

**Supplementary Figure 1 Quality control, mutation calling and mutational significance workflow**

# All Melanoma SMGs (Putative GoF)



## **Supplementary Figure 2 Expression differences between mutant vs. wild-type putative gain of function, previously unknown cancer gene, melanoma SMGs**

Mean expression fold-change differences (in TCGA samples) between mutant vs. wild-type melanoma putative gain of function (GoF) SMGs that are not in the COSMIC Cancer Gene Census or OncoKB databases. *NRAS* is included as a reference for GoF mutations (purple name). Genes highlighted by a yellow point have a statistically significant difference in expression between mutant vs. wild-type tumors.



# All Melanoma SMGs (Putative LoF)

## **Supplementary Figure 3 Expression differences between mutant vs. wild-type putative loss of function, previously unknown cancer gene, melanoma SMGs**

Mean expression fold-change differences (in TCGA samples) between mutant vs. wild-type melanoma putative loss of function (LoF) SMGs that are not in the COSMIC Cancer Gene Census or OncoKB databases. *NF1* is included as a reference for LoF mutations (purple name). Genes highlighted by a yellow point have a statistically significant difference in expression between mutant vs. wild-type tumors.



# **Supplementary Figure 4 CCFs of SMGs identified in the full cohort of 1,048 melanomas**

Density plots showing the distribution of CCFs for mutations in melanoma SMGs. Some genes are almost always clonal (e.g. *CDKN2A, EIF1AX*), while others are bimodal (e.g. *GINS1*, *EZH2*) indicating those genes may be both clonal and subclonal drivers.



**Supplementary Figure 5 Impact of hypermutated tumors on SMG analysis**

To determine the effect hypermutated tumors may have on false positives in our cohort, we classified tumors in the top 10% of mutational burden as hypermutated tumors. Here we show SMGs identified only in the (a) entire cohort ( $n = 1048$ ), (b) only when hypermutators are removed  $(n = 943)$ , and  $(c)$  both. (a) For some genes only called SMGs when including all samples, over half the mutations are from hypermutator tumors, including known cancer genes (e.g. *CHIC2*, *FAM58A*). Further, a nontrivial amount of hypermutator tumors are *NF1* melanomas (49%), and several SMGs identified only when including all tumors are driven by *NF1* melanoma (e.g. *SPRED1*, *RASA2*). (c) For all genes identified in both analyses, the fraction of mutations belonging to hypermutated tumors never exceeded 50%. However, this phenomenon was also observed in genes only identified when (a) including all samples, and (b) when removing hypermutators. Thus, the covariates included in mutational significance algorithms likely contribute more to statistical signifiance than the fraction of mutations contributed to hypermutated tumors. Indeed, the Brown's p-values of SMGs (Benjamini-Hochberg, q-value cutoff < 0.1) was not associated with the fraction of mutations contributed by hypermutated tumors (linear regression,  $p > 0.05$ , two-sided). Expanding to all genes, the fraction of mutations contributed by hypermutated tumors slowly becomes more significantly associated with higher p-values (linear regression,  $p < 0.05$ , two-sided). This is likely because hypermutated tumors comprise a large percentage of mutations for infrequently mutated genes.



## **Supplementary Figure 6 Clonal and subclonal mutations per genomic subtype**

(a) Distribution of clonal: subclonal mutation ratios per genomic subtype (Mann-Whitney U,  $p < 1.27 \times 10^{-4}$  for all pairwise). (b) Distribution of clonal (Mann-Whitney U,  $p < 5.6 \times 10^{-8}$  for all pairwise) and (c) subclonal nonsynonymous mutational burdens per genomic subtype (Mann-Whitney U,  $p < 6.0 \times 10^{-4}$  for all pairwise). The data are represented as boxplots where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points. The p-values derived from the Mann-Whitney U tests are two-sided.

