Published in final edited form as:
Cell. 2012 July 20; 150(2): 251–263. doi:10.1016/j.cell.2012.06.024.

A Landscape of Driver Mutations in Melanoma

Eran Hodis[1,2,19], **Ian R. Watson**[3,10,19], **Gregory V. Kryukov**[1,2,12], **Stefan T. Arold**[4], **Marcin Imielinski**[1], **Jean-Philippe Theurillat**[1], **Elizabeth Nickerson**[1], **Daniel Auclair**[1], **Liren Li**[3,10], **Chelsea Place**[10], **Daniel DiCara**[1], **Alex H. Ramos**[1,2], **Michael S. Lawrence**[1], **Kristian Cibulskis**[1], **Andrey Sivachenko**[1], **Douglas Voet**[1], **Gordon Saksena**[1], **Nicolas Stransky**[1], **Robert C. Onofrio**[1], **Wendy Winckler**[1], **Kristin Ardlie**[1], **Nikhil Wagle**[1,2], **Jennifer Wargo**[13], **Kelly Chong**[14], **Donald L. Morton**[15], **Katherine Stemke-Hale**[5], **Guo Chen**[6], **Michael Noble**[1], **Matthew Meyerson**[1,2,10,11], **John E. Ladbury**[4], **Michael A. Davies**[5,6], **Jeffrey E. Gershenwald**[7,8], **Stephan N. Wagner**[16], **Dave S.B. Hoon**[14], **Dirk Schadendorf**[17], **Eric S. Lander**[1,18], **Stacey B. Gabriel**[1], **Gad Getz**[1], **Levi A. Garraway**[1,2,10,11,20,*], and **Lynda Chin**[1,2,3,9,10,20,*]

[1]The Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

[2]Harvard Medical School, Boston, MA 02115, USA

[3]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[4]Department of Biochemistry and Molecular Biology and Center for Biomolecular Structure and Function, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[5]Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[6]Department of Melanoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[7]Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[8]Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[9]Institute for Applied Cancer Science, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[10]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[11]Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[12]Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA

[*]Correspondence: levi_garraway@dfci.harvard.edu (L.A.G.), lchin@mdanderson.org (L.C.).

[19]These authors contributed equally to this work

[20]These authors contributed equally to this work

ACCESSION NUMBERS

The dbGaP accession number for the exome sequence data reported in this paper is phs000452.v1.p1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and thirteen tables.

[13]Department of Surgery, Massachusetts General Hospital, Boston, MA 02114, USA

[14]Department of Molecular Oncology, John Wayne Cancer Institute, Santa Monica, CA 90404, USA

[15]Melanoma Department, John Wayne Cancer Institute, Santa Monica, CA 90404, USA

[16]Division of Immunology, Allergy and Infectious Diseases, Department of Dermatology, Medical University of Vienna and CeMM-Research, Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria

[17]Department of Dermatology, University Hospital Essen, 45122 Essen, Germany

[18]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

SUMMARY

Despite recent insights into melanoma genetics, systematic surveys for driver mutations are challenged by an abundance of passenger mutations caused by carcinogenic ultraviolet (UV) light exposure. We developed a permutation-based framework to address this challenge, employing mutation data from intronic sequences to control for passenger mutational load on a per gene basis. Analysis of large-scale melanoma exome data by this approach discovered six novel melanoma genes (*PPP6C*, *RAC1*, *SNX31*, *TACC1*, *STK19* and *ARID2*), three of which - *RAC1*, *PPP6C* and *STK19* - harbored recurrent and potentially targetable mutations. Integration with chromosomal copy number data contextualized the landscape of driver mutations, providing oncogenic insights in BRAF- and NRAS-driven melanoma as well as those without known *NRAS*/*BRAF* mutations. The landscape also clarified a mutational basis for RB and p53 pathway deregulation in this malignancy. Finally, the spectrum of driver mutations provided unequivocal genomic evidence for a direct mutagenic role of UV light in melanoma pathogenesis.

INTRODUCTION

In recent years, much has been learned about the molecular basis of melanoma genesis, progression, and response to therapy. *BRAF* V600 mutations (present in 50% of melanomas) predict clinical efficacy of RAF inhibitors such as vemurafenib; activating KIT aberrations may predict response to tyrosine kinase inhibitors such as imatinib, nilotinib, or dasatinib; and some NRAS-mutant tumors may exhibit sensitivity to MEK inhibition (reviewed in (Flaherty et al., 2012)). Other melanoma gene mutations that offer therapeutic insights include *CDNK2A* deletions, *MITF* amplification/alteration resulting in dysregulation of “druggable” anti-apoptotic proteins, and *PTEN* disruption leading to PI3 kinase/AKT activation (reviewed in (Chin et al., 2006)). The continuing discovery of recurrently mutated melanoma genes (Berger et al., 2012; Nikolaev et al., 2012; Stark et al., 2012; Wei et al., 2011), and the lack of identified driver mutations in the subtype without *NRAS* or *BRAF* mutation, suggests that genetic understanding of this malignancy remains incomplete.

While the potential of comprehensive genome sequencing for melanoma gene discovery is recognized, there is also increasing appreciation for the confounding impact of high mutational load due to UV mutagenesis. In particular, cutaneous melanomas exhibit markedly elevated base mutation rates compared to nearly all other solid tumors (Berger et al., 2012; Pleasance et al., 2010), which is almost entirely attributable to increased abundance of the cytidine to thymidine (C>T) transitions characteristic of an ultraviolet UV-light-induced mutational signature. Highly elevated somatic mutation rates that vary across genomic loci may limit the ability of statistical approaches that assume uniformity of the

basal mutation rate to distinguish genes harboring 'driver' mutations (i.e., mutations that confer or at some point conferred a fitness advantage to the tumor cell) from those with 'passenger' mutations (i.e., mutations that never conferred a fitness advantage). Although methods to account for this mutation rate heterogeneity are an active area of research (Chapman et al., 2011; Greenman et al., 2006; Lohr et al., 2012), rigorous approaches to address this challenge in melanoma have been lacking.

A related question pertains to the tumorigenic effects of UV-induced DNA damage at the nucleotide level. Epidemiological and experimental data have established a causal role for intense UV exposure during development (e.g., blistering sunburns early in life) in melanoma genesis (reviewed in (Garibyan and Fisher, 2010)). Several model systems have also linked UV-dependent tumorigenic effects to modulation of signaling pathways (e.g., enhanced gamma interferon secretion; (Zaidi et al., 2011); activation of JNK signaling pathway (Derijard et al., 1994)), thus supporting a non-mutagenic role in melanoma. Conversely, evidence for a direct UV mutagenic effect in melanoma pathogenesis has been more equivocal. For example, the recurrent base mutations that produce oncogenic NRAS and BRAF mutations are not C>T transitions indicative of UV mutagenesis. Definitive resolution of this question requires demonstration of driver mutations that are directly attributable to UV-induced damage in melanoma.

To analyze whole-exome sequencing data from 121 melanoma tumor/normal pairs, we have employed a statistical approach that infers positive selection at each gene locus based on exon/intron mutational distributions, as well as the predicted functional impact of each mutation. This approach enabled both discovery of several new cancer genes with functionally consequential (and plausibly actionable) mutations and identification of numerous driver mutations directly attributable to UV mutagenesis. In the aggregate, these results offer a comprehensive view of the landscape of driver coding mutations in human melanoma.

RESULTS

Identification of melanoma coding mutations by whole exome sequencing

Solution-phase hybrid capture and whole-exome sequencing were performed on paired tumor and normal genomic DNA obtained from 135 patients with melanoma (Table S1, S2). 103-fold mean target coverage was achieved, with 87% of bases covered at least 14-fold in the tumor and 8-fold in the normal – the threshold which offers 80% power to detect mutations with an allelic fraction of 0.3 (Carter et al., 2012). A set of 121 tumor/normal pairs (15 primary tumors, 30 metastatic samples and 76 short-term cultures derived from metastatic tumor tissue (Table S1)) were qualified for analysis. Altogether, this sample collection comprised 95 melanomas of cutaneous, 5 of acral, 2 of mucosal, 1 of uveal and 18 of unknown primary origin. Somatic copy-number aberration profiles identified expected melanoma alterations (Curtin et al., 2005; Lin et al., 2008), including gains of *MITF, TERT* and *CCND1* and deletions of *CDKN2A* and *PTEN,* among others (Figure S1A, Table S3).

Across all samples, 86,813 coding mutations were detected at a 2:1 ratio of nonsynonymous to synonymous events, consistent with a high passenger mutation load (Table S4). The median sample mutation rate was 14.4 coding mutations per megabase (lower-upper quartile range: 8.0–24.9). As expected, this rate was higher than that reported for any other tumor type, including lung cancer (Lee et al., 2010; Pasqualucci et al., 2011; Cancer Genome Atlas Research Network, 2011), and a signature of UV mutagenesis predominated (median YC>YT mutations: 82.2%; lower-upper quartile range: 73.4%–86.5%). Accordingly, 13,905 genes harbored a nonsilent mutation in at least one tumor, 9,782 genes were thus mutated in two or more tumors, 515 genes were mutated in >10% of tumors, and 78 genes were

mutated in >20% of tumors (Figure S1B). In genes mutated in >10% and >20% of samples, 85.5% and 85.2% of nonsilent coding mutations resulted from YC>YT transitions, respectively, suggesting that many high-frequency melanoma gene mutations may derive from UV-associated passenger events.

