Lawrence Berkeley National Laboratory

Title
Functional annotation of mouse mutations in embryonic stem cells using expression profiling

Permalink
https://escholarship.org/uc/item/91w732sp

Journal
Mammalian Genome, 15(1)

Authors
Symula, Derek J.
Zhu, Yiwen
Schimenti, John C.
et al.

Publication Date
2003-07-15

Functional annotation of mouse mutations in embryonic stem cells using expression profiling

Derek J. Symula[1]*, Yiwen Zhu[1], John C. Schimenti[2], Edward M. Rubin[1]

1 – Genome Sciences Department, Lawrence Berkeley National Lab, Berkeley, California 94720
2 – The Jackson Laboratory, Bar Harbor, Maine 04609

* Current address: Genomics Institute, Wadsworth Center, Albany, New York 12208

correspondence should be addressed to E.M.R.

Abstract

Expression profiling offers a potential high-throughput phenotype screen for mutant mouse embryonic stem (ES) cells. We have assessed the ability of expression arrays to distinguish among heterozygous mutant ES cell lines and to accurately reflect the normal function of the mutated genes. Two ES cell lines hemizygous for overlapping regions of mouse chromosome 5 differed substantially from the wildtype parental and from each other. Expression differences included frequent downregulation of hemizygous genes and downstream effects on genes mapping to other chromosomes. Some genes were affected similarly in each deletion line, consistent with the overlap of the deletions. To determine whether such downstream effects reveal pathways impacted by a mutation, we examined ES cell lines heterozygous for mutations in either of two well-characterized genes. A heterozygous mutation in the cell cycle regulator, cyclin D kinase 4 (*Cdk4*), affected expression of many genes involved in cell growth and proliferation. A heterozygous mutation in the ATP binding cassette transporter family A, member 1 (*Abca1*) gene altered genes associated with lipid homeostasis, the cytoskeleton, and vesicle trafficking. Heterozygous *Abca1* mutation had similar effects in liver, indicating that ES cell expression profile reflects changes in fundamental processes relevant to mutant gene function in multiple cell types.

2

Introduction

One of the more powerful strategies for functionally annotating the human genome uses mutant mice to assign function to the mouse orthologs of human genes. However, existing mutant mouse strains represent a small fraction of the predicted number of gene products in the mouse (Nadeau 2000). Several approaches currently are being applied to address this deficit, including *in vivo* mutagenesis using ethylnitrosourea (ENU) (Hrabe de Angelis et al. 2000; Nolan et al. 2000). ENU efficiently generates genome-wide mutations, generally substitutions, that may result in loss of function, gain of function, hypomorphic, hypermorphic, or antimorphic alleles (Bode 1984; Hitotsumachi et al. 1985; Rinchik and Carpenter 1999; Russell et al. 1979). Progeny of the mutagenized mice can be assessed for one or more phenotypes to identify potential mutant lines of interest or bred to produce homozygotes for subsequent recessive mutation screening. Such an approach is well-suited for producing mutants for genome annotation, as it requires no *a priori* knowledge of gene function and produces large numbers of mutant phenotypes. However, the extensive resources needed to breed and maintain the mutant pedigrees and the practical limits of *in vivo* screening constrain its throughput.

These limitations have resulted in a drive to "yeastify" mouse genetics (Chen et al. 2000a) by developing efficient *in vitro* mutagenesis methods for use in mouse ES cells. Insertional mutagenesis using retroviral (Stanford et al. 2001) or plasmid (Mitchell et al. 2001; Wiles et al. 2000; Zambrowicz et al. 1998) vectors has been used at a large scale to produce mutant ES lines. The mutant gene is identified on the based of

flanking sequence and a beta-gal reporter may be included in the vector to allow the *in vivo* expression pattern to be facilely monitored. More recently, the same chemical mutagens successfully used *in vivo* have been used to generate mutations in ES cells, resulting in similar efficiency and allele spectrum. Chemical mutagenesis in ES cells requires simply adding dilute ENU (Chen et al. 2000b) or ethylmethylsulfonate (EMS, (Munroe et al. 2000)) to the culture dish. Mutants can be screened by genotyping genes of interest. These strategies share the advantages that mutations are efficiently generated, stored, and screened prior to making mice, but they are all gene-driven. Novel genes with no functional annotation to guide analysis generally require wide-ranging *in vivo* experiments to discover their functions, limiting the advantages of *in vitro* mutagenesis. In addition, novel functions of even well-characterized genes may be overlooked. While mutant mice may be generated from the mutant ES cell lines and subsequently and screened (Munroe et al. 2000), this sacrifices some of the efficiency of *in vitro* mutagenesis. An *in vitro* phenotype screen would maximize the efficiency of ES cell mutagenesis and allow largely unbiased discovery of gene function.

