

1 Mapping chromatin interactions at melanoma susceptibility loci and cell-type
2 specific dataset integration uncovers distant gene targets of *cis*-regulation

3

4 Rohit Thakur^{1,#}, Mai Xu^{1,#}, Hayley Sowards¹, Joshua Yon¹, Lea Jessop², Timothy Myers², Tongwu
5 Zhang³, Raj Chari⁴, Erping Long^{1,5}, Thomas Rehling¹, Rebecca Hennessey¹, Karen Funderburk¹, Jinhu
6 Yin¹, Mitchell J. Machiela³, Matthew E. Johnson⁶, Andrew D. Wells⁷, Alessandra Chesi⁷, Struan F.A.
7 Grant⁷, Mark M. Iles^{8,9}, Maria Teresa Landi³, Matthew H. Law^{10,11,12}, Melanoma Meta-Analysis
8 Consortium¹³, Jiyeon Choi¹, Kevin M. Brown^{1,*}

9 ¹Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer
10 Institute, Bethesda, MD, USA

11 ²Laboratory of Genomic Susceptibility, Division of Cancer Epidemiology and Genetics, National Cancer
12 Institute, Bethesda, MD, USA

13 ³Integrative Tumor Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer
14 Institute, Bethesda, MD, USA

15 ⁴Genome Modification Core, Frederick National Lab for Cancer Research, Frederick, MD, USA

16 ⁵Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Beijing, China

17 ⁶Division of Human Genetics, Children's Hospital of Philadelphia Research Institute, Philadelphia, PA,
18 USA

19 ⁷Center for Spatial and Functional Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

20 ⁸Leeds Institute for Data Analytics, University of Leeds, Leeds, UK

21 ⁹NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

22 ¹⁰Population Health Department, QIMR Berghofer Medical Research Institute, Herston, QLD, Australia

23 ¹¹School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD,
24 Australia

25 ¹²School of Biomedical Sciences, University of Queensland, Brisbane, QLD, Australia

26

27 # Contributed equally

28 * Correspondence: kevin.brown3@nih.gov

29 **ABSTRACT**

30 Genome-wide association studies (GWAS) of melanoma risk have identified 68 independent signals at 54
31 loci. For most loci, specific functional variants and their respective target genes remain to be
32 established. Capture-HiC is an assay that links fine-mapped risk variants to candidate target genes by
33 comprehensively mapping cell-type specific chromatin interactions. We performed a melanoma GWAS
34 region-focused capture-HiC assay in human primary melanocytes to identify physical interactions
35 between fine-mapped risk variants and potential causal melanoma susceptibility genes. Overall,
36 chromatin interaction data alone nominated potential causal genes for 61 of the 68 melanoma risk
37 signals, identifying many candidates beyond those reported by previous studies. We further integrated
38 these data with cell-type specific epigenomic (chromatin state, accessibility), gene expression
39 (eQTL/TWAS), DNA methylation (meQTL/MWAS), and massively parallel reporter assay (MPRA) data to
40 prioritize potentially *cis*-regulatory variants and their respective candidate gene targets. From the set of
41 fine-mapped variants across these loci, we identified 140 prioritized candidate causal variants linked to
42 195 candidate genes at 42 risk signals. In addition, we developed an integrative scoring system to
43 facilitate candidate gene prioritization, integrating melanocyte and melanoma datasets. Notably, at
44 several GWAS risk signals we observed long-range chromatin connections (500 kb to >1 Mb) with distant
45 candidate target genes. We validated several such *cis*-regulatory interactions using CRISPR inhibition,
46 providing evidence for known cancer driver genes *MDM4* and *CBL*, as well as the SRY-box transcription
47 factor *SOX4*, as likely melanoma risk genes.

