Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma

Michael Krauthammer[1], Yong Kong[2,3], Byung Hak Ha[4], Perry Evans[1], Antonella Bacchiocchi[5], James P McCusker[1], Elaine Cheng[5], Matthew J Davis[4], Gerald Goh[6, 7], Murim Choi[6, 7], Stephan Ariyan[8], Deepak Narayan[8], Ken Dutton-Regester[9,10], Ana Capatana[1], Edna C Holman[5], Marcus Bosenberg[5], Mario Sznol[11], Harriet M Kluger[11], Douglas E Brash[5,6,12], David F Stern[1], Miguel A Materin[13], Roger S Lo[14], Shrikant Mane[6,15,16], Shuangge Ma[17], Kenneth K Kidd[6], Nicholas K. Hayward[10], Richard P Lifton[6,7], Joseph Schlessinger[4], Titus J Boggon[4], Ruth Halaban[5*]

[1]Department of Pathology, Yale University School of Medicine, New Haven, CT, 06520-8059
[2]Deparment of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT, 06520-8059
[3]W.M. Keck Foundation Biotechnology Resource Laboratory, Yale University School of Medicine, New Haven, CT, 06520-8059
[4]Department of Pharmacology, Yale University School of Medicine, New Haven, CT, 06520
[5]Department of Dermatology, Yale University School of Medicine, New Haven, CT, 06520
[6]Department of Genetics, Yale University School of Medicine, New Haven, CT, 06520
[7]Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT, 06520
[8]Department of Surgery, Yale University School of Medicine, New Haven, CT, 06520
[9]Queensland University of Technology, Brisbane, QLD, Australia
[10]Queensland Institute of Medical Research, Brisbane, QLD, Australia.
[11]Comprehensive Cancer Center Section of Medical Oncology, Yale University School of Medicine, New Haven, CT, 06520
[12]Department of Therapeutic Radiology, Yale University School of Medicine, New Haven, CT, 06520
[13]Department of Ophthalmology, Yale University School of Medicine, New Haven, CT, 06520
[14]Division of Dermatology/Department of Medicine, UCLA Jonsson Comprehensive Cancer Center, Los Angeles, CA 90095-1750.
[15]Center for Human Genetics and Genomics Yale University School of Medicine, New Haven, CT, 06520
[16]Program on Neurogenetics, Yale University School of Medicine, New Haven, CT, 06520[17]School of Public Health, Yale University School of Medicine, New Haven, CT, 06520

* Correspondence should be addressed to RH (ruth.halaban@yale.edu).

Addendum to the Online Methods

Identification of genes with significant somatic mutation burden: In order to calculate a list of significantly mutated genes, i.e., genes with more mutations than expected by the background mutation frequency, we modified a recently established protocol[1]. In essence, we used the non-silent to silent mutation ratio (NS:SN ratio), i.e., the number of mutations that cause amino acid changes over those that do not, and the silent mutation frequency, i.e., the number of silent mutations over the number of sequenced bases, to estimate the non-silent background mutation frequency. The latter is then used to determine whether some observed number of non-silent mutations in a gene is above the expected. We also used insights into melanoma-specific mutation patterns to calculate mutation frequencies based on sequence contexts, and on expression of the gene locus. We measured an increase in mutational frequency when studying non-expressed versus expressed genes, and observed that most mutations occur at cytosines in the dipyrimidine context, as clearly shown in Figure 1. Taken together, this led us to calculate the non-silent background mutation frequencies separately for expressed and non-expressed genes, and separately for the three following sequence contexts: 1) mutating Cs at dipyrimidines, 2) mutating Cs at non-dipyrimidines, and 3) mutating Ts, which stand for, respectively, mutations in cytosines with a flanking pyrimidine, mutations in cytosines without a flanking pyrimidine, and mutations in thymines with no restriction on the flanking bases. For example, a C>T mutations in a TC*G context would be counted towards mutations in the *C dipyrimidine* context, as would a G>A mutation in the CG*A context (i.e., the reverse complement). Conversely, a C>T mutation in the GC*G context would be counted towards the *C non-dipyrimidine* context.

The context-specific non-silent mutation frequency $MF_{NS,C}$ is estimated by $MF_{NS,C} = MF_{SN,C} \times NS:SN_C$ where is the context-specific silent mutation frequency, i.e., the number of silent somatic mutations in context C divided over all bases in context C with sufficient sequence depth in the exome capture region, and $NS:SN_C$ is the non-silent to silent ratio for mutations in context C (see below). We calculated $MF_{NS,C}$ for each of the three contexts, and performed, for each gene, and for each context, a binomial test for whether the observed non-silent mutations in a gene are explained by $MF_{NS,C}$, receiving 3 distinct and independent p-values for each context. We then use the Fisher's combined probability test to generate an overall p-value measuring whether the number of non-silent mutations in a gene is more than expected.

We added two additional processing steps to the basic workflow discussed above. As the NS:SN ratios vary considerably between genes, we estimated gene-specific NS:SN ratios in each of the three contexts. We proceeded as follows: we first identified all bases in a particular gene that are positioned in the context C under consideration. We then performed an in-silico

experiment where we mutated each base and recorded whether the change resulted in a non-silent change or not. The resulting ratios between non-silent and silent changes were weighted according to the observed frequencies for a particular base change. The frequencies for each base change, in each context, were calculated from the frequencies of the observed silent and non-silent base changes, with the exception of non-silent changes in the top 100 mutated genes, which may be enriched for driver mutations (the top 100 genes were determined by dividing the number of observed somatic mutations by the gene length, and ranking of the resulting ratios). We determined an overall NS:SN ratio, across the three contexts, and across all genes, of 1.93 in sun-exposed melanomas, close to the observed NS:SN ratio of 2.0. We added an additional processing step for genes, which exhibited context-specific non-silent mutation counts beyond what was expected by the exome-wide $MF_{SN,C}$. For those genes, instead of using $MF_{SN,C}$, we estimated a $MF_{SN,C,G}$ by dividing the observed silent mutations in gene G over the number of bases in context C across gene G. This adjustment is necessary to account for biases, for which we did not account for in the model, and which have not been fully studied and quantified in melanoma. Among these are replication timing, location on the chromosome and others[2], all of which may affect the number of mutations in a gene.

We then combined the p-values for the individual binomial tests across contexts, using the Fisher combined method. The resulting ranking and p-values are labeled as “Comprehensive Model”. We also calculated two more rankings: the first one did not take into account expression (“No Expression Model”), and the other one did not weigh the genes according to their silent counts, and did not take into account expression and sequence context (“Simple Model”). The latter model represents a simple weighting of the somatic mutations by gene length and a single exome-wide background mutation frequency based on a (context-independent) exome-wide NS:SN ratio. The final gene burden ranks were matched against similar ranks that were generated by excluding the top 5% of mutated samples, in order to ensure robustness of the results. Only genes that were ranked high in both lists were retained. It should be noted that for these calculations, SNVs affecting the same codon were counted as independent events.

Identification of genes with a significant number of deleterious mutations. We tabulated nonsense SNVs, splice-site variants, frame-shift InDels, and InDels with insertion or deletions of 3 or more codons, across all sun-exposed melanomas, including the unmatched samples. We used a binomial test to find genes enriched in deleterious mutations, using the exome-wide frequency of these mutations. For each highly ranked gene, we required that at least 30% of the mutations were in the matched melanoma set.

Gene expression

Whole genome gene expression was derived from hybridization to NimbleGen human whole genome expression microarrays and RNA-Seq. Array analysis was performed on 15

melanomas and four independent human melanocytes at NimbleGen Systems Iceland LLC. Vínlandsleið 2-4, 113 Reykjavik, Iceland (currently Roche Applied Science, Basel, Switzerland) and by the Yale W.M. Keck Foundation Biotechnology Resource as described[3,4]. Data from the array analysis was used to identify expressed genes in normal melanocytes and melanomas. Genes with median expression value of 550 and above were called expressed.

RNA-Seq was performed on two independent cultures of two normal human melanocytes cultures derived from newborn foreskins and adult skin. Total RNA was extracted using Trizol (Invitrogen) followed by DNase digestion and Qiagen RNeasy (Qiagen, Valencia, CA) column purification following the manufacture's protocol. The RNA integrity was verified using an Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA). One microgram of high-quality RNA was processed using an Illumina RNA-Seq sample prep kit following the manufacturer's instructions (Illumina, San Diego, CA). Final RNA-Seq libraries were sequenced at 75 bp/sequence using an GAIIx Illumina sequencer. Reads were processed with bwa and SAMtools. Mapping was performed against the reference genome. Reads were counted in bins of 100 bp, and normalized with regard to the median. To calculate the expression value for a particular RefSeq transcript, we determined the transcript exon boundaries, and summed up all bin read values for bins within the boundaries. The transcript length-normalized, and log-transformed value was used as the measure of gene expression. A two component Gaussian mixture model was fit to the data, and a lower bound for expressed genes was chosen as two standard deviations away from the higher distribution mean. The RNA-Seq data is used for identifying expressed genes in normal melanocytes for the gene burden analysis.

Supplementary Results

Sequencing statistics

Mean error rate and coverage: We first tested the sequence fidelity and read coverage of all Illumina sequencing runs. In general, there was an excellent low average sequencing error rate of 0.24%, representing the fraction of bases from sequencing reads that do not align with the reference genome. The average coverage was 65 ±14.8 independent reads per targeted base pair (minimum mean sample coverage 30, maximum mean sample coverage 93) for tumor samples sequenced with the Illumina GA IIx, and 224 ± 47 independent reads per targeted base for samples sequenced with the Illumina HiSeq 2000 (minimum sample coverage 100, maximum sample coverage 376). The % bases covered at least eight times across the capture area were 90.5% for GA IIx, and 97.2% for HiSeq 2000. We compared the mean number of somatic mutations in melanomas that have been sequenced using the GAIIx and HiSeq technologies. We found that both technologies resulted in comparable non-synonymous somatic SNV counts. Both the GAII-sequenced and HiSeq-sequenced melanoma were evenly distributed among Figure 1a (ranking of melanomas by mutation count), and both technologies contributed to

samples with counts above the 90th percentile. Below the 90th percentile, GAII-sequenced sun-exposed melanomas had a median of 123 somatic mutations, while HiSeq-sequenced melanomas had a median count of 154. The median number of SNVs in sun-shielded melanomas was 11 (GAII) and 7 (HiSeq), respectively. We compared the mutation counts in melanomas from cell lines and fresh frozen tumors. The median number of somatic mutations in sun-exposed cell lines and tumors was 138 and 168, respectively. The corresponding numbers in sun-shielded melanomas was 10 and 7.

Somatic SNV call precision based on Sanger validations: The precision analysis for our two-step somatic calling pipeline was as follows: we first established the precision of calling a tumor SNV (establishing the presence of the variant in tumor), and then determined the precision of classifying it as a somatic variant based on matched germline DNA data. We defined precision as the ratio of correctly called variants over all called variants.

Validation by Sanger sequencing revealed that of 266 SNVs that were automatically called according to the thresholds discussed above, 21 were false positives. We thus calculated a precision of 245/266 = 92.1% for calling tumor variants in Exome-Seq. In the presence of a matched germline DNA sample, we called a tumor variant as either somatic or inherited. To determine somatic call precision, we determined by Sanger sequencing how many of the somatic calls were actually inherited SNVs, or false positive tumor SNVs that are erroneously called somatic. We counted 80 tumor SNVs that were called using the thresholds above, and for which we had matched germline DNA sequencing data: 64 of those were true somatic variants, 9 were inherited variants, and 7 were false positive variants. The sequencing pipeline automatically called 59 out of the 80 SNVs as somatic. Of those, 55 were true positive somatic SNVs, one was an inherited SNV, and 3 were false positive tumor SNVs. We thus determined a somatic call precision of 55/ 59 = 93.2%.

Somatic SNV call sensitivity based on detection of SNVs in germline DNA: The true total number of somatic changes in tumor is not known, and yet there is a need to assess somatic call sensitivity, which is defined as the number of called variants over all real variants. We designed a somatic call sensitivity estimate based on our two-step somatic call procedure: First, we determined the sensitivity of detecting the presence of a variant in tumor. Then, we established the sensitivity of detecting those variants that are somatic using matched germline DNA sequencing data. For the estimation, we assumed that detection of SNVs in tumor could be equated to detecting inherited tumor SNVs given adequate tumor purity. We therefore estimated the sensitivity of detecting tumor SNVs by counting the number of known SNPs in germline DNA, which are positively called in a matched tumor, measuring a mean sensitivity of 95% across our melanomas. We then measured our ability to detect those tumor SNVs that are somatic. Using our 64 automatically called tumor SNVs that were Sanger validated (somatic) and had corresponding sequencing data in germline DNA, we correctly called 55 of those

variants as somatic, a somatic call sensitivity of 55/64=85.9%. Overall sensitivity to detect somatic variants is thus estimated to be 0.95 x 0.859 = 81.6%.

Somatic SNV call sensitivity for known melanoma driver mutations: Routine Sanger sequencing of all 147 melanomas in our Exome-Seq screens identified 50 $BRAF^{V600}$ (V600E/K/R), 25 $NRAS^{Q61}$ (Q61L/R/H), two $NRAS^{G12}$ (G12D/V), two *NRAS* (G13D/R) and one $HRAS^{Q61}$ mutations. We determined the sensitivity of Exome-Seq to identify these variants. At the thresholds discussed earlier, Exome-Seq automatically called all but two *BRAF*, and all *NRAS* and *HRAS* variants, a SNV detection sensitivity of 79/81=97.5%. For matched samples all but one BRAF, and all *NRAS* and *HRA*S mutations were correctly called somatic. One of the failed BRAF calls in tumor was due to high level of fibroblast contamination in the cell culture (80%). BRAF and NRAS mutations likely occur early in melanoma genesis, and are thought to be present across all tumor clones. Difficulties in detecting these mutations are therefore primarily caused by stromal tissue contamination, as opposed to clonal heterogeneity. The fact that most of these variants were recovered indicates that our sequencing depth and sensitivity are sufficient to retrieve variants with similar clonal distributions as BRAF and NRAS mutations.

Structural analysis

The crystal structure of $RAC1^{P29S}$ in complex with GMP-PNP is broadly unchanged from $RAC1^{WT}$ in complex with GMP-PNP both previously published[5], and described here (**Supplementary Table 10**). The major exceptions are the conformational differences in the switch I loop. Overall the structure shows RMSDs of 0.8 Å and 0.7 Å over 177 and 175 Cα atoms for chains A and B when compared to $RAC1^{WT}$, PDB ID: 1MH1[5], and smaller differences to $RAC1^{WT}$ described here (see above). Comparison of chains A and B of $RAC1^{P29S}$ shows that both chains have very similar conformation with RMSD of 0.4 Å over 175 Cα atoms. $RAC1^{P29S}$ shows good electron density throughout the structure, with average protein *B*-factors of 30.7 $Å^2$.