Expression array profiling represents a sensitive, specific, and high-throughput phenotypic assay to screen mutant ES cell clones, select mutants for *in vivo* analysis, and guide *in vivo* experiments. Expression profiling accurately distinguishes among closely related cell types (Tanaka et al. 2002), closely related leukemia subtypes (Golub et al. 1999) or breast cancer tumors with different prognoses (van de Vijver et al. 2002), and single gene mutations (Chauhan et al. 2002). Expression changes are generally relevant to the function of the mutated genes in a variety of cell types. Recent work by

Tanaka et al. ((Tanaka et al. 2002) and Ramalho-Santos et al. (Ramalho-Santos et al. 2002) revealed a set of genes expressed at unique levels in pluripotent stem cells. What is striking, though, is that these genes comprise a small fraction of the ES cell transcriptome. The majority of genes expressed by ES cells also are expressed in a wide variety of differentiated cell types, suggesting that mutations detected in ES cells frequently will lead to mutant phenotypes in adult mice. We report here that heterozygous mutants frequently alter ES cell expression profile in specific ways and that these alterations accurately reflect both the function of the mutated gene and the *in vivo* effects of the mutation.

Methods

Cell lines and mice:

The v6.4, B12, C6, 3F12 (*Abca1 +/-*), and ESVJ ES cell lines were cultured on mitotically inactivated murine embryonic fibroblasts (MEF) using standard techniques. Three independent dishes of cells were grown up to ~70-90% confluence, trypsinized, plated on gelatin (0.1%) coated plates for one hour to deplete MEFs, then cultured on gelatin coated plates 1-2 days prior to total RNA isolation. Three replicates of each feeder-independent ES cell line (obtained from Baygenomics, (Stryke et al. 2003)), E14Tg2a and XA053 (*CDK4 +/-*), were grown as described (Skarnes 2000) prior to total RNA isolation. All lines were grown in the presence of LIF and met morphological criteria for undifferentiated ES cells at the time of RNA isolation. *Abca1 +/-* mutant mice were produced under approved protocols at the Lawrence Berkeley National Laboratory and were of a mixed 129/SvJ and C57BL/6J background.

RNA isolation:

ES cell RNA was isolated from replicate 6 cm plates using Trizol (Invitrogen, Carlsbad, CA) or RNeasy (Qiagen, Valencia, CA). Liver RNA was isolated by homogenization and subsequent purification with Qiagen RNeasy.

Array protocols:

For spotted cDNA arrays, individual RNAs were reverse transcribed in the presence of dUTP-Cy5 (AmershamPharmacia), and hybridized as described (Friddle et al. 2000) against a Cy3 labeled control consisting of a pooled control of v6.4, B12, and C6 cDNA. Arrays were scanned with an Axon Genepix (Union City, CA) scanner. For experiments using AffyMetrix GeneChips (AffyMetrix, Santa Clara, CA), RNA was prepared and labeled according to the manufacturers instructions, hybridized to the Mu74Av2 chip, and scanned.

Array data analysis:

Spotted cDNA arrays were gridded using Scanalyze (Eisen et al. 1998) and expression data imported into Genespring 5.0.2 for analysis. 2661 clones with low background and control signal >150 in at least 6 arrays and test signal >150 in at least 5 arrays were chosen for analysis. The signal for each spot was normalized using intensity-dependent Lowess normalization and then log transformed data were clustered using a standard correlation. For the AffyMetrix GeneChip data, default values for the AffyMetrix MAS5 software were used to normalize data on each chip and

to identify probe sets that were present, marginal, or absent. Genespring 5.0.2 (Silicon Genetics, Foster City, CA) was used to identify genes present or marginal in at least three of the six chips in each mutant/wildtype pair. Log transformed data for the genes meeting these criteria were examined by t-test to identify genes differing between genotypes. Genes with a nominal p-value of 0.05 or less were considered to be differentially expressed. The chi-square test was used to test for differences in the numbers of differentially expressed genes meeting various criteria.

