Title

Evolution of metastasis revealed by mutational landscapes of chemically induced skin cancers

Journal

Nature Medicine, 21(12)

ISSN

1078-8956

Authors

McCreery, Melissa Q

Halliwill, Kyle D

Chin, Douglas

et al.

Publication Date

2015-12-01

DOI

10.1038/nm.3979

Peer reviewed

Published in final edited form as:

Nat Med. 2015 December ; 21(12): 1514–1520. doi:10.1038/nm.3979.

Evolution of metastasis revealed by mutational landscapes of chemically induced skin cancers

Melissa Q McCreery[1], **Kyle D Halliwill**[1], **Douglas Chin**[1], **Reyno Delrosario**[1], **Gillian Hirst**[1], **Peter Vuong**[1], **Kuang-Yu Jen**[1], **James Hewinson**[2], **David J Adams**[2], and **Allan Balmain**[1]

[1]University of California San Francisco Helen Diller Family Comprehensive Cancer Center, Cancer Research Institute, University of California, San Francisco, San Francisco, California, USA

[2]Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, UK

Abstract

Human tumors show a high level of genetic heterogeneity, but the processes that influence the timing and route of metastatic dissemination of the subclones are unknown. Here we have used whole-exome sequencing of 103 matched benign, malignant and metastatic skin tumors from genetically heterogeneous mice to demonstrate that most metastases disseminate synchronously from the primary tumor, supporting parallel rather than linear evolution as the predominant model of metastasis. Shared mutations between primary carcinomas and their matched metastases have the distinct A-to-T signature of the initiating carcinogen dimethylbenzanthracene, but non-shared mutations are primarily G-to-T, a signature associated with oxidative stress. The existence of carcinomas that either did or did not metastasize in the same host animal suggests that there are tumor-intrinsic factors that influence metastatic seeding. We also demonstrate the importance of germline polymorphisms in determining allele-specific mutations, and we identify somatic genetic alterations that are specifically related to initiation of carcinogenesis by *Hras* or *Kras* mutations. Mouse tumors that mimic the genetic heterogeneity of human cancers can aid our understanding of the clonal evolution of metastasis and provide a realistic model for the testing of novel therapies.

Next-generation sequencing technologies have revolutionized our understanding of the genomic landscapes of human cancers[1,2], but we have only a poor appreciation of the genetic determinants of metastasis, the leading cause of human cancer death[3]. Metastases

Correspondence should be addressed to A.B. (abalmain@cc.ucsf.edu).

Accession codes. Microarray data are available under Gene Expression Omnibus accession number GSE63967 and sequence data under European Nucleotide Archive study accession number ERP004081.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

M.Q.M., D.J.A. and A.B. contributed to the overall study design; M.Q.M. carried out the experiments and computational analysis, with input from K.D.H.; R.D. and D.C. carried out the animal studies; G.H., D.C. and P.V. organized the tissue bank; and K.-Y.J. did histological assessments of all of the tumors. Sequencing and MiSeq validation were carried out at the Sanger Institute under the supervision of D.J.A., with assistance from J.H. M.Q.M. and A.B. wrote the manuscript, with contributions from the other authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

may spread linearly (from one organ site to the next) or in parallel (each departing separately from the primary tumor)[4], but lack of appropriate human matched primary carcinomas and metastases has hindered analysis of these questions. We took advantage of a mouse skin tumor model that encompasses genetic and environmental factors and for which all stages, from benign lesions to metastases, are available[5]. We show here that chemically induced skin tumors display a diversity of point mutations (single-nucleotide variants, or SNVs) and gene copy number variants (CNVs) that permit a detailed analysis of clonal evolution from the initiating cell to metastasis.

To mimic the extreme germline and somatic genetic heterogeneity in human samples, we bred *Mus musculus* mice (FVB/N) with a wild-derived strain, *Mus spretus* (SPRET/EiJ), to create a heterogeneous backcross population (hereafter referred to as FVBBX) (Fig. 1a) and carried out chemical carcinogenesis with dimethylbenzanthracene (DMBA) and 12-*O*-tetradecanoylphorbol-13-acetate (TPA). Squamous carcinomas (SCCs) in this model almost always carry mutations in *Hras*[6], which is also a driver of squamous carcinogenesis in human head and neck[7] and skin[8]. To compare routes to carcinoma development that are driven by *Hras* or *Kras*[5,9], which are mutated in many human cancers[10], we also bred $Hras^{-/-}$ FVBBX mice that develop aggressive tumors carrying *Kras* mutations after DMBA-TPA treatment[5]. To further mimic human clinical practice, we surgically removed primary carcinomas, allowing continued survival and subsequent harvesting of distant metastases (Fig. 1a)[5]. We selected 103 tumor samples from 15 animals (9 wild-type and 6 $Hras^{-/-}$ mice)—for which matched normal tissue, benign papillomas, carcinomas and metastases were available—and carried out whole exome-sequencing and gene expression microarrays (Supplementary Table 1).

RESULTS

Mutation spectrum of early and late mutations

The predominant mutation found in tumors from wild-type FVBBX mice was the classical A-to-T transversion in the *Hras* gene[6,11], leading to Q61L mutations (in 58 tumors from 9 mice). By using a nested-PCR approach[12], this mutation can be detected as early as 1 week after DMBA treatment, but the skin remains morphologically normal for periods of up to 1 year in the absence of further treatment with TPA[13]. Tumors from $Hras^{-/-}$ mice had *Kras* mutations at a range of 'hot spot' sites—G12D, G13R, Q61L and Q61H (in 26 tumors from 6 mice) (Supplementary Table 1). Both *Hras* and *Kras* mutations showed a strong preference for the FVB allele in mice heterozygous at these loci (84% and 100% specificity, respectively; $P < 0.001$, Supplementary Fig. 1), consistent with previous work indicating allelic preference for *Ras* mutations in mouse models[14,15]. A single mouse carried no *Hras* or *Kras* mutations, but it had several tumors with a mutation in *Pik3ca*, which encodes phosphatidylinositol-4,5-bisphosphate-3-kinase catalytic subunit alpha, at the most common hot spot (H1047L) found in human cancers[16].

Tumors carried an average of 237 SNVs (5.2 mutations/Mb), which is similar to the mutational burden of human adenocarcinomas and SCCs of the lung[17]. Papillomas harbored fewer mutations than did age-matched primary carcinomas (172 versus 284 mutations on average, respectively; $P = 0.01$; Supplementary Table 2), but whether this is due to

differences in DNA repair, proliferation or increased genomic instability during progression is presently unclear. The genome-wide mutation spectrum across all 103 tumors showed an enrichment in DMBA signature transversion mutations (i.e., A-to-T mutations; hereafter referred to as A>T; 45% of all mutations), predominantly at 2 of the 96 possible trinucleotide contexts (Fig. 1b)[18].