The electron density for the $RAC1^{P29S}$ switch I region, GMP-PNP and Mg^{2+}, is well defined for both molecules in the asymmetric unit (**Supplementary Fig. 7**). In molecule A, hydrogen bonds of 2.9 Å between S29 carbonyl oxygen and ribose 2'-hydroxyl and 2.9 Å between G30 carbonyl oxygen and ribose 3'-hydroxyl are observed. In molecule B, hydrogen bonds of 2.9 Å and 3.1 Å between G30 carbonyl oxygen and ribose 2'-hydroxyl and 3'-hydroxyl respectively are observed, and the S29 carbonyl oxygen is 3.3 Å from the ribose 2'-hydroxyl group (**Supplementary Fig. 7**).

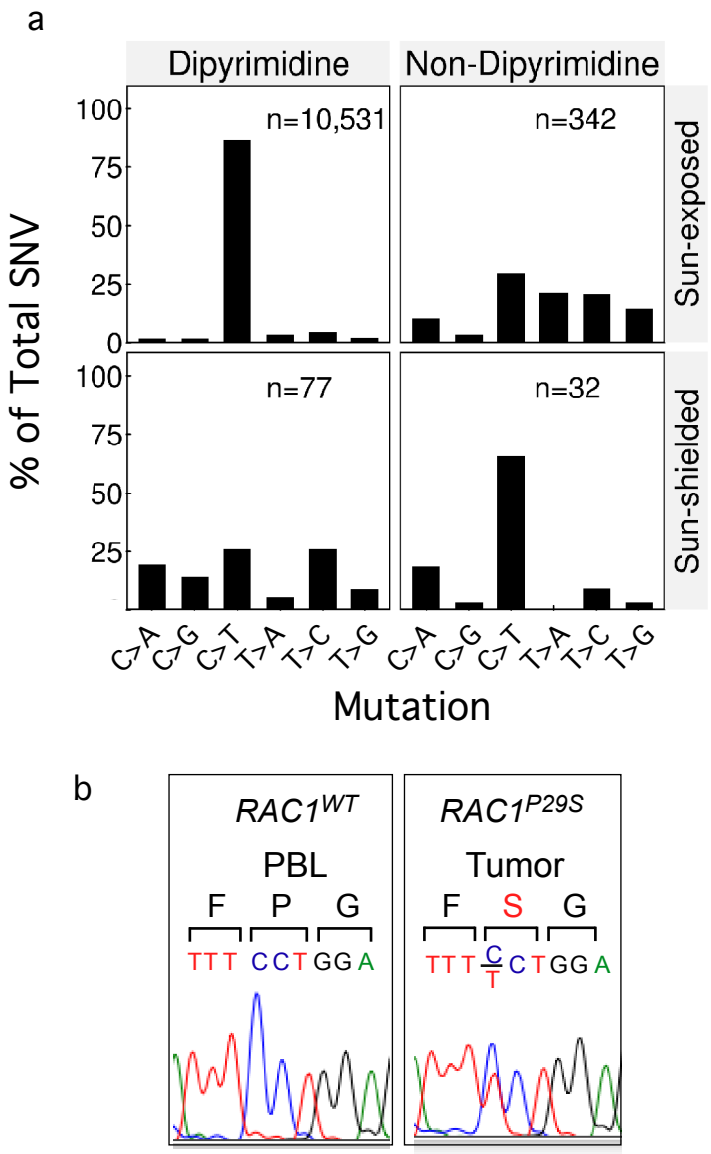
It is unusual in a crystal structure of a RHO family GTPase to observe direct hydrogen bonding interactions between backbone carbonyls of P29 and G30 and the ribose hydroxyl groups; in RHO family GTPases the ribose hydroxyl-switch I backbone interactions are usually mediated through water molecules. There are few previous examples of direct hydrogen bonding between switch I and the ribose hydroxyls for RHO family GTPases. These include a) three RND1 structures: 2CLS (RND1, unpublished), 2REX (RND1 bound with PLXNB1,

unpublished), and 3Q3J (RND1 in complex with plexin A2 RBD, unpublished); b) RHO1P – Sec3p complex from *Saccharomyces cerevisia*, PDB ID: 3A58[6]; and c) photoactivatable RAC1-LOV2 fusion protein, PDB ID: 2WKP[7]. Although these crystal structures suggest that Proline at position 29 does not absolutely preclude hydrogen bond formation between the switch I backbone and ribose hydroxyl groups, these cases are unusual for RHO GTPase family members (Supplementary Fig. 9). In contrast, for non-RHO family GTPases direct interactions between the ribose hydroxyl groups and switch I backbone are common (**Supplementary Fig. 8**). To confirm these differences we superposed a collection of GTP- or GTP-analogue-bound GTPase structures deposited in the Protein Data Bank of either RHO family GTPases or GTPases that are not members of the RHO family onto the crystal structure of $RAC1^{P29S}$, using the program TOPP[8]. The superposition illustrates that the switch I backbone conformation of $RAC1^{P29S}$ diverges from RHO family GTPases (**Supplementary Fig. 8b**) and is similar to GTPases that are not members of the RHO family (**Supplementary Fig. 8c**). The highly conserved proline residue at position 29 in RAC1, 29 in CDC42 and 31 in RHOA is not observed in most non-RHO family GTPases. Proline at this location therefore seems to stabilize the conformation of switch I and to reduce the ability of RHO family GTPases to form hydrogen bonds between the switch I peptide backbone and ribose hydroxyls. The RAC1 P29S mutation therefore releases this conformational restraint allowing altered GTPase signal transduction. Overall, the clear electron density profile of switch I in $RAC1^{P29S}$ (**Supplementary Fig. 7**) suggests that this region of the protein is stabilized by direct hydrogen bonding between the peptide backbone and ribose hydroxyls.

Supplementary References

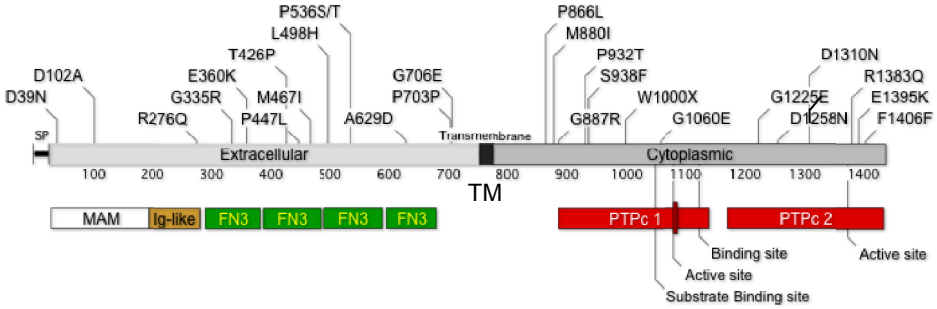
1. Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455, 1069-1075 (2008).
2. Hodgkinson, A., Chen, Y. & Eyre-Walker, A. The large-scale distribution of somatic mutations in cancer genomes. *Hum Mutat* 33, 136-143 (2012).
3. Halaban, R. *et al.* Integrative analysis of epigenetic modulation in melanoma cell response to decitabine: clinical implications. *PLoS ONE* 4, e4563 (2009).
4. Tworkoski, K. *et al.* Phosphoproteomic screen identifies potential therapeutic targets in melanoma. *Mol Cancer Res* 9, 801-812 (2011).
5. Hirshberg, M., Stockley, R.W., Dodson, G. & Webb, M.R. The crystal structure of human rac1, a member of the rho-family complexed with a GTP analogue. *Nat Struct Biol* 4, 147-152 (1997).
6. Yamashita, M. *et al.* Structural basis for the Rho- and phosphoinositide-dependent localization of the exocyst subunit Sec3. *Nat Struct Mol Biol* 17, 180-186 (2010).
7. Wu, Y.I. *et al.* A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* 461, 104-108 (2009).
8. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763 (1994).
9. Brash, D.E. *et al.* A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proc Natl Acad Sci USA* 88, 10124-10128 (1991).

10. Wennerberg, K., Rossman, K.L. & Der, C.J. The Ras superfamily at a glance. *J Cell Sci* 118, 843-846 (2005).
11. Wallace, A.C., Laskowski, R.A. & Thornton, J.M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8, 127-134 (1995).

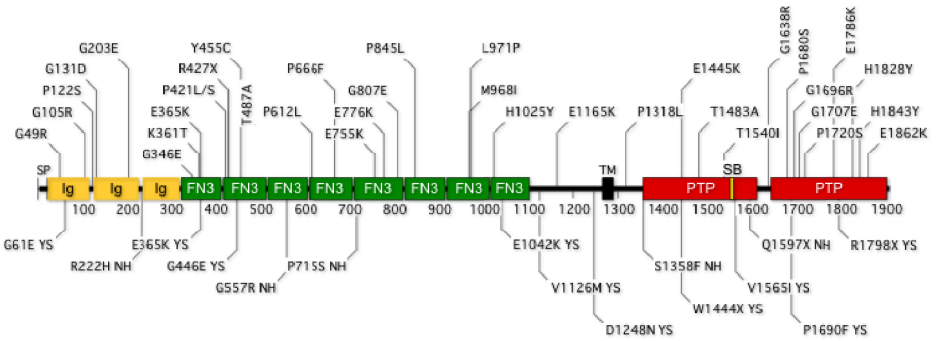


Supplementary Figure 1. Spectrum of somatic variants in introns and UV signature mutation. a, The histogram shows an excess of C>T transitions in the dipyrimidine context in sun-exposed melanomas (top) compared to sun-shielded melanomas (bottom), an indication of UV exposure and DNA damage for those melanomas or their precursors, similar to that shown for exons (Fig. 1b). b, Sanger electropherogram showing YUPROST germline DNA (PBL) compared to tumor (1.1 mm primary melanoma) at the site of the RAC1P29S mutation. The amino acid sequence of the adjacent mutation site is indicated.

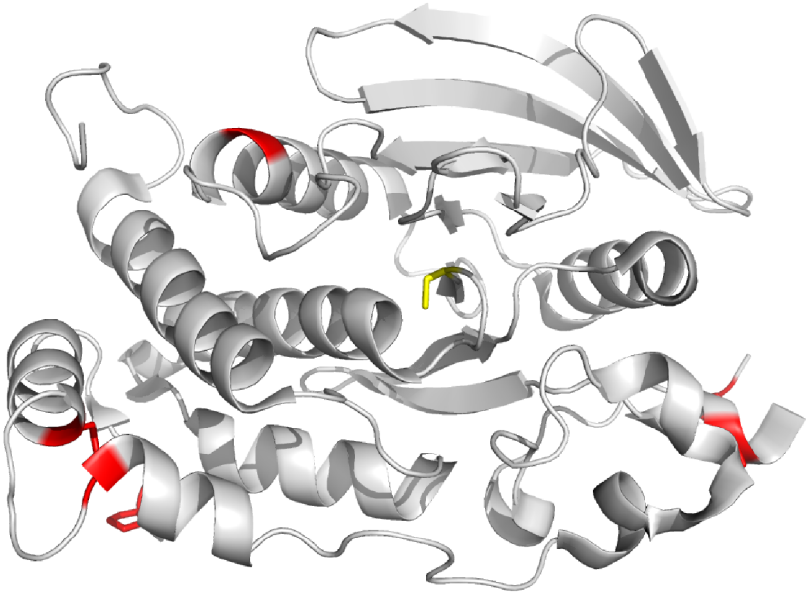
a PTPRK (protein tyrosine phosphatase, receptor type, K)



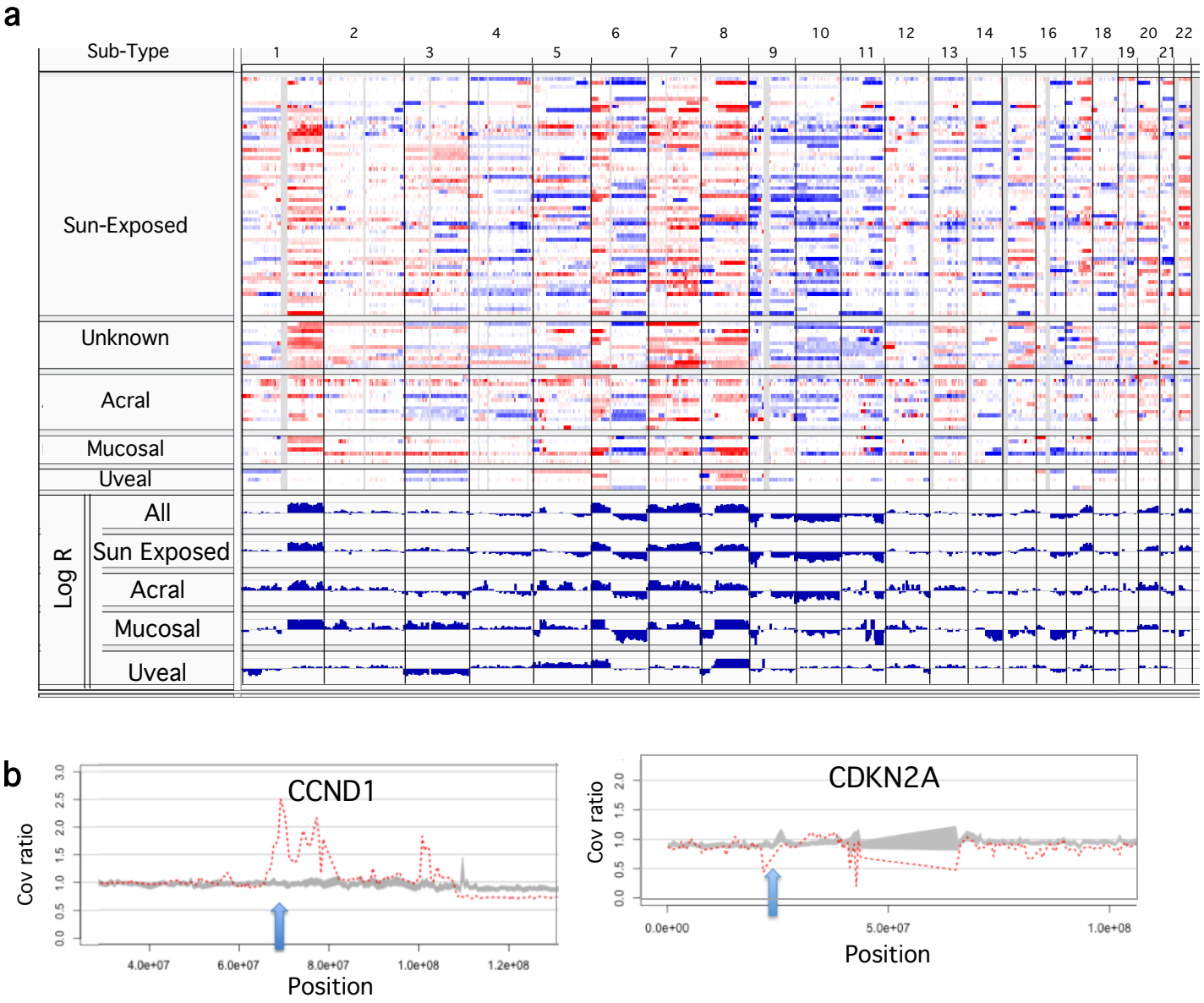
b PTPRD (protein tyrosine phosphatase, receptor type, D)