Array annotation:

To determine the biological effects of each mutation, we asked which of these categories were more frequent among a differentially expressed gene (determined by t-test) set than expected from their frequency in the appropriate transcriptome (Cho et al. 2001). Categories had to be at least two-fold more frequent and be represented by at least two genes (defined by a unique LocusLink or Unigene number) among the differentially expressed genes. Related categories were then grouped to identify broad functional themes affected in each mutant line.

Results

To determine whether heterozygous null mutations specifically affect expression profile, we compared the Hdh16C/1-C6 (C6) and Hdh16C/1-B12 (B12) deletion lines, comprising a maximum of approximately 13 and 34 Mb of mouse chromosome 5, respectively, and overlapping by approximately 5 - 7 Mb ((Celera 2001; Schimenti et al. 2000)). 2661 clones highly expressed in the experiment were chosen for further

analysis (see supplementary data Table 1). These genes are expressed in a variety of embryonic and adult cell types and have a broad range of known or inferred functions, consistent with previous gene-trap (Mitchell et al. 2001) and expression studies (Ramalho-Santos et al. 2002; Tanaka et al. 2002). Hierarchical clustering of these data revealed genes with expression levels that were highly reproducible within a cell line, but differed substantially among the different lines (Figure 1). Each mutant line clearly differed from the parental and from the other mutant line (for example, clusters denoted by open arrows in Figure 1). Several clusters of genes were similarly affected in both deletion lines relative to the wildtype (Figure 1, closed arrow), consistent with the overlap of the deletions. Quantitative RT-PCR analysis of a subset of the differentially expressed genes indicates that the differences in expression represent true differences in steady-state RNA levels (data not shown). Decreased dosage was frequently sufficient to affect expression of the 17 genes in our data set that were hemizygous in one or both deletion lines. In 12 of 34 comparisons between lines with different dosage for a given gene, the gene was significantly downregulated ($p < 0.05$, t-test) in the hemizygous line (Table 1). This frequency of down regulated genes is approximately 3-fold greater than that found among genes mapping to other regions of chromosome 5 and is statistically significant ($p < 0.05$, chi-squared test). Many other differentially expressed genes, particularly genes upregulated in the deletion lines, mapped outside the deletion regions, suggesting downstream effects on wildtype genes. These results suggest that haploinsufficiency may be sufficient to alter ES cell expression profile specifically.

The large number of hemizygous genes in the deletion lines made it difficult to determine whether hemizygosity for individual genes resulted in specific expression profile changes. We therefore examined two new ES cell lines, each heterozygous for a mutation in a single, well characterized gene. We obtained a genetrap clone from the BayGenomics consortium (Stryke et al. 2003) heterozygous for an insertion mutation in the *Cdk4* gene. These mutations are generally nulls (Mitchell et al. 2001). The *Abca1* *+/-* cell line reproduces a previously-published null mutation (McNeish et al. 2000). Three independent replicates of each mutant and three replicates of its parental line were profiled on AffyMetrix GeneChips and differentially expressed genes identified by t-test (nominal p-value $\leq$ 0.05). The expression profiles changes in each mutant ES cell line were highly specific to that line. Among approximately 6500 probe sets expressed in the *Cdk4 +/-* ES cell line and its parental, 879 probe sets, representing 850 genes were differentially expressed (supplementary data Table 2 for all differentially expressed genes). Expression of 162 probe sets (supplementary data Table 3), representing 160 genes, differed between the *Abca1 +/-* ES cell line and its parental out of approximately 6200 expressed probe sets. While 25 probe sets were differentially expressed in each mutant versus its parental, suggesting that the two mutations might have some similar effects, the effects were quantitatively distinct for all but two genes. A single gene was upregulated in each mutant relative to its parental and one downregulated in each. Quantitative RT-PCR confirmed the difference in expression for several genes in each mutant/parental pair (not shown). This analysis suggests that heterozygous mutations in *Abca1* or *Cdk4* have specific effects on expression of specific genes.

