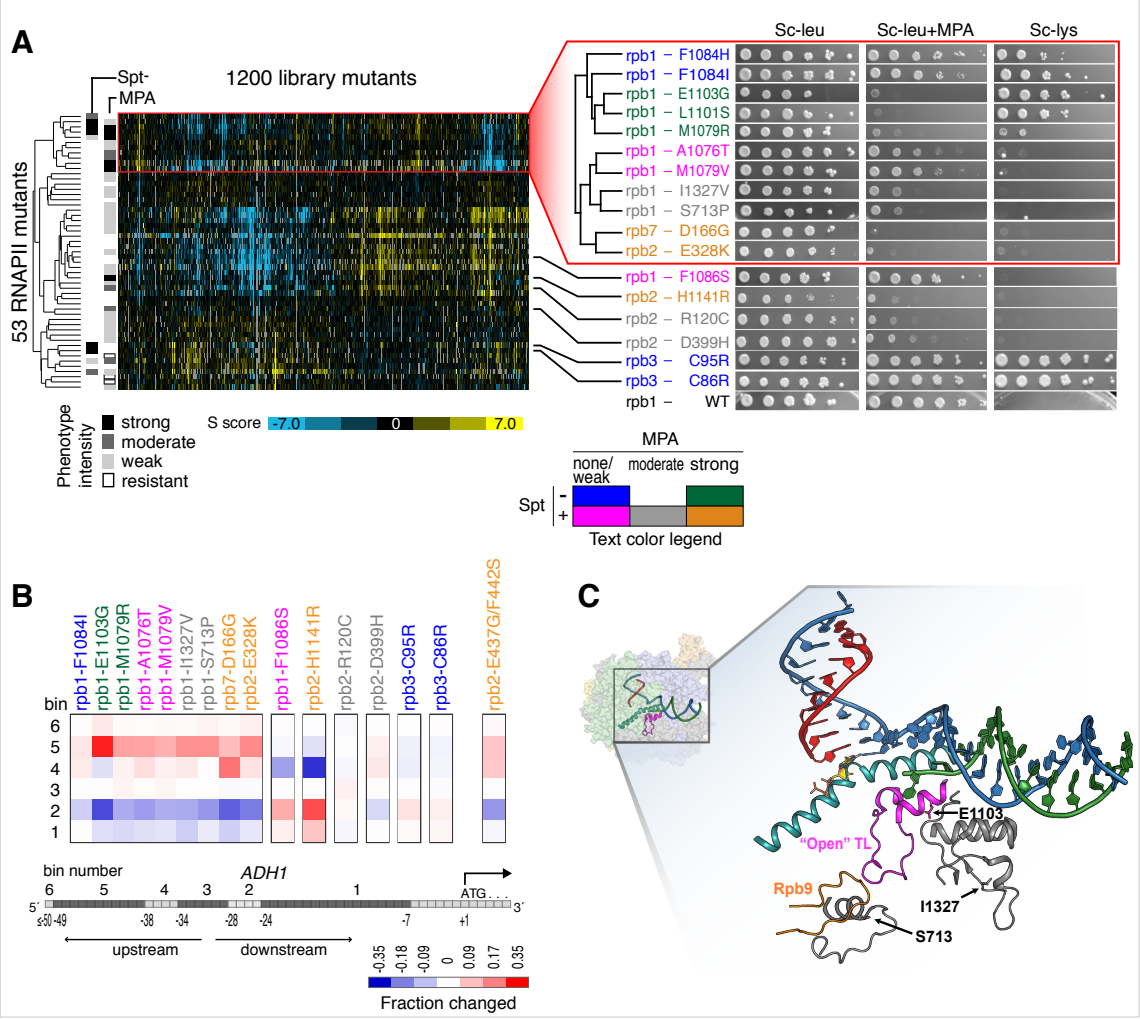


Figure 4. pE-MAP profiles differentiate between subtle changes in transcription-related phenotypes and identify RNAPII mutations that affect start site selection. A) pE-MAP clustering in relation to MPA sensitivity and Spt phenotypes of alleles. The RNAPII alleles are clustered by their pE-MAP profiles and colored based on their degrees of MPA sensitivity and Spt phenotypes (determined from the displayed spot tests). Notably, the pE-MAP clustering differentiates between subtle changes in phenotypes. **B**) Effect of RNAPII mutations on start-site selection at *ADH1* determined by primer extension analysis. The heatmap colors describe the fractional change of start site in each bin of the *ADH1* schematic

(bottom). The leftmost block of the heatmap corresponds to the 9 tested alleles from the subcluster highlighted in (**A**), all of which display an upstream start site selection compared to wild-type. The following 6 columns correspond to the mutants that cluster separately in (**A**), and these exhibit either downstream (*rpb1* F1086S, *rpb2* H1141R) or no start site shift. The rightmost column highlights the upstream start-site effect of *rpb2* E437G/F442S, a double mutant that clusters with the highlighted cluster in (**A**) (**Figure S4C, D**). **C**) Rpb1 I1327 and Rpb1 S713 connect to the trigger loop (TL, magenta) in the active site of RNAPII. Mutations in I1327 could affect the structural region of the trigger loop (E1103) via a network of loops and helices in Rpb1 (grey), while S713 is in close proximity to the TL active site in its open conformation, via a Rpb9 loop (orange). In particular, the proline substitution, S713P, could result in significant structural changes affecting the TL. Coordinates for TL, Rpb9 and the loop containing S713 are from PDB ID 1Y1V (Kettenberger et al., 2004) and all other coordinates from 2E2H (Wang et al., 2006). The bridge helix is shown in cyan, template DNA in blue, non-template DNA in green, RNA in red, and the incoming GTP base is colored by atom.

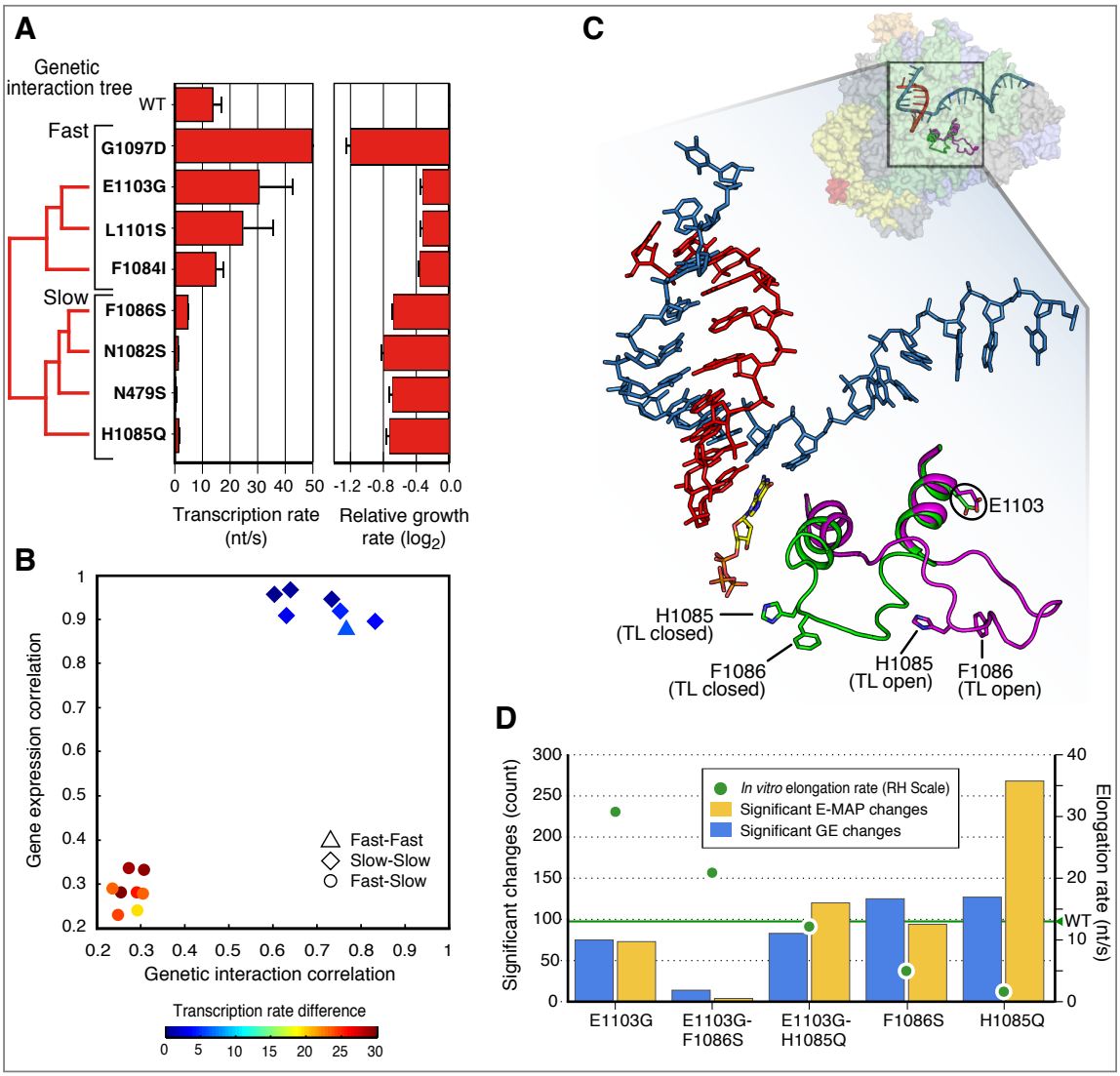


Figure 5. pE-MAP and expression profiles are indicative of biochemical activity. A) *In vitro* transcription rates from (Kaplan et al., 2012) and *in vivo* growth rates relative to wild-type for mutants in the active site of RNAPII. The dendrogram was generated via hierarchical clustering of the genetic profiles from mutants in the active site known to affect RNAPII speed (**Supp. Data 5**). Means and standard deviations of growth rates were derived from three replicates. Note that *rpb1* G1097D was too sick to be reproducibly analyzed using the E-MAP approach. **B**) *In*