Mutations in metastases that were shared with a matched primary tumor showed an enrichment of A>T DMBA signature mutations (63% of all mutations). Subsequent non-shared mutations were predominantly G-to-T mutations (hereafter referred to as G>T mutations; 50%, as compared to 14% of shared mutations; Fig. 1c). Subclonal mutations in tumors of all stages were also much more likely to be G>T than A>T substitutions (Fig. 1c). The average mutant allele fraction of A>T mutations was significantly higher than that of G>T mutations (0.265 versus 0.190, $P = 2.2 \times 10^{-16}$), consistent with A>T mutations being early and clonal, and G>T mutations occurring later and more frequently being subclonal. These data suggest that most G>T mutations are not directly induced by DMBA but arise during subsequent tumor growth owing to increased genomic instability or oxidative stress.

Evolutionary trees reveal patterns of metastatic dissemination

We constructed phylogenetic trees[19] demonstrating the evolutionary relationships among papillomas, carcinomas and metastases from each mouse (Fig. 2 and Supplementary Fig. 2). In 7 of 8 cases in which multiple metastases originated from a single primary carcinoma, the number and identity of mutations shared with the primary carcinoma were almost identical between metastases, suggesting that dissemination had occurred synchronously and in parallel. However, both primary tumor and metastases continued to evolve, and each tumor accumulated many private mutations after divergence.

Tumor cells can disseminate at an early stage, before the evolution of overt carcinomas[4], but the relationship between these early circulating cells and the outgrowth of metastasis is still unclear. In the skin model, metastasis occurred at different points during tumor evolution, in some cases diverging after many mutations had accumulated, whereas in others diverging after relatively few mutations had accumulated. In mouse 1664 (Fig. 2a), 46% of total SNVs and only 17% of G>T substitutions were shared between carcinoma A and its four metastases. In contrast, nearly 88% of SNVs and 66% of G>T substitutions were shared among carcinoma A and its four metastases in mouse 1383 (Fig. 2b). Other mice had more intermediate distributions (Fig. 2c–f and Supplementary Fig. 2). We conclude that although dissemination may be able to occur at an early stage, additional factors determine whether these circulating cells can lead to seeding and growth of metastases. Moreover, the timing of dissemination seems to be unrelated to the total mutational burden of the primary tumor.

These data also provide information on possible routes of dissemination. If metastasis occurs first to a regional lymph node and subsequently to distant sites, as proposed in the ‘linear evolution’ model[4], then the metastases should be more genetically related to each other than to the primary carcinoma. This is in fact seen in only one animal (mouse 1949, Fig. 2f). In most cases, each metastasis was genetically distinct, supporting a model of independent, parallel spreading from the primary tumor. Furthermore, metastases to lymph nodes do not

always depart the primary tumor first; in mouse 1407, the spleen metastasis departed first (Fig. 2c).

In some cases, carcinomas developed that showed no signs of metastasis, even in the presence of other metastatic carcinomas in the same animals (such as in mouse 1664, 1383, 2104 and 1949, Fig. 2; and in mouse 1508, Supplementary Fig. 2). This suggests that some carcinomas are intrinsically metastatic, whereas others are restrained from forming metastases at distant sites, possibly through immune surveillance[20], lack of appropriate 'niche' factors or other tumor-specific mechanisms[21]. Comparative analysis of mutations in metastasizing and non-metastasizing primary tumors did not reveal any obvious candidate metastasis driver mutations (see Online Methods). Gene expression analysis identified *Pdlim7* (also known as *LMP1* in humans, a known downstream target of transforming growth factor (TGF)-β1 signaling[22]) as a more highly expressed gene in metastasizing tumors than in non-metastasizing primary tumors ($P = 6.4 \times 10^{-6}$). Several other genes were suggestive of being differentially expressed but did not meet the threshold for significance, including *Cd151* ($P = 1.4 \times 10^{-4}$), a cell surface glycoprotein involved in cell adhesion, integrin trafficking and metastasis[23,24]. Expression levels of *Pdlim7* and *Cd151* are correlated with each other in carcinomas from this data set (Spearman's rank correlation coefficient (ρ) = 0.77; $P = 6.7 \times 10^{-6}$) as well as in an independent set of mouse carcinomas ($\rho = 0.52$; $P = 2.2 \times 10^{-5}$) (ref. 25) and in the Cancer Genome Atlas (TCGA) profiles of human head and neck cancers ($\rho = 0.67$) (refs. 26,27).

Mesenchymal-to-epithelial transition depends on the metastatic site

Metastases to all organs except the lung—including the lymph nodes, kidney, liver, spleen and thymus—nearly always matched the squamous or spindle morphology of the primary tumor (21 of 22 cases, 95%). Metastases to the lung were almost always squamous (7 of 9 cases, 78%; one of the seven cases had mixed SCC and spindle morphology), even when they arose from primary tumors with spindle morphology. This finding is consistent with previous observations of metastases to the lung with squamous morphology in an inbred FVB/N model[5]; however, here we were able to sequence-match these metastases to their respective primary tumors (Supplementary Table 3) and demonstrate that squamous lung metastases could arise from a primary spindle tumor, even when all other metastases from that tumor were of spindle morphology (Fig. 3 and Supplementary Table 3). These data suggest that the requirement for mesenchymal-to- epithelial transition (MET) for the outgrowth of metastases at distant sites[28], which predicts epithelial-like squamous metastases, may in fact depend on the specific organ site: MET may be favored in the lung, but not in the soft tissue sites in which spindle metastases were frequently found.

Mutations shared with human SCCs

This analysis identified a large number of recurrently mutated genes in addition to *Hras* and *Kras*, many of which are reported to be mutated in human SCCs of the head and neck, skin and lung[7,8,29]. (Fig. 4 and Supplementary Table 4). These included recurrent mutations in *Trp53*, *Fat1* and genes in the Notch signaling pathway, including *Notch1*, *Notch3* and *Trp63* (Fig. 4). Dysregulation of Notch signaling, which is involved in epithelial differentiation, has been implicated in a variety of human SCCs[7,8,29]. The three mutations we observed in

Notch1 and *Notch3* were all in the N-terminal epidermal growth factor (EGF)-like repeat domains responsible for Ca^{2+} ligand binding, and were probably inactivating mutations, as expected for SCCs, rather than activating mutations as seen in human leukemias[7,8].

Recurrent mutations were also seen in *Ep300*, *Apc*, *Ncor1*, *Syne1*, *Syne2*, *Ros1* and *Dnmt1*, all of which are mutated in human tumors. Single mutations were also seen in *Pik3ca* and *Casp8* (consistent with those seen in head and neck SCC (HNSCC)), in *Bbs9*, *Dclk1* and *Kmt2c* (consistent with those seen in cutaneous SCCs), and in *Keap1* (consistent with those seen in lung SCCs). Among these, stop-gains (mutations that create stop codons) were seen in *Apc*, *Ncor1* and *Dclk1*.

Dependence of CNVs on initiation by *Hras* or *Kras* mutations

We used the whole-exome sequencing data to construct copy number profiles for all of the tumors. Copy number gains of chromosome 7 (on which *Hras* is located) were seen in tumors from wild-type mice but not $Hras^{-/-}$ mice, pointing to *Hras* as the driving force behind whole chromosome 7 gains. Undifferentiated spindle tumors did not have chromosome 7 gains, in line with our observation that spindle carcinomas have reduced *Hras* expression[5]. Copy number gains of chromosome 6, previously noted by karyotype analysis[30], were seen in tumors from both wild-type and $Hras^{-/-}$ mice (Fig. 5a). *Kras*, *Braf*, *Craf* and *Met* all reside on chromosome 6 and are candidate drivers of chromosome 6 gains. Additionally, we observed that SCCs from wild-type mice (i.e., *Hras*-driven carcinomas) have copy number gains on chromosome 1, and these are not seen in any other tumor class. The absence of chromosome 1 gains in papillomas and spindle tumors suggests that these gains are specifically linked to the squamous papilloma-carcinoma conversion.