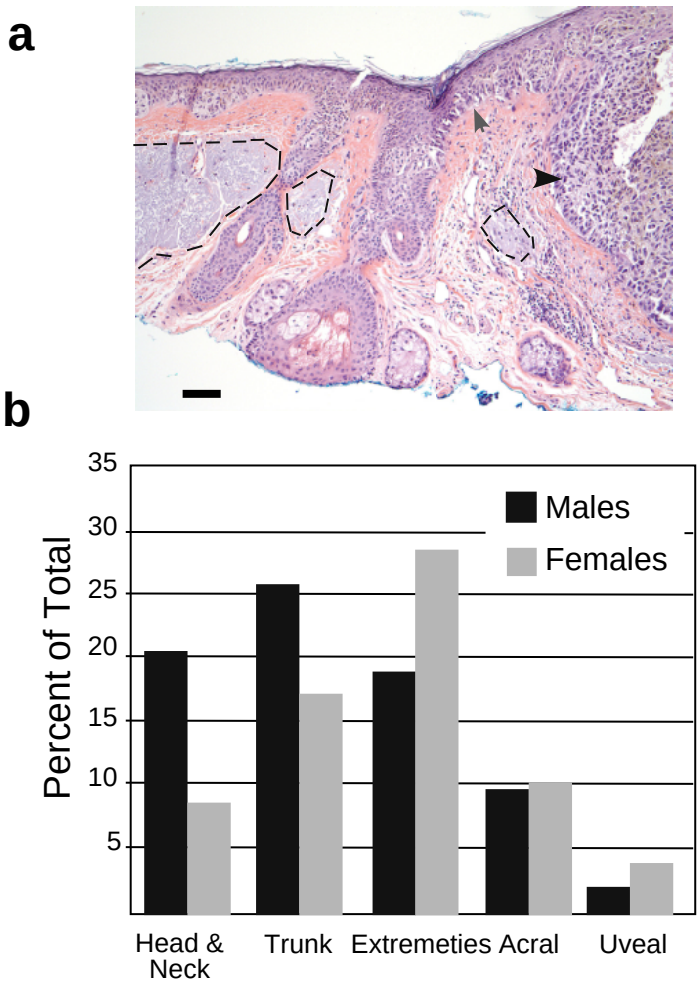
c



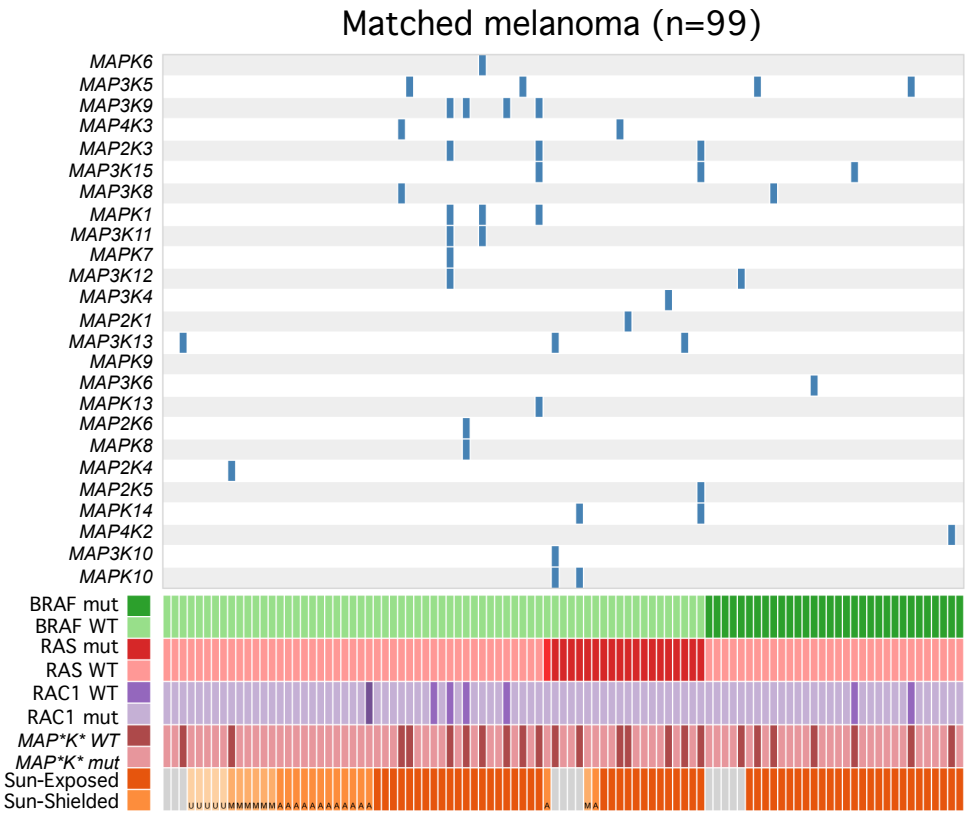
Supplementary Figure 2: PTPRK and PTPRD mutations, and PTPRK crystal structure. a, and b show schematic representations of PTPRK and PTPRD depicting functional domains and the position of mutations in the complete cohort (147 matched and unmatched melanoma samples). The published mutations in PTPRD are shown below the bar with initials of the communicating author (YS, Yardena Samuels; NH, Nicholas Hayward. c, Crystal structure of phosphatase domain 1 from PTPRK, PDB ID: 2C7S. The locations of P866L, M880I, G887R, P932T, S938F and G1060E mutations are shown in red. The active site cysteine is shown in yellow.



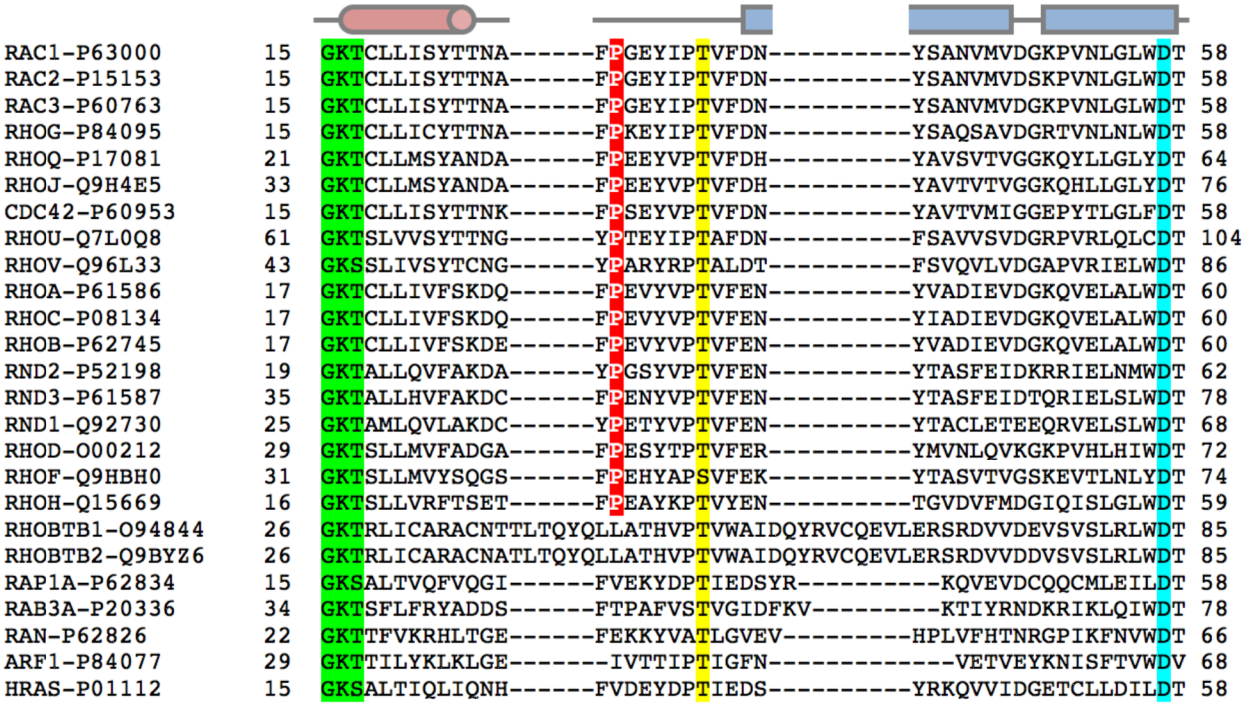
Supplementary Figure 3: Somatic copy number alterations (SCNAs) in matched melanomas. a, Heat map showing genomic aberrations across matched melanoma samples. Gain and loss regions are indicated in red and blue, respectively. The x-axis represents genomic position beginning with chromosome 1p and ending with chromosome 22 (numbered 1-22, top). The mean log fold change (R) across melanomas is featured in the bottom tracks. b, Composite plots of the normalized coverage ratio between tumor and normal samples for the overall chromosome using all 99 matched samples. Left and right show amplification in CCND1 and deletion in CDKN2A. Red is the mean of samples with the CNV; the grey shaded area is the 95% confidence interval of the mean coverage ratios of the samples without the CNV at each position. Blue arrows indicate approximate location of gene.



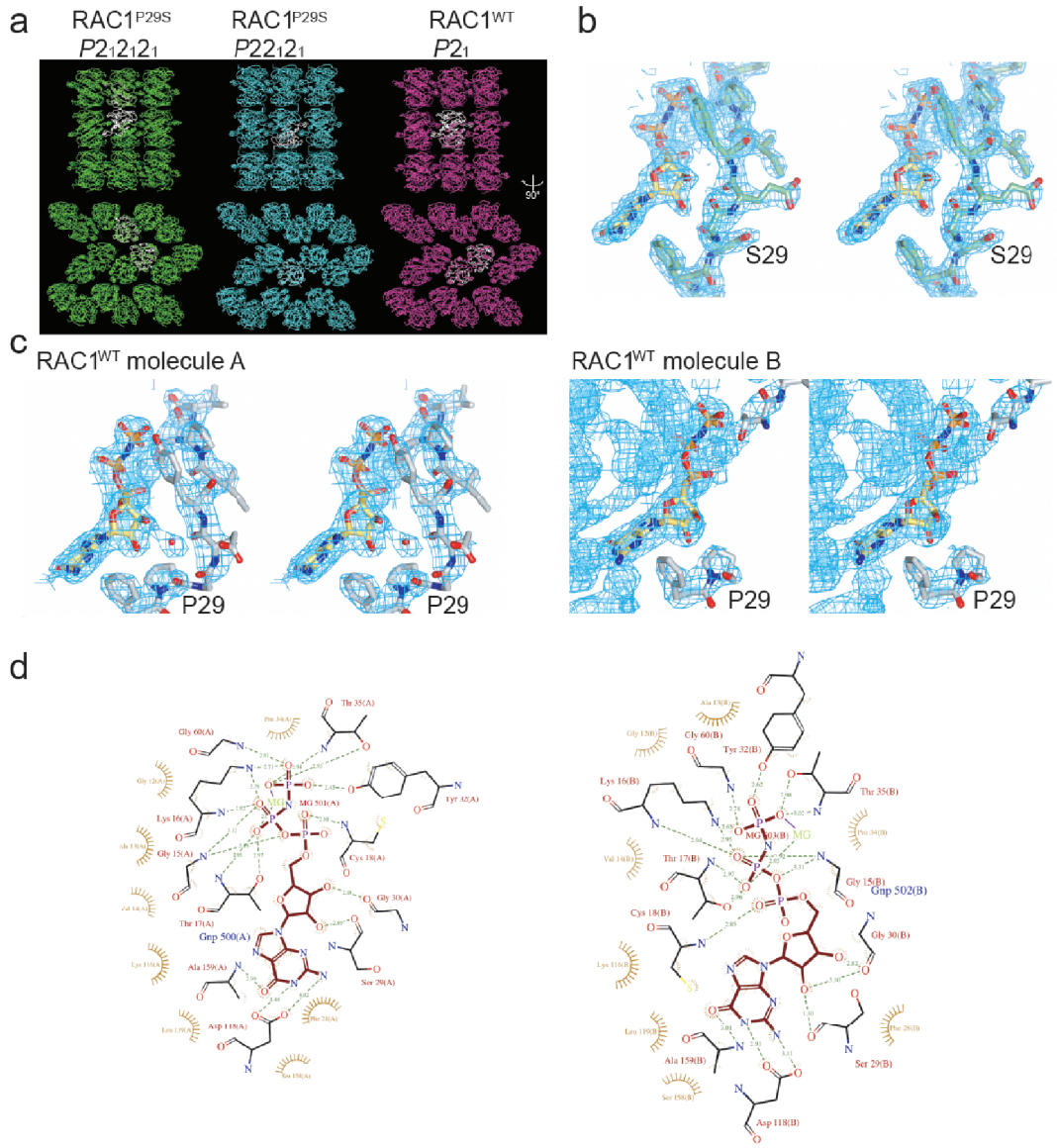
Supplementary Figure 4. Sun damage and frequency of melanoma body sitedistribution. a, YUVEME primary melanoma with RAC1P29S mutation showing extensive solar elastosis (marked with broken lines). Grey and dark arrows point at the malignant sections. Scale bar represents 0.1 mm. b, Frequency distribution of melanomas by location and sex in the Yale cohort. The data represent 312 patients composed of 144 males and 87 females. Fisher's exact test shows that the head and neck lesions are particularly significantly enriched in men compared to women with p values as follows: Head and neck: 0.0043; Trunk: 0.097; Legs and arms: 0.055; Acral: 1.0; Ocular (uveal): 0.49.