To determine whether the differences in expression profile observed in each mutant line were likely caused by each mutation, we first examined our data for expression of "signature" genes (Tanaka et al. 2002) recently shown to distinguish ES cells from embryonic fibroblasts and trophoblast stem cells (Figure 2). ES cells may differentiate, altering gene expression (Rathjen and Rathjen 2001). While our cell lines all had morphology typical of undifferentiated ES cells, it is possible that one or more lines had differentiated sufficiently to affect expression profile without affecting morphology. Of 26 signature genes expressed in the *Cdk4 +/-* cell line and its parental, two were differentially expressed ($p<0.05$, t-test). Of 25 such genes expressed in the *Abca1 +/-* and its parental, a single gene differed ($p<0.05$, t-test). This analysis indicates that signature genes are no more likely to be differentially expressed than other genes ($p>0.05$ for each, chi-square test), suggesting that differences in pluripotency contribute minimally to differences in expression profile in these experiments.

To further explore the link between mutant gene function and expression profile, we next asked whether the expression differences in a mutant ES cell line were consistent with the functions of the mutant gene. To systematically determine the biological roles of each mutant gene, we identified the appropriate Gene Ontology categories for biological process, molecular function, and cellular component (Ashburner et al. 2000; CONSORTIUM 2003) represented by each differentially expressed gene. Table 2 shows groups of Gene Ontology categories overrepresented among the genes differentially expressed in the *Cdk4 +/-* data and represented by at

10

least four genes. Many of these genes are associated with cell growth or proliferation,

consistent with the function of Cdk4 in cell cycle progression (Tsutsui et al. 1999).

These genes include many that function directly in cell cycle regulation, including *Cdk4*

itself (downregulated in two independent probe sets), *Cdk7*, and the *Cdk4* inhibitor

Cdkn2a, or in checkpoints such as DNA damage (*Gas1*, *Gas5*) or metabolic status

(*Prkag1*, *Prkab1*). Genes involved in chromatin structure, an important regulator of

gene expression during cell cycle progression (Zhang et al. 2000), mitochondrial energy

production, and insulin like growth factor binding are impacted as well. We also noted a

group of genes related to differentiation, but many of these genes are also involved in

cell proliferation (*Ctgf* and *Opn*), expansion of B-cell (*Vpreb1*) or T-cell (*Wnt4*)

precursors, or apoptosis during development (*Spp1*, *Col18A1*, *Gna13*, *Pbx1*). These

genes likely occur in our data set due to cell growth-related categories that do not meet

our criteria for overrepresentation. These expression changes suggest altered

regulation of the cell growth and proliferation in the *Cdk4 +/-* cell line due to decreased

Cdk4 expression, accurately reflecting the biological effects of the heterozygous *Cdk4*

mutation in ES cells.

Several biological themes predominate in the *Abca1 +/-* ES cell data set. Table 3

shows genes grouped by overrepresented Gene Ontology categories, in this case

groups represented by at least three unique genes (24 genes total). Consistent with the

role of ABCA1 in cholesterol transport, the cholesterol biosynthetic enzymes lanosterol

14alpha-demethylase (*Cyp51*) and mevalonate kinase (*Mvk*) are significantly

downregulated in *Abca1 +/-* relative to its parental, as is the fatty acid oxidation gene,

phytanoyl-CoA hydroxylase. In contrast to *Cdk4*, however, loss of one *Abca1* allele

does not significantly affect net steady state expression of the *Abca1* gene itself. Other

changes are also consistent with the observed trafficking of Abca1 through the cell

(Neufeld et al. 2001). A large number of genes related to the cytoskeleton are affected,

particularly those coding for proteins that interact with the actin cytoskeleton, including

the LIM-domain proteins, *Pdlim3* and *Csrp1*. A third group of affected genes relates to

vesicle or protein movement within the cell, likely through interactions with the

cytoskeleton. Genes involved in apoptosis or functioning as chaperones also meet our

criteria (Table 3). This expression profile suggests abnormal lipid homeostasis and

cytoskeleton-associated trafficking in these cells due to heterozygosity for *Abca1*.