To understand what might be driving the copy number gains of chromosome 1, we searched for genes on chromosome 1 that were expressed at significantly higher levels in squamous tumors that had chromosome 1 copy number gains as compared to those that did not (Fig. 5a). *Nuak2*, which encodes an AMP protein kinase (AMPK)-related kinase, showed the greatest significant change in expression (1.7-fold increase, $P = 5.3 \times 10^{-5}$), and changes in expression of *Itpkb* (which encodes inositol 1,4,5-trisphosphate-3-kinase B) and *Cflar* (which encodes caspase 8 and FADD-like apoptosis regulator) were also significant ($P = 5.1 \times 10^{-6}$ for *Itpkb;* $P = 1.1 \times 10^{-4}$ for *Cflar*; Supplementary Table 5). *Nuak2* expression levels were strongly correlated with *Hras* expression levels in tumors from wild-type mice in this study ($\rho = 0.63$, $P = 1.9 \times 10^{-9}$), as well as in an independently derived data set[25] ($\rho = 0.58$, $P = 1.3 \times 10^{-6}$). *Nuak2* has been implicated in human cancer development by copy number gains in human melanoma[31], in breast and liver carcinomas[26,27] and, to a lesser extent, in HNSCC and lung SCC[7,8], and is therefore a possible therapeutic target. However, other genes, including those shown in Supplementary Table 5, may also have a role in promoting the development of *Hras*-driven squamous carcinomas.

Among the most frequent copy number changes in aggressive human cancers are deletions at the *CDKN2A* locus on chromosome 9p and amplification or copy number gains of *MET*[29,32], both of which we also observed in our cohort. *Cdkn2a* was deleted in 52 tumors (50%), including nearly all spindle tumors (27 of 29 tumors, 93%) (Fig. 5b,c)[5,33]. *Met* showed copy number gains in 47 tumors (46%), as the result of both focal amplifications (6

tumors) and of whole gains of chromosome 6 (41 tumors). Both the occurrence and co-occurrence of these lesions increased dramatically with tumor stage: both events were rare in papillomas, but in SCCs, 63% of tumors had either a *Cdkn2a* loss or a *Met* gain, and 25% had both. In spindle tumors, nearly all tumors had a *Cdkn2a* loss or a *Met* gain, and 59% of tumors had both (Fig. 5c).

Acquisition of SNVs precedes CNVs during progression

To verify the timing of SNV and CNV acquisition during tumor progression, we sequenced a set of 18 early, small (2–4 mm in diameter) FVB/N papillomas at 12–14 weeks after tumor initiation (9 from wild-type mice and 9 from *Hras*$^{-/-}$ mice). The overall mutation spectrum was highly similar to that in the tumors from the FVBBX population (Fig. 1 and Supplementary Fig. 3). Notably, these early papillomas had already acquired a similar proportion of G>T mutations to that observed in late papillomas from FVBBX mice, suggesting that most G>T mutations are acquired relatively early in the life of the tumor, during the phase of proliferation and chronic inflammation that is induced by the tumor-promoting TPA. Analysis of CNVs in early papillomas identified copy number gains of chromosome 7, which contains the mutant *Hras* gene, exclusively in papillomas from wild-type mice, as the only recurrent and significant early event (Supplementary Fig. 4). The striking absence of additional copy number events in these early tumors suggests that the acquisition of CNVs seen in the original cohort is a later event that is associated with tumor progression.

DISCUSSION

We have exploited a classical skin tumor model in a genetically heterogeneous mouse population to investigate the evolutionary history of metastatic tumors using whole-exome sequencing of multiple lesions from the same animals. The ability to distinguish the parental alleles in each mouse shows that both *Hras* and *Kras* mutations are highly allele specific, thus demonstrating the important role of germline polymorphisms in determining mutation selection. Furthermore, the nature of the initiating event (*Hras* or *Kras* mutation) influences the subsequent genomic changes that take place, making this model useful for identifying potential drug candidates, such as *Nuak2,* that may be activated in tumors with *Hras* mutations.

The use of specific initiating mutagens combined with clonal analysis has enabled us to distinguish early from late mutations. Whereas the initiating agent DMBA causes primarily A>T mutations across the genome, tumors accumulate G>T transversions during the early phase of TPA treatment, presumably owing to the induction of oxidative stress pathways that are linked to inflammation or to Ras signaling[34,35]. The high frequency of point mutations in these skin tumors, which contrasts with the paucity of informative mutations in genetically engineered mouse (GEM) models[36,37], provides the opportunity to map, in great detail, the sequential mutations that take place during tumor evolution.

Our data support the ‘parallel evolution’ model, in which each metastasis is seeded directly by the primary tumor rather than via a lymph node[4,21,38], as the major route to metastasis. Only one mouse (mouse 1949, Fig. 2f) showed evidence in favor of the ‘linear’ model (Fig.

2 and Supplementary Fig. 2). Notably, the number of shared mutations between primary and metastatic carcinomas in these mice was highly variable, suggesting that the seeding could occur either early or late in primary tumor evolution. Dissemination from the primary tumor can therefore occur very early and probably continuously[4], but the capacity of disseminated tumor cells to grow out into full metastases is controlled by separate, presently unknown, factors (possibly secreted 'niche factors')[39,40] that allow progressive growth at distant sites. However, the presence of multiple primary carcinomas in some animals, only one of which showed clear signs of metastasis, suggests that seeding is also controlled by tumor-intrinsic factors, which may include the repression or absence of neo-antigens that would stimulate recognition by the immune system[41].

Our study design also allowed us to compare primary carcinomas that had metastasized with other carcinomas from the same animal that had not. The availability of these unique matched samples allowed us to perform gene expression analyses to identify a network of genes that were more highly expressed in metastasizing carcinomas than in their non-metastasizing counterparts, including *Pdlim7* and *Cd151*, which have been previously linked to integrin signaling and tumor invasion[23,24]. Further functional studies of these candidate genes in this model will be the subject of future investigations.

Another recent study of the evolution of metastasis in a mouse model of small-cell lung cancer, using next-generation sequencing, reached the conclusion that lymph nodes may be the gateway to distant metastases[37], in agreement with the linear model. However, evidence for linear seeding was based mostly on a primary lung cancer and two metastases, one in the lymph node and one in the liver, from a single mouse. The reasons for the different conclusions from this study are unclear, but they may be due to the relative numbers of samples analyzed or to the use of a GEM model with very few nonsynonymous point mutations. During the revision stage of this manuscript, another study was published that used sequencing to identify relationships between primary skin tumors and their metastases[42]. In this study, the conclusions regarding tumor spread were, again, made on the basis of one informative mouse with two metastases. Thus, no clear conclusions can be drawn from these studies regarding the predominance of the linear versus parallel evolution models of metastasis.