vitro transcription rate difference (nt/s) between pairs of active site mutants in relation to their genetic and gene expression profile correlations. **C**) The trigger loop (TL), shown in green (closed conformation) and magenta (open conformation) couples recognition of the correct RNA base with phosphodiester bond catalysis. Residues H1085 and F1086 reside in the catalytic site of the TL, while E1103 is part of the distal flanking alpha helix that structurally constrains the TL in open conformations. Template DNA is shown in blue, the growing RNA strand in red and the incoming GTP base is colored by atom. Coordinates for TL in open formation are from PDB ID 1Y1V and all other coordinates are from 2E2H. **D**) Counts of significant genetic interactions in the complete genetic profiles, or gene expression changes larger than 1.7 fold in the genome-wide expression profiles, of the indicated RNAPII mutants. Significantly positive and negative genetic interactions were defined as those higher than the 97.5 percentile (pE-MAP score > 3.3) or lower than the 2.5 percentile (pE-MAP score < -5.1) of the complete pE-MAP, respectively. *In vitro* transcription rates are indicated on the right-hand scale.

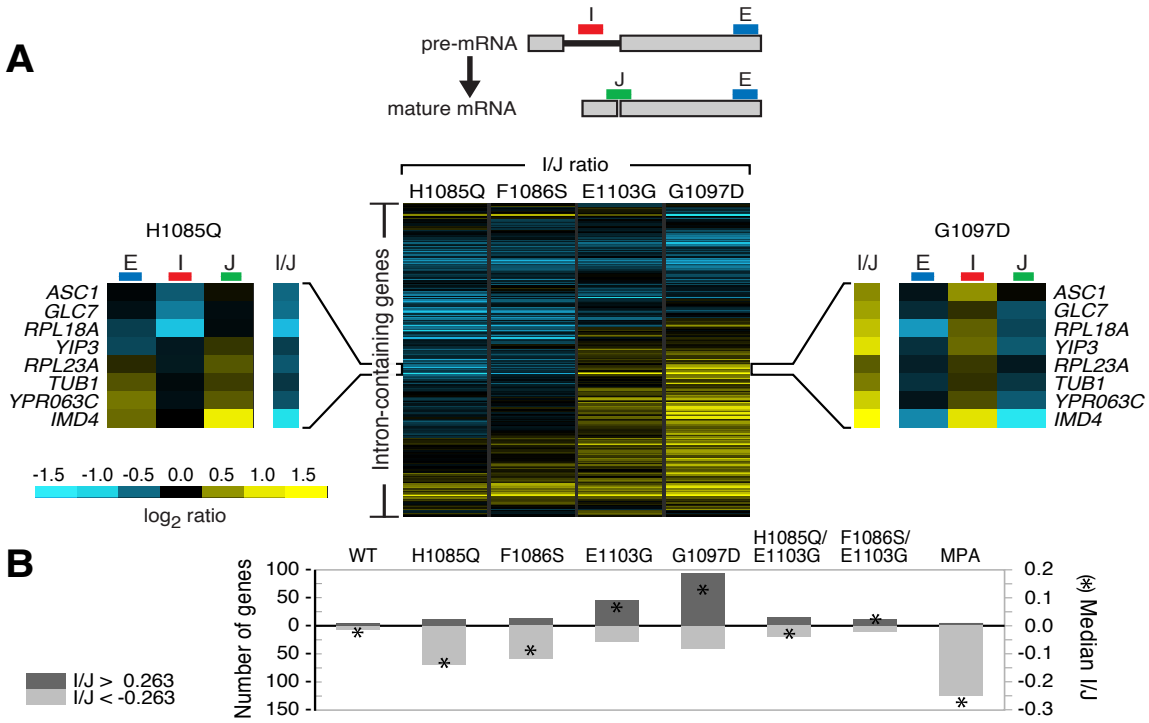
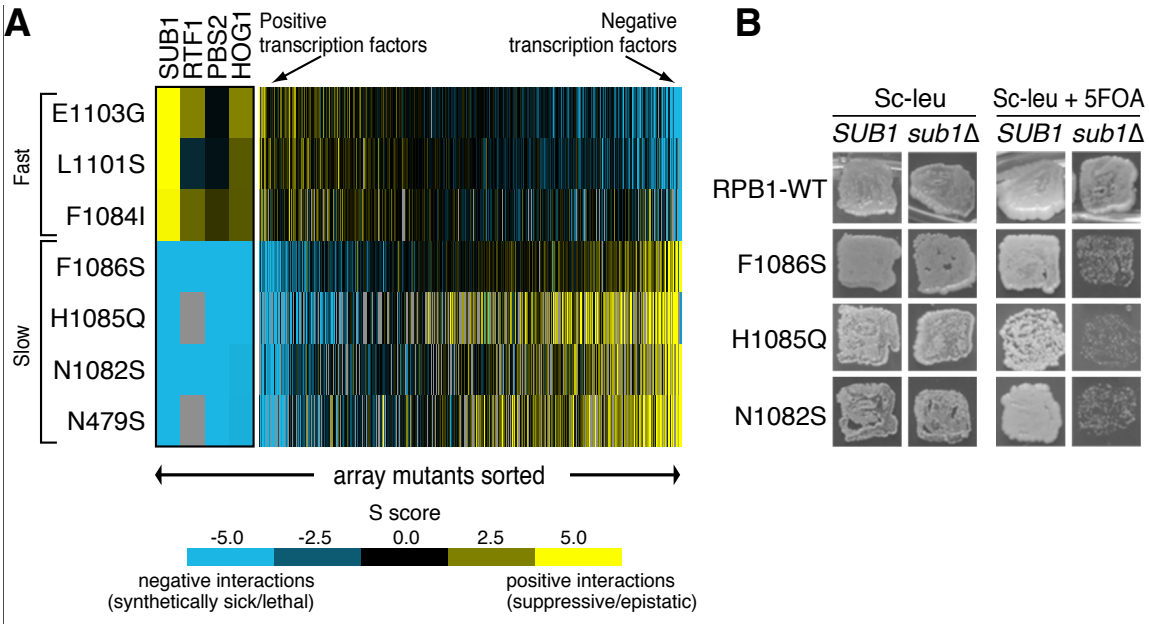
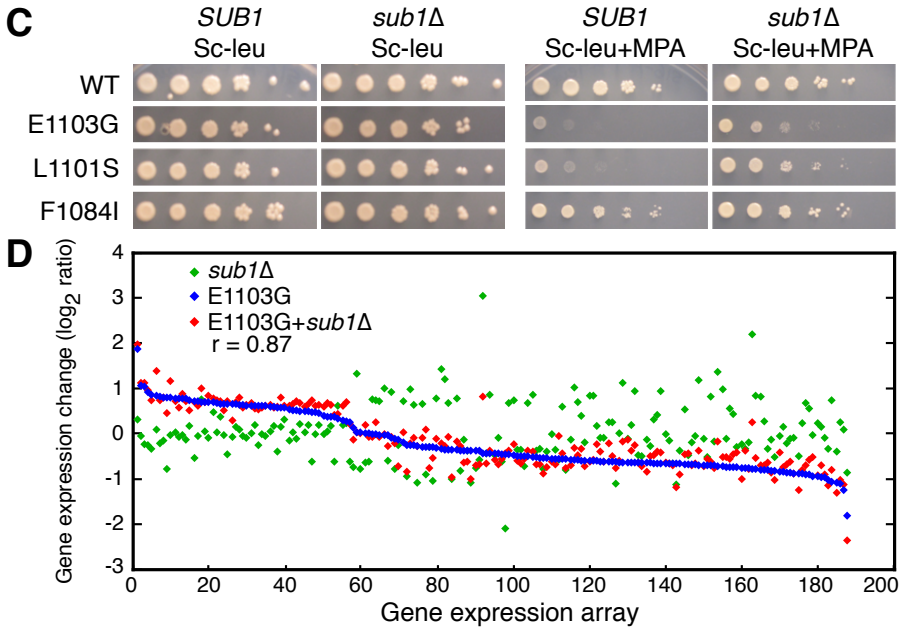


Figure 6. Effects of altering RNAPII transcription rate on *in vivo* splicing efficiency. **A**) Top panel shows schematic of the microarray for each intron-containing gene in yeast: probe I (intron) hybridizes to pre-mRNA, J (junction) to mature mRNA, and E (exon) to both. The center panel displays a heat map of Intron/Junction ratios for the slow (*rpb1* H1085Q and F1086S) and fast (*rpb1* E1103G and G1097D) mutants, corresponding to the enrichment of pre-mRNA over mature mRNA. The side-panels highlight a subset of genes that behave reciprocally in fast and slow RNAPII backgrounds. **B**) Number of genes exhibiting a >20% change in I/J ratio (bars, left-hand scale), and median I/J ratio (asterisks, right-hand scale), for a wild-type-versus-wild-type negative control, the indicated slow and fast *rpb1* single mutants, mutually- suppressive *rpb1* double mutants, and wild-type treated with MPA for 10 minutes.