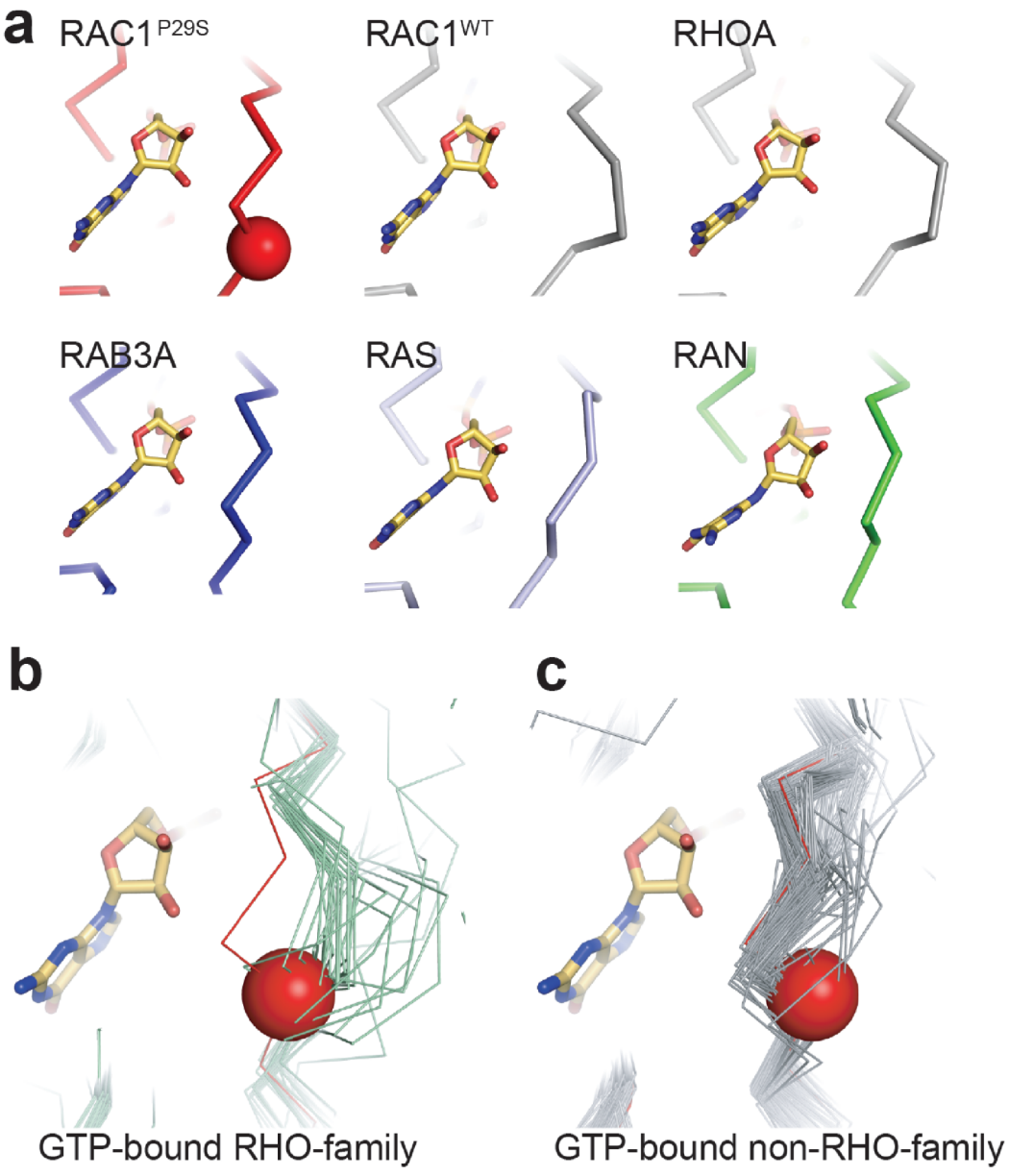
Supplementary Figure 5: Genes with somatic *MAPK* SNVs. The figure shows the occurrence of somatic MAPK SNVs across all matched melanoma samples. Blue rectangles indicate samples with somatic mutations



Supplementary Figure 6. Alignment of human RHO-family GTPases. The figure shows alignment of human RHO-family GTPases. The switch I loop region is shown. The conserved proline corresponding to codon 29 in RAC1 is highlighted in red. The Swiss-Prot ID for each protein is indicated. Representative non-RHO-family GTPases are shown. Conserved residues of the G1, G2 and G3 elements are also highlighted in green, yellow and blue, respectively[10]. Alignment made using ClustalW. Secondary structure elements for RAC1^{P29S} crystal indicated, α-helix as a cylinder, β-strand as blue rectangle, loop as line.



Supplementary Figure 7. Analysis of the Switch I region of RAC1. a, Lattices of RAC1^{P29S} and RAC1^{WT} crystals. Asymmetric unit colored white. The RAC1^{P29S} and RAC1^{WT} crystals pack in a very similar fashion. **b**, Stereoview of $2F_o$-F_c electron density for the switch I region of RAC1^{P29S} contoured at 1σ (blue) and 2σ (light blue). For clarity, electron density is clipped at 2 Å from either GMP-PNP or the switch I region. **c,** Stereoview of $2F_o$-F_c electron density for the switch I region of RAC1^{WT}. The switch I regions of both molecules (A and B) of the asymmetric unit are shown. The switch I loop shows poor electron density in molecule A and is not visible in molecule B. The wild-type crystal structure clearly shows that the switch I region of RAC1^{WT} is conformationally flexible and that this is not due to crystal packing effects. Maps for molecule A are contoured and clipped as per panel B. Maps for molecule B are contoured as per panel B and are clipped at 20 Å from GMP-PNP. **d,** Ligplot[11] diagrams for GMP-PNP bound to molecules A (left) and B (right) of the $P2_12_12_1$ crystal structure of RAC1^{P29S}.



Supplementary Figure 8. Please see next page for the legend

Supplementary Figure 8. Switch I conformations. **a,** Comparison of carbon-alpha trace for representative GTP-bound, or GTP-analogue-bound GTPase crystal structures superposed onto the $RAC1^{P29S}$ crystal structure. $RAC1^{WT}$, RHOA1 (1A2B), RAB3A (3RAB), RAS (5P21) and RAN (1RRP) shown. PDB ID in parentheses. Location of RAC1 P29S is shown as a red sphere. **b**, Superposition of representative GTP-bound, or GTP-analogue-bound, RHO family GTPases onto the crystal structure of $RAC1^{P29S}$ shows that for the switch I loop of $RAC1^{P29S}$ is conformationally divergent. Crystal structures, 1A2B, 1AM4, 1CEE, 1CXZ, 1E0A, 1GWN, 1I4T, 1KMQ, 1M7B, 1NF3, 1RYH, 1S1C, 1Z2C, 2ATX, 2FJU, 2GCO, 2GCP, 2IC5, 2ODB, 2OV2, 2QME, 2QRZ, 2RMK, 2V55, 2W2V, 2W2X, 2WKQ, 3EG5 and 3KZ1 are shown in light green. $RAC1^{P29S}$ shown in red. Location of RAC1 P29S is shown as a red sphere. RHO-family GTPases that have a similar conformation to $RAC1^{P29S}$ are discussed in the **Supplemental Text** and are not shown here. **c,** Superposition of representative GTP-bound, or GTP-analogue-bound, non-RHO-family GTPases onto the crystal structure of $RAC1^{P29S}$ shows that the switch I loop of $RAC1^{P29S}$ adopts a similar, RAS-like, conformation. Crystal structures 121P, 1AGP, 1C1Y, 1CLU, 1CTQ, 1EK0, 1G17, 1GNP, 1GNQ, 1GNR, 1GUA, 1HE8, 1HUQ, 1IAQ, 1IBR, 1JAH, 1JAI, 1K5D, 1K8R, 1KY2, 1LF0, 1LFD, 1N6H, 1N6L, 1N6N, 1N6O, 1N6P, 1N6R, 1NVU, 1NVW, 1NVX, 1OIW, 1P2S, 1P2T, 1P2U, 1P2V, 1PLJ, 1PLK, 1QBK, 1QRA, 1R2Q, 1RRP, 1RVD, 1T91, 1TU3, 1U8Y, 1UAD, 1VG0, 1X3S, 1XCM, 1XTR, 1XTS, 1YHN, 1YU9, 1YVD, 1YZK, 1YZL, 1YZN, 1YZQ, 1YZT, 1YZU, 1Z06, 1Z07, 1Z08, 1Z0J, 1Z0K, 1ZBD, 1ZC3, 1ZC4, 1ZW6, 221P, 2BME, 2C5L, 2CL0, 2CL6, 2CL7, 2D7C, 2EVW, 2EW1, 2F9M, 2FFQ, 2FG5, 2G6B, 2GIL, 2GZD, 2GZH, 2HV8, 2OCB, 2RAP, 2RGA, 2RGB, 2RGC, 2RGD, 2RGE, 2RGG, 2UZI, 2VH5, 2X19, 2ZET, 3A6P, 3BBP, 3BC1, 3CWZ, 3DDC, 3E5H, 3GFT, 3GJX, 3I3S, 3K8Y, 3K9L, 3K9N, 3KKM, 3KKN, 3KKO, 3L8Y, 3L8Z, 3LAW, 3LBH, 3LBI, 3LBN, 3M1I, 3MJH, 3NBY, 3NBZ, 3NC0, 3NC0, 3NC1, 3NKV, 3OES, 3OIU, 3OIV, 3OIW, 3PIR, 3PIT, 3QBT, 3RAB, 3RAP, 3RAP, 421P, 521P, 5P21, 621P, 6Q21, 721P and 821P are shown in grey. $RAC1^{P29S}$ shown in red. Location of RAC1 P29S is shown as a red sphere. Figure made using Pymol (www.pymol.org).

Supplementary Table 1: Characteristics of Melanoma Samples

	Matched	Unmatched	Total
Melanoma	99	48	147
Normal	99		99
Cells/Snap frozen tumors			
Frozen Tissue	68	35	103
Cell Culture	31	13	44
Gender			
Male	62	26	88
Female	37	22	59
Tumor			
Primary	35	13	48
Metastasis	64	35	99
Type			
Hair-bearing skin (Sun-exposed)	61	36	97
Acral	14	3	17
Mucosal	7	1	8
Uveal	5	1	6
Unknown Primary	12	7	19

Gene	Mutation	Total	%	Cutaneous	%	Sun-shielded	%
BRAF	V600E/K/R	49	33.3%	41	42.3%	0	0.0%
NRAS	Q61H/K/L//R	24	16.3%	19	19.6%	1	3.2%
	G12D/V	2	1.4%	1	1.0%	1	3.2%
	G13R	1	0.7%	1	1.0%	0	0.0%
HRAS	Q61H	1	0.7%	0	0.0%	1	3.2%
Total *RAS*		28	19.0%	21	21.6%	3	9.7%
*BRAF/NRAS***	V600K/G13D	1	0.7%	1	1.0%	0	0.0%
*NRAS/KIT***	Q61K/L160V	1	0.7%	1	1.0%	0	0.0%
KIT	K642E	1	0.7%	0	0.0%	1	3.2%
	L576P	2	1.4%	0	0.0%	2	6.5%
	N822Y	1	0.7%	0	0.0%	1	3.2%
	V559D	1	0.7%	0	0.0%	1	3.2%
Total *KIT*		5	3.4%	0	0.0%	5	16.1%
Wild Type*		63	42.9%	33	34.0%	23	74.2%

* Four samples showed additional somatic BRAF missense mutations: N581S, N581T, S732F, and P75L.

** Samples showed non-exclusive mutations.

Supplementary Table 2: Summary of clinical, pathological, and mutation status of melanoma samples

Sample	Dataset	Cell line / tumor	Mutation Status	Primary or Metastasis	Type	Location primary	Location of tumor excised	Sex	Age at Resection
YUAKER	matched	T	WT	primary	sun-exposed	head/neck	head/neck	M	85
YUAVEY	matched	C	*NRAS-Q61R*	metastasis	sun-exposed	extremity	extremity	F	58
YUBAN	matched	T	*BRAF-V600K*	metastasis	sun-exposed	head/neck	trunk	M	65
YUBARON	matched	T	WT	primary	sun-exposed	head/neck	head/neck	M	68
YUBATIK	unmatched	T	*BRAF-V600K*	metastasis	sun-exposed	head/neck	head/neck	M	71
YUBEM	matched	T	WT	primary	sun-exposed	head/neck	head/neck	F	58
YUBER	matched	T	WT	metastasis	sun-exposed	head/neck	lung	M	83
YUBOO	matched	T	WT	metastasis	uveal	choroid	head/neck	F	56
YUBOT	matched	T	*NRAS-Q61R*	primary	acral	heel	sole	F	80
YUBRA	unmatched	T	*BRAF-V600E*	metastasis	unknown	head/neck	head/neck	M	40
YUBRO	unmatched	T	WT	metastasis	sun-exposed	extremity	extremity	M	64
YUBRUSE	matched	C	*KIT-L576P*	metastasis	acral	sole	lymph node	M	88
YUCHIME	matched	T	*BRAF-V600K*	primary	sun-exposed	trunk	trunk	M	83
YUCHUFA	matched	C	*NRAS-Q61R*	primary	sun-exposed	extremity	extremity	F	73
YUCINJ	unmatched	C	*BRAF-V600E*	metastasis	sun-exposed	extremity	extremity	F	30
YUCIR	unmatched	T	WT	metastasis	unknown	unknown	lymph	M	50
YUCLAT	matched	C	*BRAF-V600K*	metastasis	sun-exposed	trunk	trunk	M	65
YUCOMO	unmatched	T	WT	primary	sun-exposed	head/neck	trunk	M	71
YUCOT	unmatched	C	*BRAF-V600E*	metastasis	sun-exposed	head/neck	extremity	F	33
YUCRATE	matched	T	WT	primary	acral	sole	sole	F	69
YUCRENA	matched	C	WT	primary	uveal	choroid	choroid	M	55
YUCROFT	matched	T	*BRAF-V600E*	primary	sun-exposed	trunk	trunk	M	54
YUDAB	matched	T	WT	primary	sun-exposed	head/neck	head/neck	F	73
YUDARE	matched	T	*BRAF-V600E*	metastasis	sun-exposed	trunk	lymph	M	58
YUDATE	matched	C	WT	metastasis	sun-exposed	trunk	trunk	M	81
YUDEDE	matched	C	*NRAS-Q61H*	metastasis	sun-exposed	extremity	extremity	M	83
YUDEXA	matched	T	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	M	82
YUDIALE	matched	T	*BRAF-V600E*	metastasis	unknown	unknown	lymph	M	53
YUDUTY	matched	C	*NRAS-Q61H*	metastasis	unknown	unknown	trunk	F	80
YUEGO	unmatched	T	*NRAS-Q61K*	metastasis	sun-exposed	head/neck	trunk	M	66
YUESTA	unmatched	T	WT	primary	sun-exposed	extremity	extremity	F	51
YUFARCI	matched	T	*NRAS-Q61K*	metastasis	unknown	unknown	head/neck	M	54
YUFAR	unmatched	T	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	M	48
YUFISO	matched	T	*HRAS-Q61H*	primary	acral	subungual	subungual	M	43
YUFLA	matched	T	*NRAS-Q61K*	metastasis	sun-exposed	head/neck	head/neck	M	52
YUFOLD	matched	T	*NRAS-Q61L*	primary	sun-exposed	extremity	extremity	F	92

Supplementary Table 2: Summary of clinical, pathological, and mutation status of melanoma samples