We next asked whether the effects of the heterozygous *Abca1* mutation in ES

cells were consistent with the effects of this mutation *in vivo*. Such a relationship must

exist for at least some differentiated cell types if ES cell expression files are to be used

to identify mutations of biological interest. Of 5000 expressed probe sets, 195 probe

sets, representing 192 genes, differed between *Abca1 +/-* and *+/+* liver (supplementary

data Table 4). Only a handful of these genes were differentially expressed in *Abca1 +/-*

ES cells as well (not shown), suggesting that the mutation has different effects in ES

cells and liver. It was possible, though, that the mutation impacted the same pathways,

but different genes, in different tissues. We therefore determined whether the same

Gene Ontology categories overrepresented in the *Abca1 +/-* ES cell data were likely to

be overrepresented in the *Abca1 +/-* liver data. The 19 Gene Ontology categories

overrepresented in the liver data included genes associated with the cytoskeleton,

vesicle trafficking, lipid homeostasis, and apoptosis (Table 4). 31% of categories (5 of 16) overrepresented in the *Abca1 +/-* ES cell data are also overrepresented in the liver data, including those associated with the cytoskeleton, vesicle trafficking, the peroxisome, and defense response (Figure 3). The shared categories are highly specific; only 4% of categories (2 of 49) overrepresented in the *Cdk4 +/-* ES cell data are also overrepresented in the liver data, a significant difference from the relationship between the *Abca1 +/-* data sets (p=0.006, Fisher's exact test). No categories were overrepresented in both ES cell data sets. These results support a causal relationship between *Abca1* heterozygous mutation and the expression changes listed in Tables 3 and 4 and suggest that the *Abca1 +/-* ES cell expression profile accurately predicts elements of the *Abca1 +/-* expression profile *in vivo*.

Discussion

The lack of a high-throughput phenotype screening strategy has limited the potential of powerful *in vitro* genetic strategies. Our data suggest that expression array profiling may functionally categorize mutant ES clones on a large scale, thereby efficiently identifying mutant phenotypes of interest *in vitro* and focusing *in vivo* experiments.

The effects of heterozygous mutations on expression profile have rarely been examined. Chauhan *et al* used expression profiling of the eye lens of heterozygous null *Pax6* mutant mice to show that expression of several hundred genes is affected by heterozygosity for the mutation. Some of the affected genes are known targets of *Pax6*;

others are targets of transcription factors that are themselves regulated by *Pax6* (Chauhan et al. 2002). However, a semi-dominant mutation that strongly impacts net expression of the mutant gene and the heterozygous mutant phenotype may represent an uncommon class, as many mutations show no overt heterozygous mutant phenotype. For example, heterozygous *Cdk4* mutant embryonic fibroblasts proliferate normally, while homozygotes show delayed entry into S-phase of the cell cycle (Rane et al. 1999). Despite this lack of an obvious heterozygous *in vivo* phenotype, the *Cdk4* mutation detectably altered expression of both the *Cdk4* gene and specific genes with similar functions in ES cells. *Abca1* represents another class of mutations, as heterozygotes have phenotypes intermediate between wildtype and the homozygous mutant (Clee et al. 2000). Net steady state *Abca1* expression is nearly normal in *Abca1* +/- ES cells and liver, but this is still sufficient to cause multiple downstream effects on pathways relevant to ABCA1 function. Although we could not link individual hemizygous genes in the B12 and C6 deletion lines to altered expression of specific downstream genes, many of the hemizygous genes were downregulated. Downregulated genes likely impacted expression profile, as with *Cdk4* +/-, and some hemizygous genes with statistically normal expression, as with *Abca1*, may also have impacted expression. These data suggest that heterozygous mutations, including many types of point mutations generated by chemical mutagenesis, will frequently impact ES cell expression profile.