# *BRAF* Melanoma SMGs



# **Supplementary Figure 7 Expression differences between mutant vs. wild-type, previously unknown cancer gene,** *BRAF* **melanoma SMGs**

Mean expression fold-change differences (in TCGA samples) between mutant vs. wild-type *BRAF* melanoma SMGs that are not in the COSMIC Cancer Gene Census or OncoKB databases. *BRAF*, *NRAS*, and *NF1* are included as references for gain of function (GoF) and loss of function (LoF) mutations (purple names). Genes with a mean fold-change difference above 0 indicate higher expression in mutant samples compared to wild-type (i.e. GoF mutations). Genes with a mean fold-change difference below 0 indicate lower expression in mutant samples compared to wild-type (i.e. LoF mutations). Genes highlighted by a yellow point have a statistically significant difference in expression between mutant vs. wild-type tumors.



# **CCFs of** *BRAF* **SMGs**

Density plots showing the distribution of CCFs for mutations in *BRAF* SMGs. Some genes are almost always clonal (e.g. *STAT4, DDX3X*), while others are bimodal (e.g. *STK19*, *ZFP91*) indicating those genes may be both clonal and subclonal drivers.



# **Supplementary Figure 9 SMGs in the** *BRAF* **melanoma subtype**

(a) CoMut plot focusing on the TGF-β pathway associated SMGs (*MECOM*, *BMP5*) identified exclusively in *BRAF* melanomas. *MECOM* is an antagonist of the TGF-β pathway (specifically with the SMAD genes), as is *BMP5* (Alliston *et al*., 2005; Bramlage *et al*., 2011).



#### **Supplementary Figure 10** *BRAF* **V600E vs. V600K mutational burden when stratifying by age**

(a) Age at diagnosis was significantly older in *BRAF* V600K patients compared to V600E patients (Mann-Whitney U,  $p =$ 1.07 x 10-5). (b) Nonsynonymous mutational load was significantly elevated in *BRAF* V600K melanomas compared to V600E melanomas (Mann-Whitney U, 1.2 x 10-13). (c) Even when stratifying by age there is still a significant increase in mutations in V600K tumors compared to V600E tumors (15-44 yrs old: 34.68 mut/Mb vs. 6.69 mut/Mb, Mann-Whitney U,  $p = 0.024$ ; 45-59 yrs old: 15.88 mut/Mb vs. 6.68 mut/Mb, Mann-Whitney U,  $p = 0.007$ ; 60-74 yrs old: 11.39 mut/Mb vs. 6.1 mut/Mb, Mann-Whitney U,  $p = 0.0002$ ; 75-87 yrs old: 8.29 mut/Mb vs. 4.57 mut/Mb, Mann-Whitney U,  $p = 0.03$ ). The data are represented as boxplots where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points. The p-values derived from the Mann-Whitney U tests are two-sided.



# **Supplementary Figure 11 Global CNA properties of** *BRAF* **V600E and V600K samples**

(a) *BRAF* V600E tumors experience significantly more copy number events than V600K tumors (49.5 vs. 42, Mann-Whitney U, p = 0.027,). (b) Likewise, V600E tumors also have a significantly higher proportion of the genome altered compared to V600K tumors (54.2% vs. 42.5%, Mann-Whitney U,  $p = 0.028$ ). The data are represented as boxplots where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points. The p-values derived from the Mann-Whitney U tests are two-sided.



**Supplementary Figure 12 GISTIC2.0 amplification and deletion peaks for** *BRAF* **V600E and V600K melanomas**

(a) Significant amplification regions in *BRAF* V600E and *BRAF* V600K melanomas (Benjamini-Hochberg, q-value cutoff < 0.1). (b) Significant deletion regions in *BRAF* V600E and *BRAF* V600K melanomas (Benjamini-Hochberg, q-value  $cutoff < 0.1$ ). The complete list of peaks and the genes they contain can be found in SupplementaryTable 12. Supplementary Table 12 also contains annotations on what genes contained in the peaks are in the CGC and OncoKB, and what genes were called a SMG in the same genomic subtype.