A
SUB1
RTF1
PBS2
HOG1
Positive transcription factors
Negative transcription factors
Fast
E1103G
L1101S
F1084I
Slow
F1086S
H1085Q
N1082S
N479S
array mutants sorted
S score
-5.0
-2.5
0.0
2.5
5.0
negative interactions (synthetically sick/lethal)
positive interactions (suppressive/epistatic)
B
Sc-leu
Sc-leu + 5FOA
SUB1 sub1Δ
SUB1 sub1Δ
RPB1-WT
F1086S
H1085Q
N1082S


C
SUB1 Sc-leu
sub1Δ Sc-leu
SUB1 Sc-leu+MPA
sub1Δ Sc-leu+MPA
WT
E1103G
L1101S
F1084I
D
sub1Δ
E1103G
E1103G+sub1Δ
r = 0.87
Gene expression change (log2 ratio)
Gene expression array

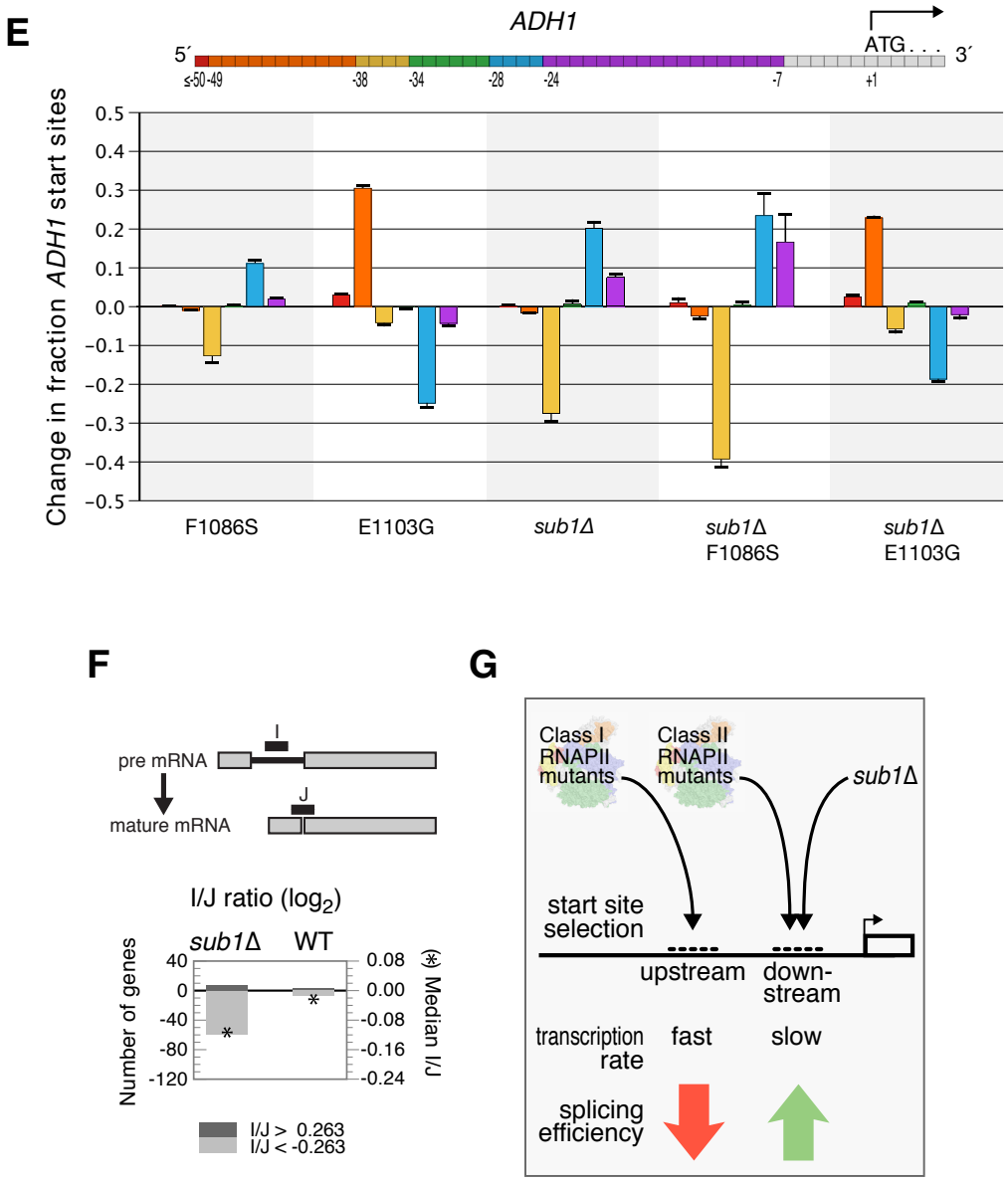


Figure 7. Genetic interaction patterns with fast and slow RNAPII mutants reveal Sub1 as a transcription factor that regulates start site selection and influences mRNA splicing. A) Genetic profiles of array mutants, sorted on the difference between their average interaction with fast and slow RNAPII mutants (**Supp. Data 8**). **B**) Patch tests examining the effect of *sub1*Δ on slow trigger loop mutants. In the left panel *RPB1* wild-type is present on a *CEN URA3* plasmid and the

strains are not sensitive to *sub1*Δ. In the right panel, only slow RNAPII is present and synthetic sickness with *sub1*Δ is observed. Mutant *rpb1* was expressed on a *CEN LEU2* plasmid. **C**) Spot tests examining the effect of *sub1*Δ on fast mutants in absence (left) and presence (right) of MPA. **D**) Comparison of *sub1*Δ effect on gene expression in E1103G (difference between red and blue) and wild-type (difference between green and y=0) (**Supp. Data 3**). The *x*-axis represent all array transcripts that exhibit a 1.5 fold or stronger expression change in at least one of the three mutants. The transcripts are sorted by magnitude expression change in E1103G. **E**) Primer extension analysis at *ADH1* to map transcription start-sites for F1086S, E1103G and s*ub1*Δ mutations (**Figure S7B**). The colors of the bars correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to wild-type. Means and standard deviations were derived from five independent experiments. **F**) Splicing microarray analysis of *sub1*Δ, as described in **Figure 6** (**Supp. Data 7**). Number of genes exhibiting a >20% change in I/J ratio (bars, left-hand scale), and median I/J ratio (asterisks, right-hand scale), for *sub1*Δ and wild-type-versus-wild-type negative control. **G**) Model for the effect of Sub1 and RNAPII activity on start site selection and splicing. Fast RNAPII mutations (Class I) result in upstream transcription start and diminished splicing efficiency, whereas *sub1*Δ or slow RNAPII mutations (Class II) shift transcription start downstream and enhance splicing.

Supplementary Figures

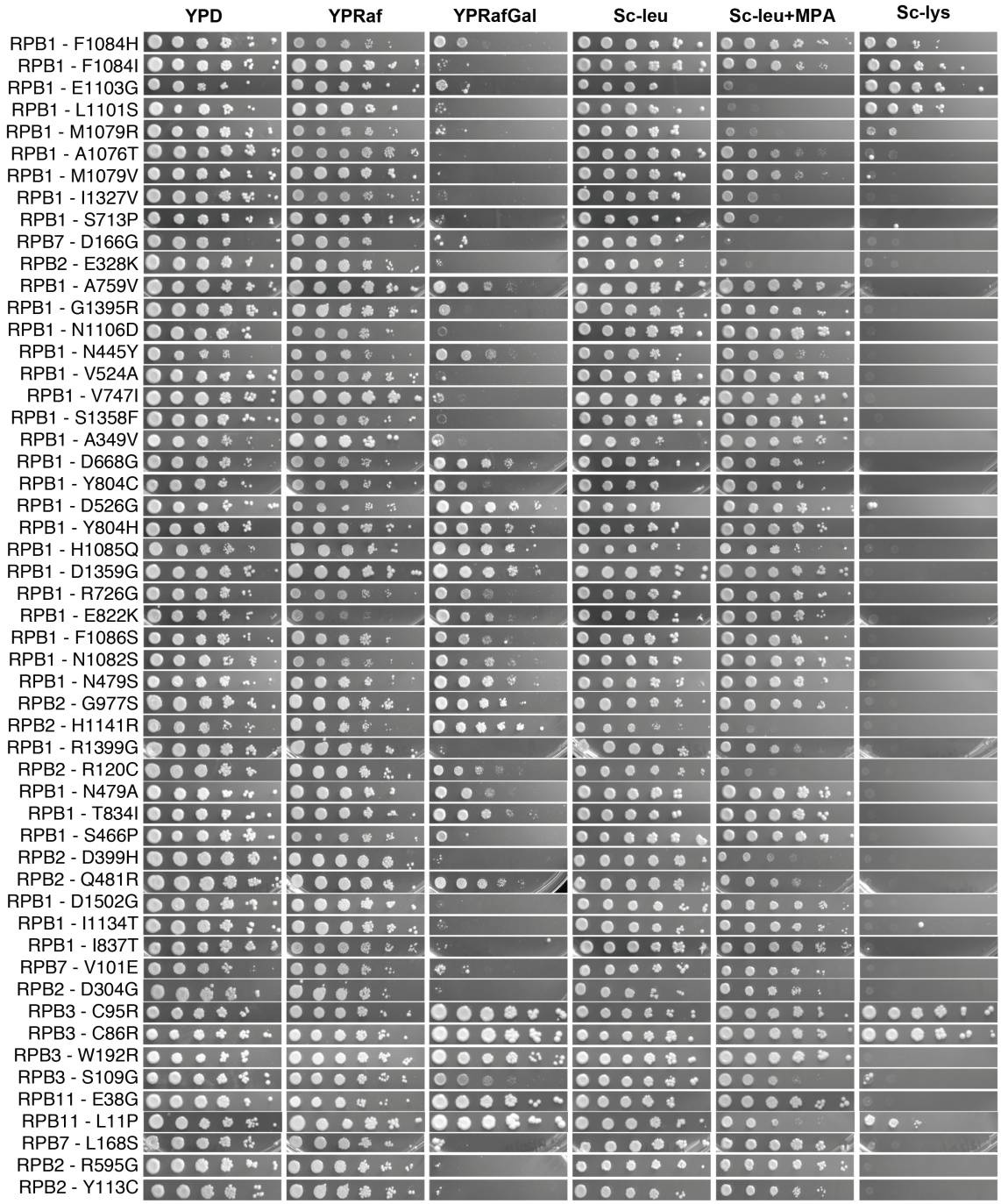


Figure S1. Complete collection of spot tests for identification of Gal^{R}, MPA sensitivity and Spt^{-} phenotypes. Spot tests on relevant media were carried out to identify RNAPII mutants that exhibit at least one of the three phenotypes, as

described in **Figure 1** and **Methods**. The order of RNAPII mutants follows that of the clustered pE-MAP (**Figure S2**, **Supp. Data 2**).