Sample	Dataset	Cell line / tumor	Mutation Status	Primary or Metastasis	Type	Location primary	Location of tumor excised	Sex	Age at Resection
YUFOR	unmatched	T	*NRAS-Q61L*	metastasis	sun-exposed	extremity	lymph	F	90
YUFUZZ	matched	T	WT	metastasis	unknown	unknown	lymph	F	85
YUGADID	matched	T	WT	metastasis	uveal	choroid	extremity	F	75
YUGAFFE	matched	C	*BRAF-V600K*	metastasis	sun-exposed	head/neck	lymph	M	39
YUGAL	unmatched	T	WT	metastasis	sun-exposed	head/neck	head/neck	M	80
YUGANK	unmatched	C	*NRAS-Q61K*	metastasis	sun-exposed	head/neck	brain	M	78
YUGASP	matched	T	*NRAS-Q61L*	metastasis	sun-exposed	extremity	lymph	F	88
YUGELU	unmatched	T	*NRAS-Q61L*	metastasis	sun-exposed	head/neck	head/neck	M	84
YUGISMO	matched	T	*NRAS-Q61K, KIT-L160V*	metastasis	sun-exposed	trunk	trunk	M	73
YUGLAUD	unmatched	T	*NRAS-Q61K*	primary	sun-exposed	trunk	trunk	M	57
YUGLIDE	matched	C	WT	primary	uveal	choroid	choroid	F	33
YUGOE	matched	C	*NRAS-G12V*	metastasis	sun-exposed	head/neck	head/neck	M	54
YUGONZO	matched	T	WT	primary	sun-exposed	extremity	extremity	F	61
YUHAY	unmatched	T	*BRAF-V600E*	primary	sun-exposed	extremity	extremity	F	34
YUHEF	matched	C	WT	metastasis	sun-exposed	head/neck	trunk	M	53
YUHIMO	matched	T	WT	metastasis	acral	heel	extremity	M	57
YUHOIN	matched	C	WT	metastasis	mucosal	nasal cavity	lymph	F	59
YUHOOD	matched	T	WT	metastasis	mucosal	gingiva	trunk	M	75
YUHUY	matched	T	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	M	64
YUIRI	matched	T	WT	metastasis	acral	sole	sole	F	25
YUISKIA	matched	T	WT	metastasis	acral	toe	trunk	M	69
YUJUMP	matched	T	WT	primary	sun-exposed	extremity	extremity	M	58
YUKADI	matched	C	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	M	56
YUKARN	unmatched	C	*BRAF-V600K, NRAS-G13D*	metastasis	sun-exposed	trunk	lymph	F	86
YUKAY	unmatched	T	WT	metastasis	unknown	unknown	trunk	F	47
YUKERE	unmatched	T	WT	primary	mucosal	vulva	vulva	F	89
YUKIL	matched	T	*BRAF-V600E*	metastasis	sun-exposed	head/neck	extremity	M	73
YUKLAB	matched	T	WT	metastasis	sun-exposed	unknown	trunk	M	84
YUKNOLL	unmatched	T	WT	metastasis	unknown	unknown	liver	F	81
YUKOLI	unmatched	C	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	M	53
YUKOS	unmatched	T	WT	primary	sun-exposed	trunk	trunk	M	87
YUKRIN	matched	T	WT	metastasis	sun-exposed	trunk	brain	F	37
YULAC	unmatched	C	*BRAF-V600K*	metastasis	sun-exposed	trunk	head/neck	F	65
YULAN	matched	T	WT	metastasis	sun-exposed	head/neck	lymph	M	81

Supplementary Table 2: Summary of clinical, pathological, and mutation status of melanoma samples

Sample	Dataset	Cell line / tumor	Mutation Status	Primary or Metastasis	Type	Location primary	Location of tumor excised	Sex	Age at Resection
YULAPE	matched	C	*NRAS-Q61H*	metastasis	unknown	unknown	trunk	M	75
YULAXER	matched	C	*NRAS-Q61K*	metastasis	sun-exposed	trunk	lymph	F	60
YULLON	matched	T	*NRAS-Q61H*	metastasis	unknown	unknown	brain	F	56
YULOMA	matched	C	*NRAS-Q61R*	metastasis	sun-exposed	extremity	trunk	M	62
YULOVY	unmatched	C	*NRAS-Q61L*	primary	sun-exposed	extremity	extremity	F	83
YULVER	matched	T	WT	metastasis	sun-exposed	head/neck	lung	M	64
YUMAC	unmatched	C	*BRAF-V600K*	metastasis	sun-exposed	extremity	extremity	M	53
YUMAN	unmatched	T	*NRAS-G13R*	metastasis	sun-exposed	trunk	lymph	F	77
YUMER	matched	T	WT	metastasis	sun-exposed	head/neck	head/neck	M	94
YUMINE	unmatched	C	*BRAF-V600E*	metastasis	unknown	unknown	liver	F	59
YUMOOK	matched	T	*BRAF-V600E*	primary	sun-exposed	extremity	extremity	F	75
YUMOYA	matched	T	*BRAF-V600E*	primary	sun-exposed	trunk	trunk	M	58
YUMUDE	matched	T	*KIT-V559D*	primary	acral	sole	sole	M	79
YUMUT	matched	C	*BRAF-V600E*	metastasis	sun-exposed	extremity	trunk	M	44
YUNACK	matched	T	*BRAF-V600E*	primary	sun-exposed	trunk	trunk	F	59
YUNELU	unmatched	T	*NRAS-Q61R*	metastasis	sun-exposed	extremity	extremity	F	81
YUNEON	matched	T	WT	primary	acral	sole	sole	M	68
YUNICA	matched	T	WT	metastasis	mucosal	vulva	liver	F	63
YUNOCA	matched	T	WT	primary	mucosal	nasal cavity	mucosal	F	53
YUNUFF	matched	C	*BRAF-V600E*	metastasis	sun-exposed	trunk	pleural cavity	F	59
YUNUVO	matched	T	WT	primary	acral	sole	sole	M	64
YUPADI	unmatched	T	WT	primary	sun-exposed	head/neck	head/neck	M	86
YUPAF	unmatched	T	*BRAF-V600K*	metastasis	unknown	unknown	chest wall	M	63
YUPAL	matched	T	WT	primary	sun-exposed	head/neck	head/neck	M	64
YUPANG	unmatched	T	WT	metastasis	acral	sole	sole	M	81
YUPAT	matched	T	WT	metastasis	sun-exposed	trunk	lung	M	52
YUPEET	matched	C	*BRAF-V600E*	primary	sun-exposed	trunk	trunk	M	54
YUPER	matched	T	*BRAF-V600E*	metastasis	sun-exposed	trunk	extremity	M	63
YUPORCH	matched	T	WT	metastasis	unknown	unknown	intestine	F	75
YUPROST	matched	T	WT	primary	sun-exposed	head/neck	head/neck	F	86
YUPRO	matched	T	*BRAF-V600K*	primary	sun-exposed	extremity	extremity	M	53
YUPYKO	matched	T	WT	primary	sun-exposed	head/neck	head/neck	M	63
YURDE	matched	C	*BRAF-V600E*	metastasis	sun-exposed	sun-exposed	trunk	M	55
YURED	matched	C	*BRAF-V600E*	metastasis	sun-exposed	extremity	trunk	F	67
YURIDA	matched	T	*NRAS-Q61R*	metastasis	sun-exposed	head/neck	head/neck	M	61
YURIF	matched	C	*BRAF-V600K*	metastasis	sun-exposed	extremity	extremity	M	53

Supplementary Table 2: Summary of clinical, pathological, and mutation status of melanoma samples

Sample	Dataset	Cell line / tumor	Mutation Status	Primary or Metastasis	Type	Location primary	Location of tumor excised	Sex	Age at Resection
YURIMO	unmatched	T	WT	primary	sun-exposed	head/neck	head/neck	M	87
YURKEN	matched	C	*BRAF-V600E*	metastasis	sun-exposed	trunk	lymph	F	52
YUROO	unmatched	T	WT	metastasis	sun-exposed	unknown	extremity	M	64
YURTHE	matched	T	WT	metastasis	unknown	unknown	lymph	F	81
YURUB	matched	T	KIT-N822Y	metastasis	acral	toe	lymph	M	71
YURUS	matched	T	WT	primary	sun-exposed	extremity	extremity	M	90
YUSAG	matched	T	*KIT-K642E*	metastasis	acral	finger	finger	F	77
YUSAN	matched	C	*NRAS-G12D*	metastasis	mucosal	vulva	trunk	F	58
YUSARI	matched	C	*BRAF-V600E*	metastasis	sun-exposed	head/neck	pleural fluid	M	49
YUSAT	unmatched	T	*BRAF-V600E*	metastasis	sun-exposed	head/neck	head/neck	F	31
YUSCH	unmatched	T	*KIT-L576P*	metastasis	acral	thumb	thumb	M	64
YUSCO	matched	T	*BRAF-V600E*	primary	sun-exposed	trunk	lymph	F	65
YUSEL	matched	T	*BRAF-V600K*	metastasis	unknown	unknown	lymph	M	53
YUSIK	unmatched	C	*BRAF-V600E*	metastasis	sun-exposed	extremity	trunk	F	50
YUSOT	matched	T	WT	metastasis	mucosal	nasal cavity	gall bladder	F	70
YUSTE	unmatched	C	*BRAF-V600E*	metastasis	sun-exposed	extremity	extremity	F	66
YUSUBA	matched	C	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	F	37
YUSWI	matched	C	*BRAF-V600E*	metastasis	sun-exposed	trunk	small intestine	M	57
YUTALO	unmatched	T	*BRAF-V600E*	metastasis	sun-exposed	trunk	trunk	F	46
YUTAZI	unmatched	T	WT	metastasis	unknown	unknown	brain	F	51
YUTEL	unmatched	T	WT	primary	sun-exposed	extremity	extremity	F	81
YUTEPA	matched	T	*NRAS-Q61R*	primary	sun-exposed	trunk	trunk	M	56
YUTER	unmatched	C	*NRAS-Q61L*	metastasis	sun-exposed	extremity	extremity	M	73
YUTOGS	unmatched	T	WT	primary	sun-exposed	head/neck	head/neck	M	50
YUTOLL	unmatched	T	BRAF-V600K	metastasis	sun-exposed	trunk	trunk	M	67
YUTRE	matched	T	WT	metastasis	acral	thumb	thumb	M	81
YUTRIP	matched	T	WT	primary	acral	sole	sole	M	78
YUVAIL	matched	T	*BRAF-V600K*	primary	sun-exposed	trunk	trunk	F	62
YUVAN	unmatched	T	WT	metastasis	uveal	choroid	trunk	F	68
YUVEDO	matched	T	WT	primary	uveal	choroid	trunk	M	48
YUVIL	unmatched	T	WT	metastasis	sun-exposed	extremity	extremity	F	50
YUWAGE	matched	C	WT	primary	sun-exposed	head/neck	trunk	M	86
YUWALI	matched	C	*BRAF-V600E*	metastasis	unknown	unknown	lymph	M	41
YUWAND	matched	T	WT	primary	sun-exposed	head/neck	head/neck	M	78
YUWHIM	matched	C	*BRAF-V600E*	metastasis	unknown	unknown	trunk	M	65
YUWIC	unmatched	T	WT	primary	acral	thumb	thumb	M	40

Supplementary Table 2: Summary of clinical, pathological, and mutation status of melanoma samples

Sample	Dataset	Cell line / tumor	Mutation Status	Primary or Metastasis	Type	Location primary	Location of tumor excised	Sex	Age at Resection
YUWISE	matched	T	WT	metastasis	mucosal	anus	pleural fluid	F	63
YUXALT	unmatched	T	WT	primary	sun-exposed	head/neck	head/neck	M	86
YUZEAL	unmatched	C	*BRAF-V600R*	metastasis	sun-exposed	trunk	pleural fluid	M	78
YUZEST	matched	T	*BRAF-V600E*	metastasis	unknown	unknown	trunk	M	55
YUZINO	matched	T	*NRAS-Q61L*	metastasis	sun-exposed	extremity	extremity	F	80

Supplementary Table 3: Mutation rates by sequence context

Sequence context	Number of sites per exome	Mutation frequency
Any	2.98E+07	1.92E-05
TC*	2.88E+06	5.53E-05
CC*	3.18E+06	2.59E-05
AC*	2.33E+06	5.64E-06
GC*	2.70E+06	6.48E-06
Other dinucleotides	1.87E+07	1.61E-05
TTTC*CT	3.49E+04	1.93E-04
CTTC*CT	3.59E+04	1.41E-04
ATTC*CT	1.84E+04	1.26E-04
GTTC*CT	1.65E+04	1.28E-04
TTTC*GT	3.38E+03	5.83E-04
CTTC*GT	5.08E+03	4.12E-04
ATTC*GT	2.54E+03	4.07E-04
GTTC*GT	2.42E+03	3.93E-04
Other hexanucleotides	2.97E+07	1.86E-05

The mutation-site motif was determined from the frequency of individual sequence contexts flanking high quality somatic mutations in the exome capture region.

Supplementary Table 4: Novel recurrent somatic SNVs across the melanoma exome screen

Gene Symbol	Accession	Chr	Position	Reference Genotype	Variant Genotype	Amino Acid Change	Total	Expression	COSMIC	Phylop Cons. Score	p-fam domain
RAC1	NM_018890.3	chr7	6426892	C/C	C/T	P29S	7	+	Y	6.25	Ras
DBC1	NM_014618.2	chr9	121929759	C/C	C/T T/T	R630Q	6	+	Y	5.76	
CAPN6	NM_014289.3	chrX	110496372	G/G	G/A A/A	R124C	4	+	N	4.17	Peptidase_C2
LOXHD1	NM_144612.6	chr18	44114381	G/G	G/A	R1377W	4	-	N	2.84	PLAT
OR4N2	NM_001004723.1	chr14	20295729	G/G	G/A	G41E	4	-	N	1.97	7tm_1
OR5T2	NM_001004746.1	chr11	56000423	C/C	C/T	G80E	4	-	Y	4.50	7tm_1
PCDHGA1	NM_018912.2	chr5	140711128	C/C	C/T	R293C	4	+	Y	0.43	Cadherin
PPP6C	NM_001123355.1	chr9	127912080	G/G	G/A	R301C	4	+	N	3.81	
PRIMA1	NM_178013.3	chr14	94187873	C/C	C/T	E127K	4	-	Y	2.74	
RGS7	NM_002924.4	chr1	241262011	G/G	G/A	R44C	4	+	N	4.94	DEP
SERPINA10	NM_016186.2	chr14	94754734	C/C	C/T	G294E	4	-	N	2.16	Serpin
SNCAIP	NM_005460.2	chr5	121786742	C/C	C/T	P734S	4	-	N	3.91	
TMC5	NM_001105248.1	chr16	19485576	G/G	G/A	E690K	4	-	N	0.28	Sugar_tr
TP53	NM_001126112.1	chr17	7574003	G/G	G/A A/A	R342X	4	+	Y	0.84	
ACVR1C	NM_145259.2	chr2	158395120	G/G	G/A	R441X	3	+	N	1.10	
ADCY8	NM_001115.2	chr8	131792904	C/C	C/T	G1163E	3	-	N	5.79	Guanylate_cyc
ANK3	NM_020987.3	chr10	62023696	C/C	C/T	G199E	3	-	N	5.94	
C15orf2	NM_018958.2	chr15	24922056	C/C	C/T	R348X	3	+	N	-0.03	
C1orf150	NM_145278.3	chr1	247712498	G/G	G/A	G2E	3	-	N	2.20	
C1orf168	NM_001004303.4	chr1	57233561	G/G	G/A	S335L	3	-	N	2.97	
C6	NM_001115131.1	chr5	41199882	G/G	G/A	R145C	3	-	N	1.76	Ldl_recept_a
C6	NM_001115131.1	chr5	41161920	G/G	G/A	R445X	3	-	N	1.06	
CCDC60	NM_178499.3	chr12	119866561	G/G	G/A	R55Q	3	+	N	2.79	
CD1C	NM_001765.2	chr1	158261016	G/G	G/A	E52K	3	-	N	0.94	MHC_I
CDH6	NM_004932.2	chr5	31317540	C/C	C/T	S524L	3	-	N	3.45	Cadherin
CDH9	NM_016279.3	chr5	26885885	C/C	C/T	D574N	3	-	N	4.03	Cadherin
CFHR3	NM_021023.5	chr1	196748927	G/G	G/A A/A	R85K	3	+	N	0.24	
CYP2C8	NM_000770.3	chr10	96798741	G/G	G/A	P402S	3	-	N	0.27	p450
DBC1	NM_014618.2	chr9	122075525	C/C	C/T	E37K	3	+	N	3.69	
DGKI	NM_004717.2	chr7	137206693	G/G	G/A	R723C	3	+	N	3.40	
DNAH5	NM_001369.2	chr5	13753355	C/C	C/T	R3620Q	3	-	N	-0.93	