48 **INTRODUCTION**

49 Melanoma is the deadliest form of skin cancer and originates from melanocytes. Multiple genome-wide
50 association studies (GWAS) of melanoma risk have been conducted¹⁻⁹, with the most recent meta-
51 analysis of 36,760 melanoma cases identifying 54 loci and 68 independent signals¹⁰. Despite this success,

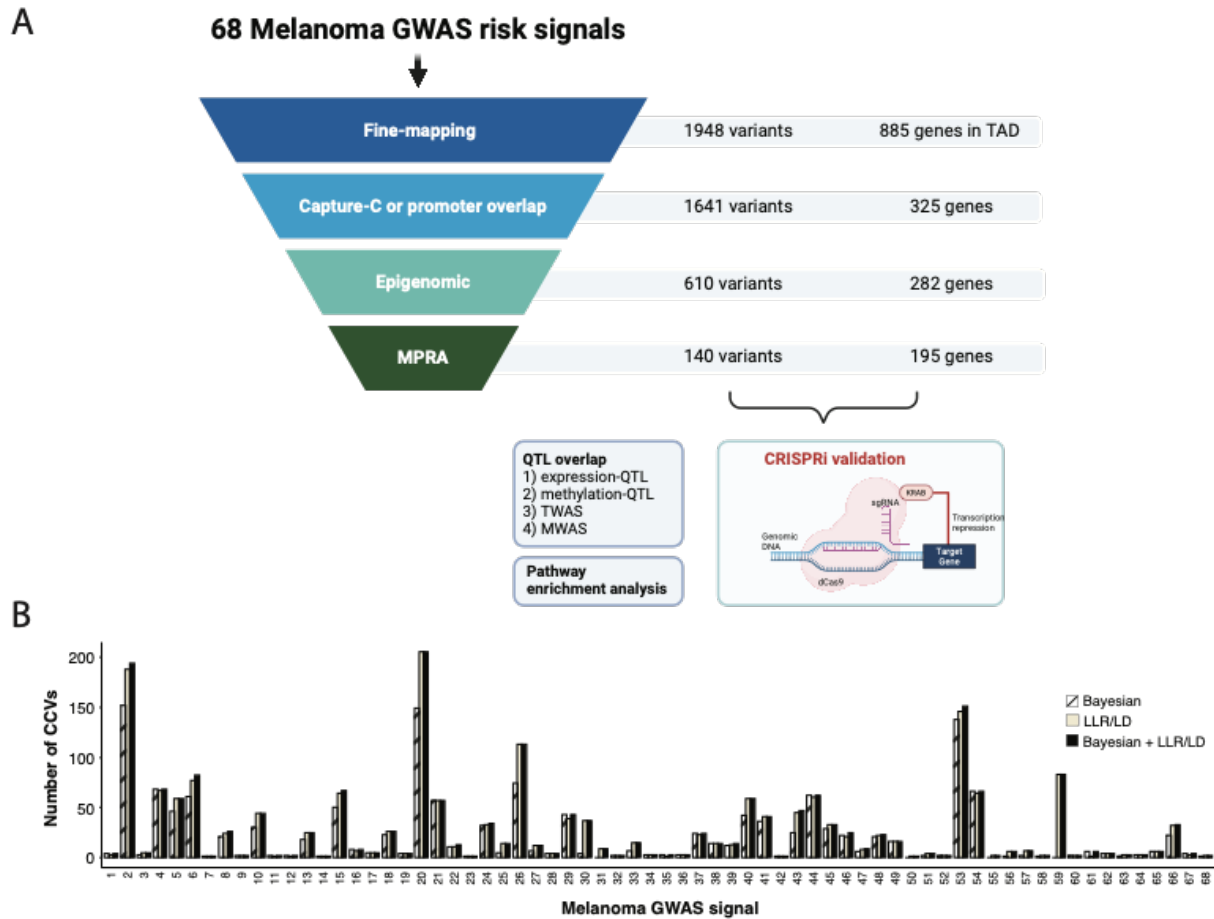
52 significant challenges lie in pinpointing the functional variants and causal genes at most of these GWAS
53 risk loci¹¹⁻¹³. Most loci associated with complex traits, including those for melanoma risk, do not harbor
54 risk-associated protein-coding variants¹¹⁻¹³. Instead, these loci may potentially function through genetic
55 variants located in *cis*-regulatory regions such as enhancers and gene promoters, influencing target gene
56 expression in an allele-specific manner¹¹⁻¹³. Consistent with this, genetic variants associated with
57 complex traits are often found to be enriched at genomic regions annotated as regulatory elements¹⁴⁻¹⁸.
58 For many loci, the lead reported variant is not necessarily the functional variant, as each locus may
59 harbor multiple risk-associated variants that are in linkage-disequilibrium (LD) with the unknown causal
60 variant or variants¹². This often makes it difficult to statistically distinguish the causal risk variant(s) from
61 LD passengers. Furthermore, given that enhancers may function over long distances, the nearest gene to
62 a GWAS risk signal is not necessarily the target gene and therefore, genes other than the nearest gene in
63 the region must be considered plausible targets^{11,18,19}.