We next sought to identify genes showing statistical evidence for positive selection for nonsilent mutations, which is a challenging task in the context of melanoma's high and heterogeneous basal mutation rate. To illustrate the problem introduced by regional heterogeneity in basal mutation rates, we defined significantly mutated genes by using a standard analytical approach that assumes a uniform basal mutation rate across the exome (controlling for trinucleotide context), as published previously (Ding et al., 2008; Getz et al., 2007; Kan et al., 2010; Stransky et al., 2011; Cancer Genome Atlas Research Network, 2011). This analysis produced a long list of genes ($n = 544$) with nonsilent mutation frequencies exceeding the exomic average, thus considered "significantly" mutated (Figure S1C). Many of these genes showed high silent mutation rates (correlation of significance rank with silent mutation rate: $R = 0.29$, $p < 2.2 \times 10^{-16}$, Pearson's; Figure S1D), suggesting locally elevated basal mutation rates. Furthermore, numerous genes were found to have minimal expression levels in melanoma based on cross-cohort analysis of published RNA sequencing (RNA-seq) data (correlation of significance rank with expression level: $R = -0.08$ and $p = 4.4 \times 10^{-16}$) (Berger et al., 2010), which is consistent with published studies showing that genes with lower expression levels tend to harbor increased somatic mutation rates (Chapman et al., 2011; Pleasance et al., 2010). We also observed an anticorrelation between gene expression and mutation rate in our data ($R = -0.10$ and $p = 4.4 \times 10^{-16}$). Together, these results highlighted the challenge of detecting positive selection in the setting of variable basal mutation rates. Such loci may accumulate frequent somatic mutations unrelated to positive selection but were nonetheless deemed significant by statistical approaches that assume uniform basal mutation rates. Conversely, genes present in loci with low basal mutation rates may accumulate few mutations. Here, evidence of positive selection will only become apparent after accounting for this reduced mutation rate. The high mutational burden linked to UV exposure further exacerbates this problem in melanoma by making heterogeneity in locus-specific mutation rates even more pronounced.

Systematic inference of positive selection at putative melanoma gene loci

To more accurately ascertain positive selection in melanoma genomes, we leveraged sequence data from flanking intronic regions and other untranslated (UTR) DNA segments that are captured alongside exonic targets during hybrid selection to define the local base mutation rate. We reasoned that any DNA sequence situated immediately adjacent to an exon is likely subjected to similar mutagenic and repair processes as the exonic sequence. Indeed, gene-specific intronic and exonic mutation rates correlated in our dataset across several orders of magnitude (R=0.35, $p < 2.2 \times 10^{-16}$; Figure S1E). However, unlike nonsilent mutations in their exonic counterparts, mutations in intronic and UTR sequences are more likely to exist under neutral selective pressure. Thus, mutation data obtained from these flanking regions should offer a means by which locus-specific mutation rates might be inferred.

A gene that contains a high frequency of nonsilent exonic mutations and a low frequency of synonymous or intronic/UTR mutations exhibits presumptive evidence of positive selection during tumor evolution. Such a mutational pattern may signify the presence of *bona fide* driver mutations in melanoma. On the other hand, a gene with a high frequency of both nonsilent exonic mutations and synonymous, intronic, and/or UTR mutations is less likely to contain mutations that experienced positive selection during tumorigenesis (Figure 1A). Based on these principles, a null model of the distribution of all mutations across the gene (exon, intron and UTR) may be generated by random permutation of the locations of all

observed mutations (Figure 1B). This null model is computed per-sample, as the locus basal mutation rate may vary across samples. A permutation test may then be performed to assess the statistical significance of any set-wide observation in the gene compared to the null model generated from all individual sample null models.

Employing this framework, we assessed the statistical significance of the set-wide 'functional mutation burden': the number of samples harboring a nonsilent mutation of predicted functional consequence (Adzhubei et al., 2010). Eleven genes were found to harbor a statistically significant functional mutation burden ($q \leq 0.2$, Benjamini-Hochberg (Benjamini, 1995)) (Figure 1C; Table S5; Figure S1F). These included six well-known cancer genes (*BRAF, NRAS, PTEN, TP53, p16INK4a* [transcript of the *CDKN2A* gene locus], and *MAP2K1*) and five new candidates (*PPP6C, RAC1, SNX31, TACC1*, and *STK19*). Manual review and mass spectrometric genotyping of observed mutations at these loci confirmed high nonsynonymous:synonymous mutation ratios and low rates of silent mutations (Figure 1C; Table S5, S6). Contrary to the output from the initial significance analysis (described above), an overall bias toward lowly expressed genes was no longer evident ($R=-0.04$, $p=6.2 \times 10^{-6}$). As a control, we performed an analogous assessment of the 'synonymous mutation burden', and no genes were identified as statistically significant by this analysis (Figure 1C; Table S7). Thus, the incorporation of exonic and nonexonic mutational data identified multiple loci that showed evidence of positive selection and hence may contain genes that harbor driver mutations.

Novel melanoma genes are linked to known cancer relevant pathways

The five novel candidate genes harboring putative somatic driver mutations (*PPP6C, RAC1, SNX31, TACC1* and *STK19)* had not previously been recognized as significantly mutated in melanoma. *PPP6C* encodes for the catalytic subunit of the heterotrimeric PP6 protein phosphatase complex (Stefansson et al., 2008). Reports have implicated PPP6C as a tumor suppressor due to its role in regulation of cell cycle and mitosis. PP6 negatively regulates levels of the melanoma oncogene *CCND1* during G1 phase of the cell cycle (Stefansson and Brautigan, 2007) and is the major T-loop phosphatase for the mitotic kinase, Aurora A (AurA), amplified in a number of human cancers (Lens et al., 2010; Zeng et al., 2010). In the discovery set of 121 tumor/normal pairs, 11 melanomas (9%) harbored nonsynonymous *PPP6C* mutations, 10 of which were predicted to be homozygous events based on high mutant allele frequencies. 60% of these *PPP6C* mutations clustered within a 12 amino acid region flanking an arginine at codon 264 (four R264C mutations, two S270L mutations and one P259S mutation; Figure 2A). When mapped onto the structure of the PP2A catalytic subunit (~60% sequence homology to PPP6C; Figure 2A, Figure S2A), the PPP6C mutations localized to highly conserved regions. In particular, R264 participates in multiple salt bridge interactions at the interface between the catalytic and regulatory subunits (Cho and Xu, 2007). The *PPP6C* R264C mutation was found at a frequency of 3% in an extension set of 63 melanoma samples (Table S8). Homozygous hot spot (defined as same amino acid change in >3% of samples in the discovery set of 121) mutations in R264 and nearby residues may result in altered interactions between the PPP6C catalytic subunit and its regulatory partners. The clustered mutation pattern and relative paucity of nonsense mutations or frame-shift indels are characteristic of gain-of-function mutations suggesting that dysregulation of this protein phosphatase's function may contribute to melanoma biology.

Mutations in *STK19* (a predicted kinase with unknown function) exhibited a hot spot pattern in melanoma. Six nonsynonymous *STK19* mutations were detected in five tumors (4%) in the discovery set (Figure 2B), four of which were located at D89 (D89N) with an immediately adjacent additional mutation (P90L). D89N mutation showed a consistent frequency (5%) in a melanoma extension set of 59 tumors (Table S8). The pattern and

significance of its somatic hot spot point mutations are strong genomic evidence in support of *STK19* as a novel cancer gene.

In contrast, *TACC1* and *SNX31* exhibited a distributed pattern of mutational events. *SNX31* encodes the poorly characterized protein sorting nexin 31. Mutations tended to occur within the protein and lipid interaction band4.1/ezrin/radixin/moesin (FERM)-like domain of SNX31 (Figure 2C), with one mutation in the domain occurring in two separate melanoma cases and over 60% of nonsilent mutations occurring in a 48-residue window in this 440-residue protein. SNX31 has been reported to bind active guanosine triphosphate (GTP)-loaded H-Ras, but not inactive guanosine diphosphate (GDP)-loaded H-Ras likely through its FERM-like domain (Ghai et al., 2011), suggesting a potential role for SNX31 as a Ras effector protein.

TACC1, encoding transforming acidic coiled-coil protein 1, has been reported to stimulate the Ras and PI3K pathways and to promote transformation in mice upon overexpression (Cully et al., 2005). *TACC1* is mutated in eight melanomas (7%) in the discovery set, with mutations occurring predominately near the C terminus of the protein, in or near the conserved TACC domain (Figure 2D). TACC1 is known to interact with AurA (Conte et al., 2003; Delaval et al., 2004), which is notable in the context of PPP6C's function as an AurA phosphatase (Zeng et al., 2010).

Finally, 5% of discovery set melanomas harbored nonsilent mutations in *RAC1*, a RAS-related member of the Rho subfamily of GTPases. RAC1 functions as a molecular switch, cycling between active GTP-bound and inactive GDP-bound states through large conformation changes near the nucleotide-binding site, localized to the switch I and II regions. Its best-characterized function is regulation of cytoskeleton rearrangement, and thus it plays important roles in cellular adhesion, migration, and invasion (Jaffe and Hall, 2005). Overexpression has been reported in a number of malignancies (Karlsson et al., 2009). *RAC1* mutations in our melanomas exhibited a hot spot pattern, with all six mutations effecting the same nucleotide change (Figure 3A). This c.85C > T transition, resulting in a P29S amino acid change, is the most frequent hot spot mutation after those in *BRAF* and *NRAS* (Table S9). Including verification data from two independent extension sets ($n = 59$ and $n = 175$), the prevalence of *RAC1* P29S hot spot mutation in melanoma was validated to be 3.9% (14/355 patients; Tables S8 and S10). In addition, mutations in homologous residues in *RAC2* (P29L) and *RHOT1* (P30L) were also found (Figure 3B), highlighting the importance of the P29 residue as a possible codon targeted by hot spot mutations in Rho family GTPases. We also observed a known RAS family-activating mutation (G12D) (reviewed in Malumbres and Barbacid, 2003) in a gene encoding for another Rho family GTPase member, *CDC42* (Figure 3B). Together, these mutational data implicate the Rho family members as melanoma oncogenes.