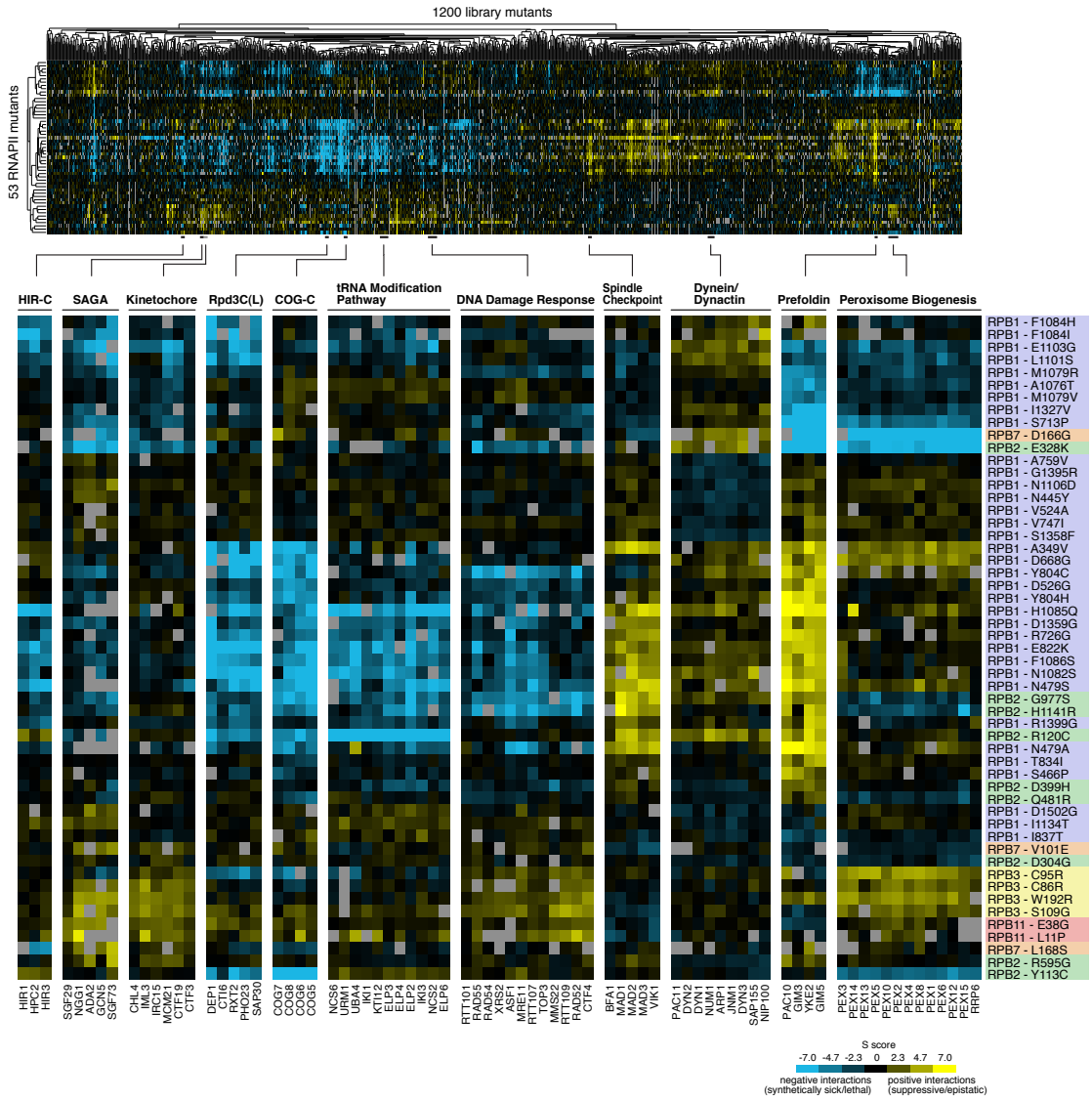


Figure S2. A high-resolution genetic interaction map of RNAPII point mutants. The 53 point mutants described in **Figure 1B** were genetically screened against a library of 1200 mutants carrying single gene deletions or DAmP alleles (**Table S2**). The resulting pE-MAP reports on a total of 59,534 genetic interactions between single residue point mutations and deletions of non-essential genes or DAmP alleles of essential genes. The dendrograms organize the mutants functionally and were generated by 2-D hierarchical clustering. Several representative clusters of library mutants that belong to the same complex or pathway are highlighted. The

background colors of RNAPII mutant labels correspond to subunit and match the color scheme in **Figure 1**.

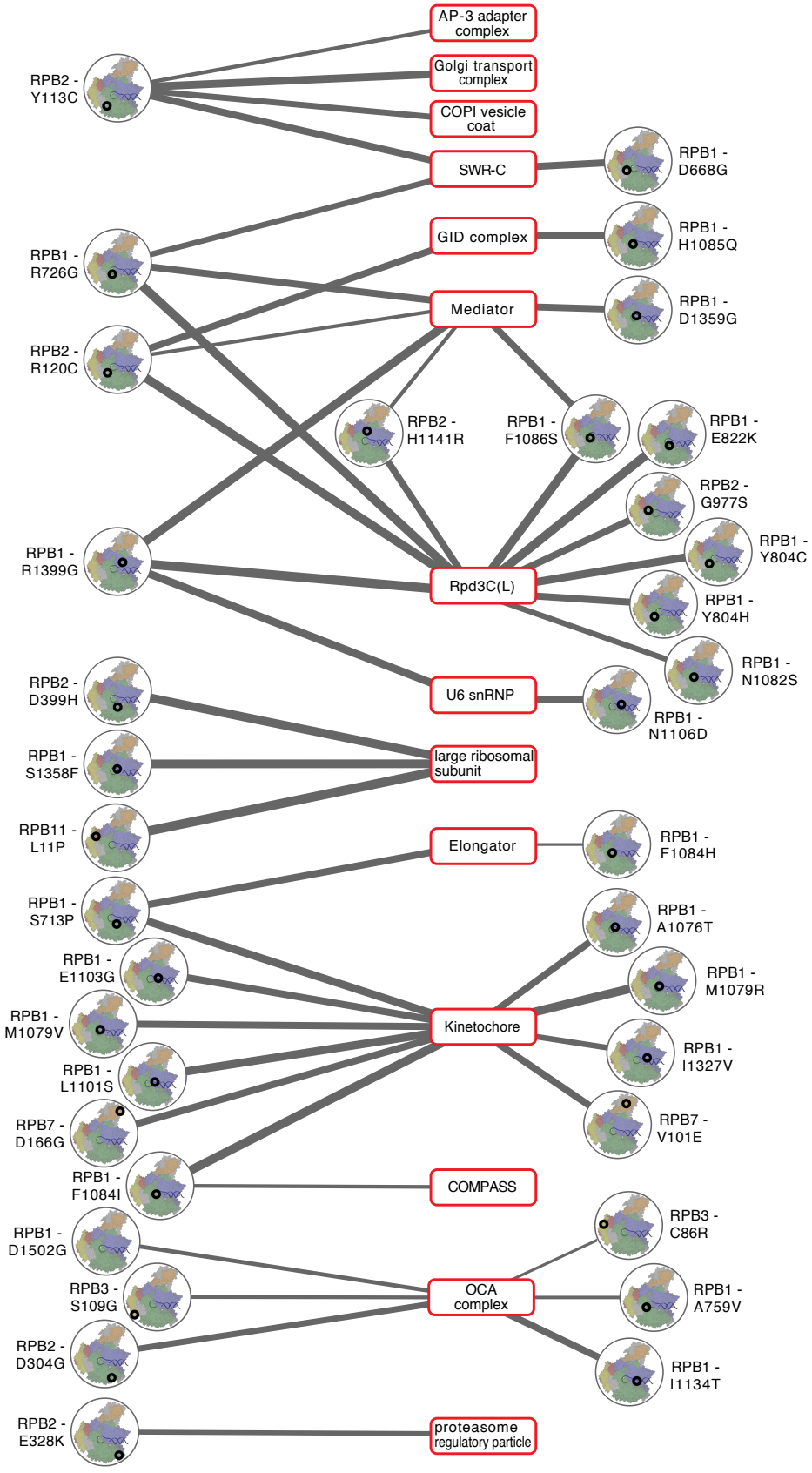

AP-3 adapter complex
Golgi transport complex
COPI vesicle coat
SWR-C
GID complex
Mediator
Rpd3C(L)
U6 snRNP
large ribosomal subunit
Elongator
Kinetochore
COMPASS
OCA complex
proteasome regulatory particle
RPB2 - Y113C
RPB1 - R726G
RPB2 - R120C
RPB1 - R1399G
RPB2 - D399H
RPB1 - S1358F
RPB11 - L11P
RPB1 - S713P
RPB1 - E1103G
RPB1 - M1079V
RPB1 - L1101S
RPB7 - D166G
RPB1 - F1084I
RPB1 - D1502G
RPB3 - S109G
RPB2 - D304G
RPB2 - E328K
RPB1 - D668G
RPB1 - H1085Q
RPB1 - D1359G
RPB2 - H1141R
RPB1 - F1086S
RPB1 - E822K
RPB2 - G977S
RPB1 - Y804C
RPB1 - Y804H
RPB1 - N1082S
RPB1 - N1106D
RPB1 - F1084H
RPB1 - A1076T
RPB1 - M1079R
RPB1 - I1327V
RPB7 - V101E
RPB3 - C86R
RPB1 - A759V
RPB1 - I1134T