64 One commonly used post-GWAS approach for identifying target genes is colocalization of the GWAS
65 signal with those from quantitative trait locus (QTL) datasets generated from disease- or trait-relevant
66 cell types or tissues²⁰⁻²². Colocalization of melanoma GWAS with multiple QTL types derived from
67 expression and methylation data from human primary melanocytes identified at least one colocalizing
68 QTL for less than half (39%) of melanoma GWAS loci^{23,24}. While QTLs derived from expression data
69 nominate specific gene candidates for a given locus, meQTLs do not necessarily directly implicate
70 specific genes. Long and colleagues compiled available QTL data with a custom massively parallel
71 reporter assay (MPRA) of fine-mapped melanoma-associated variants and still only linked MPRA-positive
72 variants to genes for roughly 50% of loci²⁵. Given many GWAS risk variants have relatively small effects
73 on disease risk, this lack of colocalizing QTLs could be explained by limited statistical power in small QTL
74 studies^{11,12,26-28}. Alternatively, there is growing evidence that many variants associated with complex

75 traits may function in a context- and/or state-specific manner which may be missed when using QTL
76 data from cells in a steady-state or MPRA assays in specific cell systems^{22,29-31}.

77 Enhancer elements regulate gene expression via physical interactions with promoter elements and can
78 thus regulate expression of distant genes via long-range three-dimensional chromatin interaction.
79 Methods to characterize chromatin conformation, including HiC-based methods³²⁻³⁷ have emerged as
80 powerful approaches to map such interactions at GWAS risk loci^{34,35,38-43}. One of these methods, capture-
81 HiC⁴³⁻⁴⁶, utilizes capture baits targeting regions of interest, often gene promoters or GWAS signals. To
82 date, targeted cell-type specific chromatin interactions have not been evaluated across all genome-wide
83 significant melanoma risk loci, but the utility of this approach has been demonstrated in establishing
84 *AHR* as a functionally-validated ultraviolet B (UVB)-responsive melanoma susceptibility gene⁴⁷. In this
85 study, we performed a GWAS region-specific capture-HiC assay, baiting the entire regions of association
86 for the 68 melanoma GWAS risk signals (locus and signal numbering is listed in **Table S1**) to
87 comprehensively map cell-type specific chromatin interactions between fine-mapped risk variants and
88 potential target genes in human primary melanocytes. We integrated capture-HiC data with fine-
89 mapping, as well as cell-type specific epigenomic (chromatin state, accessibility), gene expression
90 (eQTL/TWAS), DNA methylation (meQTL/MWAS), and high-throughput screening (massively parallel
91 reporter assays; MPRA) data to prioritize variants and their respective candidate gene target(s) for *cis*-
92 regulation (**Figure 1A**).

93 Our approach nominated potential causal genes for the vast majority of melanoma risk signals,
94 identifying many plausible candidates beyond those reported by previous studies^{10,23,24,48}. Notably, we
95 identify multiple candidate genes previously identified as somatically altered in melanoma and pan-
96 cancer tumor analyses.



97

98 **Figure 1. Schematic of data integration of capture-HiC data with orthogonal data to prioritize**
 99 **candidate causal variants and genes.** (A) Schematic summary of this study utilizing an integrative
 100 analysis approach to identify candidate causal variants (CCVs) and target candidate genes at the 68
 101 melanoma GWAS risk signals. We performed GWAS region-specific capture-HiC assay, baiting the entire
 102 region of association for the 68 melanoma GWAS risk signals to comprehensively map chromatin
 103 interactions. Subsequently we utilized this dataset to link fine-mapped risk variants to candidate target
 104 genes. We integrated fine-mapping with observed chromatin interactions, further overlaying cell-type
 105 specific epigenomic (chromatin state, accessibility) and high-throughput reporter assay screening
 106 (massively parallel reporter assays, MPRA) datasets to prioritize likely functional variants and respective
 107 candidate gene target(s) for *cis*-regulation. Finally, we validated candidate genes nominated at multiple
 108 loci via CRISPR inhibition system. (B) Summary of fine-mapped credible causal variants (CCVs) using
 109 Bayesian, LLR/LD, or both criteria at the 68 melanoma GWAS signals (a key to numbered loci is provided
 110 in **Table S1**). The black bar shows the union of fine-mapped variants identified by Bayesian and LLR/LD
 111 approaches.