#### **Supplementary Figure 13 Expression differences between mutant vs. wild-type, previously unknown cancer gene, (***N)RAS* **melanoma SMGs**

Mean expression fold-change differences (in TCGA samples) between mutant vs. wild-type (*N)RAS* melanoma SMGs that are not in the COSMIC Cancer Gene Census or OncoKB databases. *BRAF*, (*N)RAS*, and *NF1* are included as references for gain of function (GoF) and loss of function (LoF) mutations (purple names). Genes with a mean fold-change difference above 0 indicate higher expression in mutant samples compared to wild-type (i.e. GoF mutations). Genes with a mean fold-change difference below 0 indicate lower expression in mutant samples compared to wild-type (i.e. LoF mutations). Genes highlighted by a yellow point have a statistically significant difference in expression between mutant vs. wild-type tumors.



# **Supplementary Figure 14 CCFs of** *(N)RAS* **SMGs**

Density plots showing the distribution of CCFs for mutations in *(N)RAS* SMGs. Some genes are almost always clonal (e.g. *CDKN2A, RB1*), while others are bimodal (e.g. *IARS2*, *LONP2*) indicating those genes may be both clonal and subclonal drivers.



# **Supplementary Figure 15 Mutations in BAF/PBAF complex genes identified as SMGs in (***N)RAS* **melanomas**

(a) CoMut plot of (*N)RAS* subtype BAF/PBAF complex SMGs in *BRAF* melanomas. *ARID2, ARID1A,* and *BRD7* were identified as SMGs in the *BRAF* subtype, although at lower frequencies and statistical significance. (b) CoMut plot of (*N)RAS* subtype BAF/PBAF complex SMGs in *NF1* melanomas. Although a higher proportion of *NF1* melanomas harbored mutations in these genes compared to (*N)RAS* melanomas, only *ARID2* was identified as significantly mutated. Further, the majority of mutations in *NF1* melanomas are not putative loss of function (nonsense mutations, splice-site variants and indels).



**Supplementary Figure 16 Clinical characteristics and SMGs in** *NF1* **melanomas**

(a) Distribution of age at diagnosis between the genomic subtypes. *NF1* melanomas are associated with significantly older age at diagnosis compared to the other genomic subtypes (Mann-Whitney U,  $p < 0.028$  pairwise for all, two-sided). The data is represented as a boxplot where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points. (b) The co-mutation plot of *NF1* RASopathy SMGs and the novel RAS-associated SMG *RASSF2*, including the annotation of missense and inactivating *NF1* mutations. Loss of function mutations in the RASopathy genes (*RASA2* and *SPRED1*) were never observed in the same tumor, as were loss of function mutations between *SPRED1* and *RASSF2*. One tumor harbored loss of function mutations in both *RASA2* and *RASSF2*.

# *NF1* Melanoma SMGs



#### **Supplementary Figure 17 Expression differences between mutant vs. wild-type, previously unknown cancer gene,** *NF1* **melanoma SMGs**

Mean expression fold-change differences (in TCGA samples) between mutant vs. wild-type *NF1* melanoma SMGs that are not in the COSMIC Cancer Gene Census or OncoKB databases. *BRAF*, *NRAS*, and *NF1* are included as references for gain of function (GoF) and loss of function (LoF) mutations (purple names). Genes with a mean fold-change difference above 0 indicate higher expression in mutant samples compared to wild-type (i.e. GoF mutations). Genes with a mean fold-change difference below 0 indicate lower expression in mutant samples compared to wild-type (i.e. LoF mutations). Genes highlighted by a yellow point have a statistically significant difference in expression between mutant vs. wild-type tumors.



#### **Supplementary Figure 18 CCFs of** *NF1* **SMGs**

Density plots showing the distribution of CCFs for mutations in *NF1* SMGs. Some genes are almost always clonal (e.g. *FAM58A, RASSF2*), while others are bimodal (e.g. *RAC1*, *MSH6*) indicating those genes may be both clonal and subclonal drivers.



# **Supplementary Figure 19 SMGs in TWT Melanomas**

CoMut plot of SMGs in the cohort of only cutaneous TWT melanomas.