The parallel evolution of metastases observed in this study is supported by a substantial body of clinical data showing that lymph node removal has no impact on patient survival[43–45]. The observation that primary and metastatic tumors in some mice share a very low number of mutations supports prior observations that early-stage tumors can disseminate even before reaching overt malignancy[4,38]. Our data indicate that not only must these tumors have begun dissemination early but that they also acquired the ability to seed metastases, probably before the primary tumor advanced to a clinically recognizable carcinoma stage. Clearly, efforts to prevent dissemination or metastatic seeding in such cases would be futile, and other therapeutic approaches would be required. The availability of mouse models of cancer that mimic both the genomic heterogeneity and natural evolution of metastasis, such as that described here, will provide the opportunity to address these questions in more detail in a preclinical setting.

ONLINE METHODS

Mouse backcross and carcinogenesis

Wild-type male SPRET/EiJ *M. spretus* mice were crossed to female *Hras*$^{-/-}$ FVB/N *M. musculus* mice to generate *Hras*-heterozygous F1s. Female F1s were crossed with male *Hras*$^{-/-}$ or *Hras*-heterozygous FVB/N *M. musculus* mice to generate FVBBX mice. Fifteen FVBBX mice were used in this experiment, 13 males and 2 females, chosen based on the availability of papilloma, primary carcinoma and metastasis samples for sequencing. Early papillomas were harvested from four female and three male FVB/N *M. musculus* mice, either wild type or *Hras*$^{-/-}$, at 12–14 weeks after initiation with DMBA, at a size of 2–4 mm in diameter.

Mice were shaved and treated with 25 μg DMBA (Sigma-Aldrich) (dissolved in 200 μl acetone) at 8 weeks, and they subsequently received TPA (Sigma-Aldrich) (200 μL of a 10^{-4} M solution in acetone) twice a week for 20 weeks, following an established chemical carcinogenesis protocol[5]. Carcinomas were surgically resected when they reached a size of >1 cm in longest diameter. Mice were sacrificed when disease progressed, as per animal care requirements. At sacrifice, papillomas and carcinomas were removed from skin, and all internal tumors were resected. Matched normal skin was also taken. A section of each primary carcinoma and metastasis was embedded in paraffin at the time of removal, for histological evaluation. The remaining tumor was flash frozen and stored at −80 °C, as was normal skin. No randomization or blinding was used in these experiments, as all mice underwent an identical treatment protocol. All animal experiments were approved by the University of California San Francisco (UCSF) Laboratory Animal Resource Center. Assessment of mouse tumor burden and size was carried out according to the UCSF Institutional Animal Care and Use Committee guidelines.

Histological classification

Formalin-fixed paraffin-embedded tumor was sectioned to 6 μm and H&E stained for histologic assessment.

Nucleic acid extraction

DNA was extracted from flash-frozen tumor and skin samples. Frozen tissue was ground and digested in 3 mg proteinase K (Bioline) overnight. DNA was separated by addition of phenol-chloroform and use of 5 PRIME Phase Lock Gel Heavy Tubes (Fisher Scientific), precipitated with isopropanol, washed with 70% ethanol and resuspended in nuclease-free water. Concentration and quality were determined by Nanodrop spectrophotometry and by PicoGreen (Invitrogen).

RNA was extracted from flash-frozen tumor and skin samples using TriZol reagent (Invitrogen) and cleaned with Qiagen RNeasy kit, according to the manufacturers' instructions. Concentration was determined by Nanodrop spectrophotometry, and RNA integrity number (RIN) was determined by an Agilent 2100 Bioanalyzer. All samples used for microarray analysis had a minimum RIN score of 6.5.

Exome sequencing, alignment and quality control

Exome enrichment and sequencing genomic libraries were prepared using the Illumina Paired-End Sample Prep Kit, following the manufacturer's instructions. Enrichment was performed as described previously[46] using the Agilent SureSelect Mouse All Exon kit according to the manufacturer's recommended protocol. Each exome was sequenced using a 76-bp paired-end protocol on the Illumina platform (GAII or HiSeq2000).

Tumor .bam files were aligned to the GRCm38/mm10 version of the *M. musculus* genome using BWA (version 0.5.10) (ref. 47). After alignment, duplicates were marked and mate information was fixed using Picard (version 1.72; http://picard.sourceforge.net/). We then recalibrated base quality score and realigned reads around insertions and deletions (indels) using Genome Analysis Toolkit (GATK) (version 1.5–9) (ref. 48). Finally, alignment and coverage metrics were collected using Picard. We sequenced an average of 66 million unique on-target reads per tumor. Targeted bases were sequenced to a mean depth of 50, and 78% of targeted bases were sequenced to 20× coverage or greater. There were no significant differences in depth of coverage or proportion of regions covered to 20× between tumor types.

Variant calling

SNVs were called using somatic variant detection program MuTect (version 1.1.4) (ref. 49). Each tumor was called against its matched normal tissue, and calls were filtered against a database of known *M. musculus* FVB and *M. spretus* SPRET germline single-nucleotide polymorphisms (SNPs) available at ftp://ftp-mouse.sanger.ac.uk/ (release 1303, mgp.v3), as well as against a panel of normal skins from this experiment. MuTect was run with the 'required_maximum_alt_allele_mapping_quality_score' parameter set to 60, as the mixed genetic background of our FVBBX cross resulted in a number of poorly mapping reads we wished to exclude. Results were further filtered to calls with a minimum read depth of 10 and with a minimum mutant allele fraction of 10%. Variants were annotated using Annovar (downloaded on 2/13/2014)[50], and these annotations were used as the basis for assessing exonic variants as synonymous, nonsynonymous, stop-gain or stop-loss. Mutations were compared to mutations observed in human cancers using the COSMIC database (version 68, http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/download)[51]. Genes were selected for display (Fig. 4) based on whether they were commonly mutated in HNSCC or cutaneous skin SCC and on genes on a previous published list of cancer driver genes[52].

Copy number calling was done with CNVkit[53], and a tumor's copy number status was called against a panel of normal skins from this experiment. Raw copy number gain was adjusted to remove dilution from contaminating normal cells, with contamination determined by PyClone estimates (see below). Whole chromosome gains were considered to be those in which over 90% of the chromosome was gained and the weighted-mean amplification $\log_2$ was greater than 0.4, implying trisomy in at least 30% of cells. Focal copy number events were determined as those for which the average copy number change for all exons of the gene had a $\log_2$ increase of 0.6 or a $\log_2$ decrease of −1 (corresponding to a 50% gain or loss of alleles).

Mutation validation

To validate our mutation calls, we sequenced 205 SNVs to a median depth of 7,080 reads using MiSeq. We obtained an 86% validation rate.