This approach relies on a robust connection between expression profile and mutant gene function. However, an expression profile may also include components

that are not related to or not specific to the mutant gene, complicating interpretation of the mutant profile. It is possible, for example, that genetic and epigenetic events accumulated and were selected for in each cell line during antibiotic selection or subsequent independent passages, impacting expression profile (Gitan et al. 2002) or pluripotency (and thus expression profile) independently of the mutant genes. Several factors argue that such events contribute minimally to differences in expression profile in our experiments. Differences in differentiation state apparently do not significantly impact expression profile, as noted above. The expression profiles for the *Cdk4*, *Abca1,* B12, and C6 mutant ES cell lines are each distinct, despite the fact that all were subjected to neomycin selection. Many of the expression changes in the *Cdk4 +/-* and *Abca1 +/-* ES cell lines also are consistent with previously described functions of the mutated genes. Furthermore, the similar pathways affected in *Abca1 +/-* ES cells and *Abca1 +/-* liver strongly argue that other sources of variation do not mask the biological effects of the mutation. Expression profile might also include the effects of non-specific stresses. The B12 and C6 ES cell line profiles, caused by hemizygosity for hundreds of genes, might be expected to be heavily impacted by such effects. That these two ES cell lines differ from each other suggests that these profiles are not simply generic effects of haploinsufficiency for a large number of genes and that any non-specific effects present do not mask unique, gene-specific effects. Thus, expression profiles specifically reflect the effects of mutant genes.

As pluripotent, embryonic cells, ES cells differ in several respects from the differentiated cell types present in the whole animal. Will the expression profile of this

unique cell type accurately predict the impact of a mutation *in vivo*? *Cdk4* is an

important regulator of the transition from G1 to S phase of the cell cycle in somatic cells

(Burdon et al. 2002). Cdk4 protein interacts with D-type cyclins and the resulting kinase

phosphorylates Rb, releasing the cell from G1 and initiating expression of genes

necessary for S phase. The Cdkn2a gene product may bind *Cdk4*, preventing its

interaction with the type D cyclins and delaying cell cycle progression (Serrano et al.

1993). In contrast, Cdk2/Cyclin E is considered the primary regulator of exit from G1 in

ES cells; Cdk4/Cyclin D is thought to be relatively less important (Burdon et al. 2002).

In the present study, however, heterozygosity for *Cdk4* impacts expression of genes

involved in cell proliferation, including *Cdkn2a* (Table 2), *Myc*, *Myb*, *Brca2*, *Bub1*, *Cyclin

A2*, *Cyclin G2*, and *Cdc7l1* (supplementary data table 2). *Cdk4* function appears to

differ in ES cells and somatic cells, yet the impact of the heterozygous *Cdk4* mutation in

our study is entirely consistent with *Cdk4* function in multiple cell types.

      Abca1 has been intensively studied for its role in cellular cholesterol

homeostasis. Lack of human ABCA1 leads to Tangier Disease, a disorder

characterized by decreased plasma high density lipoprotein levels due to defective

efflux of cholesterol from cells (Oram and Lawn 2001). Mutations in mouse *Abca1* have

similar effects (McNeish et al. 2000). It has gradually become clear that Abca1 affects

diverse processes, including endocytosis (Zha et al. 2001), phagocytosis (Hamon et al.

2000), and even apoptosis (Buechler et al. 2002a). Abca1 protein also actively traffics

through several organelles (Neufeld et al. 2001), a process frequently requiring

interaction with the actin cytoskeleton (Kanzaki and Pessin 2002), and has recently

been shown to bind beta-2-syntrophin, a PDZ-domain protein involved in linking proteins to the actin cytoskeleton (Buechler et al. 2002b). Lack of *Abca1* affects expression of genes such as *Cdc42*, a member of the Rho family of GTPases involved in vesicle trafficking and actin cytoskeleton organization (Hirano et al. 2000). In fact, heterologous expression of human *ABCA1* in HEK293 cells alters the cytoskeleton of these cells (Tsukamoto et al. 2001). Thus, while loss of Abca1 in mouse liver results in widespread, high magnitude changes in expression of genes involved in lipid homeostasis, haploinsufficiency for *Abca1* appears to preferentially affect at the RNA level other functions dependent on *Abca1*. That similar effects occur in *Abca1 +/-* liver indicates that they are a function of *Abca1* genotype, rather than cell type. There is no reason to believe that the striking relationship between liver and ES cells found for *Abca1* in our study is unique to either this tissue type or this gene. These results strongly suggest that mutant ES cell expression profiles may functionally categorize mutants *in vitro*.

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25, 25-29

Bode VC (1984) Ethylnitrosourea mutagenesis and the isolation of mutant alleles for specific genes located in the T region of mouse chromosome 17. Genetics 108, 457-70.