# TWT Melanoma SMGs



#### **Supplementary Figure 20 Expression differences between mutant vs. wild-type, previously unknown cancer gene, TWT melanoma SMGs**

Mean expression fold-change differences (in TCGA samples) between mutant vs. wild-type TWT melanoma SMGs that are not in the COSMIC Cancer Gene Census or OncoKB databases. *BRAF*, *NRAS*, and *NF1* are included as references for gain of function (GoF) and loss of function (LoF) mutations (purple names). Genes with a mean fold-change difference above 0 indicate higher expression in mutant samples compared to wild-type (i.e. GoF mutations). Genes with a mean fold-change difference below 0 indicate lower expression in mutant samples compared to wild-type (i.e. LoF mutations). Genes highlighted by a yellow point have a statistically significant difference in expression between mutant vs. wild-type tumors.



# **Supplementary Figure 21 CCFs of TWT SMGs**

Density plots showing the distribution of CCFs for mutations in TWT SMGs. Some genes are almost always clonal (e.g. *RQCD1, GNA11*), while others are bimodal (e.g. *SF3B1*, *DDX59*) indicating those genes may be both clonal and subclonal drivers.



**Supplementary Figure 22 GISTIC2 workflow for calling focal regions enriched in amplifications and deletions**



#### **Supplementary Figure 23 GISTIC2.0 amplification and deletion peaks for each of the genomic subtypes**

We used GISTIC2.0 to identify regions selectively targeted by somatic CNAs in each of the genomic subtypes. (a) A total of 26, 29, 14 and 16 significant focal amplification peaks, and (b) 16, 14, 9, and 7 significant focal deletion peaks, were identified in *BRAF*, (*N)RAS*, *NF1* and TWT melanomas, respectively (Benjamini-Hochberg, q-value cutoff < 0.1). Several of these peaks were in regions containing CGC and OncoKB genes (Supplementary Table 12).



#### **Supplementary Figure 24 CNA and fusion events per genomic subtype**

(a) There was significant heterogeneity in the number of CNAs between the genomic subtypes (Kruskal-Wallis,  $p = 7.78$  x 10-8, two-sided), ranging from 75 events in TWT melanomas to 47 events in *BRAF* melanomas. However, there was no difference in the proportion of the genome altered by CNA events between the subtypes (Kolmogorov-Smirnov,  $p > 0.05$ , two-sided). (b) The occurrence of gene fusions also differed significantly between the subtypes (Kruskal-Wallis,  $p =$ 0.006, two-sided), ranging from 6 fusion events in TWT melanomas to 2 fusion events in *BRAF* melanomas. The data are represented as boxplots where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points.



# **Supplementary Figure 25 NMF validation of deconstructSigs results on the entire cohort via SomaticSignatures**

Signature<br>Signature

Signature<br>Signature

Signature

**22**<br>27

Signature<br>Signature

Signature

O1 S.  $\frac{1}{4}$  $\vec{q}$  $\vec{\omega}$  $\overrightarrow{L}$ 

 $\circ$ 

 $\rightarrow$ 

(a) NMF statistics for the entire cohort of melanomas. (b) Cosine similarity between COSMIC signatures and signatures decomposed via NMF for the entire cohort of melanomas. Signature S1 had the highest cosine similarity with signature 7 (UV exposure), signature S2 had the highest cosine similarity with signature 11 (exposure to alkylating agents), and signature S3 had the highest cosine similarity with signature 1 (spontaneous deamination of 5-methylcytosine).

Signature 10<br>Signature 28

Signature 9<br>Signature 21

Signature 12 Signature Signature Signature Signature Signature 16 Signature Signature Signature Signature

29  $\vec{\infty}$ Z4 ω S1

S2

S3

bs

 $\infty$ ည္က

 $\overline{\phantom{a}}$ 

0.2

0.4

0.6

0.8

b)

Signature 11 Signature 23 Signature 19 Signature 30 Signature 2 Signature 7 Signature 1 Signature 6 Signature 5 Signature 20 Signature 14 Signature 15 Signature 13 Signature 17 Signature 27 Signature 22 Signature 28 Signature 10 Signature 21 Signature 9 Signature 12 Signature 29 Signature 18 Signature 24 Signature 3 Signature 16 Signature 4 Signature 26 Signature 8 Signature 25