RAC1(P29S) mutation is a gain-of-function oncogenic event

To explore possible consequences of the *RAC1* P29S mutation, we conducted homology modeling based on crystal structures of the 97% amino-acid-sequence-identical RAC3 in GDP-bound and GTP/PAK1-bound conformations (Figure S4A, B). In the GDP-bound state, P29 is found in a hydrophobic pocket in switch I, and S29 is predicted to be energetically less favorable due to its lack of shape complementarity, reduced hydrophobicity and unfavorable proximity of the serine hydroxyl oxygen to adjacent hydrophobic residues (Figure 4A, bottom left and right panels). In the GTP-bound state, the packing of the switch 1 loop is less compact (Fig. 4B, top left and right panels). The energetic advantage of having the wild-type P29 rather than the mutant S29 is therefore lost. Conversely, S29 is predicted to engage in hydrogen bonds with the polar side and main chains of E31, which would stabilize the GTP bound form (Figure 4B, bottom left and right

panels). Furthermore, the P29S mutant is predicted to gain more entropy upon transitioning from the GDP to the GTP-bound form than wild-type, because in the GDP-bound state switch 1 is tethered to the protein core, whereas in the GTP-bound state, switch 1 flexibility is restricted by P29 (Figure S4C). These observations suggested that p.P29S likely destabilizes RAC1's inactive GDP-bound state and favors its active GTP-bound state.

Because active, GTP-loaded RAC1 is known to interact with the p21-binding domain (PBD) of p21-activated protein kinase 1 (PAK1) to regulate downstream events relevant for tumorigenesis, PAK1 PBD pull-down assays can be employed to measure GTP-bound RAC1. In HEK 293FT cells, PAK1 PBD pull-down revealed a significantly higher fraction of RAC1(P29S) in the GTP-loaded active state when compared to wild-type (Figure 4C, compare lanes 2 and 3). As expected, a constitutively active RAC1(Q61L) mutant was found in a robust GTP-loaded fraction (Figure 4C, compare lanes 2 to 4 and 5). In the presence of exogenous GDP, RAC1(P29S) demonstrated an attenuated shift to the inactive, GDP-bound form, which was in accordance with the structural prediction (Figure 4D, compare lanes 1 and 2 to 4 and 5). Importantly, the increase in GTP-loaded RAC1(P29S) was also evident in immortalized human melanocytes stably expressing oncogenic NRAS or BRAF (Figure 4E and F). Together, the biochemical and structural results support the conclusion that the *RAC1* P29S mutation is activating, rendering RAC1 preferentially in an active, GTP-bound state.

Predicted loss-of-function melanoma gene mutations

Mutations in putative tumor suppressor genes that result in protein truncation may carry a higher likelihood of conferring a fitness advantage to the tumor cell compared to the effect of missense mutations in the same genes. As the permutation-based framework described above modeled basal mutation rates without regard to functional consequence of mutations, we next employed it to detect genes with a higher 'loss-of-function (LoF) mutation burden' than expected by chance. LoF mutations were defined as nonsense, splice site and frame-shift events. Both $p16^{INK4a}$ and *ARID2* showed statistically significant LoF burden ($q \leq 0.2$; Table S11), with $p16^{INK4a}$ LoF mutations in 14 discovery samples (12%) and *ARID2* LoF mutations in 9 samples (7%).

All nonsense mutations in *ARID2*, which encoded a component of the SWI/SNF chromatin-remodeling complex, were predicted to yield truncated variants lacking the C2H2 Zn-finger motifs required for DNA binding (5% of samples) (Figure 5A), which is reminiscent of the inactivating *ARID2* mutations found in hepatitis-C-virus-associated hepatocellular carcinomas (Li et al., 2011). Although *ARID2* has not been previously identified as significantly mutated in melanoma, singleton mutations—all nonsense events—have been reported in three studies (Nikolaev et al., 2012; Stark et al., 2012; Wei et al., 2011). A targeted search for LoF mutations in other components of the SWI/SNF complex identified three nonsense mutations in *ARID1B* (a gene that also had a significant LoF burden, though it did not pass correction for multiple hypothesis testing in our discovery set), three in *ARID1A*, and one in *SMARCA4* (Figure 5B). Thus, 13% (16/121) of the discovery samples harbored a LoF mutation in a component of the SWI/SNF complex, suggesting a role for dysregulation of chromatin remodeling in melanomagenesis.

A landscape of driver mutations in melanoma

The identification of known and novel drivers in this study provided a global view of melanoma gene mutations. By cross-referencing all observed mutations to recurrently mutated base pairs ($n \geq 20$) reported in the COSMIC database (Forbes et al., 2011) we augmented this view with rare driver events whose low frequency precluded statistical identification. This identified driver mutations in *CTNNB1*, *PIK3CA*, $p14^{ARF}$ (alternative

transcript of the *CDKN2A* gene locus), *EZH2*, *IDH1*, *GNA11*, *KIT*, *HRAS* and *WT1* (Figure 6A; Table S12). To provide a fuller context to the landscape, focal amplifications or deletions of signature melanoma genes, such as amplifications in *CCND1*, *KIT*, *CDK4* and *TERT* and deletions in *CDKN2A* and *PTEN*, were delineated in the same set of samples.

Integrating these mutational and copy number data, we mapped the spectrum of driver genes in Figure 6A. As expected, 83% (100/121) of melanoma samples harbored either a hot spot or a COSMIC-recurrent mutation (referred to hereafter as 'highly recurrent' mutations) in *NRAS* (n=27) or *BRAF* (n=73) in a mutually exclusive fashion ($p = 3 \times 10^{-14}$, Fisher's exact). The two cases with co-occurring *BRAF* and *NRAS* mutations harbored either a non-V600 BRAF mutation together with an oncogenic NRAS mutation, or an NRAS mutation not known to be oncogenic together with an activating BRAF mutation. Nearly 44% (32/73) of melanomas with highly recurrent mutations in *BRAF* harbored a *PTEN* mutation or focal deletion, conversely *PTEN* was altered in only 4% (1/27) of melanomas with highly recurrent mutations in *NRAS* ($p = 4.9 \times 10-5$) (Figure S6A). Significance of these mutational patterns was confirmed by a pairwise search across all genes in Figure 6A ($q \leq 0.2$; Table S13).

The melanoma discovery set included 21 tumors without highly recurrent mutations in either *BRAF* or *NRAS* ("*BRAF*/*NRAS* wild-type" samples) (Figure 6B). A search for genes mutated in at least 25% of these samples and ranked among the top 50 genes by functional mutation burden identified *NF1*. A significant enrichment of *NF1* mutations was observed in this subset; putative loss-of-function *NF1* mutations occurred in 5 of 21 of these tumors (25%) compared to 2 of the remaining 100 samples (2%) ($p = 5.8 \times 10^{-3}$) (Figure 6B). Given the role of NF1 as a negative regulator of RAS signaling (Vigil et al., 2010), these results suggest that *NF1* inactivation may confer aberrant mitogen-activated protein kinase (MAPK) pathway activation in these *BRAF*/*NRAS* wild-type samples. In addition, an activating *HRAS* G13I mutation, an activating *CRAF* E478K mutation (Emuss et al., 2005), and two *MAP2K1* mutations were observed in *BRAF*/*NRAS* wild-type samples that were also *NF1* wild-type (Figure 6B). Of the 13 remaining *BRAF*/*NRAS* wild-type samples, 1 harbored an activating *KIT* V559A mutation, 6 (1 of which was acral and 1 of which was mucosal) showed focal amplification of *KIT*, *CCND1*, and/or CDK4, and 1 (a uveal melanoma) possessed an activating *GNA11* Q209L mutation. Altogether, known melanoma driver events spanned 81% (17/21) of cases that lacked highly recurrent *NRAS* or *BRAF* mutations (Figure 6B), providing a unified view of driver mutations in this subtype of melanomas.

CDKN2A is a well-known melanoma tumor suppressor gene that encodes for two tumor suppressor proteins through alternative splicing: $p16^{INK4a}$, a cyclin-dependent kinase inhibitor that activates RB through negative regulation of CDK4, and $p14^{ARF}$, which activates p53 through inhibition of its major negative regulator, MDM2 (Chin et al., 2006). The *p16*^{INK4a} transcript was mutated in over 20% of our discovery set, with 14 out of 25 mutations being putative LoF events. Coupled with one splice site and two nonsense mutations in *RB1* as well as three R24 activating mutations in *CDK4*, we estimated that the cell cycle checkpoint was deregulated directly through somatic mutations of its core components in at least 24% (29/121) of samples. Most of the melanoma cases harboring p53 mutation (19% in the discovery set) were without concurrent mutation in $p14^{ARF}$ or $p16^{INK4a}$ (Figure S6B). Taken together, these data support the consensus view that genetic pressure to mutate p53 directly in melanoma is reduced due to frequent deletion of the *CDKN2A* locus, and show that p53 mutation is prevalent in a subset of melanoma without $p14^{ARF}$ mutation.

Finally, LoF mutations in members of the SWI/SNF complex, together with COSMIC-recurrent mutations in *EZH2* and *IDH1*, were found in 17% (20/121) of melanomas, providing genomic evidence that chromatin-modifying proteins and epigenetic regulators contribute to melanoma genesis or progression.

The role of UV mutagenesis in the advent of melanoma driver gene mutations

Next, we systematically addressed the direct effect of mis-repair of UV-induced DNA damage as a cause of melanoma driver mutations, namely C>T (by UVB) or G>T (by UVA). Specifically, we assessed the distribution of mutations attributable to UV-induced DNA damage among the driver mutations. Out of the 262 driver mutations in 21 genes defined by our analysis, 46% were caused by C>T (37%) or G>T (9%) mutations characteristic of UVB/UVA–induced mutations. This percentage increased to 67% (103/150) when driver mutations in BRAF or NRAS were excluded.

TP53 possessed the greatest number of total putative UV-induced mutations among mutated melanoma genes identified in this study (Figure 7A), challenging the dogma that often cites its wild-type status as characteristic of human melanomas (Chin et al., 2006; Flaherty et al., 2012). Presumed UV-induced LoF mutations in known melanoma tumor suppressors (*PTEN, p14ARF, p16INK4a*) were also evident. Newly discovered significantly mutated genes *ARID2, PPP6C, SNX31 and TACC1* each had a high fraction of mutations attributed to C>T transitions, suggesting a possible role in UVB-induced melanomagenesis.