Supplementary Table 4: Novel recurrent somatic SNVs across the melanoma exome screen

Gene Symbol	Accession	Chr	Position	Reference Genotype	Variant Genotype	Amino Acid Change	Total	Expression	COSMIC	Phylop Cons. Score	p-fam domain
DNAH5	NM_001369.2	chr5	13885322	C/C	C/T	G920E	3	-	N	0.23	
DNAH5	NM_001369.2	chr5	13692194	G/G	G/A	R4592X	3	-	N	4.14	
DNAH5	NM_001369.2	chr5	13753599	G/G	G/A	R3539C	3	-	N	6.31	
DYNC1I1	NM_004411.4	chr7	95726852	C/C	C/T	R629C	3	+	N	5.14	
E2F1	NM_005225.2	chr20	32267771	G/G	G/A	S121F	3	+	N	4.16	
GABRB2	NM_021911.2	chr5	160886715	C/C	C/T	D125N	3	-	N	5.91	Neur_chan_LB
GIMAP7	NM_153236.3	chr7	150217300	G/G	G/A	E80K	3	+	N	2.87	AIG1
GPR20	NM_005293.2	chr8	142366972	G/G	G/A	A351V	3	-	N	1.36	
GRID2	NM_001510.2	chr4	94344033	G/G	G/A	E487K	3	+	N	5.95	Lig_chan-Glu_
ISX	NM_001008494.1	chr22	35478537	C/C	C/T	R86C	3	-	N	1.47	Homeobox
KCNH7	NM_033272.3	chr2	163302566	G/G	G/A	P506S	3	+	N	6.13	Ion_trans
KCNH7	NM_033272.3	chr2	163241264	C/C	C/T	D966N	3	+	N	-1.20	
KCNT2	NM_198503.2	chr1	196398861	C/C	C/T	R222Q	3	-	N	5.56	Ion_trans_2
KIAA1324L	NM_001142749.2	chr7	86509846	C/C	C/T	E1011K	3	+	N	4.15	
KL	NM_004795.3	chr13	33628324	G/G	G/A	E414K	3	-	N	4.37	
LRP2	NM_004525.2	chr2	170014006	C/C	C/T	G3965E	3	+	N	2.89	
MSR1	NM_138715.2	chr8	16026278	C/C	C/T	E107K	3	-	Y	0.31	
MSR1	NM_138715.2	chr8	16026295	C/C	C/T	G101E	3	-	N	0.42	
NELL1	NM_006157.3	chr11	21581830	C/C	C/T	P628S	3	+	N	1.86	EGF_CA
NETO1	NM_138966.3	chr18	70526090	C/C	C/T	G147E	3	-	N	5.82	CUB
NR3C2	NM_000901.4	chr4	149356367	G/G	G/A	S549F	3	+	N	2.77	
NRG3	NM_001165972.1	chr10	84744883	G/G	G/A	R537Q	3	+	N	1.83	Neuregulin
OPN5	NM_181744.3	chr6	47759679	G/G	G/A	R131Q	3	-	N	5.55	7tm_1
OR13C8	NM_001004483.1	chr9	107332146	G/G	G/A	G233E	3	-	Y	3.64	7tm_1
OR4A15	NM_001005275.1	chr11	55135452	G/G	G/A	M31I	3	-	Y	2.73	
OR4K1	NM_001004063.2	chr14	20404475	C/C	C/T	S217F	3	-	N	1.92	7tm_1
OR4M1	NM_001005500.1	chr14	20249284	C/C	C/T	S268F	3	+	N	3.21	7tm_1
OR6C1	NM_001005182.1	chr12	55714592	C/C	C/T	S70L	3	-	N	0.96	7tm_1
PAH	NM_000277.1	chr12	103288633	C/C	C/T	E78K	3	-	N	4.19	ACT
PRSS58	NM_001001317.3	chr7	141955085	C/C	C/T	E76K	3	-	N	3.83	Trypsin
RICTOR	NM_152756.3	chr5	38950699	G/G	G/A	S1084L	3	+	N	5.51	

Supplementary Table 4: Novel recurrent somatic SNVs across the melanoma exome screen

Gene Symbol	Accession	Chr	Position	Reference Genotype	Variant Genotype	Amino Acid Change	Total	Expression	COSMIC	Phylop Cons. Score	p-fam domain
RNF217	NM_152553.2	chr6	125397950	C/C	C/T	R185C	3	+	N	4.29	
RQCD1	NM_005444.1	chr2	219449406	C/C CC/CC	CC/TT C/T	P131L	3	+	N	6.04	Rcd1
S100A7	NM_002963.3	chr1	153430314	C/C	C/T	G92R	3	-	N	0.45	
SLC22A2	NM_003058.3	chr6	160671634	G/G	G/A A/A	R207C	3	-	N	3.66	
SLC27A6	NM_014031.3	chr5	128368943	G/G	G/A	D610N	3	-	N	0.63	
STAC	NM_003149.1	chr3	36484932	G/G	G/A A/A	R63Q	3	-	N	1.68	
TRIM58	NM_015431.3	chr1	248039730	G/G	G/A	G467E	3	+	N	-0.03	
TSHZ2	NM_173485.5	chr20	51871450	G/G	G/A	E485K	3	+	N	4.25	
TUBA3C	NM_006001.2	chr13	19751395	C/C	C/T	R243Q	3	+	N	1.36	Tubulin
UPB1	NM_016327.2	chr22	24909338	G/G	G/A	R169Q	3	-	N	5.62	CN_hydrolase
WDR49	NM_178824.3	chr3	167293784	C/C	C/T	M136I	3	-	N	2.40	
WNK3	NM_020922.4	chrX	54263821	G/G	G/A A/A	S1393L	3	+	N		
ZNF385D	NM_024697.2	chr3	21462821	C/C	C/T	R358Q	3	-	N	2.99	
ZNF536	NM_014717.1	chr19	30936347	G/G	G/A	M626I	3	-	N	0.42	

Supplementary Table 5: Significantly mutated genes in sun-exposed melanomas

a) Comprehensive Model (BH P-value <0.05)

	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
1	*BRAF*	+	2290	28	36	0	0.25	3.43	5.46	6.65E-47	3.48E-43	1.05E-42
2	*NRAS*	+	565	13	13	0	0.54	3.63	4.96	1.27E-21	4.99E-18	1.99E-17
3	*DCC*	+	4171	21	35	6	0.09	2.80	4.25	2.02E-12	6.34E-09	3.17E-08
4	*FAM5C*	-	1886	14	19	2	0.16	2.84	4.39	1.63E-10	4.28E-07	2.57E-06
5	*ADAM7*	-	2213	14	20	2	0.20	0.95	1.30	1.92E-09	3.88E-06	3.02E-05
6	*TNC*	+	5819	11	20	2	0.05	2.21	3.72	1.98E-09	3.88E-06	3.1E-05
7	*TP53*	+	1112	9	9	0	0.56	1.27	2.63	5.04E-08	8.8E-05	7.92E-04
8	*PTPRK*	+	4118	12	17	1	0.18	3.25	4.95	1.37E-07	2.16E-04	2.16E-03
9	*PPP6C*	+	928	8	9	0	0.33	3.06	4.10	1.61E-07	2.17E-04	2.52E-03
10	*TLR4*	+	1802	8	10	2	0.00	1.03	0.71	1.65E-07	2.17E-04	2.6E-03
11	*DSG4*	-	2816	13	22	3	0.14	1.50	1.98	2E-07	2.42E-04	3.15E-03
12	*CD163L1*	+	3877	15	24	2	0.04	0.56	0.79	2.85E-07	3.2E-04	4.48E-03
13	*FAM83B*	-	2172	12	18	2	0.17	1.78	2.71	3.24E-07	3.39E-04	5.09E-03
14	*TNR*	-	3753	13	22	6	0.14	2.62	3.24	4.36E-07	4.28E-04	6.85E-03
15	*GRM3*	+	2124	12	15	2	0.27	3.11	3.72	6.43E-07	5.95E-04	0.01
16	*C6*	-	2726	15	18	3	0.06	1.57	1.42	1.21E-06	1.05E-03	0.02
17	*GRIN3A*	-	2727	9	15	3	0.00	2.70	4.16	2.4E-06	1.98E-03	0.04
18	*CASR*	-	2487	10	14	5	0.07	2.69	2.82	6.3E-06	4.23E-03	0.1
19	*OR4K15*	-	741	7	8	0	0.13	0.78	1.29	6.45E-06	4.23E-03	0.1
20	*SI*	-	5435	16	30	5	0.17	1.72	2.57	9.5E-06	5.79E-03	0.15
21	*TPTE*	-	1512	12	12	2	0.08	0.59	1.06	9.58E-06	5.79E-03	0.15
22	*CD163*	-	3000	13	21	6	0.10	1.27	1.32	1.33E-05	7.44E-03	0.21
23	*C15orf2*	+	2248	13	17	4	0.12	-0.29	-0.16	1.37E-05	7.44E-03	0.22
24	*WDR49*	-	2031	9	12	1	0.17	1.22	1.98	1.55E-05	8.1E-03	0.24
25	*SLC15A2*	+	2186	11	14	1	0.07	2.02	3.52	1.67E-05	8.19E-03	0.26
26	*C1orf168*	-	1914	12	14	4	0.14	0.61	0.60	2.05E-05	8.96E-03	0.32
27	*CAPZA3*	-	606	8	8	1	0.25	1.19	2.43	2.96E-05	0.01	0.47
28	*RAC1*	+	607	6	6	0	0.17	3.48	6.27	3.12E-05	0.01	0.49
29	*MAGEC1*	+	2245	8	16	3	0.06	0.13	0.56	4.29E-05	0.02	0.67

Supplementary Table 5: Significantly mutated genes in sun-exposed melanomas

	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
30	*GJB2*	-	505	4	4	0	0.00	2.82	4.02	4.57E-05	0.02	0.72
31	*CDH9*	-	2125	11	14	3	0.21	2.46	4.15	4.58E-05	0.02	0.72
32	*ARMC4*	-	3050	13	17	5	0.06	1.92	2.39	5.76E-05	0.02	0.9
33	*USH1C*	-	2545	10	14	1	0.07	2.24	2.32	5.95E-05	0.02	0.94
34	*FAM49A*	-	972	7	9	1	0.00	3.32	4.48	6.45E-05	0.02	1
35	*FSTL5*	-	2367	9	13	2	0.15	2.50	3.47	7.39E-05	0.02	1
36	*JAKMIP2*	+	2356	12	14	4	0.14	3.19	4.80	9E-05	0.03	1
37	*EYA2*	-	1615	8	11	2	0.18	2.76	3.34	1.49E-04	0.04	1
38	*ZNF385D*	-	1125	9	10	1	0.00	2.92	3.12	1.79E-04	0.05	1

b) No Expression Model (top 50 genes)

Rank	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
1	*BRAF*	+	2290	28	36	0	0.25	3.43	5.46	5.03E-51	3.95E-47	7.91E-47
2	*NRAS*	+	565	13	13	0	0.54	3.63	4.96	4.99E-21	2.61E-17	7.84E-17
3	*FAM5C*	-	1886	14	19	2	0.16	2.84	4.39	9.94E-14	3.12E-10	1.56E-09
4	*ADAM7*	-	2213	14	20	2	0.20	0.95	1.30	5.11E-13	1.34E-09	8.04E-09
5	*DSG4*	-	2816	13	22	3	0.14	1.50	1.98	4.28E-11	9.61E-08	6.73E-07
6	*CD163L1*	+	3877	15	24	2	0.04	0.56	0.79	8.24E-11	1.55E-07	1.29E-06
7	*TNR*	-	3753	13	22	6	0.14	2.62	3.24	9.78E-11	1.55E-07	1.54E-06
8	*SI*	-	5435	16	30	5	0.17	1.72	2.57	2.31E-10	3.3E-07	3.63E-06
9	*C6*	-	2726	15	18	3	0.06	1.57	1.42	6.99E-10	8.03E-07	1.1E-05
10	*FAM83B*	-	2172	12	18	2	0.17	1.78	2.71	7.16E-10	8.03E-07	1.12E-05
11	*GRM3*	+	2124	12	15	2	0.27	3.11	3.72	1.03E-09	1.08E-06	1.62E-05
12	*CD163*	-	3000	13	21	6	0.10	1.27	1.32	5.32E-09	5.22E-06	8.36E-05
13	*C15orf2*	+	2248	13	17	4	0.12	-0.29	-0.16	1.51E-08	1.39E-05	2.37E-04
14	*GRIN3A*	-	2727	9	15	3	0.00	2.70	4.16	1.85E-08	1.62E-05	2.91E-04
15	*C1orf168*	-	1914	12	14	4	0.14	0.61	0.60	6.56E-08	5.33E-05	1.03E-03
16	*SLC15A2*	+	2186	11	14	1	0.07	2.02	3.52	6.79E-08	5.33E-05	1.07E-03
17	*CASR*	-	2487	10	14	5	0.07	2.69	2.82	7.31E-08	5.47E-05	1.15E-03
18	*MAGEC1*	+	2245	8	16	3	0.06	0.13	0.56	7.66E-08	5.47E-05	1.2E-03