Figure S3. Functional connections between RNAPII mutants and protein complexes. Functional associations between RNAPII mutants and protein complexes were determined as described in **Figure 3** and **Methods**. Edge widths correspond to the statistical significance of connections, and only edges with a false discovery rate (FDR) <0.1 are displayed (**Table S4**).

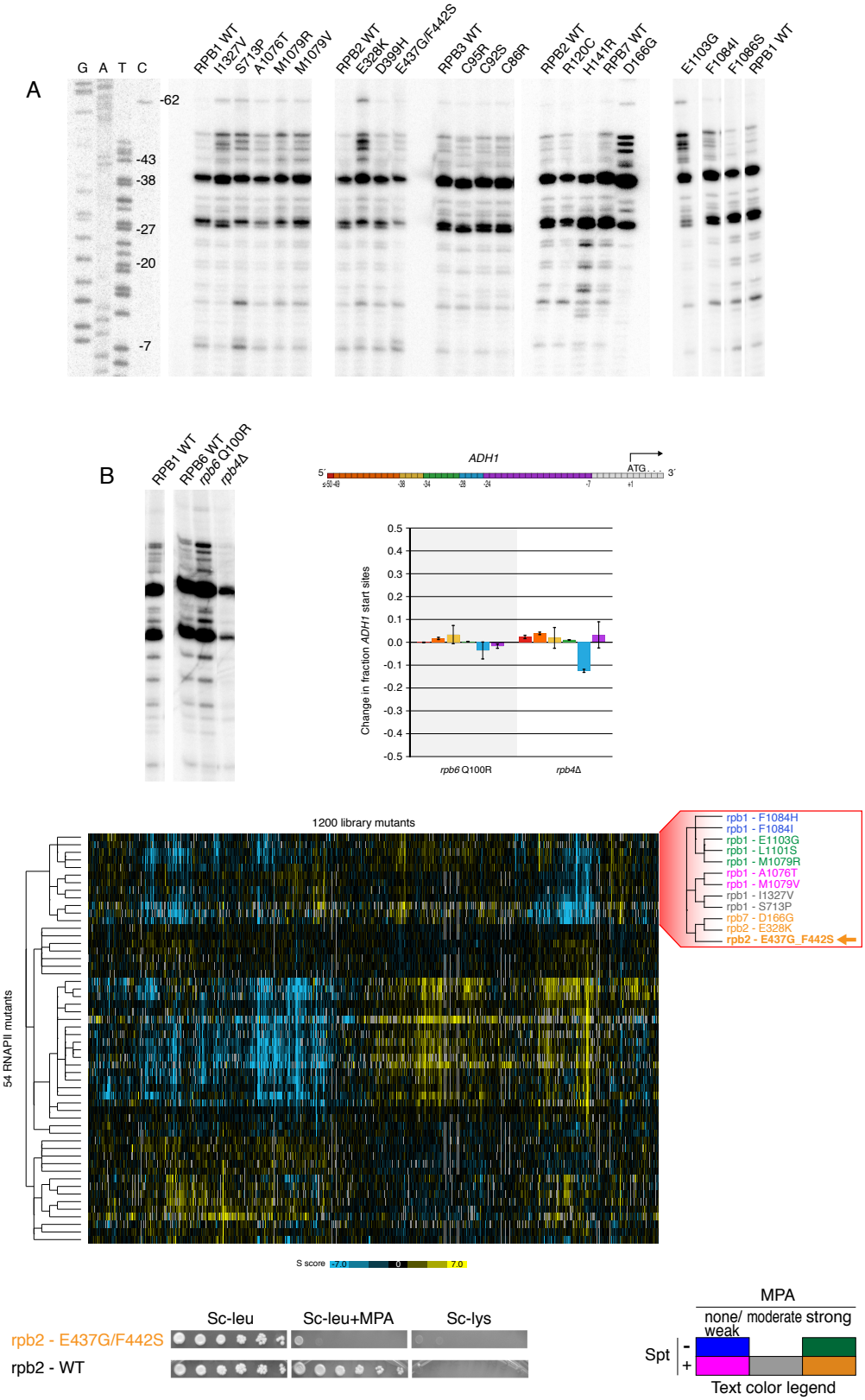

A
G A T C
RPB1 WT
I1327V
S713P
A1076T
M1079R
M1079V
RPB2 WT
E328K
D399H
E437G/F442S
RPB3 WT
C95R
C92S
C86R
RPB2 WT
R120C
H141R
RPB7 WT
D166G
E1103G
F1084I
F1086S
RPB1 WT
-62
-43
-38
-27
-20
-7
B
RPB1 WT
RPB6 WT
rpb6 Q100R
rpb4Δ
ADH1
5'
ATG
3'
Change in fraction ADH1 start sites
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
rpb6 Q100R
rpb4Δ
C
1200 library mutants
54 RNAPII mutants
rpb1 - F1084H
rpb1 - F1084I
rpb1 - E1103G
rpb1 - L1101S
rpb1 - M1079R
rpb1 - A1076T
rpb1 - M1079V
rpb1 - I1327V
rpb1 - S713P
rpb7 - D166G
rpb2 - E328K
rpb2 - E437G_F442S
S score
-7.0
0
7.0
D
Sc-leu
Sc-leu+MPA
Sc-lys
rpb2 - E437G/F442S
rpb2 - WT
MPA
none/ weak
moderate
strong
Spt
-
+
Text color legend

Figure S4. Primer extension analysis at *ADH1* to identify the effect of RNAPII mutations on start-site selection, and data relating to *rpb2* E437G/F442S. A) Sequencing gel for primer extension analysis. Lanes 1-4 correspond to Sanger dideoxy sequencing reactions as reference ladders, and the following lanes carry reverse transcription products from RNAPII alleles as indicated (**Table S6**). **B**) Primer extension analysis at *ADH1* to map transcription start sites for *rpb4Δ* and *rpb6* Q100R mutations. The colors of the bars correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to wild-type. Means and standard deviations were derived from three independent experiments. **C**) Clustering of *rpb2* E437G/F442S in pE-MAP. **D**) MPA sensitivity and Spt phenotypes of *rpb2* E437G/F442S.

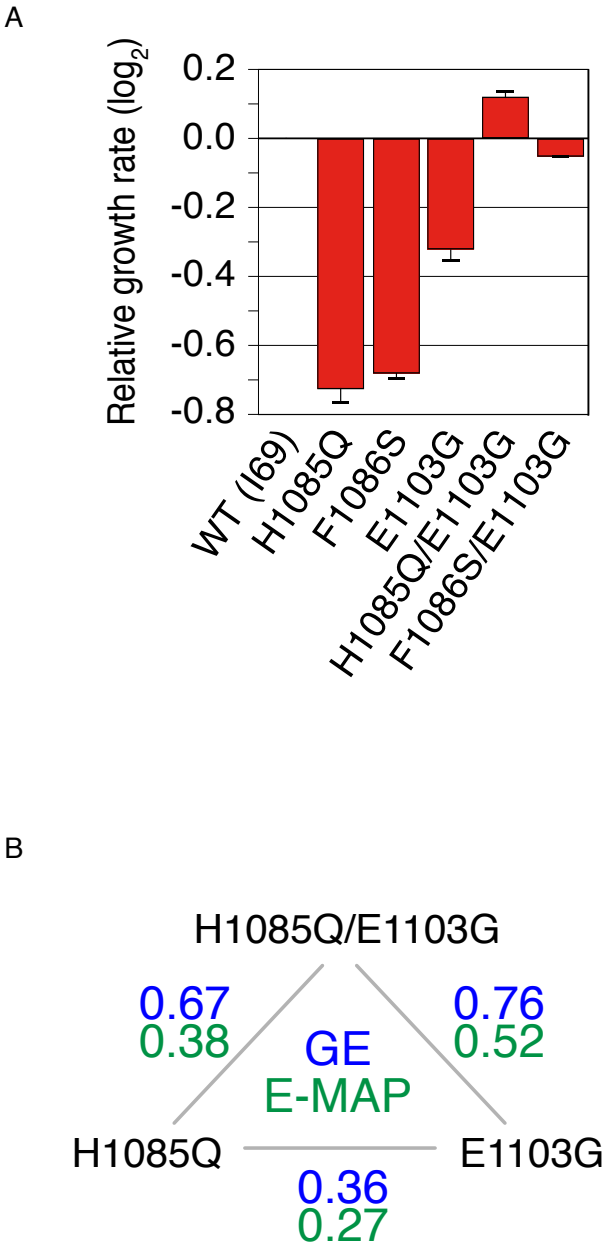


Figure S5. RNAPII trigger loop double and single mutant growth rates, and profile correlations between *rpb1* E1103G, *rpb1* H1085Q and *rpb1* E1103G/H1085Q. A) Growth rates relative to wild-type of *rpb1* E1103G, *rpb1* F1086S, *rpb1* H1085Q, *rpb1* E1103G/F1086S and *rpb1* E1103G/H1085Q. Means and standard deviations were derived from three replicates. **B**) pE-MAP and GE profile correlations between *rpb1* E1103G, *rpb1* H1085Q and *rpb1* E1103G/H1085Q.