The majority of known activating mutations in the MAPK pathway, which include *BRAF* (c.1799T > A encoding V600E) ($n = 63$), *NRAS* (c.182A > T, Q61L and c.182A > G, Q61R) ($n = 16$), KIT (c.1676T > C, V559A) ($n = 1$), and *GNA11* (c.626A > T, Q209L) ($n = 1$), do not appear attributable to direct UV-induced damage (Figure 7A). There are possible exceptions in mutations in *BRAF*, in which all dinucleotide mutations include a C > T transition, including V600E (c.1799–1800TG > AA) ($n = 1$), V600K (c.1798–1799GT > AA) ($n = 7$), V600R (c.1798–1799GT > AG) ($n = 1$), and L597S (c.1789–1790CT > TC) ($n = 1$), that could be attributed to UVB-induced mutagenesis. There are also mutations in *RAS*, including *NRAS* Q61K (c.181C > A and c.180–181AC > CA) ($n = 9$), Q61R (c.181–182CA > AG) ($n = 1$), G12D (c.35G > A) ($n = 1$), and HRAS G13I (c.37–38GG > AT) ($n = 1$), which may result from UVA- and UVB-induced damage.

Four genes, *RAC1, STK19, FBXW7* and *IDH1*, all possessed a relative percentage of C>T mutations above the exome-wide per-sample median (~83%). Notably, the hot spot mutations *PPP6C* R264C, *STK19* D89N, and *RAC1* P29S are each mediated solely by presumptive UVB damage (Figure 7B). Given evidence that P29S renders RAC1 preferentially in GTP-bound form, leading to downstream activation of PAK signaling (Figure 4), our data revealed the first example of a hot-spot-activating mutation in a melanoma gene attributable to direct UVB-mediated damage, providing definitive evidence for UV mutagenesis in melanoma pathogenesis.

DISCUSSION

We described here a permutation-based framework (available for download at http://www.broadinstitute.org/cancer/cga/InVEx) that leverages intron and UTR sequences in a gene locus to control for gene-specific basal mutation rates, a conceptual advance that represents a natural but important evolution of prior works. Pioneering studies have led to increasing appreciation of the confounding effects of variable regional basal mutation rates, motivating refinements such as gene-specific basal mutation rate calculations based on synonymous mutations, binning genes based on expression levels to correct for correlation between expression and mutation rate, and within-gene permutation tests to assess positional

clustering and evolutionary conservation of mutated residues (Chapman et al., 2011; Ding et al., 2008; Greenman et al., 2006; Kan et al., 2010; Lohr et al., 2012). We expect that future research will account for within-gene variation in the basal mutation rate and, with enough data, per-base basal mutation rates can eventually be inferred. Local rate-altering events will also need to be considered, for example somatic rearrangements have recently been reported to elevate the local mutation rate (Nik-Zainal et al., 2012). Our results should motivate refinement of the standard exon-capture bait set to expressly target a portion of intron/UTR segments for use in basal mutation rate modeling. With whole genome sequence data, which fully covers introns and UTRs, our approach can be more robust, offering increased statistical power.

Although we have assessed the significance of a gene's functional mutation burden using PolyPhen-2 (Adzhubei et al., 2010) in this study (as well as LoF mutation burden), other mutation scoring algorithms (Cooper and Shendure, 2011) may too prove useful. Increased cohort sizes (which will emerge in the fullness of time through The Cancer Genome Atlas [TCGA] and other large-scale efforts) will give sufficient power to evaluate the significance of the more naïve 'nonsilent mutation burden', which does not depend on functional prediction scores. More broadly, this methodology of modeling locus-specific basal mutation rates in combination with optional functional weighting can improve the identification of driver mutations in nonexonic genomic regions predicted to experience positive selection, such as conserved regulatory domains.

Although mutation prevalence of the novel melanoma genes identified herein is relatively low, their importance extends beyond melanoma as underscored by cross-tumor relevance and protein family recurrence. For example, *RAC1* P29S mutation has been reported in a head and neck tumor (Stransky et al., 2011) and a breast tumor (Forbes et al., 2011); furthermore, homologous P29 mutations in other Rho family members were observed in melanoma (Figure 3). The appearance of singleton known activating mutations in our cohort, such as those seen in *HRAS, GNA11* and *KIT*, predicts that larger sequencing studies will uncover additional melanoma genes, reaffirming the importance of systematic resequencing in statistically powered sets of human cancers.

Finally, while sun exposure has been shown to be a leading risk factor for melanoma (Garibyan and Fisher, 2010), it has been perplexing that the most prevalent UVB radiation-induced genetic change — the transition of a cytosine to a thymidine, accounting for >70% of nucleotide substitutions — has not been shown to be the molecular basis for known oncogenic mutations in melanoma, including *BRAF* V600E and *NRAS* Q61L/R. The identification of statistically significant hot spot mutations in *RAC1*, *STK19* and *PPP6C* resulting from C>T transitions offers missing genomic evidence linking UVB mutagenesis mechanistically to this malignancy.

EXPERIMENTAL PROCEDURES

Clinical samples

All melanoma samples analyzed in this study were collected and sequenced under Institution Review Board approved protocols (MIT/COUHES # 110700457). The DNeasy Tissue Kit or the QIAmp DNA Mini kit (Qiagen, Valencia, CA) was used to extract genomic DNA from tissues. The Puregene DNA purification kit (Gentra Systems, Minneapolis, MN) was used to extract genomic DNA from short-term cultures. All DNA samples were subjected to quality assessment.

DNA-Library Preparation, Whole-Genome Sequencing, and Assembly

Exome capture and library construction was performed as in (Gnirke et al., 2009), adapted for production-scale exome capture. Libraries were sequenced on Illumina HiSeq 2000 machines, generating 2 × 76 bp paired-end reads. Sequencing data obtained from the Illumina pipeline were processed by the Picard pipeline (http://picard.sourceforge.net/).

High-density SNP Arrays

DNA samples were hybridized to Affymetrix SNP Array 6.0 genome-wide human SNP microarrays (Affymetrix, Santa Clara, CA) and chromosomal copy number segments were determined as described previously (Cancer Genome Atlas Research Network, 2008). A gene was identified as focally amplified/deleted if a segment above absolute value 0.6 of length ≤5Mb intersected the gene. Significantly recurrent amplifications and deletions were identified using GISTIC (Beroukhim et al., 2007).

Exome Quality Assessment

Samples with nonaberrant copy number profiles and normal samples with aberrant copy number profiles were removed from analysis. Each lane from a tumor/normal pair was crosschecked to have the same SNP fingerprint as each other lane from that pair; non-matching lanes were removed from analysis. Cross-contamination was estimated using ContEst (Cibulskis et al., 2011) (Table S1B). Samples with greater than 10% contamination were excluded from further consideration.

Identification of Somatic Substitutions and Indels

Somatic base pair substitutions were identified using MuTect (http://www.broadinstitute.org/cancer/cga/MuTect) and somatic small indels were identified using Indelocator (http://www.broadinstitute.org/cancer/cga/Indelocator), as in previous reports (Stransky et al., 2011). Identified somatic mutations were annotated for effect of the mutation on the protein product using Oncotator, a comprehensive parsing script for mutation annotation (http://www.broadinstitute.org/cancer/cga/Oncotator/). Each of the above algorithms or scripts were executed within the Broad Firehose infrastructure (http://www.broadinstitute.org/cancer/cga/Firehose).

Mutational Significance Assuming Uniform Background Mutation Rate

An initial attempt at mutational significance analysis assuming a uniform background mutation rate was performed using the per-sample version of MutSig described in the supplement of (Getz et al., 2007).

Statistical Determination of Positive Selection for Non-Silent Mutations

For each gene with at least one observed somatic mutation, the observed ‘mutation burden’ score was calculated (see below for three such score definitions). Mutations were permuted randomly across the gene’s covered base pairs, respecting tri-nucleotide context, and the mutation burden score of the randomized instance was calculated. Up to 10^8 random instances were generated and scored. The fraction of mutation burden scores for random instances that were equal to or greater than the observed burden defined the *p*-value.

(1) Functional mutation burden: mutations were weighted with their PolyPhen-2 p-value (Adzhubei et al., 2010). Frame-shift indels, nonsense and splice site mutations, and mutations at a nucleotide mutated ≥5 times in COSMIC (Forbes et al., 2011) were given a weight of 1. The mutation with the largest weight was identified in each sample and the sum of these ‘largest weights’ was defined as the functional mutation burden. (2) Synonymous mutation burden: the number of samples with ≥1 synonymous mutation. (3) Loss-of-

function (LoF) mutation burden: the number of samples with ≥1 nonsense mutation, frame-shift indel, or splice site mutation. (To increase statistical power, we assessed excess LoF mutation burden above 2).

The source code for this method, termed InVEx (for 'Introns Vs Exons'), is available at http://www.broadinstitute.org/cancer/cga/InVEx.

Mutation Validation and Extension

Mass spectrometric genotyping (Sequenom) on melanoma samples and accompanying normal tissue was performed as previously described (Stransky et al., 2011; Thomas et al., 2007). MassEXTEND® primers were designed using MassARRAY® Assay Design Software from Sequenom, Inc. to generate allele-specific products.

Homology Modeling and Structural Analysis

The structural analysis compared wild-type and P29S mutants of both GDP-bound apo-RAC1 and GTP-bound RAC1 in complex with the PAK1 Cdc42/Rac interactive binding (CRIB) domain. Crystallographic models for RAC1 exist for the GTP-bound state (1MH1) and for a particular Zn-bound trimeric version of GDP-RAC1 (2P2L). However, a GTP and PAK1 CRIB–bound crystal structure exists for RAC3 (2QME, 97% identical to RAC1 for all residues included in the crystal structure; Figure S4). GDP-RAC3 has also been crystallized (2G0N). RAC1 and RAC3 structures are highly similar and superimpose with a root mean square distance (rmsd) of 1.1 Å and 0.9 Å for GDP and GTP bound forms, respectively. To nonetheless avoid any influence of local structural distortions due to the Zn-bound trimeric conformation of the GDP-RAC1 structure, a homology model of RAC1 was built based on GDP-RAC3, and this model was compared with a homology model of GTP-RAC1 bound to PAK1 CRIB. Homology models were built using Swiss-Model (Arnold et al., 2006).