PCR and sequencing

Primers for each candidate SNP were designed as nested pairs using Primer3 (refs. 54,55). External amplicons were fixed at 400–800 bp, and internal amplicons were fixed at 250–300 bp. Internal primer pairs had partial Illumina adaptor sequences added to allow the construction of barcoded Illumina libraries. Tail sequences are: left adaptor 5′-ACACTCTTTCCCTACAC GACGCTCTTCCGATCT-3′, right adaptor 5′-TCGGCATTCCTGCTGAACC GCTCTTCCGATCT-3′. PCR was performed using hot start Taq DNA polymer-ase according to the manufacturer's instructions (Thermo Scientific). PCR1 (external primers) used a touchdown PCR approach for increased specificity. The products of PCR1 were diluted 100× and used as a template for PCR2, which added the partial Illumina adaptor tails. Candidate SNP PCRs for the same tumor were then pooled, and each tumor was barcoded with a different full-length Illumina adaptor barcode in a short PCR3 (15 cycles), so that each tumor could be decoded after sequencing. Resulting tumor-specific PCRs were pooled, cleaned up with 0.8× AmpureXP beads (Beckman Coulter), quantified and run on an Illumina MiSeq 150-bp PE run.

Data analysis

Illumina adaptor sequence and low-quality bases (<20) were trimmed from reads using cutadapt[56] (https://cutadapt.readthedocs.org/en/stable/). Reads were mapped to the GRCm38 build of the mouse genome using BWA mem[47] (http://bio-bwa.sourceforge.net/). Because amplicon sequencing was used, duplicate reads were not marked. Samtools mpileup and bcftools call[57] (http://www.htslib.org/) were used to output reference or alternate base calls and allele depths for the candidate positions. Samples were analyzed as groups of matched, normal skin and tumor samples. SNVs were considered validated if they contained at least 70 reads supporting the expected mutation.

Mutation context

For mutation spectrum analysis, SNVs in all tumors were annotated with 1 of 96 possible trinucleotide-context substitutions (six types of substitutions × four possible flanking 5′-bases × four possible flanking 3′-bases), and counts of each mutation context were summed.

Clonality

Clonality analysis was performed with PyClone (version 0.12.7) (ref. 58). For inputs to PyClone, reference and variant read depths were taken from the MuTect output, and copy number at each locus was determined from the CNVkit output. Clonal and subclonal mutation clusters were determined from the PyClone results table and compared with the PyClone cellular frequency plots for confirmation. Clonal mutations were those present in all tumor cells; subclonal mutations were those present in only a fraction of tumor cells.

Phylogenies

To build phylogenetic trees, absolute distance matrices were calculated based on the presence of mutations in the sample, which were based on filtered MuTect calls. Rooted trees were built with the use of the 'analyses of phylogenetics and evolution' (APE) package and the Manhattan calculation method implemented in R version 2.15. Metastases were matched to primary tumors on the basis of shared mutations.

Identification of early and late mutations

Early and late mutations were identified with two methods. In the first method, mutations from 26 metastases with matched primaries were divided into those that were shared with the primary (early) and those that were not shared (later). Tumors used in this method included a representation of SCCs and spindle tumors and of *Hras*-driven and *Kras*-driven tumors. In the second method, mutations in 12 tumors, in which multiple clones were identified with PyClone, were divided into fully clonal mutations (i.e., those present in all tumor cells) (early) and subclonal mutations (i.e., those present in only a fraction of tumor cells) (late). Tumors used in this method included a representation of papillomas, SCCs and spindle tumors; *Hras*-driven and *Kras*-driven tumors; and primary tumors and metastases.

Evaluation of candidate 'metastasis genes'

To assess whether we had any candidate 'metastasis genes' that could explain why certain primary tumors metastasized whereas others did not, we first examined genes that were recurrently mutated in metastasizing carcinomas. We found 22 genes that were mutated more than once in a metastasizing carcinoma; however, 19 of these were also mutated in non-metastasizing carcinomas at a similar frequency. The remaining three genes–*Foxn4*, *Scn9a* and *Sspo*—were found mutated in one or more papillomas and thus also did not present good candidates. We found that *Pdlim7*, *Tecpr2* and *Cd151* were expressed at higher levels in metastasizing carcinomas, although, of these, expression of *Pdlim7* was significant after a multiple-test correction ($P = 6.4 \times 10^{-6}$). None of these genes carried any mutations in any primary or metastatic carcinomas. No genes were expressed at significantly lower levels in metastasizing carcinomas.

Allele specificity analysis

To determine the allele specificity of *Hras* and *Kras* mutations, 75-bp reads that contained both the mutation and nearby SNPs were extracted from the .bam file using mpileup[56], and the genotype of mutated allele was determined. For codon 61 mutations in *Hras*, the SNPs used to determine genotype were chr7:141192537, A (FVB) or G (SPRET) and chr7:141192567, G (FVB) or A (SPRET). For codon 61 mutations in *Kras*, the SNPs used were chr6:145234318, T (FVB) or C (SPRET) and chr6:145234388 G (FVB) or A (SPRET). For mutations in *Kras* codons 12 and 13, the SNPs used were chr6:145246755, G (FVB) or A (SPRET) and chr6:145246782 C (FVB) or T (SPRET). Significance was evaluated using a chi-squared test.

Expression microarrays, normalization and differential expression

Gene expression was quantified using Affymetrix MoGene ST 1.1 arrays hybridized on an Affymetrix GeneTitan instrument. Affymetrix MoGene arrays were normalized using the oligo package[59], and a probe database was prepared for FVBBX mice to avoid probes which intersect known FVB/N or SPRET/EiJ SNPs (http://davidquigley.com/equalizer.html)[60]. Differential expression for genes on chromosome 1 and for comparison of metastasizing versus non- metastasizing carcinomas was done using the siggenes package in R, with a false discovery rate threshold of 0.01. Microarray data is available under GEO accession number GSE63967 and sequence data under ENA study accession number ERP004081.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by US National Cancer Institute (NCI) grants UO1 CA84244 (A.B.), UO1 CA141455 (A.B.), UO1 CA176287 (A.B.) and RO1CA184510 (A.B.), and the Barbara Bass Bakar Distinguished Professorship in Cancer Genetics. K.D.H. was supported by the US National Institutes of Health training grant T32 GM007175 and is currently supported by an NCI F31 National Research Service Award (NRSA). D.J.A. is supported by Cancer Research UK and the Wellcome Trust. We thank our colleagues for their help and comments in refining this study and manuscript. We would also like to thank R. Kettleborough and R. White (both at the Wellcome Trust Sanger Institute) for assistance with MiSeq analysis and E. Wu, K. Copren and the UCSF Helen Diller Family Comprehensive Cancer Cancer Center Genome Analysis Core for assistance with Affymetrix microarrays and with sequencing.