Supplementary Table 5: Significantly mutated genes in sun-exposed melanomas

Rank	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
19	*CADM2*	-	1243	9	12	1	0.00	2.95	3.59	8.23E-08	5.62E-05	1.29E-03
20	*ARMC4*	-	3050	13	17	5	0.06	1.92	2.39	1.01E-07	6.16E-05	1.59E-03
21	*TPTE*	-	1512	12	12	2	0.08	0.59	1.06	1.01E-07	6.16E-05	1.59E-03
22	*MYOCD*	-	2351	11	15	3	0.33	1.88	2.69	1.09E-07	6.32E-05	1.71E-03
23	*USH1C*	-	2545	10	14	1	0.07	2.24	2.32	2.16E-07	1.13E-04	3.4E-03
24	*JAKMIP2*	+	2356	12	14	4	0.14	3.19	4.80	3.44E-07	1.75E-04	5.41E-03
25	*CDH9*	-	2125	11	14	3	0.21	2.46	4.15	3.9E-07	1.9E-04	6.13E-03
26	*OR4K15*	-	741	7	8	0	0.13	0.78	1.29	3.99E-07	1.9E-04	6.27E-03
27	*COL4A5*	+	4913	19	35	5	0.23	2.37	3.25	4.58E-07	2.12E-04	7.2E-03
28	*WDR49*	-	2031	9	12	1	0.17	1.22	1.98	6.01E-07	2.7E-04	9.44E-03
29	*TP53*	+	1112	9	9	0	0.56	1.27	2.63	6.28E-07	2.74E-04	9.87E-03
30	*COL14A1*	+	5307	15	24	5	0.13	2.27	3.18	1.11E-06	4.58E-04	0.02
31	*CAPZA3*	-	606	8	8	1	0.25	1.19	2.43	1.17E-06	4.71E-04	0.02
32	*DNAH3*	+	11243	18	25	11	0.04	2.30	2.92	1.23E-06	4.82E-04	0.02
33	*FSTL5*	-	2367	9	13	2	0.15	2.50	3.47	1.41E-06	5.39E-04	0.02
34	*FAM49A*	-	972	7	9	1	0.00	3.32	4.48	2.41E-06	9.01E-04	0.04
35	*PPP6C*	+	928	8	9	0	0.33	3.06	4.10	2.58E-06	9.42E-04	0.04
36	*EYA2*	-	1615	8	11	2	0.18	2.76	3.34	2.98E-06	1.02E-03	0.05
37	*TBX15*	-	1243	9	11	3	0.09	3.30	3.90	3.49E-06	1.17E-03	0.05
38	*ZNF385D*	-	1125	9	10	1	0.00	2.92	3.12	3.9E-06	1.26E-03	0.06
39	*TGM3*	-	2045	11	12	3	0.17	1.29	1.88	4.96E-06	1.56E-03	0.08
40	*C9*	-	1563	8	11	1	0.09	0.84	0.61	5.92E-06	1.82E-03	0.09
41	*PCDH15*	+	5159	12	20	4	0.00	2.09	2.48	6.95E-06	2.06E-03	0.11
42	*ENPEP*	-	2734	11	14	4	0.07	1.83	1.94	7.3E-06	2.09E-03	0.11
43	*CYP2C18*	-	1451	8	10	2	0.30	1.00	0.67	7.93E-06	2.22E-03	0.12
44	*KCNH7*	+	3443	12	15	3	0.00	3.02	2.92	8.06E-06	2.22E-03	0.13
45	*MECOM*	+	2568	11	14	1	0.07	2.90	4.16	1.08E-05	2.74E-03	0.17
46	*KIAA2022*	+	2769	12	13	3	0.15	2.00	3.02	2.72E-05	5.77E-03	0.43
47	*SIGLEC12*	-	1570	8	10	4	0.00	-0.01	0.03	2.75E-05	5.77E-03	0.43
48	*PTPRK*	+	4118	12	17	1	0.18	3.25	4.95	3.24E-05	6.37E-03	0.51
49	*RBP3*	-	2708	7	10	3	0.00	1.79	2.58	3.48E-05	6.76E-03	0.55

Supplementary Table 5: Significantly mutated genes in sun-exposed melanomas

Rank	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
50	*IFNA16*	-	290	3	5	0	0.00	-0.03	0.31	3.71E-05	6.85E-03	0.58

c) Simple Model (top 50 genes)

Rank	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
1	*BRAF*	+	2290	28	36	0	0.25	3.43	5.46	1.34E-57	7.01E-54	2.1E-53
2	*DCC*	+	4171	21	35	6	0.09	2.80	4.25	3.36E-44	1.06E-40	5.28E-40
3	*COL4A5*	+	4913	19	35	5	0.23	2.37	3.25	1.55E-38	4.06E-35	2.43E-34
4	*SI*	-	5435	16	30	5	0.17	1.72	2.57	4.7E-35	1.05E-31	7.38E-31
5	*ANK3*	-	10038	20	37	24	0.00	2.90	3.29	5.9E-34	1.16E-30	9.27E-30
6	*COL3A1*	+	4194	17	31	3	0.06	2.56	3.98	9.33E-33	1.63E-29	1.47E-28
7	*SCN10A*	-	5528	15	30	13	0.07	2.03	2.61	3.47E-32	5.45E-29	5.45E-28
8	*PTPRD*	+	5305	17	27	10	0.19	3.51	4.39	8.62E-29	1.04E-25	1.36E-24
9	*RP1*	-	4311	12	24	7	0.04	0.94	1.49	6.91E-28	7.24E-25	1.09E-23
10	*ADAM7*	-	2213	14	20	2	0.20	0.95	1.30	1.11E-27	1.09E-24	1.75E-23
11	*DSG4*	-	2816	13	22	3	0.14	1.50	1.98	1.47E-27	1.36E-24	2.31E-23
12	*SPHKAP*	-	3683	15	24	8	0.08	1.38	1.79	1.96E-27	1.62E-24	3.08E-23
13	*RELN*	-	9997	16	32	13	0.06	3.00	3.29	2.7E-27	2.02E-24	4.25E-23
14	*FAM5C*	-	1886	14	19	2	0.16	2.84	4.39	7.72E-27	5.51E-24	1.21E-22
15	*ADAMTS2(*	+	5489	15	25	5	0.16	2.00	3.23	3.38E-26	2.31E-23	5.31E-22
16	*COL5A1*	+	5440	15	29	10	0.03	2.55	2.89	7.66E-26	5.02E-23	1.2E-21
17	*CD163L1*	+	3877	15	24	2	0.04	0.56	0.79	1.56E-25	9.8E-23	2.45E-21
18	*CSMD2*	+	10321	17	32	21	0.06	2.86	3.63	2.92E-25	1.76E-22	4.59E-21
19	*XDH*	-	3943	14	23	10	0.04	2.41	3.23	1.32E-24	7.7E-22	2.08E-20
20	*MUC17*	+	8441	16	25	3	0.12	-0.16	-0.42	2.39E-24	1.34E-21	3.76E-20
21	*FAM83B*	-	2172	12	18	2	0.17	1.78	2.71	2.52E-24	1.37E-21	3.96E-20
22	*ADAMTS1&*	-	3586	11	21	10	0.10	2.38	2.88	1.12E-23	5.88E-21	1.76E-19
23	*CD163*	-	3000	13	21	6	0.10	1.27	1.32	1.77E-23	9E-21	2.79E-19
24	*COL14A1*	+	5307	15	24	5	0.13	2.27	3.18	2.28E-23	1.12E-20	3.59E-19
25	*COL11A1*	-	5557	15	26	6	0.08	3.12	4.28	2.65E-23	1.26E-20	4.17E-19

Supplementary Table 5: Significantly mutated genes in sun-exposed melanomas

Rank	Gene Symbol	Expression	Effective Length	Mutated Sample Count	Nonsynonymous SNV Count	Synonymous SNV Count	SNV LOH Fraction	Mean Gene PhyloP	Mean SNV PhyloP	Pvalue	BH Pvalue	BF Pvalue
26	*C15orf2*	+	2248	13	17	4	0.12	-0.29	-0.16	3.14E-23	1.45E-20	4.94E-19
27	*TNR*	-	3753	13	22	6	0.14	2.62	3.24	3.82E-23	1.71E-20	6E-19
28	*ADAMDEC*	-	1412	11	15	11	0.07	0.86	1.15	2.56E-22	1.12E-19	4.02E-18
29	*NRAS*	+	565	13	13	0	0.54	3.63	4.96	4.32E-22	1.84E-19	6.79E-18
30	*C6*	-	2726	15	18	3	0.06	1.57	1.42	8.19E-22	3.39E-19	1.29E-17
31	*USH2A*	+	14158	17	29	14	0.14	1.62	1.77	8E-21	3.14E-18	1.26E-16
32	*SLCO1B3*	+	2001	13	15	5	0.07	0.90	0.82	1.07E-20	4.12E-18	1.69E-16
33	*MAGEC1*	+	2245	8	16	3	0.06	0.13	0.56	1.58E-20	5.91E-18	2.48E-16
34	*PCDH15*	+	5159	12	20	4	0.00	2.09	2.48	4.25E-20	1.52E-17	6.68E-16
35	*FAT4*	+	6880	13	22	6	0.09	2.41	3.64	9.39E-20	3.28E-17	1.48E-15
36	*GRID2*	+	2837	13	17	6	0.12	3.20	3.35	2.45E-19	8.19E-17	3.85E-15
37	*ARMC4*	-	3050	13	17	5	0.06	1.92	2.39	2.85E-19	9.33E-17	4.48E-15
38	*APOB*	-	10020	16	24	12	0.08	1.21	1.27	3.53E-19	1.13E-16	5.55E-15
39	*SPATA18*	+	1541	10	14	3	0.00	1.28	2.59	3.82E-19	1.2E-16	6E-15
40	*GRM3*	+	2124	12	15	2	0.27	3.11	3.72	6.67E-19	2.02E-16	1.05E-14
41	*C1orf168*	-	1914	12	14	4	0.14	0.61	0.60	4.07E-18	1.12E-15	6.39E-14
42	*CDH6*	-	2154	12	15	7	0.13	2.78	3.05	4.62E-18	1.23E-15	7.25E-14
43	*RFX6*	-	2639	10	15	4	0.07	2.54	2.64	4.78E-18	1.25E-15	7.51E-14
44	*CDH9*	-	2125	11	14	3	0.21	2.46	4.15	5.69E-18	1.47E-15	8.95E-14
45	*C1orf173*	-	3550	15	18	11	0.17	0.71	0.71	9.42E-18	2.35E-15	1.48E-13
46	*GRIN3A*	-	2727	9	15	3	0.00	2.70	4.16	1.35E-17	3.31E-15	2.12E-13
47	*DNAH3*	+	11243	18	25	11	0.04	2.30	2.92	1.94E-17	4.69E-15	3.05E-13
48	*TPTE*	-	1512	12	12	2	0.08	0.59	1.06	2.6E-17	5.92E-15	4.09E-13
49	*SLC15A2*	+	2186	11	14	1	0.07	2.02	3.52	2.72E-17	6.1E-15	4.27E-13
50	*PTPRK*	+	4118	12	17	1	0.18	3.25	4.95	2.93E-17	6.5E-15	4.61E-13

Supplementary Table 6: Somatic mutations in genes coding MAPKs

Gene Symbol	Accession	Chr	Position	Reference Genotype	Variant Genotype	Amino Acid Change	Total	Expression	COSMIC	Phylop Cons. Score	p-fam domain
MAP2K1	NM_002755.3	chr15	66729162	C/C	C/T	P124S	2	+	N	5.92	Pkinase
MAP2K3	NM_145109.2	chr17	21202225	C/C	C/T	T51I	1	+	N	4.58	
MAP2K3	NM_145109.2	chr17	21201778	G/G	G/A	A35T	1	+	N	1.34	
MAP2K3	NM_145109.2	chr17	21201737	G/G	G/T	R21M	1	+	N	0.13	
MAP2K3	NM_145109.2	chr17	21206510	G/G	G/A	E178K	1	+	N	6.06	Pkinase
MAP2K4	NM_003010.2	chr17	12016578	C/C	C/A	D238E	1	+	N	-0.02	Pkinase
MAP2K4	NM_003010.2	chr17	11958245	C/C	C/T	P52L	1	+	N	5.01	
MAP2K5	NM_145160.2	chr15	68020260	C/C	C/T	L351F	1	+	N	4.41	Pkinase
MAP2K6	NM_002758.3	chr17	67517227	C/C	C/T	S174F	1	+	N	4.13	Pkinase
MAP3K10	NM_002446.3	chr19	40710502	C/C	C/T	P325L	1	+	N	3.80	Pkinase
MAP3K11	NM_002419.3	chr11	65375153	C/C	C/G	E402Q	1	+	N	5.41	
MAP3K11	NM_002419.3	chr11	65373447	G/G	G/A	S570F	1	+	N	4.03	
MAP3K12	NM_001193511.1	chr12	53880776	T/T	T/C	M134V	1	+	N	4.50	
MAP3K12	NM_001193511.1	chr12	53880918	C/C	C/A	Q86H	1	+	N	0.69	
MAP3K12	NM_001193511.1	chr12	53878135	C/C	C/T	R385Q	1	+	N	5.79	Pkinase
MAP3K13	NM_004721.3	chr3	185155318	T/T	T/C	F187L	1	+	N	5.06	Pkinase
MAP3K13	NM_004721.3	chr3	185183626	G/G	G/A	E494K	1	+	N	4.16	
MAP3K13	NM_004721.3	chr3	185167771	T/T	T/C	V365A	1	+	N	5.11	Pkinase
MAP3K15	NM_001001671.3	chrX	19428074	C/C	C/T	W572X	1	+	N	3.88	
MAP3K15	NM_001001671.3	chrX	19389588	C/C	C/T	D1057N	1	+	N	3.87	
MAP3K15	NM_001001671.3	chrX	19392708	G/G	G/A	S887F	1	+	N	1.24	Pkinase
MAP3K15	NM_001001671.3	chrX	19379650	C/C	C/T	W1247X	1	+	N	4.93	
MAP3K4	NM_005922.2	chr6	161514074	G/G	G/A	E1112K	1	+	N	6.26	
MAP3K4	NM_005922.2	chr6	161455358	G/G	G/T	E74X	1	+	N	4.33	
MAP3K5	NM_005923.3	chr6	136901529	G/G	G/A	R1143W	1	+	N	2.09	
MAP3K5	NM_005923.3	chr6	136958516	C/C	C/T	E655K	1	+	N	5.25	
MAP3K5	NM_005923.3	chr6	137019697	G/G	A/A	L246F	1	+	N	4.10	
MAP3K5	NM_005923.3	chr6	136980475	C/C	C/T	E470K	1	+	N	5.92	
MAP3K5	NM_005923.3	chr6	136972229	C/C	C/T	V561I	1	+	N	5.04	
MAP3K6	NM_004672.3	chr1	27682167	G/G	G/A	R1261C	1	+	N	0.08	
MAP3K8	NM_005204.2	chr10	30736798	G.T/G.T	G.T/T.A	D142X	1	-	N	2.19	Pkinase