Cell Culture

Human primary melanocytes (pMEL/hTERT/CDK4(R24C)/p53DD) expressing either BRAF(V600E) (pMEL-BRAF) or NRAS(G12D) (pMEL-NRAS) have been previously described (Garraway et al., 2005; Scott et al., 2011). HEK 293FT cells were obtained from Life Technologies (Grand Island, NY). All cells were maintained in Dulbecco's modified Eagle's medium (Cellgro, Manassas, VA) in 10% heat inactivated fetal bovine serum (FBS) at 37°C in a humidified 5% CO^2 atmosphere.

Plasmids

pcDNA3-EGFP-RAC1 (wild-type, T17N, and Q61L) were obtained from Addgene (Plasmids 13719, 13720, and 13721) courtesy of Klaus Hahn (Kraynov et al., 2000). pcDNA3-EGFP-RAC1 P29S was generated using QuickChange Lightning Site-Directed Mutagenesis (Stratagene, Santa Clara, CA) according to the manufacturer's instruction.

RAC1 Activation Assay

Equal amounts of pcDNA3-EGFP-RAC1 plasmids were transiently transfected with Lipofectamine 2000 reagent (Invitrogen) and 48 h post-transfection RAC1 activation assay was performed according to the manufacturer's protocol (Cell Biolabs, Inc.). Briefly, cells growing in monolayers were lysed in 10 cm tissue culture plates, cell lysates were cleared by centrifugation, and protein concentrations were determined by DC Protein Assay (BioRad). Lysates were diluted to equal concentrations, and RAC1 pull-down assays were performed with equal amounts of protein using GST fusion proteins containing the p21-binding domain (PBD) of p21-activated protein kinase 1 (PAK1) coupled to glutathione

agarose beads for 1 h. Pull-downs in the presence of exogenous GDP/GTPγS were performed according to manufacturer's instructions followed by Western analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank L. Ambrogio and E. Bevilacuqa for management of sequencing data production. We thank A. McKenna, L. Zou, S.L. Carter, P. Stojanov, P. Lin, L. Lichtenstein, and the rest of the Broad Cancer Genome Analysis group. We thank C.Z. Zhang and C. Johannessen for helpful discussion, as well as the members of the Garraway and Chin labs. This work was supported by the NHGRI Large Scale Sequencing Program; grant U54 HG003067 to the Broad Institute (PI, E.S.L.); the Melanoma Research Alliance; The University of Texas MD Anderson Cancer Center Melanoma Specialized Programs of Research Excellence and Melanoma Informatics, Tissue Resource, and Pathology (core grant P50 CA93459) (PI, J.E.G., M.A.D.); and the NCI Support Grant (CA-16672). I.R.W. is a recipient of the Canadian Institutes of Health Research Fellowship. J.-P.T. is supported by the Swiss National Science Foundation (PASMP3_134379/1). J.E.L. is supported by The G. Harold and Leila Y. Mathers Charitable Foundation. S.N.W. was supported by the FWF-Austrian Science Fund (L590-B12). L.A.G. is supported by the NIH New Innovator Award (DP2OD002750), NCI (R33CA126674), the Melanoma Research Alliance, and the Starr Cancer Consortium. L.C. is supported by NCI RO1 (R01 CA093947), TCGA GDAC (U24 CA143845), and the Melanoma Research Alliance. L.C. is a recipient of the Milestein Innovation Award in Melanoma Research and is a CPRIT Scholar. N.W. is a consultant and shareholder in Foundation Medicine. L.A.G. is an equity holder and consultant in Foundation Medicine, a consultant to Novartis and Millennium/Takeda, and a recipient of a grant from Novartis.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nature methods. 2010; 7:248–249. [PubMed: 20354512]

Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006; 22:195–201. [PubMed: 16301204]

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995; 57:289–300.

Berger MF, Hodis E, Heffernan TP, Deribe YL, Lawrence MS, Protopopov A, Ivanova E, Watson IR, Nickerson E, Ghosh P, et al. Melanoma genome sequencing reveals frequent PREX2 mutations. Nature. 2012; 485:502–506. [PubMed: 22622578]

Berger MF, Levin JZ, Vijayendran K, Sivachenko A, Adiconis X, Maguire J, Johnson LA, Robinson J, Verhaak RG, Sougnez C, et al. Integrative analysis of the melanoma transcriptome. Genome research. 2010; 20:413–427. [PubMed: 20179022]

Beroukhim R, Getz G, Nghiemphu L, Barretina J, Hsueh T, Linhart D, Vivanco I, Lee JC, Huang JH, Alexander S, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:20007–20012. [PubMed: 18077431]

Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA. Absolute quantification of somatic DNA alterations in human cancer. Nature biotechnology. 2012

Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M, et al. Initial genome sequencing and analysis of multiple myeloma. Nature. 2011; 471:467–472. [PubMed: 21430775]

Chin L, Garraway LA, Fisher DE. Malignant melanoma: genetics and therapeutics in the genomic era. Genes Dev. 2006; 20:2149–2182. [PubMed: 16912270]

Cho US, Xu W. Crystal structure of a protein phosphatase 2A heterotrimeric holoenzyme. Nature. 2007; 445:53–57. [PubMed: 17086192]

Cibulskis K, McKenna A, Fennell T, Banks E, DePristo M, Getz G. ContEst: estimating cross-contamination of human samples in next-generation sequencing data. Bioinformatics. 2011; 27:2601–2602. [PubMed: 21803805]

Conte N, Delaval B, Ginestier C, Ferrand A, Isnardon D, Larroque C, Prigent C, Seraphin B, Jacquemier J, Birnbaum D. TACC1-chTOG-Aurora A protein complex in breast cancer. Oncogene. 2003; 22:8102–8116. [PubMed: 14603251]

Cooper GM, Shendure J. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. Nature reviews Genetics. 2011; 12:628–640.

Cully M, Shiu J, Piekorz RP, Muller WJ, Done SJ, Mak TW. Transforming acidic coiled coil 1 promotes transformation and mammary tumorigenesis. Cancer research. 2005; 65:10363–10370. [PubMed: 16288026]

Curtin JA, Fridlyand J, Kageshita T, Patel HN, Busam KJ, Kutzner H, Cho KH, Aiba S, Brocker EB, LeBoit PE, et al. Distinct sets of genetic alterations in melanoma. The New England journal of medicine. 2005; 353:2135–2147. [PubMed: 16291983]

Delaval B, Ferrand A, Conte N, Larroque C, Hernandez-Verdun D, Prigent C, Birnbaum D. Aurora B - TACC1 protein complex in cytokinesis. Oncogene. 2004; 23:4516–4522. [PubMed: 15064709]

Derijard B, Hibi M, Wu IH, Barrett T, Su B, Deng T, Karin M, Davis RJ. JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain. Cell. 1994; 76:1025–1037. [PubMed: 8137421]

Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, Sougnez C, Greulich H, Muzny DM, Morgan MB, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008; 455:1069–1075. [PubMed: 18948947]

Emuss V, Garnett M, Mason C, Marais R. Mutations of C-RAF are rare in human cancer because C-RAF has a low basal kinase activity compared with B-RAF. Cancer research. 2005; 65:9719–9726. [PubMed: 16266992]

Flaherty KT, Hodi FS, Fisher DE. From genes to drugs: targeted strategies for melanoma. Nature reviews Cancer. 2012; 12:349–361.

Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic acids research. 2011; 39:D945–950. [PubMed: 20952405]

Garibyan L, Fisher DE. How sunlight causes melanoma. Curr Oncol Rep. 2010; 12:319–326. [PubMed: 20623386]

Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S, Beroukhim R, Milner DA, Granter SR, Du J, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature. 2005; 436:117–122. [PubMed: 16001072]

Getz G, Hofling H, Mesirov JP, Golub TR, Meyerson M, Tibshirani R, Lander ES. Comment on “The consensus coding sequences of human breast and colorectal cancers”. Science. 2007; 317:1500. [PubMed: 17872428]

Ghai R, Mobli M, Norwood SJ, Bugarcic A, Teasdale RD, King GF, Collins BM. Phox homology band 4.1/ezrin/radixin/moesin-like proteins function as molecular scaffolds that interact with cargo receptors and Ras GTPases. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:7763–7768. [PubMed: 21512128]

Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nature biotechnology. 2009; 27:182–189.

Greenman C, Wooster R, Futreal PA, Stratton MR, Easton DF. Statistical analysis of pathogenicity of somatic mutations in cancer. Genetics. 2006; 173:2187–2198. [PubMed: 16783027]

Jaffe AB, Hall A. Rho GTPases: biochemistry and biology. Annu Rev Cell Dev Biol. 2005; 21:247–269. [PubMed: 16212495]

Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D, Stern HM, Yue P, Haverty PM, Bourgon R, Zheng J, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. Nature. 2010; 466:869–873. [PubMed: 20668451]

Karlsson R, Pedersen ED, Wang Z, Brakebusch C. Rho GTPase function in tumorigenesis. Biochim Biophys Acta. 2009; 1796:91–98. [PubMed: 19327386]

Kraynov VS, Chamberlain C, Bokoch GM, Schwartz MA, Slabaugh S, Hahn KM. Localized Rac activation dynamics visualized in living cells. Science. 2000; 290:333–337. [PubMed: 11030651]

Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, Stinson J, Yue P, Zhang Y, Pant KP, Bhatt D, et al. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature. 2010; 465:473–477. [PubMed: 20505728]

Lens SM, Voest EE, Medema RH. Shared and separate functions of polo-like kinases and aurora kinases in cancer. Nature reviews Cancer. 2010; 10:825–841.