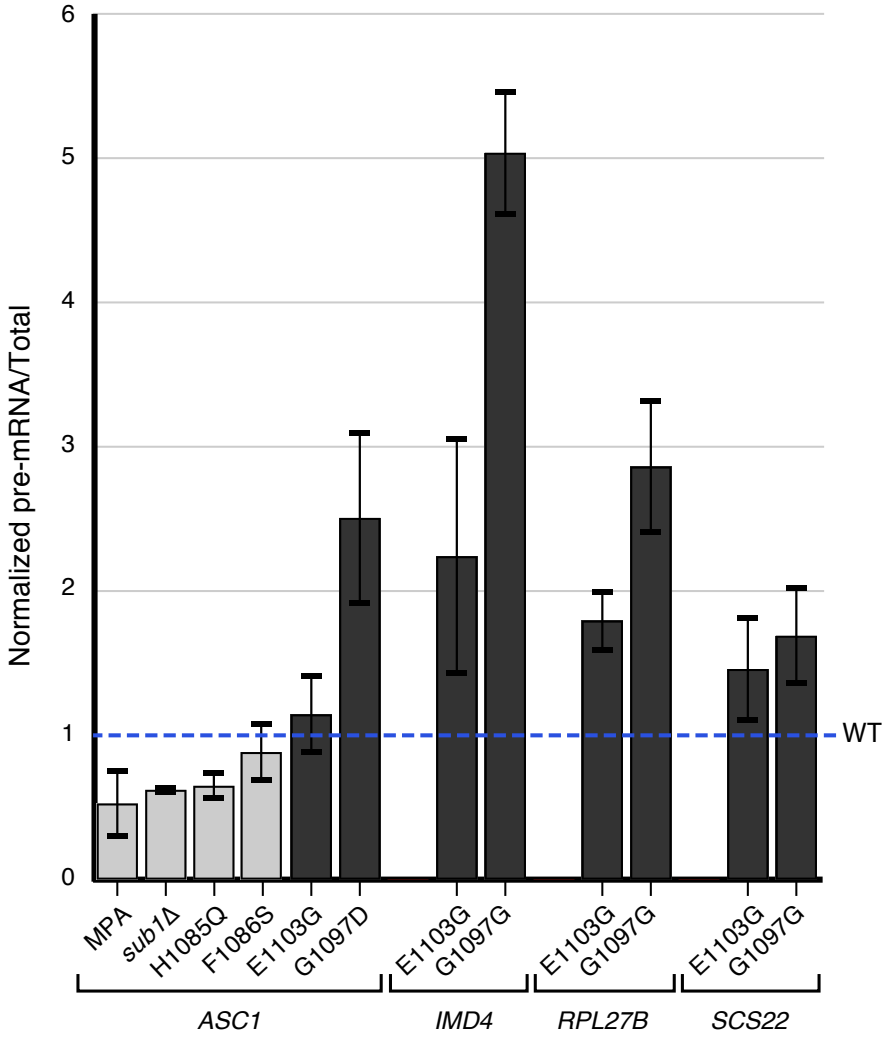


Figure S6. Splicing efficiency measured by qPCR to confirm trends discovered by splicing microarrays. Pre-mRNA and total mRNA for transcripts from *ASC1*, *IMD4*, *RPL27B* and *SCS22* were measured in the indicated strains and conditions. Shown are pre-mRNA/total mRNA ratios, normalized to that of an untreated wild-type. Light grey bars are data points that exhibited improved splicing in the microarray; dark grey bars are data points that exhibited a splicing defect in the microarray. Error bars represent standard deviation of 2-8 biological replicates. "MPA" refers to a 10-minute treatment of a wild-type strain with mycophenolic acid.

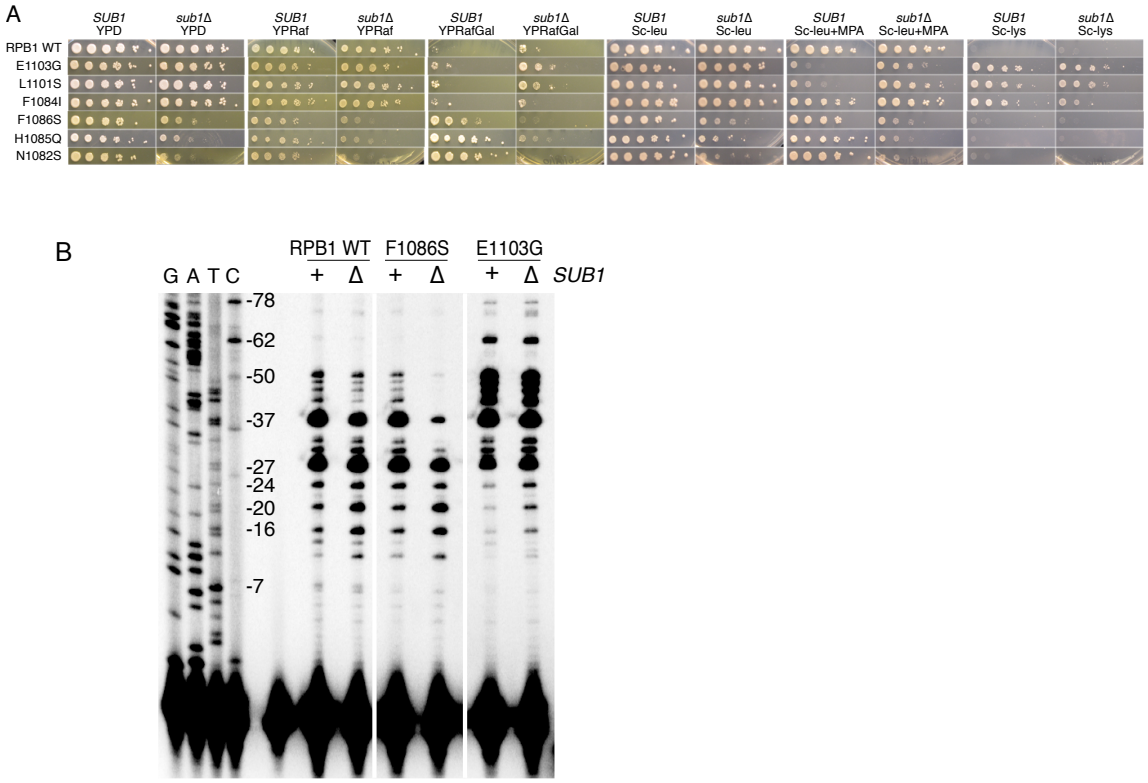


Figure S7. Spot tests to determine Gal^R, MPA sensitivity and Spt$^-$ phenotypes of *sub1Δ* with fast and slow RNAPII alleles, and primer extension to identify the effects of *sub1*Δ on start-site selection. A) Spot tests on relevant media were carried out to determine the effects of *sub1*Δ on the transcription-related phenotypes of fast and slow RNAPII alleles, as described in **Figure 1** and **Methods**. **B)** Primer extension analysis at *ADH1* to identify the effects of *sub1*Δ on start-site selection. Lanes 1-4 correspond to Sanger dideoxy sequencing reactions as reference ladders and the following lanes carry reverse transcription products from RNAPII mutants in *SUB1* and *sub1Δ* background as indicated (**Table S6**).