References

1. Chang K, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013; 45:1113–1120. [PubMed: 24071849]
2. Ciriello G, et al. Emerging landscape of oncogenic signatures across human cancers. Nat Genet. 2013; 45:1127–1133. [PubMed: 24071851]
3. Brabletz T, Lyden D, Steeg PS, Werb Z. Roadblocks to translational advances on metastasis research. Nat Med. 2013; 19:1104–1109. [PubMed: 24013756]
4. Klein CA. Parallel progression of tumor and metastases. Nat Rev Cancer. 2009; 9:302–312. [PubMed: 19308069]
5. Wong CE, et al. Inflammation and Hras signaling control epithelial-mesenchymal transition during skin tumor progression. Genes Dev. 2013; 27:670–682. [PubMed: 23512660]
6. Quintanilla M, Brown K, Ramsden M, Balmain A. Carcinogen-specific mutation and amplification of Ha-*ras* during mouse skin carcinogenesis. Nature. 1986; 322:78–80. [PubMed: 3014349]
7. Stransky N, et al. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011; 333:1157–1160. [PubMed: 21798893]
8. Pickering CR, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. Clin Cancer Res. 2014; 20:6582–6592. [PubMed: 25303977]
9. Ise K, et al. Targeted deletion of the *Hras* gene decreases tumor formation in mouse skin carcinogenesis. Oncogene. 2000; 19:2951–2956. [PubMed: 10871846]
10. Bos JL. *Ras* oncogenes in human cancer: a review. Cancer Res. 1989; 49:4682–4689. [PubMed: 2547513]
11. Balmain A, Pragnell I. Mouse skin carcinomas induced *in vivo* by chemical carcinogens have a transforming Harvey-*ras* oncogene. Nature. 1983; 303:72–74. [PubMed: 6843661]

12. Finch JS, Albino HE, Bowden GT. Quantitation of early clonal expansion of two mutant 61st codon c-Ha-*ras* alleles in DMBA/TPA-treated mouse skin by nested PCR/RFLP. Carcinogenesis. 1996; 17:2551–2557. [PubMed: 9006088]
13. Loehrke H, et al. On the persistence of tumor initiation in two-stage carcinogenesis on mouse skin. Carcinogenesis. 1983; 4:771–775. [PubMed: 6407773]
14. Nagase H, Mao J, Balmain A. Allele-specific *Hras* mutations and genetic alterations at tumor susceptibility loci in skin carcinomas from interspecific hybrid mice. Cancer Res. 2003; 63:4849–4853. [PubMed: 12941805]
15. Chen B, You L, Wang Y, Stoner DG, You M. Allele-specific activation and expression of the *Kras* gene in hybrid mouse lung tumors induced by chemical carcinogens. Carcinogenesis. 1994; 15:2031–2035. [PubMed: 7923598]
16. Samuels Y, et al. High frequency of mutations of the *PIK3CA* gene in human cancers. Science. 2004; 304:554. [PubMed: 15016963]
17. Lawrence MS, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013; 499:214–218. [PubMed: 23770567]
18. Alexandrov LB, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500:415–421. [PubMed: 23945592]
19. Campbell PJ, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature. 2010; 467:1109–1113. [PubMed: 20981101]
20. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol. 2002; 3:991–998. [PubMed: 12407406]
21. Giancotti FG. Mechanisms governing metastatic dormancy and reactivation. Cell. 2013; 155:750–764. [PubMed: 24209616]
22. Lin Z, et al. LMP1 regulates periodontal ligament progenitor cell proliferation and differentiation. Bone. 2010; 47:55–64. [PubMed: 20348040]
23. Sadej R, Grudowska A, Turczyk L, Kordek R, Romanska HM. CD151 in cancer progression and metastasis: a complex scenario. Lab Invest. 2014; 94:41–51. [PubMed: 24247563]
24. Takeda Y, et al. Diminished metastasis in tetraspanin CD151–knockout mice. Blood. 2011; 118:464–472. [PubMed: 21536858]
25. Quigley DA, et al. Network analysis of skin tumor progression identifies a rewired genetic architecture affecting inflammation and tumor susceptibility. Genome Biol. 2011; 12:R5. [PubMed: 21244661]
26. Cerami E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2:401–404. [PubMed: 22588877]
27. Gao J, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6:pl1. [PubMed: 23550210]
28. Tsai JH, Donaher JL, Murphy DA, Chau S, Yang J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell. 2012; 22:725–736. [PubMed: 23201165]
29. The Cancer Genome Atlas. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012; 489:519–525. [PubMed: 22960745]
30. Aldaz CM, Trono D, Larcher F, Slaga T, Conti C. Sequential trisomization of chromosomes 6 and 7 in mouse skin premalignant lesions. Mol Carcinog. 1989; 2:22–26. [PubMed: 2499343]
31. Namiki T, et al. AMP kinase–related kinase NUAK2 affects tumor growth, migration and clinical outcome of human melanoma. Proc Natl Acad Sci USA. 2011; 108:6597–6602. [PubMed: 21460252]
32. The Cancer Genome Atlas. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014; 511:543–550. [PubMed: 25079552]
33. Linardopoulos S, et al. Deletion and altered regulation of $p16^{INK4a}$ and $p15^{INK4b}$ in undifferentiated mouse skin tumors. Cancer Res. 1995; 55:5168–5172. [PubMed: 7585567]
34. Huber KVM, et al. Stereospecific targeting of MTH1 by (*S*)-crizotinib as an anticancer strategy. Nature. 2014; 508:222–227. [PubMed: 24695225]

35. Gad H, et al. MTH1 inhibition eradicates cancer by preventing sanitation of the dNTP pool. Nature. 2014; 508:215–221. [PubMed: 24695224]
36. Westcott PMK, et al. The mutational landscapes of genetic and chemical models of *Kras*-driven lung cancer. Nature. 2015; 517:489–492. [PubMed: 25363767]
37. McFadden DG, et al. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. Cell. 2014; 156:1298–1311. [PubMed: 24630729]
38. Klein CA. Selection and adaptation during metastatic cancer progression. Nature. 2013; 501:365–372. [PubMed: 24048069]
39. Peinado H, Lavotshkin S, Lyden D. The secreted factors responsible for pre-metastatic niche formation: old sayings and new thoughts. Semin Cancer Biol. 2011; 21:139–146. [PubMed: 21251983]
40. Cox TR, et al. The hypoxic cancer secretome induces pre-metastatic bone lesions through lysyl oxidase. Nature. 2015; 522:106–110. [PubMed: 26017313]
41. Mittal D, Gubin MM, Schreiber RD, Smyth MJ. New insights into cancer immunoediting and its three component phases—elimination, equilibrium and escape. Curr Opin Immunol. 2014; 27:16–25. [PubMed: 24531241]
42. Nassar D, Latil M, Boeckx B, Lambrechts D, Blanpain C. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. Nat Med. 2015; 21:946–954. [PubMed: 26168291]
43. Engel J, Emeny RT, Hölzel D. Positive lymph nodes do not metastasize. Cancer Metastasis Rev. 2012; 31:235–246. [PubMed: 22198520]
44. Fisher B, et al. Twenty-five–year follow-up of a randomized trial comparing radical mastectomy, total mastectomy and total mastectomy followed by irradiation. N Engl J Med. 2002; 347:567–575. [PubMed: 12192016]
45. Veronesi U, Marubini E, Mariani L, Valagussa P, Zucali R. The dissection of internal mammary nodes does not improve the survival of breast cancer patients. 30-year results of a randomised trial. Eur J Cancer. 1999; 35:1320–1325. [PubMed: 10658521]
46. Varela I, et al. Exome sequencing identifies frequent mutaiton of the SWI/SNF complex gene *PBRM1* in renal carcinoma. Nature. 2011; 469:539–542. [PubMed: 21248752]
47. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]
48. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43:491–498. [PubMed: 21478889]
49. Cibulskis K, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013; 31:213–219. [PubMed: 23396013]
50. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]
51. Forbes SA, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 2015; 43:D805–D811. [PubMed: 25355519]
52. Vogelstein B, et al. Cancer genome landscapes. Science. 2013; 339:1546–1558. [PubMed: 23539594]
53. Talevich E, Shain AH, Bastian BC. CNVkit: Copy number detection and viziualization for targeted sequencing using off-target reads. BioRxiv. 2014
54. Untergasser A, et al. Primer3—new capabilities and interfaces. Nucleic Acids Res. 2012; 40:e115. [PubMed: 22730293]
55. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics. 2007; 23:1289–1291. [PubMed: 17379693]
56. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17:10.
57. Li H, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]
58. Roth A, et al. PyClone: statistical inference of clonal population structure in cancer. Nat Methods. 2014; 11:396–398. [PubMed: 24633410]

59. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics. 2010; 26:2363–2367. [PubMed: 20688976]

60. Quigley D. Equalizer reduces SNP bias in Affymetrix microarrays. BMC Bioinformatics. 2015; 16:238. [PubMed: 26223252]

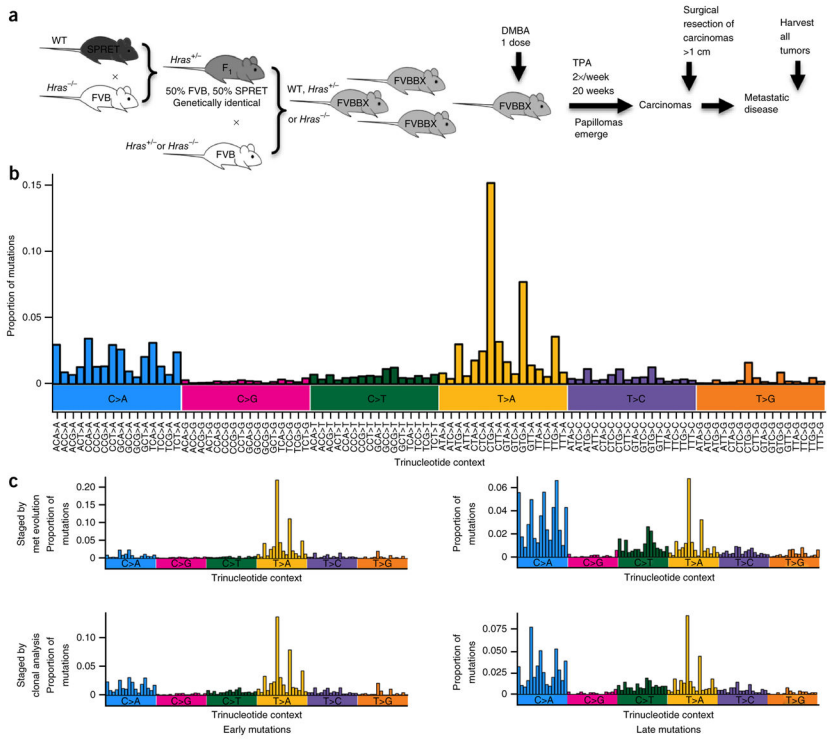

Figure 1.

Chemically induced tumors carry a mutation signature of the carcinogen DMBA. (**a**) Scheme for the generation of genetically unique backcrossed (FVBBX) mice and of metastases. FVB/N *Hras*$^{-/-}$mice were crossed with SPRET/EiJ mice and the resulting offspring were crossed to inbred wild-type (WT), *Hras*$^{+/-}$ or *Hras*$^{-/-}$ FVB/N mice to generate a backcross population with all three *Hras* genotypes. Tumors were induced in FVBBX mice, carcinomas were resected upon reaching >1 cm in the longest diameter, and mice were allowed to progress to metastatic disease. (**b**) Frequencies of mutations observed in each of 96 possible trinucleotide mutation contexts for all of the mutations across all tumors. Trinucleotide contexts, arranged on the *x* axis, are grouped by the base pair change of the mutation. Peaks are observed at two specific contexts, CTG to CAG and GTG to GAG. The same pattern is also observed when only nonsynonymous mutations are considered (data not shown). (**c**) Frequencies of 'early mutations' (left graphs) or 'late mutations' (right graphs) in metastases, as classified by two strategies. Top, mutations were classified by whether they are shared with the ancestral primary tumor (early, upper left) or not shared (late, upper right). Bottom, mutations in tumors of all stages were classified on the basis of whether they were fully clonal (i.e., present in all tumor cells; early; lower left) or subclonal (i.e., present only in a fraction of tumor cells; late; lower right). Early mutations show an enrichment of A>T mutations (equivalent to T>A mutations on the opposite strand), and late mutations show a higher proportion of G>T mutations (equivalent to C>A mutations on the opposite strand).

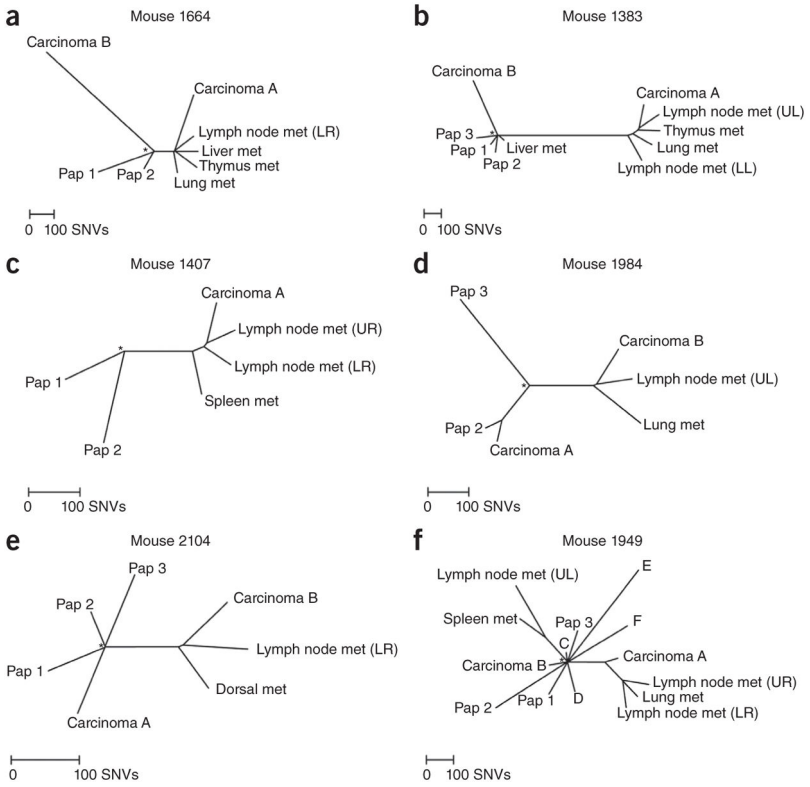

Figure 2.