Li M, Zhao H, Zhang X, Wood LD, Anders RA, Choti MA, Pawlik TM, Daniel HD, Kannangai R, Offerhaus GJ, et al. Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. Nature genetics. 2011; 43:828–829. [PubMed: 21822264]

Lin WM, Baker AC, Beroukhim R, Winckler W, Feng W, Marmion JM, Laine E, Greulich H, Tseng H, Gates C, et al. Modeling genomic diversity and tumor dependency in malignant melanoma. Cancer research. 2008; 68:664–673. [PubMed: 18245465]

Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C, Cruz-Gordillo P, Knoechel B, Asmann YW, Slager SL, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:3879–3884. [PubMed: 22343534]

Malumbres M, Barbacid M. RAS oncogenes: the first 30 years. Nature reviews Cancer. 2003; 3:459–465.

Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, et al. Mutational Processes Molding the Genomes of 21 Breast Cancers. Cell. 2012; 149:979–993. [PubMed: 22608084]

Nikolaev SI, Rimoldi D, Iseli C, Valsesia A, Robyr D, Gehrig C, Harshman K, Guipponi M, Bukach O, Zoete V, et al. Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma. Nature genetics. 2012; 44:133–139. [PubMed: 22197931]

Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, Wells VA, Grunn A, Messina M, Elliot O, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nature genetics. 2011; 43:830–837. [PubMed: 21804550]

Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordonez GR, Bignell GR, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. 2010; 463:191–196. [PubMed: 20016485]

Scott KL, Nogueira C, Heffernan TP, van Doorn R, Dhakal S, Hanna JA, Min C, Jaskelioff M, Xiao Y, Wu CJ, et al. Proinvasion metastasis drivers in early-stage melanoma are oncogenes. Cancer Cell. 2011; 20:92–103. [PubMed: 21741599]

Stark MS, Woods SL, Gartside MG, Bonazzi VF, Dutton-Regester K, Aoude LG, Chow D, Sereduk C, Niemi NM, Tang N, et al. Frequent somatic mutations in MAP3K5 and MAP3K9 in metastatic melanoma identified by exome sequencing. Nature genetics. 2012; 44:165–169. [PubMed: 22197930]

Stefansson B, Brautigan DL. Protein phosphatase PP6 N terminal domain restricts G1 to S phase progression in human cancer cells. Cell Cycle. 2007; 6:1386–1392. [PubMed: 17568194]

Stefansson B, Ohama T, Daugherty AE, Brautigan DL. Protein phosphatase 6 regulatory subunits composed of ankyrin repeat domains. Biochemistry. 2008; 47:1442–1451. [PubMed: 18186651]

Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, Kryukov GV, Lawrence MS, Sougnez C, McKenna A, et al. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011; 333:1157–1160. [PubMed: 21798893]

Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008; 455:1061–1068. [PubMed: 18772890]

Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011; 474:609–615. [PubMed: 21720365]

Thomas RK, Baker AC, Debiasi RM, Winckler W, Laframboise T, Lin WM, Wang M, Feng W, Zander T, MacConaill L, et al. High-throughput oncogene mutation profiling in human cancer. Nature genetics. 2007; 39:347–351. [PubMed: 17293865]

Vigil D, Cherfils J, Rossman KL, Der CJ. Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? Nature reviews Cancer. 2010; 10:842–857.

Wei X, Walia V, Lin JC, Teer JK, Prickett TD, Gartner J, Davis S, Stemke-Hale K, Davies MA, Gershenwald JE, et al. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. Nature genetics. 2011; 43:442–446. [PubMed: 21499247]

Zaidi MR, Davis S, Noonan FP, Graff-Cherry C, Hawley TS, Walker RL, Feigenbaum L, Fuchs E, Lyakh L, Young HA, et al. Interferon-gamma links ultraviolet radiation to melanomagenesis in mice. Nature. 2011; 469:548–553. [PubMed: 21248750]

Zeng K, Bastos RN, Barr FA, Gruneberg U. Protein phosphatase 6 regulates mitotic spindle formation by controlling the T-loop phosphorylation state of Aurora A bound to its activator TPX2. J Cell Biol. 2010; 191:1315–1332. [PubMed: 21187329]

HIGHLIGHTS

- Landscape of driver mutations by exon sequencing of 121 melanoma tumor/normal pairs
- Method for detecting genes with driver mutations in high mutation rate setting
- *PPP6C*, *RAC1*, *SNX31*, *TACC1*, *STK19* and *ARID2* are significantly mutated melanoma genes
- Signature spectrum of UV mutagenesis accounts for 46% of driver mutations found

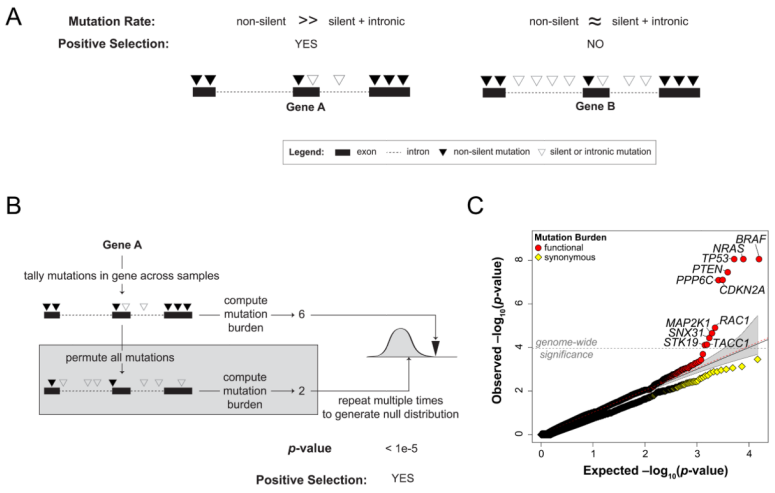

Figure 1. Detection of Positive Selection for Non-Silent Mutations

(A) Gene A locus displaying a greater rate of nonsilent mutation compared to silent and intronic mutation rate (left) indicative of positive selection for nonsilent mutations, and Gene B locus displaying approximately equivalent rates of nonsilent mutation and silent/intronic mutation (right) indicative of a nonsilent mutation rate that matches the basal locus mutation rate. **(B)** Schema of permutation-based framework for identifying genes harboring positively-selected nonsilent mutations. **(C)** Q-Q plot of functional mutation burden test (λ=1.02) and synonymous mutation burden test across all genes with at least one mutation in the set of 121 sequenced samples. Dashed line indicates $q \leq 0.2$ for the functional mutation burden test. Grey-shaded area represents 95% confidence interval for expected p-values. (Please see Figure S1).

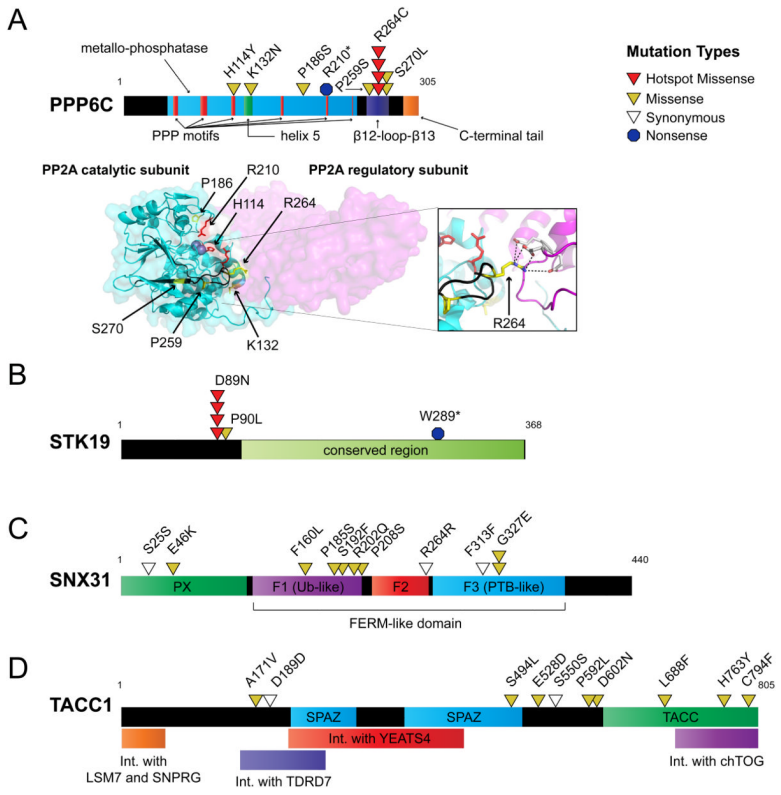

Figure 2. Significantly Mutated Genes *PPP6C, STK19, SNX31, TACC1*

(A–D) Schematic diagram of domains and mutations of *PPP6C*, *STK19*, *SNX31* and *TACC1*. (**A, bottom panel**) Structure of the PPP6C homologous protein, PP2A (PDB: 2IAE), with mapped PPP6C somatic mutations (all mutated residues are conserved between the two proteins except for PPP6C S270, which maps to PP2AC A274). Salt bridge interactions represented by dashed lines in zoom image. PPP motifs: protein phosphatase; Ub-like: ubiquitin-related fold; PTB-like: PTB: phosphotyrosine-binding; PX: Phox homology; FERM-like domain: Band 4.1 (F), Ezrin (E), Radixin (R), and Moesin (M); SPAZ: Ser-Pro Azu-1 motif; TACC: transforming acidic coiled-coil. (Please see Figure S2).