Buechler C, Bared SM, Aslanidis C, Ritter M, Drobnik W, et al. (2002a) Molecular and functional interaction of the ATP-binding cassette transporter A1 with Fas-associated death domain protein. J Biol Chem 277, 41307-10

Buechler C, Boettcher A, Bared SM, Probst MC Schmitz G (2002b) The carboxyterminus of the ATP-binding cassette transporter A1 interacts with a beta2-syntrophin/utrophin complex. Biochem Biophys Res Commun 293, 759-65.

Burdon T, Smith A Savatier P (2002) Signalling, cell cycle and pluripotency in embryonic stem cells. Trends Cell Biol 12, 432-8.

Chauhan BK, Reed NA, Zhang W, Duncan MK, Kilimann MW, et al. (2002) Identification of genes downstream of Pax6 in the mouse lens using cDNA microarrays. J Biol Chem 277, 11539-48.

Chen Y, Schimenti J Magnuson T (2000a) Toward the yeastification of mouse genetics: chemical mutagenesis of embryonic stem cells. Mamm Genome 11, 598-602.

Chen Y, Yee D, Dains K, Chatterjee A, Cavalcoli J, et al. (2000b) Genotype-based screen for ENU-induced mutations in mouse embryonic stem cells. Nat Genet 24, 314-7.

Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, et al. (2001) Transcriptional regulation and function during the human cell cycle. Nat Genet 27, 48-54.

Clee SM, Kastelein JJ, van Dam M, Marcil M, Roomp K, et al. (2000) Age and residual cholesterol efflux affect HDL cholesterol levels and coronary artery disease in ABCA1 heterozygotes. J Clin Invest 106, 1263-70.

Eisen MB, Spellman PT, Brown PO Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95, 14863-8.

Friddle CJ, Koga T, Rubin EM Bristow J (2000) Expression profiling reveals distinct sets of genes altered during induction and regression of cardiac hypertrophy. Proc Natl Acad Sci U S A 97, 6745-50.

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286, 531-7.

Hamon Y, Broccardo C, Chambenoit O, Luciani MF, Toti F, et al. (2000) ABC1 promotes engulfment of apoptotic cells and transbilayer redistribution of phosphatidylserine. Nat Cell Biol 2, 399-406.

Hirano K, Matsuura F, Tsukamoto K, Zhang Z, Matsuyama A, et al. (2000) Decreased expression of a member of the Rho GTPase family, Cdc42Hs, in cells from Tangier disease - the small G protein may play a role in cholesterol efflux. FEBS Lett 484, 275-9.

Hitotsumachi S, Carpenter DA Russell WL (1985) Dose-repetition increases the mutagenic effectiveness of N-ethyl-N- nitrosourea in mouse spermatogonia. Proc Natl Acad Sci U S A 82, 6619-21.

Hrabe de Angelis MH, Flaswinkel H, Fuchs H, Rathkolb B, Soewarto D, et al. (2000) Genome-wide, large-scale production of mutant mice by ENU mutagenesis. Nat Genet 25, 444-7.

Kanzaki M Pessin JE (2002) Caveolin-associated filamentous actin (Cav-actin) defines a novel F-actin structure in adipocytes. J Biol Chem 277, 25867-9

McNeish J, Aiello RJ, Guyot D, Turi T, Gabel C, et al. (2000) High density lipoprotein deficiency and foam cell accumulation in mice with targeted disruption of ATP-binding cassette transporter-1. Proc Natl Acad Sci U S A 97, 4245-50.

Mitchell KJ, Pinson KI, Kelly OG, Brennan J, Zupicich J, et al. (2001) Functional analysis of secreted and transmembrane proteins critical to mouse development. Nat Genet 28, 241-9.

Munroe RJ, Bergstrom RA, Zheng QY, Libby B, Smith R, et al. (2000) Mouse mutants from chemically mutagenized embryonic stem cells. Nat Genet 24, 318-21.

Nadeau JH (2000) Muta-genetics or muta-genomics: the feasibility of large-scale mutagenesis and phenotyping programs. Mamm Genome 11, 603-7.

Neufeld EB, Remaley AT, Demosky SJ, Stonik JA, Cooney AM, et al. (2001) Cellular localization and trafficking of the human ABCA1 transporter. J Biol Chem 276, 27584-90.