Signature 11

Signature

ಜ  $\vec{6}$ မ္ထ  $\mathbf{v}$ 

Signature Signature Signature Signature 7 Signature Signature



**Supplementary Figure 26 DSB repair deficiency - loss of heterozygosity (LoH) score**

(a) Distribution of copy number LoH events in signature 3 (yellow) and non-signature 3 (purple) melanomas in the entire cohort. This satisfies the test used in Abkevich et al., 2012 and Timms et al., 2014, as the distribution was significantly different via the Kolmogorov-Smirnov test ( $p = 0.005$ , two-sided) and univariate logistic regression ( $p = 5.34 \times 10^{-5}$ ). (b) Density plot of copy number LoH events in the entire cohort. (c) Distribution of copy number LoH events in signature 3 and non-signature 3 melanomas in TWT melanomas ( $p = 0.015$ , Kolmogorov-Smirnov, two-sided;  $p = 0.077$ , univariate logistic regression, two-sided). (d) Density plot of copy number LoH events in the TWT melanomas.



**Supplementary Figure 27 DSB repair deficiency - allelic telomeric imbalance score**

(a) Distribution of copy number TAI events in signature 3 (yellow) and non-signature 3 (purple) melanomas in the entire cohort. This satisfies the test used in Birkbak et al., as the distribution was significantly different via a Mann-Whitney U test ( $p = 4.40 \times 10^{-5}$ , two-sided). (b) Density plot of copy number TAI events in the entire cohort. (c) Distribution of copy number TAI events in signature 3 and non-signature 3 melanomas in TWT melanomas (Mann-Whitney U,  $p = 1.8 \times 10^{-3}$ , two-sided). (d) Density plot of copy number TAI events in the TWT melanomas. In (a) and (c) the data is represented as a boxplot where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points.



## **Supplementary Figure 28 DSB repair deficiency - large scale transitions**

(a) Distribution of copy number LST events in signature 3 (yellow) and non-signature 3 (purple) melanomas in the entire cohort. This satisfies the test used in Popova et al., as the distribution was significantly different via a Mann-Whitney U test ( $p = 6.82$  x 10<sup>-3</sup>, two-sided). (b) Density plot of copy number LST events in the entire cohort. (c) Distribution of copy number LST events in signature 3 and non-signature 3 melanomas in TWT melanomas (Mann-Whitney U,  $p = 0.056$ , two-sided). (d) Density plot of copy number LST events in the TWT melanomas. In (a) and (c) the data is represented as a boxplot where the middle line is the median, the lower and upper edges of the box are the first and third quartiles, the whiskers represent the interquartile range (IQR) multiplied by 1.5, and beyond the whiskers are outlier points.



**Supplementary Figure 29 Associations between mutational signatures and scarHRD scores**

(a-c) Signature 3 was the only mutational signature to be associated with all three scarHRD copy number event scores (loss of heterozygosity, allelic telomeric imbalance, large scale transitions), and (d-f) this relationship still held when excluding acral and mucosal melanomas, which are enriched in copy number alterations compared to cutaneous melanomas. The dashed lines represent p-value cutoffs of 0.05.





Missense Mutation

(a) Top: Lollipop plot of somatic mutations in *ATM* for TWT melanomas. The splice-site variant in the FAT domain of *ATM* was exclusive to a TWT melanoma with signature 3. Bottom: Lollipop plot of somatic mutation in *ATM* for non-TWT melanomas. A signature 3 non-TWT melanoma also had a splice-site variant in the FAT domain of *ATM*. (b) Lollipop plot of somatic mutations in *BLM* for TWT melanomas. The missense mutation in the HRDC domain of *BLM*  was exclusive to a TWT tumor with signature 3.

**0 500 1000 1420**

Pairwise Relatedness



**Supplementary Figure 31 Relatedness between normal samples**

To prevent duplicate mutation calls from the same patient influencing our analyses, we used Somalier to determine the relatedness between normal samples in our cohort. Samples from the same patient would have a relatedness value very close to 1. Opacity was used to show the density of points.