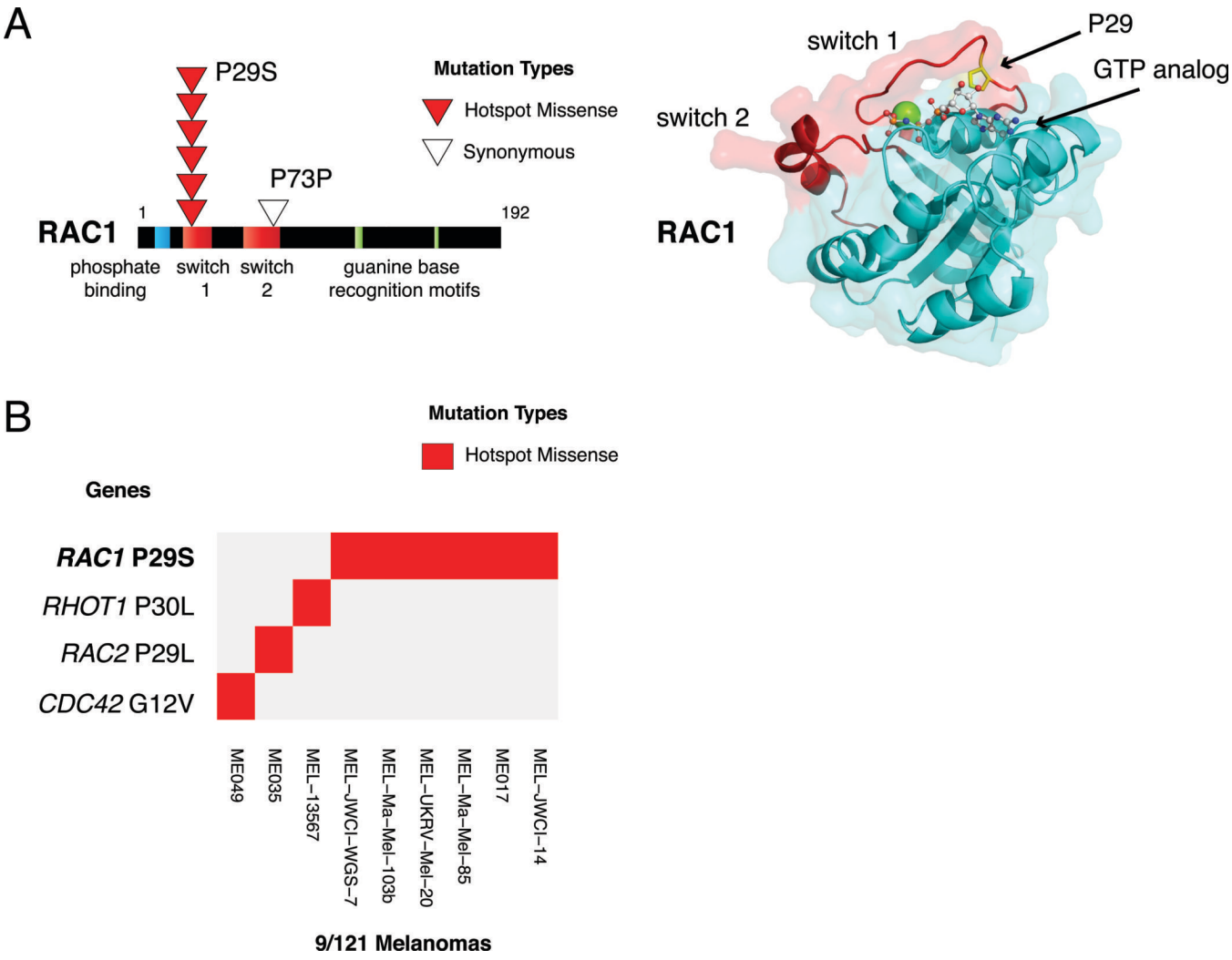

Figure 3. RAC1 hot spot Mutation in Switch I Implicates Rho Family of GTPases in Melanoma

(A) Schematic diagram (left) and image of RAC1 crystallographic model (right) (PDB: 1MH1) with P29 shown. **(B)** Distribution of P29-homologous or known-activating mutations in Rho family members, *RAC1*, *RAC2*, *RHOT1*, and *CDC42*.

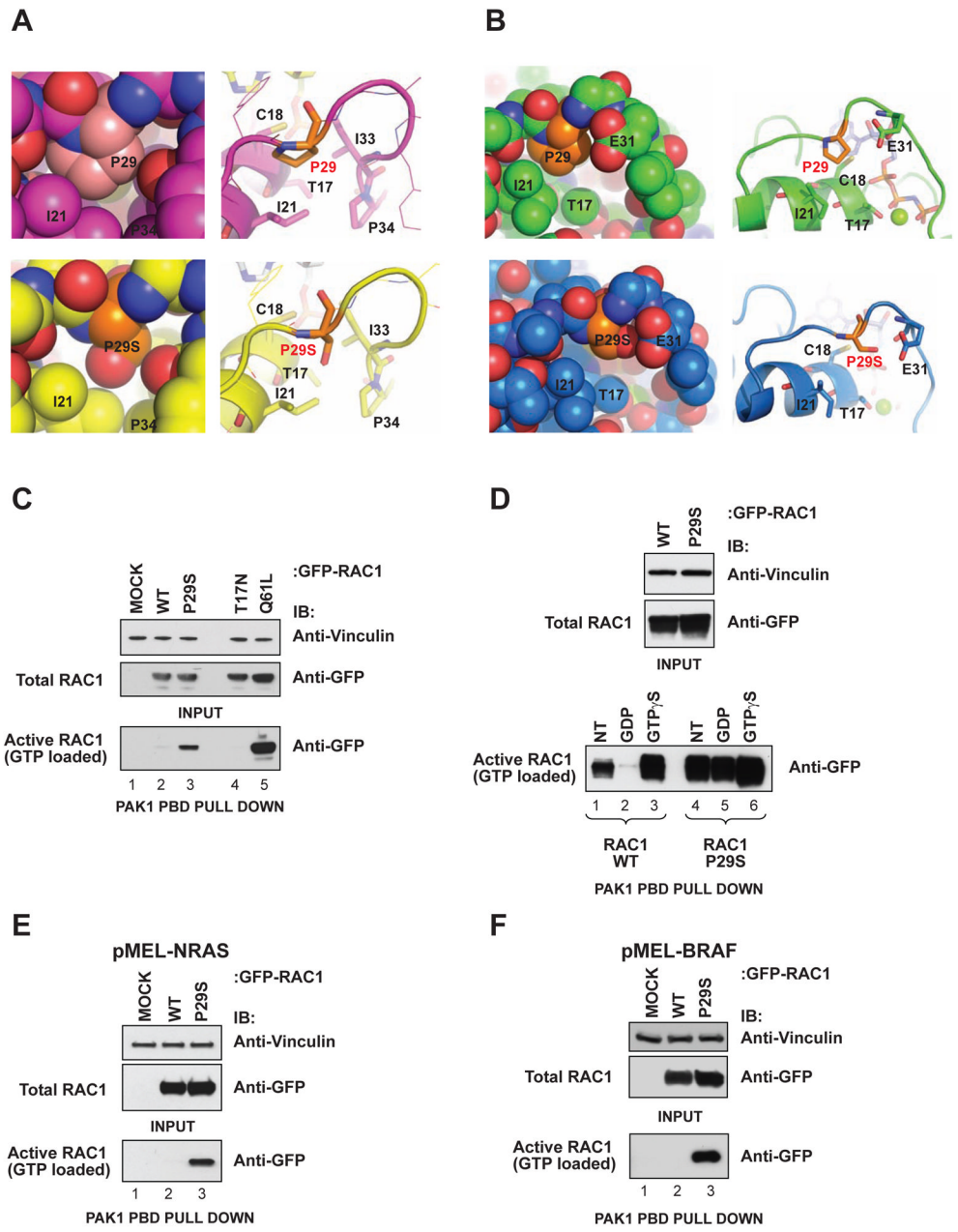

Figure 4. *RAC1* P29S Is Activating

(A) Homology model (based on PDB entry 2G0N) zoom images onto P29 (top panels) and P29S (bottom panels) in the GDP-bound form are shown. **(B)** Homology model (based on PDB entry 2QME) zoom images onto P29 (top panels) and P29S (bottom panels) in the GTP-bound form are shown. Relevant amino acids are highlighted in both sphere (left panels) and cartoon representation (right panels). **(C)** GFP-tagged RAC1 GTP-bound status assayed by p21-binding domain (PBD) of p21-activated protein kinase 1 (PAK1) pull downs in HEK293FT cells (T17N: dominant negative; Q61L: constitutively active); **(D)** In presence of exogenous GDP or GTPγS (NT: No Treatment; GTPγS: non-hydrolysable GTP analog); **(E)** Following transfection of immortalized melanocytes (pMEL) stably expressing mutant forms of NRAS or (**F**) BRAF. (Please see Figure S4).