Phylogenetic trees reveal evolutionary relationship between tumors. (**a**) In mouse 1664, carcinoma A had four metastases (met) that diverged early (sharing an average of 46% of mutations; 83 mutations), whereas carcinoma B did not form any detectable metastases. Distant metastases to the lung and thymus did not show evidence of disseminating from a lymph node, but rather appeared to diverge from the primary carcinoma at approximately the same time as the lymph node metastases. The asterisk (*) denotes normal tissue (root of tree). Pap 1 and Pap 2 are independently arising papillomas, and LR is a lower right lymph node metastasis (under the rear right limb). (**b**) In mouse 1383, carcinoma A had four metastases that diverged relatively late and approximately synchronously, and which shared an average of 88% of their mutations with the primary carcinoma. UL and LL designate lymph node metastases under the upper (front) left and lower (rear) left limbs, respectively. (**c**) In mouse 1407, two lymph node metastases departed synchronously from carcinoma A and a spleen metastasis diverged slightly earlier than the lymph node metastases. The primary carcinoma and its metastases shared 153 mutations, corresponding to an average of 69% of their total mutations. UR designates a lymph node metastasis at the upper (front) right limb. (**d**) In mouse 1984, a lymph node metastasis and a lung metastasis diverged from carcinoma B synchronously, sharing 160 mutations with the primary carcinoma. (**e**) In mouse 2104, a lymph node metastasis and a dorsal metastasis shared 118 mutations with carcinoma B, indicating that the metastases diverged synchronously. (**f**) In mouse 1949, two lymph node metastases and a lung metastasis shared 167 mutations with carcinoma A (average of 52% of total mutations) and an additional 90 mutations with one another, suggesting that one of the metastases may have given rise to the others. This mouse provides

the only counter-example to parallel evolution that we observed. In all panels, scale bars represent 100 SNVs.

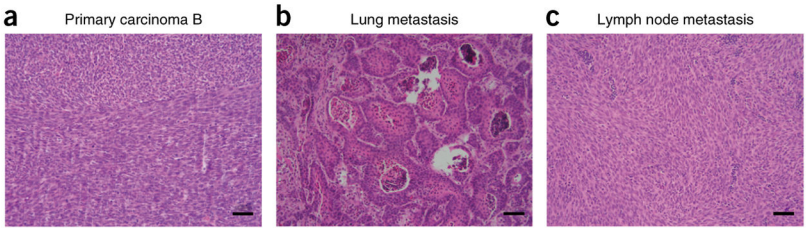

Figure 3.

Histological analysis of primary tumors and metastases from mouse 1984. (**a**–**c**) H&E-stained samples of primary carcinoma B showing spindle tumor cells arranged in fascicles (**a**), lung metastasis arising from carcinoma B showing squamous histology (**b**) and lymph node metastasis arising from carcinoma B showing spindle histology (**c**). Images are shown at 200× magnification. Scale bars, 50 μm.

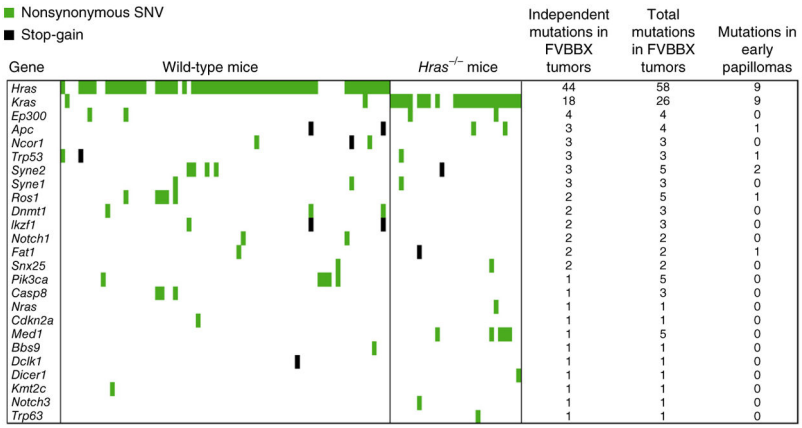

Figure 4.

Mutated genes in carcinogen-induced tumors. Nonsynonymous (green bars) and stop-gain (black bars) mutations in genes observed to be mutated in human cancer. Tumors are arranged on the horizontal axis by mouse genotype (wild type or *Hras*$^{-/-}$). Independent mutations are defined as mutations in tumors that show no phylogenetic relationship (on the basis of shared mutations). Number of mutations observed in early papillomas is also shown.

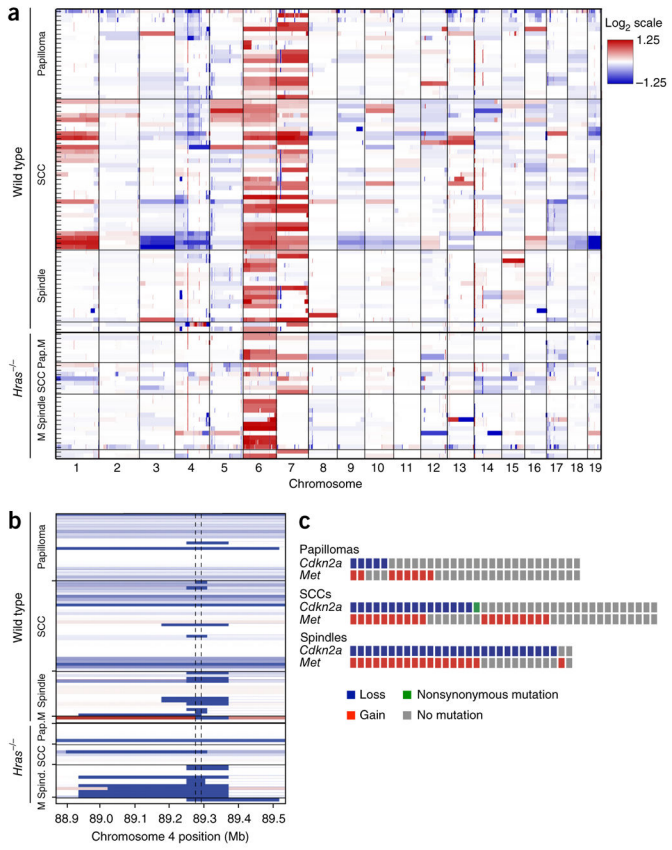

Figure 5.

Gene CNVs in mouse skin tumors. (**a**) CNVs are shown on each chromosome, with *Ras* genotype and tumor morphology shown on the *y* axis. Chromosome numbers are on *x* axis, and chromosomes are arranged with proximal end on the left and distal ends on the right. The red and blue scale represents copy number gains and losses, respectively. (**b**) Focal *Cdkn2a* deletions are visible in many tumors, frequently less than 1 Mb in size. The vertical dashed lines indicate the boundaries of the *Cdkn2a* gene, and the blue horizontal lines show the extent of the deletion. (**c**) *Cdkn2a* losses and *Met* amplifications in samples displayed by tumor type. Co-occurrence of these events increases substantially with tumor progression, from papillomas (7% co-occurrence) to SCCs (25% co-occurrence) to spindle tumors (59% co-occurrence). Each box represents one tumor, and tumors with shared CNVs are arranged vertically. Pap., papilloma; Spind., spindle morphology; M, mixed SCC and spindle morphology.