Supplementary Table 6: Somatic mutations in genes coding MAPKs

Gene Symbol	Accession	Chr	Position	Reference Genotype	Variant Genotype	Amino Acid Change	Total	Expression	COSMIC	Phylop Cons. Score	p-fam domain
MAP3K8	NM_005204.2	chr10	30749739	G/G	G/A	G460R	1	-	N	3.84	
MAP3K9	NM_033141.2	chr14	71199675	G/G	G/A	P818L	1	-	N	1.45	
MAP3K9	NM_033141.2	chr14	71205034	G/G	G/A	A591V	1	-	N	2.71	
MAP3K9	NM_033141.2	chr14	71209269	G/G	G/C	L456V	1	-	N	1.78	
MAP3K9	NM_033141.2	chr14	71227765	C/C	C/T	E319K	1	-	N	5.94	Pkinase
MAP3K9	NM_033141.2	chr14	71197422	G/G	G/A	P1011L	1	-	N	5.81	
MAP4K2	NM_004579.3	chr11	64564613	G/G	G/A	T443M	1	+	N	1.11	
MAP4K3	NM_003618.2	chr2	39479010	G/G	G/A	S853L	1	+	N	3.85	CNH
MAP4K3	NM_003618.2	chr2	39535101	T/T	T/G	T368P	1	+	N	2.57	
MAP4K3	NM_003618.2	chr2	39505579	A/A	A/G	L588P	1	+	N	4.64	CNH
MAP4K3	NM_003618.2	chr2	39559079	G/G	G/A	S170F	1	+	N	6.10	Pkinase
MAPK1	NM_138957.2	chr22	22127166	T/T	T/A	D321V	1	+	N	4.91	
MAPK1	NM_138957.2	chr22	22142605	G/G	G/A	S266F	1	+	N	6.35	Pkinase
MAPK1	NM_138957.2	chr22	22142599	G/G	G/A	P268L	1	+	N	6.35	Pkinase
MAPK10	NM_138982.2	chr4	87019707	G/G	G/A	R258C	1	+	N	6.30	Pkinase
MAPK10	NM_138982.2	chr4	86989058	C/C	C/T	E285K	1	+	N	6.06	Pkinase
MAPK13	NM_002754.3	chr6	36106731	C/C	C/T	P306L	1	+	N	2.98	Pkinase
MAPK14	NM_139012.2	chr6	36040751	G/G	G/A	R136Q	1	+	N	6.31	Pkinase
MAPK14	NM_139012.2	chr6	36063799	G/G	G/A	G240R	1	+	N	5.50	Pkinase
MAPK6	NM_002748.3	chr15	52356254	A/A	A/G	K408R	1	+	N	4.58	
MAPK6	NM_002748.3	chr15	52353594	CC/CC	CC/TT	P322L	1	+	N	5.85	
MAPK6	NM_002748.3	chr15	52350880	G/G	G/T	V251L	1	+	N	3.07	Pkinase
MAPK6	NM_002748.3	chr15	52356250	G/G	G/T	E407X	1	+	N	5.56	
MAPK7	NM_139034.2	chr17	19286249	C/C	C/T	P763S	1	+	N	2.47	
MAPK8	NM_139047.1	chr10	49617959	C/C	C/T	S97F	1	+	N	4.50	Pkinase
MAPK9	NM_139070.2	chr5	179688748	G/G	G/A	S129F	1	+	N	5.96	Pkinase

Supplementary Table 7: Genes with significant numbers of deleterious mutations across all sun-exposed melanomas (n=97)

Gene Symbol	Length (bp)	Number of melanomas with deleterious mutations			Pval	BH Pval	BF Pval	Expression
		Nonsense Mutations	Frame Shift InDels	Splice Site Variants				
TP53	1112	6	2	2	3.29E-15	5.17E-11	5.17E-11	+
NF1	7866	8	2	2	2.35E-12	1.23E-08	3.69E-08	+
ARID2	4426	6	5	0	6.73E-12	2.65E-08	1.06E-07	+
DCC	4171	9	0	0	1.98E-08	6.21E-05	3.11E-04	+
ZNF560	1759	5	0	1	3.36E-07	8.81E-04	5.28E-03	+
FAM49A	972	4	0	1	4.45E-07	9.98E-04	6.99E-03	-
SLC22A25	1546	5	0	0	4.22E-06	8.22E-03	0.07	-
FAM58A	715	4	0	0	4.71E-06	8.22E-03	0.07	+
ME1	1695	3	0	1	6.56E-06	0.01	0.1	+
TGM3	2045	3	0	2	1.61E-05	0.02	0.25	-

Supplementary Table 8: Genes with multiple somatic mutations in sun-shielded melanomas

Gene Symbol	Accession	Acral Melanomas	Mucosal Melanomas	Uveal Melanomas
ARFGEF2	NM_006420.2	YUIRI(P1649P)	YUSAN(A751A)	
COL2A1	NM_001844.4	YUISKIA(G1209S)		YUBOO(S1001L)
DYNC1I1	NM_004411.4	YUBOT(R629C) YUWIC(R629C) YUSAG(V506G)		
GNA11	NM_002067.2			YUCRENA(Q209L) YUVEDO(Q209L) YUBOO(R183C)
GYS1	NM_002103.4	YUMUDE(I524V)		YUBOO(R192C)
SZT2	NM_015284.2	YUNUVO(D374N)		YUBOO(R1207Q)
KIT	NM_000222.2	YURUB(N822Y) YUBRUSE(L576P) YUSAG(K642E) YUSCH(L576P) YUMUDE(V559D)		
SPAG17	NM_206996.2	YUISKIA(A2T)		YUBOO(R830C)

* YUVEDO has only 2 somatic mutations.

** The R629C mutation was also identified in YUDUTY, a melanoma of unknown primary.

Supplementary Table 9: Copy number gains and losses in melanomas

Chr	Band	CNV Status	All Samples	Sun Exposed	Acral	Mucosal	Uveal	Unknown	Candidate genes with significant numbers of SCNA events
chr1	q31	gain	11	7	0	2	0	2	*ASPM-PTPRC*
chr1	q42	gain	5	4	0	1	0	0	*LYST*
chr5	p13	gain	12	3	6	3	0	0	*NADKD1-NIPBL-NUP155-PDZD2-RANBP3L-RICTOR*
chr7	q11	gain	5	4	0	1	0	0	*GTF2IRD2*
chr7	q34	gain	8	5	1	1	0	1	*ADCK2-BRAF*
chr8	q24	gain	9	5	2	1	2	1	*ASAP1-ATAD2-KIFC2-PTK2*
chr11	q13	gain	9	1	4	3	0	1	*ACER3-CAPN5-CCND1-CTTN-SHANK2*
chr11	q14	gain	10	3	3	3	0	1	*AQP11-INTS4-RSF1-TMEM135*
chr12	q14	gain	5	0	3	1	0	1	*CDK4*
chr20	q13	gain	10	5	2	1	1	2	*PRIC285-SYCP2*
chr5	q31	loss	4	4	0	0	0	0	*PCDHB8*
chr8	p23	loss	7	5	1	0	1	1	*DEFA3-SPAG11B*
chr9	p13	loss	4	3	0	0	0	1	*ZNF658*
chr9	p21	loss	27	17	3	2	0	5	*CDKN2A-DMRTA1-ELAVL2-MTAP*
chr9	p24	loss	9	6	0	0	0	3	*CBWD1*
chr10	p12	loss	17	9	3	1	0	4	*LYZL1-MRC1L1*
chr10	q11	loss	19	11	3	1	0	4	*AGAP4-ANXA8L2-FRMPD2-PTPN20B*
chr10	q23	loss	8	6	0	1	0	1	*PTEN*
chr10	q26	loss	4	1	3	0	0	0	*TACC2*
chr14	q11	loss	4	4	0	0	0	0	*POTEG*
chr14	q24	loss	7	5	1	0	0	1	*ACOT1*
chr17	q11	loss	7	2	2	1	1	2	*ATAD5-LRRC37B*
chr19	q13	loss	7	4	2	0	1	1	*FCGBP-KIR2DL1-KIR2DL3-KIR3DL3*

Supplementary Table 10: Information on patients with the RAC1^{P29S} mutation

Clinical and genetic information of melanomas with the RAC1^{P29S} mutation

Melanoma	Gender/ Age	Breslow thickness (mm)	Primary Location	Tumor Analyzed	Final Stage	*BRAF*	*NRAS*	*CDKN2A*
DF-T	M/79	Unknown	Shoulder	M	IV	V600E	WT	WT
YUBRO-T	M/56	0.85	Forearm	M	IV	WT	WT	WT
YUCAV-T	M/61	2.85	Forearm	P	IIB	WT	Q61K	WT
YUCOW-T	M/69	1.5	Mid back	P	IB	V600E	WT	WT
YUFAR-T	M/48	0.91	Chest	P	IV	V600E	WT	T79X/T[U]
YUFIC-C	M/65	2.24	Leg	M	IV	WT	Q61R	WT
YUGOV-T	M/60	2.1	Scapula	M	IV	V600K	WT	WT
YUHEF-C	M/52	1.7	Scalp	M	IV	WT	WT	A57X/A[S]
YUKAT-T	M/78	NA	Unknown	M	IV	WT	WT	WT
YUKLAB-T	M/84	NA	Unknown	M	IV	WT	WT	WT
YULAN-T	M/81	7.5	Scalp	M	IV	WT	WT	WT
YUMCE-T	M/81	MIS	Upper arm	M	IV	WT	WT	WT
YUNACK-T	F/59	22.0	Trunk	P	I	V600E	WT	WT
YUPROST-T	F/86	1.1	Neck	P	IIB	WT	WT	WT
YURIF-C	M/52	3.0	Thigh	M	IV	V600K	WT	R62K/R[S]
YUSOC-C	M/98	1.2	Cheek	P	II	WT	WT	T79X/T[U]
YUSUKA-T	M/89	1.42	Ear	P	IB	WT	WT	WT
YUTOGS-T1	M/50	3.5	Scalp	M	IV	WT	WT	WT
YUVEME-T	M/78	1.0	Cheek	M	IV	WT	WT	WT
YUWIA-T	M/83	0.9	Calf	M	IIIB	WT	Q61K	WT
C021-C	M/37	UK	Shoulder	M	UK	WT	WT	NA
C083-C	M/38	UK	Unknown	M	UK	WT	Q61L	NA
D26-C	M/55	UK	Unknown	M	UK	WT	WT	NA
MM96L-C2	F/67	UK	Scapula	M	UK	V600E	WT	WT

T and C designate snap-frozen tumors and cultured melanoma cells, respectively. P and M denotes primary and metastatic melanoma, respectively. MIS, Melanoma in situ; NA, not applicable; UK, unknown.

[1] The RAC1 mutation was identified in the primary and metastatic lesions of this patient.

[2] Homozygous for RAC1^{P29S} due to copy neutral LOH spanning the locus.

[S] Somatic CDKN2A mutation.

[U] Unknown because germline DNA is not available.

Supplementary Table 11: RAC1 crystal structure data collection and refinement statistics

	RAC1^{P29S} **PDB ID: 3SBD**	**RAC1^{P29S}** **PDB ID: 3SBE**	**RAC1 wild-type** **PDB ID: 3TH5**
Data collection			
Space group	$P2_12_12_1$	$P22_12_1$	$P2_1$
X-ray source and detector	RIGAKU 007 HF Saturn 944+ CCD	RIGAKU 007 HF Saturn 944+ CCD	APS NECAT-E ADSC Q315
Wavelength (Å)	1.5418	1.5418	0.97921
Cell: a, b, c (Å)	50.3, 80.0, 94.9	40.6, 51.9, 99.3	40.9, 97.9, 51.7
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 96.6, 90
Resolution range (Å)	20.0 - 2.1 (2.17 - 2.1)	20.0 - 2.6 (2.7 - 2.6)	30.0 - 2.3 (2.38 - 2.3)
No. of unique reflections	23040	6926	18169
Completeness (%)	98.3 (86.4)	99.9 (100.0)	99.0 (99.3)
R_{sym} (%)	12.9 (64.7)	11.2	5.2 (29.5)
Mn $I/\sigma I$	10.1 (1.5)	13.0 (2.0)	23.1 (3.9)
Redundancy	5.6 (2.4)	6.6 (6.7)	3.5 (3.4)
Refinement statistics			
Resolution range (Å)	19.7 - 2.1 (2.15 - 2.1)	19.9 - 2.6 (2.7 - 2.6)	20.0 - 2.3 (2.36 - 2.3)
R-factor (%)			
Working set	23.8 (29.1)	23.1 (30.2)	22.9 (33.3)
Test set	28.5 (39.3)	29.6 (42.3)	26.4 (42.1)
Free *R* reflections (%)	5.2 (4.9)	4.7 (4.9)	5.1 (5.0)
Free *R* reflections, no.	1158 (63)	325 (23)	910 (63)
Residues built	A/1 - 177 B/2 - 177	0 - 177	A/2 - 177 B/2 - 29, 35 -177
No. water molecules	146	23	23
Mean *B*-factor (Å^2) Protein / GMP-PNP / Mg / H_20	30.7 / 25.7 / 24.5 / 30.8	58.4 / 47.9 / 42.0 / 54.0	60.7 / 49.9 / 50.7 / 58.1
Model statistics			
RMSD bond lengths (Å)	0.014	0.014	0.007
RMSD bond angles (°)	1.576	1.521	1.106
Ramachandran plot (%) favored/ allowed/disallowed	97.1 / 2.9 / 0	94.9 / 5.1 / 0	97.1 / 2.9 / 0

Supplementary Table 12: Primers and oligos used in the studies

Gene	Backward primer
RAC1 P29S[1]	F1: 5'- ACCTAAACAGAATGTGATGGCTCC -3' R1: 5'- GGTCAAAGAAATGTGAAACCCG -3'
	F2: 5'-TGGTGATAAAGGGTTATAGAAAACA-3' R2: 5'- CAGCAAAACAAATGGTCAAA-3'
RAC1 Q220X/R301C[2]	F: 5'- TGCTAATGCCTGGAGATACTGTACC -3 R: 5'- TCGTTCTGGGAGGAATAACACG -3'
RAC1-P29S[3]	F:GTTACACAACCAATGCATTTTCTGGAGAATATATCCCTACTG 3' R:5' CAGTAGGGATATATTCTCCAGAAAATGCATTGGTTGTGTAAC 3'
RAC1-P29S[4]	F: 5'- TCAAGTGTGTGGTGGTGGGAG -3' B: R:5'- TTTGCGGATAGGATAGGGGGCG -3'
RAC1-F28L[5]	5' CAGTTACACAACCAATGCACTTCCTGGAGAATATATCCC 3' 5' GGGATATATTCTCCAGGAAGTGCATTGGTTGTGTAACTG 3'

[1]The F1/B1 primers were used for the TaqMan® assay (Applied Biosystems, Carlsbad, California) and targeted amplification for Sanger sequencing to assess and validate *RAC1*^{P29S} mutation. The F2/B2 primers were used to assess the *RAC1*^{P29S} mutation in 76 melanoma cell lines in the Oncogenomics Laboratory, Queensland Institute of Medical Research by Sanger sequencing using BigDye Terminator v3.1 chemistry on a 3730xl DNA Analyzer (Applied Biosystems).

[2]These primers were used to PCR amplify and subclone Q220X/R301C mutation region from cDNA was The amplified PCR fragments were cloned into the pCR4-TOPO TA cloning vector (Invitrogen). One Shot TOP10 (Invitrogen) competent E. coli cells were transformed with the TOPO cloning reaction following the

[3]These primers were used for site-directed mutagenesis to generate the P29S mutation in plasmids encoding RAC1 with the QuikChange kit (Stratagene, La Jolla, CA).

[4]These primers were used to validate the mutations in the vector.

[5]These primers were used to for site-directed mutagenesis to generate the F28L mutation in plasmids encoding RAC1 with the QuikChange kit.