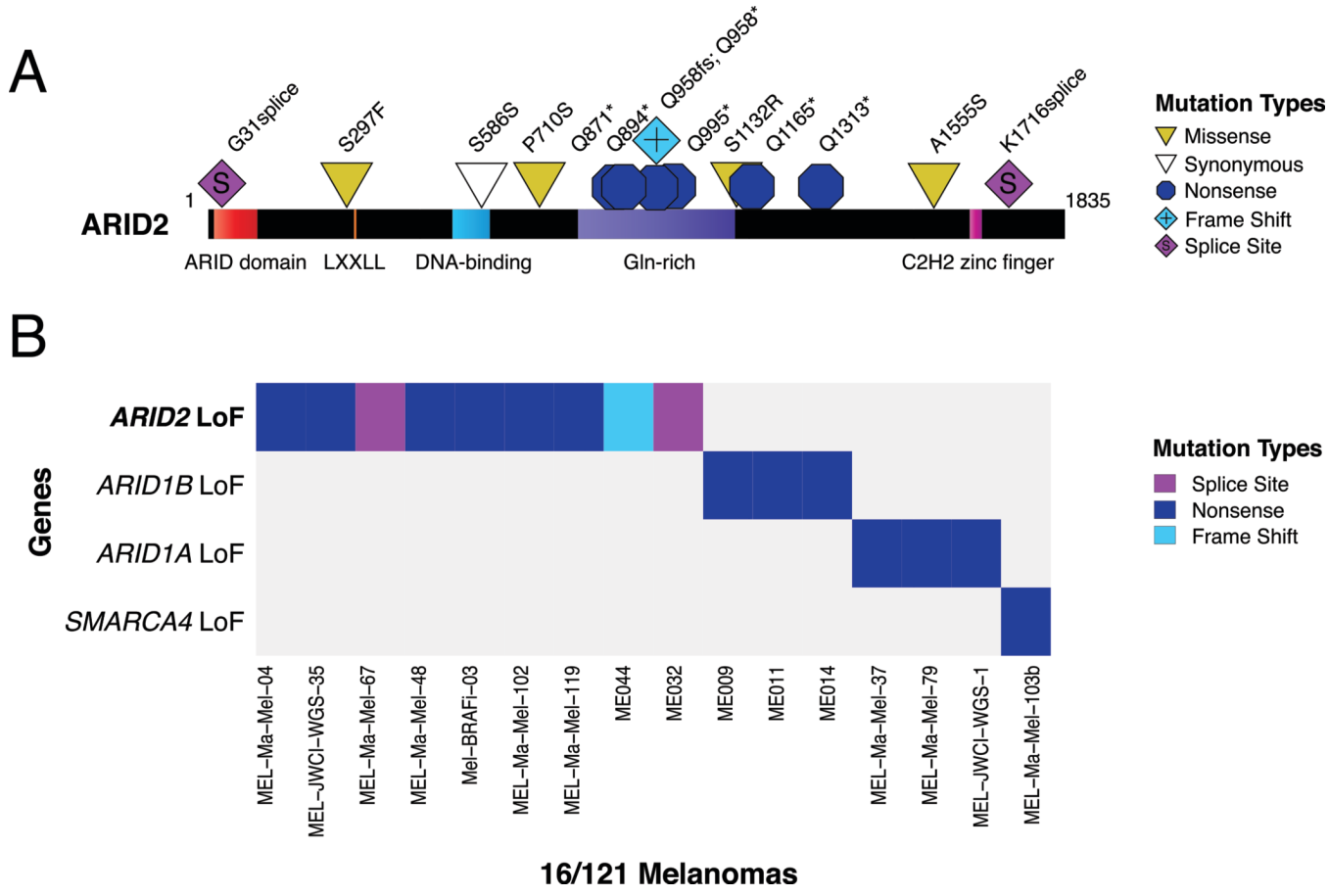

Figure 5. Loss-of-Function Mutations in *ARID2*

(A) Schematic diagram of domains and mutations in *ARID2*. **(B)** Loss-of-function (nonsense, frame-shift indel, splice site) mutations in *ARID2*, *ARID1B*, *ARID1A*, *SMARCA4* across sequenced samples.

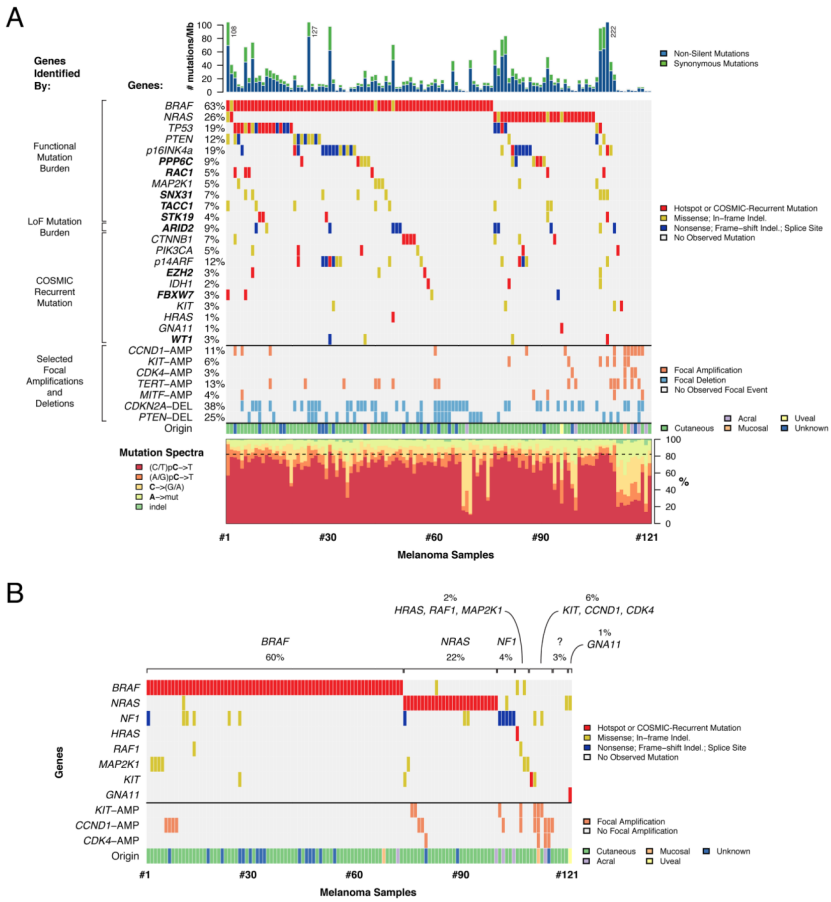

Figure 6. Landscape of Driver Mutations in Melanoma

(A) (Top Panel) Per-sample mutation rate. **(Middle Panel)** Color-coded matrix of individual mutations and copy number alterations. In cases where multiple mutations per gene were found in a sample, only one mutation is shown, with preference given to LoF (nonsense/splice/frameshift) mutations and then hot spot/COSMIC-recurrent mutations. Final row indicates primary origin of melanoma. **(Bottom Panel)** Mutation spectra of all samples. **(B)** Distribution of selected mutations and copy number amplifications in *BRAF*, *NRAS*, *NF1*, *HRAS*, *RAF1*, *MAP2K1*, *KIT*, *GNA11*, *CCND1* and *CDK4* are shown across all samples. (Please see Figure S6).

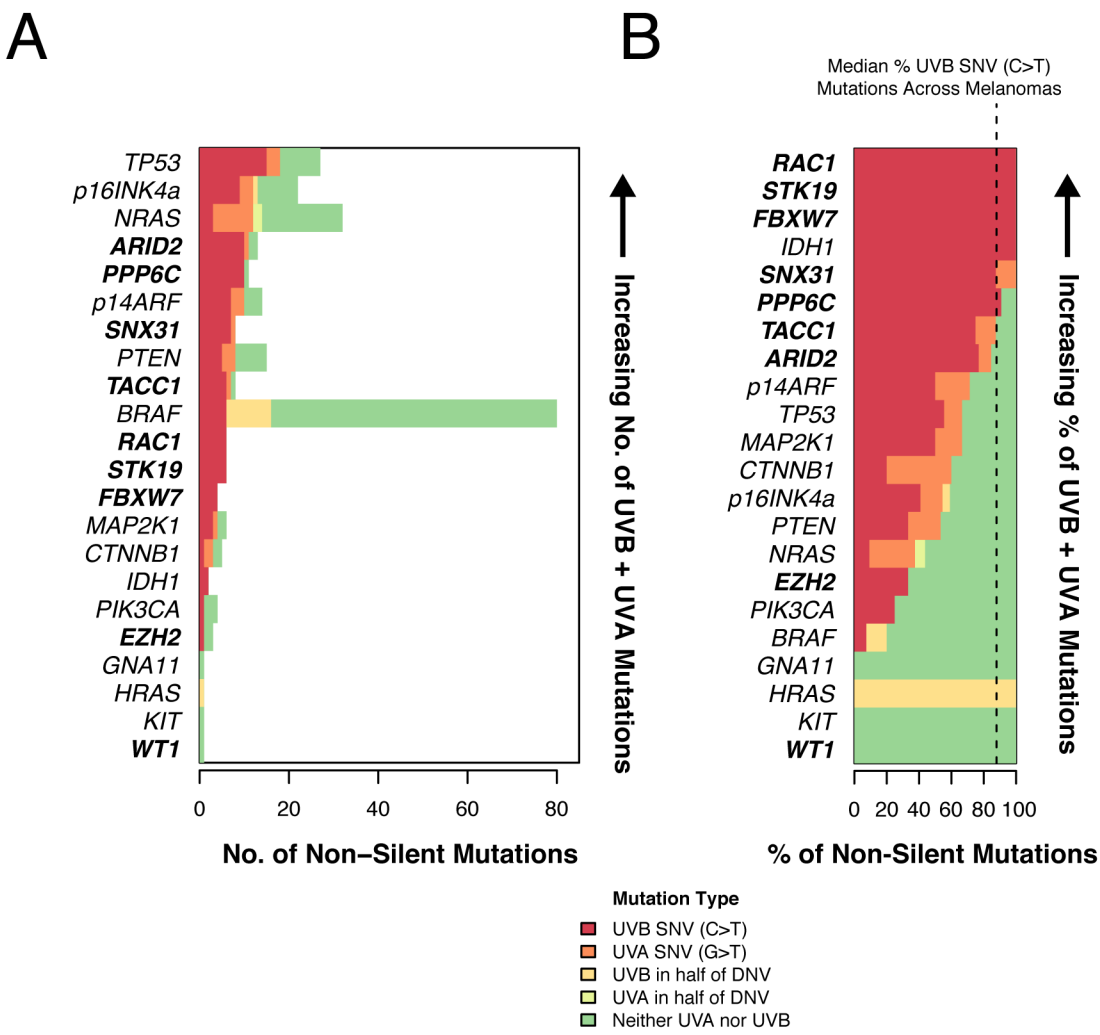

Figure 7. Signature of UV Mutagenesis Across Driver Mutations

(A) Total number and **(B)** % of driver mutations caused by UVB single nucleotide variant (SNV) (C>T), UVA SNV (G>T), UVB in half of dinucleotide variant (DNV) (NC>NT; CN>TN) and UVA in half of DNV (NG>NT; GN>TN) are indicated. Dotted line indicates exome-wide sample median % UVB SNV (C>T).