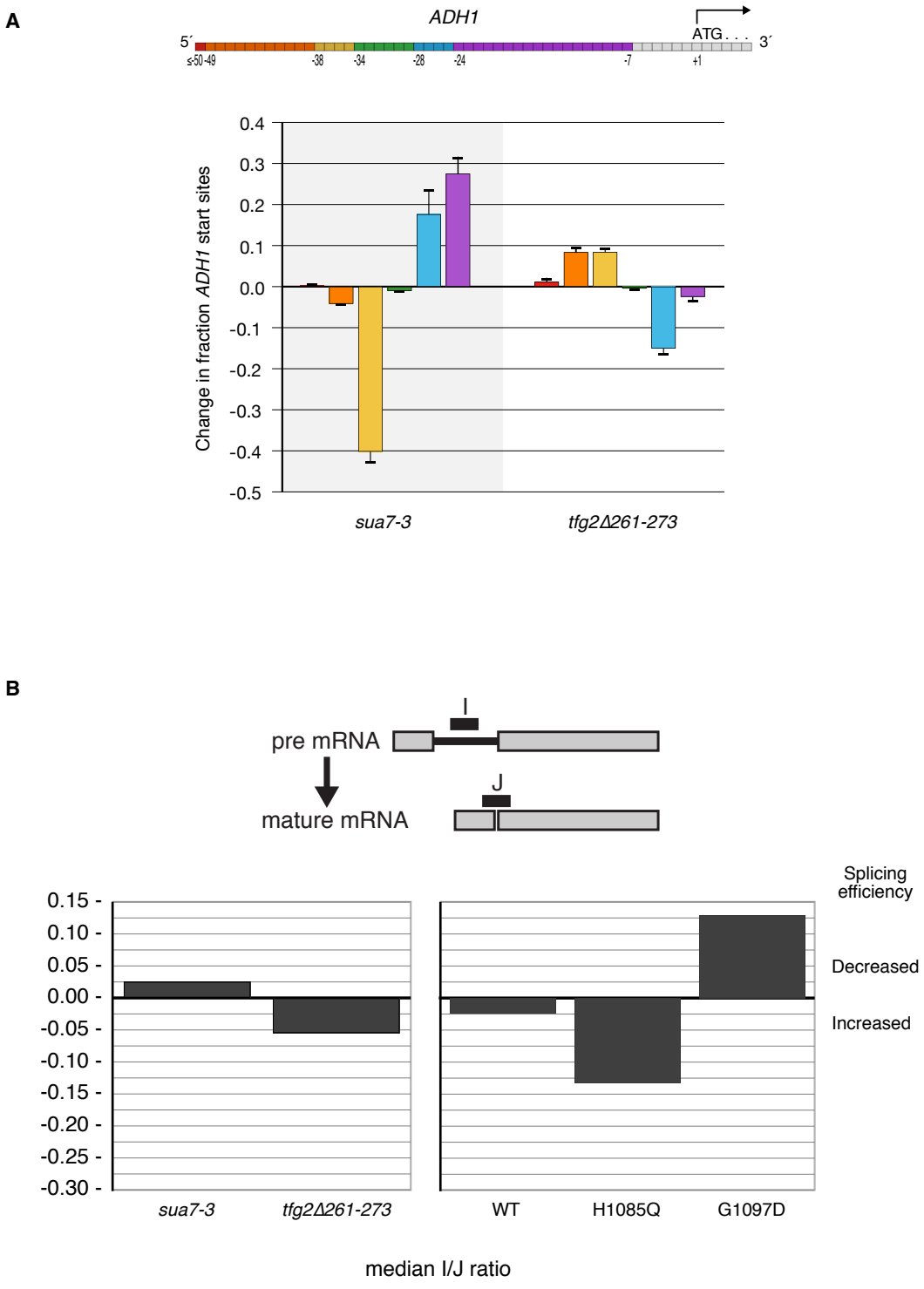


Figure S8. Start-site selection and splicing efficiency of TFIIB and TFIIF mutants. A) Primer extension analysis at *ADH1* to map transcription start-sites for mutations in TFIIB (*sua7-3*) and TFIIF (*tfg2Δ261-273*). The colors of the bars

correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to wt. Means and standard deviations were derived from three independent experiments. **B**) TFIIB (*sua7-3*) and TFIIF (*tfg2Δ261-273*) mutants were analyzed on splicing microarrays, as described in **Figure 6** (**Supp. Data 7**). Median I/J ratios for the complete profiles of the TFII mutants are shown in the left panel, with wild-type, a slow and a fast *rpb1* mutant included for reference in the right panel.

Supplementary Tables and Data

For supplementary tables and data, contact Hannes Braberg: hbraberg@gmail.com

References

Aguilar, P.S., Frohlich, F., Rehman, M., Shales, M., Ulitsky, I., Olivera-Couto, A., Braberg, H., Shamir, R., Walter, P., Mann, M.*, et al.* (2010). A plasma-membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking. Nat Struct Mol Biol *17*, 901-908.

Amberg, D.C., Burke, D.J., and Strathern, J.N. (2005). Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual, 2005 Edition, 2005 edn (Cold Spring Harbor, NY, Cold Spring Harbor Press).

Aitken, S., Alexander, R.D., and Beggs, J.D. (2011). Modelling reveals kinetic advantages of co-transcriptional splicing. PLoS Comput Biol *7*, e1002215.

Alexander, R.D., Innocente, S.A., Barrass, J.D., and Beggs, J.D. (2010). Splicing-dependent RNA polymerase pausing in yeast. Mol Cell *40*, 582-593.

Archambault, J., and Friesen, J.D. (1993). Genetics of eukaryotic RNA polymerases I, II, and III. Microbiol Rev *57*, 703-724.

Arribere, J.A., and Gilbert, W.V. (2013). Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. Genome Res *Advance online publication May 2, 2013*.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) *57*, 289-300.

Bandyopadhyay, S., Mehta, M., Kuo, D., Sung, M.K., Chuang, R., Jaehnig, E.J., Bodenmiller, B., Licon, K., Copeland, W., Shales, M.*, et al.* (2010). Rewiring of genetic networks in response to DNA damage. Science *330*, 1385-1389.

Bassik, M.C., Kampmann, M., Lebbink, R.J., Wang, S., Hein, M.Y., Poser, I., Weibezahn, J., Horlbeck, M.A., Chen, S., Mann, M.*, et al.* (2013). A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell *152*, 909-922.

Beltrao, P., Cagney, G., and Krogan, N.J. (2010). Quantitative genetic interactions reveal biological modularity. Cell *141*, 739-745.

Benschop, J.J., Brabers, N., van Leenen, D., Bakker, L.V., van Deutekom, H.W., van Berkum, N.L., Apweiler, E., Lijnzaad, P., Holstege, F.C., and Kemmeren, P. (2010). A consensus of core protein complex compositions for Saccharomyces cerevisiae. Mol Cell *38*, 916-928.

Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol *17*, 251-256.

Bergkessel, M., Whitworth, G.B., and Guthrie, C. (2011). Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA *17*, 1461-1478.

Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol *154*, 164-175.

Butland, G., Babu, M., Diaz-Mejia, J.J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A.G., Pogoutse, O.*, et al.* (2008). eSGA: E. coli synthetic genetic array analysis. Nat Methods *5*, 789-795.

Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K.M. (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. Mol Cell *40*, 571-581.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Chen, H.T., Warfield, L., and Hahn, S. (2007). The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. Nat Struct Mol Biol *14*, 696-703.

Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M.*, et al.* (2010). Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J *29*, 717-726.

Clark, T.A., Sugnet, C.W., and Ares, M., Jr. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science *296*, 907-910.

Collins, S.R., Kemmeren, P., Zhao, X.C., Greenblatt, J.F., Spencer, F., Holstege, F.C., Weissman, J.S., and Krogan, N.J. (2007a). Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics *6*, 439-450.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Collins, S.R., Roguev, A., and Krogan, N.J. (2010). Quantitative genetic interaction mapping using the E-MAP approach. Methods Enzymol *470*, 205-231.

Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol *7*, R63.

Conesa, C., and Acker, J. (2010). Sub1/PC4 a chromatin associated protein with multiple functions in transcription. RNA Biol *7*, 287-290.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S.*, et al.* (2010). The genetic landscape of a cell. Science *327*, 425-431.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science *292*, 1863-1876.

Cramer, P., Pesce, C.G., Baralle, F.E., and Kornblihtt, A.R. (1997). Functional association between promoter structure and transcript alternative splicing. Proc Natl Acad Sci U S A *94*, 11456-11460.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell *12*, 525-532.

de la Mata, M., Munoz, M.J., Allo, M., Fededa, J.P., Schor, I.E., and Kornblihtt, A.R. (2011). RNA Polymerase II Elongation at the Crossroads of Transcription and Alternative Splicing. Genet Res Int *2011*, 309865.

DeRisi, J.L., Iyer, V.R., and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science *278*, 680-686.

Edwards, A.M., Kane, C.M., Young, R.A., and Kornberg, R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. J Biol Chem *266*, 71-75.

Eichner, J., Chen, H.T., Warfield, L., and Hahn, S. (2010). Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. EMBO J *29*, 706-716.

Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S.*, et al.* (2009). Functional organization of the S. cerevisiae phosphorylation network. Cell *136*, 952-963.

Gaj, T., Gersbach, C.A., and Barbas III, C.F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol *Advance online publication May 9, 2013*.

Garcia, A., Collin, A., and Calvo, O. (2012). Sub1 associates with Spt5 and influences RNA polymerase II transcription elongation rate. Mol Biol Cell *23*, 4297-4312.

Ghazy, M.A., Brodie, S.A., Ammerman, M.L., Ziegler, L.M., and Ponticelli, A.S. (2004). Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. Mol Cell Biol *24*, 10975-10985.

Giardina, C., and Lis, J.T. (1993). DNA melting on yeast RNA polymerase II promoters. Science *261*, 759-762.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876-1882.

Goel, S., Krishnamurthy, S., and Hampsey, M. (2012). Mechanism of start site selection by RNA polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIIH. J Biol Chem *287*, 557-567.

Greger, I.H., and Proudfoot, N.J. (1998). Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. EMBO J *17*, 4771-4779.

Grunberg, S., Warfield, L., and Hahn, S. (2012). Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. Nat Struct Mol Biol *19*, 788-796.

Guthrie, C., and Fink, G.R. (2002). Guide to Yeast Genetics and Molecular and Cell Biology (Academic Press).

Hampsey, M., Na, J.G., Pinto, I., Ware, D.E., and Berroteran, R.W. (1991). Extragenic suppressors of a translation initiation defect in the cyc1 gene of Saccharomyces cerevisiae. Biochimie *73*, 1445-1455.