Nolan PM, Peters J, Strivens M, Rogers D, Hagan J, et al. (2000) A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. Nat Genet 25, 440-3.

Oram JF Lawn RM (2001) ABCA1. The gatekeeper for eliminating excess tissue cholesterol. J Lipid Res 42, 1173-9.

Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC Melton DA (2002) "Stemness": transcriptional profiling of embryonic and adult stem cells. Science 298, 597-600.

Rane SG, Dubus P, Mettus RV, Galbreath EJ, Boden G, et al. (1999) Loss of Cdk4 expression causes insulin-deficient diabetes and Cdk4 activation results in beta-islet cell hyperplasia. Nat Genet 22, 44-52.

Rathjen J Rathjen PD (2001) Mouse ES cells: experimental exploitation of pluripotent differentiation potential. Curr Opin Genet Dev 11, 587-94.

Rinchik EM Carpenter DA (1999) N-ethyl-N-nitrosourea mutagenesis of a 6- to 11-cM subregion of the Fah- Hbb interval of mouse chromosome 7: Completed testing of 4557 gametes and deletion mapping and complementation analysis of 31 mutations. Genetics 152, 373-83.

Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, et al. (1979) Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. Proc Natl Acad Sci U S A 76, 5818-9.

Schimenti JC, Libby BJ, Bergstrom RA, Wilson LA, Naf D, et al. (2000) Interdigitated deletion complexes on mouse chromosome 5 induced by irradiation of embryonic stem cells. Genome Res 10, 1043-50.

Serrano M, Hannon GJ Beach D (1993) A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature 366, 704-7.

Skarnes WC (2000) Gene trapping methods for the identification and functional analysis of cell surface proteins in mice. Methods Enzymol 328, 592-615

Stanford WL, Cohn JB Cordes SP (2001) Gene-trap mutagenesis: past, present and beyond. Nat Rev Genet 2, 756-68.

Stryke D, Kawamoto M, Huang CC, Johns SJ, King LA, et al. (2003) BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Res 31, 278-81.

Tanaka TS, Kunath T, Kimber WL, Jaradat SA, Stagg CA, et al. (2002) Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. Genome Res 12, 1921-8.

Tsukamoto K, Hirano K, Tsujii K, Ikegami C, Zhongyan Z, et al. (2001) ATP-binding cassette transporter-1 induces rearrangement of actin cytoskeletons possibly through Cdc42/N-WASP. Biochem Biophys Res Commun 287, 757-65.

van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, et al. (2002) A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347, 1999-2009.

Wiles MV, Vauti F, Otte J, Fuchtbauer EM, Ruiz P, et al. (2000) Establishment of a gene-trap sequence tag library to generate mutant mice from embryonic stem cells. Nat Genet 24, 13-4.

Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C, et al. (1998) Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature 392, 608-11.

Zha X, Genest J, Jr. McPherson R (2001) Endocytosis is enhanced in Tangier fibroblasts: possible role of ATP- binding cassette protein A1 in endosomal vesicular transport. J Biol Chem 276, 39476-83.

Zhang HS, Gavin M, Dahiya A, Postigo AA, Ma D, et al. (2000) Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. Cell 101, 79-89.

Figure 1. Hierarchical clustering of expression arrays data for C6, B12, and v6.4 ES cell lines. Normalized expression ratios were clustered using standard correlation coefficient for three replicates of each cell line. Differences in color denote differences in expression.

Figure 2. Relative expression of genes associated with pluripotency. Mutant/parental expression ratio for 26 genes shown by Tanaka et al. (Tanaka et al. 2002) to be significantly upregulated in ES cells relative to embryonic fibroblasts and embryonic trophoblast stem cells. Asterisks denote genes significantly different in a mutant vs. its parental (p<0.05, t-test).

Figure 3. Functional consequences of heterozygous mutation in *Cdk4* or *Abca1*. Venn diagram of Gene Ontology categories overrepresented among differentially expressed genes in the *Abca1 +/-* ES cell, *Abca1 +/-* liver data, and the *Cdk4 +/-* ES cell data and represented by at least two genes. Categories overrepresented in both *Abca1 +/-* data sets are shown to the left of the Venn diagram.