Harel-Sharvit, L., Eldad, N., Haimovich, G., Barkai, O., Duek, L., and Choder, M. (2010). RNA polymerase II subunits link transcription and mRNA decay to translation. Cell *143*, 552-563.

Hazelbaker, D.Z., Marquardt, S., Wlotzka, W., and Buratowski, S. (2013). Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. Mol Cell *49*, 55-66.

He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. Nature *495*, 481-486.

Hoppins, S., Collins, S.R., Cassidy-Stone, A., Hummel, E., Devay, R.M., Lackner, L.L., Westermann, B., Schuldiner, M., Weissman, J.S., and Nunnari, J. (2011). A mitochondrial-focused genetic interaction map reveals a scaffold-like complex required for inner membrane organization in mitochondria. J Cell Biol *195*, 323-340.

Horn, T., Sandmann, T., Fischer, B., Axelsson, E., Huber, W., and Boutros, M. (2011). Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. Nat Methods *8*, 341-346.

Howe, K.J., Kane, C.M., and Ares, M., Jr. (2003). Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA *9*, 993-1006.

Ideker, T., and Krogan, N.J. (2012). Differential network biology. Mol Syst Biol *8*, 565.

Ip, J.Y., Schmidt, D., Pan, Q., Ramani, A.K., Fraser, A.G., Odom, D.T., and Blencowe, B.J. (2011). Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. Genome Res *21*, 390-401.

Kahm, M., Hasenbrink, G., Lichtenberg-Frate, H., Ludwig, J., and Kschischo, M. (2010). grofit: Fitting Biological Growth Curves with R. J Stat Softw *33*, 1-21.

Kaplan, C.D. (2013). Basic mechanisms of RNA polymerase II activity and alteration of gene expression in Saccharomyces cerevisiae. Biochim Biophys Acta *1829*, 39-54.

Kaplan, C.D., Holland, M.J., and Winston, F. (2005). Interaction between transcription elongation factors and mRNA 3'-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus. J Biol Chem *280*, 913-922.

Kaplan, C.D., Jin, H., Zhang, I.L., and Belyanin, A. (2012). Dissection of Pol II Trigger Loop Function and Pol II Activity-Dependent Control of Start Site Selection In Vivo. PLoS Genet *8*, e1002627.

Kaplan, C.D., Larsson, K.M., and Kornberg, R.D. (2008). The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. Mol Cell *30*, 547-556.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., *et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol Cell *16*, 955-965.

Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purtov, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N., and Kashlev, M. (2008). Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol Cell *30*, 557-566.

Knaus, R., Pollock, R., and Guarente, L. (1996). Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-activator PC4. EMBO J *15*, 1933-1940.

Kuehner, J.N., and Brow, D.A. (2006). Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. J Biol Chem *281*, 14119-14128.

Larson, M.H., Zhou, J., Kaplan, C.D., Palangat, M., Kornberg, R.D., Landick, R., and Block, S.M. (2012). Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. Proc Natl Acad Sci U S A *109*, 6555-6560.

Laufer, C., Fischer, B., Billmann, M., Huber, W., and Boutros, M. (2013). Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. Nat Methods *10*, 427-431.

Lehner, B., Crombie, C., Tischler, J., Fortunato, A., and Fraser, A.G. (2006). Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. Nat Genet *38*, 896-903.

Lenstra, T.L., Benschop, J.J., Kim, T., Schulze, J.M., Brabers, N.A., Margaritis, T., van de Pasch, L.A., van Heesch, S.A., Brok, M.O., Groot Koerkamp, M.J.*, et al.* (2011). The specificity and topology of chromatin interaction pathways in yeast. Mol Cell *42*, 536-549.

Lin, Y.Y., Kiihl, S., Suhail, Y., Liu, S.Y., Chou, Y.H., Kuang, Z., Lu, J.Y., Khor, C.N., Lin, C.L., Bader, J.S.*, et al.* (2012). Functional dissection of lysine deacetylases reveals that HDAC1 and p300 regulate AMPK. Nature *482*, 251-255.

Malagon, F., Kireeva, M.L., Shafer, B.K., Lubkowska, L., Kashlev, M., and Strathern, J.N. (2006). Mutations in the Saccharomyces cerevisiae RPB1 gene conferring hypersensitivity to 6-azauracil. Genetics *172*, 2201-2209.

Martin, C., and Young, R.A. (1989). KEX2 mutations suppress RNA polymerase II mutants and alter the temperature range of yeast cell growth. Mol Cell Biol *9*, 2341-2349.

Mason, P.B., and Struhl, K. (2005). Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. Mol Cell *17*, 831-840.

Ohkuni, K., and Kitagawa, K. (2011). Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr Biol *21*, 1695-1703.

Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F., and Boeke, J.D. (2004). A robust toolkit for functional profiling of the yeast genome. Mol Cell *16*, 487-496.

Perales, R., and Bentley, D. (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. Mol Cell *36*, 178-191.

Pikielny, C.W., and Rosbash, M. (1985). mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. Cell *41*, 119-126.

Pinto, I., Ware, D.E., and Hampsey, M. (1992). The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. Cell *68*, 977-988.

Pinto, I., Wu, W.H., Na, J.G., and Hampsey, M. (1994). Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. J Biol Chem *269*, 30569-30573.

Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS Biol *5*, e90.

Ranish, J.A., and Hahn, S. (1991). The yeast general transcription factor TFIIA is composed of two polypeptide subunits. J Biol Chem *266*, 19320-19327.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J.*, et al.* (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science *322*, 405-410.

Roguev, A., Talbot, D., Negri, G.L., Shales, M., Cagney, G., Bandyopadhyay, S., Panning, B., and Krogan, N.J. (2013). Quantitative genetic-interaction mapping in mammalian cells. Nat Methods *10*, 432-437.

Roguev, A., Wiren, M., Weissman, J.S., and Krogan, N.J. (2007). High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nat Methods *4*, 861-866.

Rojas-Duran, M.F., and Gilbert, W.V. (2012). Alternative transcription start site selection leads to large differences in translation activity in yeast. RNA *18*, 2299-2305.

Rosonina, E., Willis, I.M., and Manley, J.L. (2009). Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. Mol Cell Biol *29*, 2308-2321.

Ryan, C.J., Roguev, A., Patrick, K., Xu, J., Jahari, H., Tong, Z., Beltrao, P., Shales, M., Qu, H., Collins, S.R.*, et al.* (2012). Hierarchical Modularity and the Evolution of Genetic Interactomes across Species. Mol Cell *46*, 691-704.

Schmitt, M.E., Brown, T.A., and Trumpower, B.L. (1990). A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae. Nucleic Acids Res *18*, 3091-3092.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F.*, et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

Schuldiner, M., Collins, S.R., Weissman, J.S., and Krogan, N.J. (2006). Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods *40*, 344-352.

Shaw, R.J., and Reines, D. (2000). Saccharomyces cerevisiae transcription elongation mutants are defective in PUR5 induction in response to nucleotide depletion. Mol Cell Biol *20*, 7427-7437.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Sikorski, T.W., Ficarro, S.B., Holik, J., Kim, T., Rando, O.J., Marto, J.A., and Buratowski, S. (2011). Sub1 and RPA associate with RNA polymerase II at different stages of transcription. Mol Cell *44*, 397-409.

Simchen, G., Winston, F., Styles, C.A., and Fink, G.R. (1984). Ty-mediated gene expression of the LYS2 and HIS4 genes of Saccharomyces cerevisiae is controlled by the same SPT genes. Proc Natl Acad Sci U S A *81*, 2431-2434.

Spencer, F., Gerring, S.L., Connelly, C., and Hieter, P. (1990). Mitotic chromosome transmission fidelity mutants in Saccharomyces cerevisiae. Genetics *124*, 237-249.

Tan, Q., Prysak, M.H., and Woychik, N.A. (2003). Loss of the Rpb4/Rpb7 subcomplex in a mutant form of the Rpb6 subunit shared by RNA polymerases I, II, and III. Mol Cell Biol *23*, 3329-3338.

Tavenet, A., Suleau, A., Dubreuil, G., Ferrari, R., Ducrot, C., Michaut, M., Aude, J.C., Dieci, G., Lefebvre, O., Conesa, C.*, et al.* (2009). Genome-wide location analysis reveals a role for Sub1 in RNA polymerase III transcription. Proc Natl Acad Sci U S A *106*, 14265-14270.

Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M.*, et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Trcek, T., Larson, D.R., Moldon, A., Query, C.C., and Singer, R.H. (2011). Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. Cell *147*, 1484-1497.

Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L.*, et al.* (2008). High-throughput, quantitative analyses of genetic interactions in E. coli. Nat Methods *5*, 781-787.

van de Peppel, J., Kemmeren, P., van Bakel, H., Radonjic, M., van Leenen, D., and Holstege, F.C. (2003). Monitoring global messenger RNA changes in externally controlled microarray experiments. EMBO Rep *4*, 387-393.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell *127*, 941-954.

Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S.*, et al.* (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell *32*, 735-746.

Winston, F., and Sudarsanam, P. (1998). The SAGA of Spt proteins and transcriptional analysis in yeast: past, present, and future. Cold Spring Harb Symp Quant Biol *63*, 553-561.

Wu, W.H., Pinto, I., Chen, B.S., and Hampsey, M. (1999). Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. Genetics *153*, 643-652.

Yuen, K.W., Warren, C.D., Chen, O., Kwok, T., Hieter, P., and Spencer, F.A. (2007). Systematic genome instability screens in yeast and their potential relevance to cancer. Proc Natl Acad Sci U S A *104*, 3925-3930.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

06/24/2013

Date