112 METHODS

113 *Fine-mapping of melanoma risk signals*

114 We performed statistical fine-mapping for the 54 GWAS loci (harboring 68 independent genome-wide
115 significant signals) described by Landi and colleagues¹⁰ (**Table S1**) using an inclusive strategy, selecting
116 any variant identified by any one of multiple approaches. Firstly, we used a combination of log-
117 likelihood ratio (LLR) and linkage disequilibrium (LD) based cut-offs, similar to that performed for a
118 recent melanoma GWAS massively parallel reporter assay (MPRA) study²⁵. The melanoma GWAS
119 summary data was from the fixed-effect inverse-variance weighted meta-analysis of the full set of
120 confirmed and self-reported melanoma cases and controls as previously described¹⁰.

121 Specifically, we selected variants that met any one of the following criteria:

- 122 a) Variants with log likelihood ratio (LLR) <1:100 relative to the lead variant for the primary signal
123 of each GWAS locus.
- 124 b) Variants that were not genotyped or successfully imputed in the GWAS (including
125 insertion/deletion variants not assessed in the Haplotype Reference Consortium imputation
126 panel) that had LD $r^2 > 0.8$ (1000 Genomes Project, Phase 3, Version 5, EUR population)⁴⁹ with
127 the primary lead variant. These variants were identified using the LDlinkR package⁵⁰⁻⁵².
- 128 c) For secondary risk signals at a locus that were identified through conditional analysis within 1
129 Mb of a primary lead SNP¹⁰ (irrespective of LLR), we relied on LD-based fine-mapping, selecting
130 all variants with LD $r^2 > 0.8$ (based on 1000 Genomes Project, Phase 3, Version 5, EUR
131 population)⁴⁹ to the leading variant at the independent risk signal.

132 These LLR and LD based data were used for the capture-HiC bait design as well as for identifying credible
133 causal variants.

134 We also used the Bayesian deterministic approximation of posteriors approach (as implemented in the
135 DAP-G software tool)^{53,54}. For each locus, we defined the region of association by identifying the set of
136 variants with $LLR < 1:1000$, ordered the variants based on increasing chromosomal position, and selected
137 the median position to create a +/-500 kilobase (kb) fine-mapping window. Fine-mapping windows were
138 visually inspected manually and adjusted to +/- 1.5 Mb for four loci where 500 kb was insufficient to
139 capture all of the association signal (5p15.3, 16q24.3, 11q14.3, and 20q11.22). The test statistic (Z-score)
140 for each variant from the GWAS summary statistics and the LD matrix (pre-computed using $n \sim 337,000$
141 unrelated British-ancestry individuals from the UK Biobank^{55,56}, `s3://broad-alkesgroup-ukbb-
142 ld/UKBB_LD/`) were used for the analysis. We set the maximum number of causals at each locus as 5,
143 with exception of the *9p21.3* locus (locus number 27, risk signal numbers 30-35, **Table S1**), where the
144 number of causals was set to six to account for the six independent genome-wide significant signals at
145 this locus. Note, while we allowed for multiple causals/credible sets, for the purpose of fine-mapping we
146 only retained variants within individual credible sets that directly correspond to each of the 68
147 independent genome-wide significant GWAS signals.

148 ***GWAS conditional analysis***

149 To identify independent risk-associated signals at the *MDM4* locus, we performed conditional and joint
150 association analyses of melanoma GWAS summary data¹⁰ using the Genome-wide Complex Trait
151 Analysis (GCTA, v1.94.1)⁵⁷ COnditional and JOint association (COJO) module⁵⁸, default settings, and a
152 genomic window of chr1: 203021577-206021577 (hg19). We used an LD reference population of 5,000
153 individuals selected randomly from the UKBB population determined to be European by PCA (LD_{EUR});
154 variants were converted to best-guess genotype (threshold 0.3) followed by data cleaning for
155 missingness > 3%, HWE $p < 1 \times 10^{-6}$, and MAF < 0.001. Separate analyses were performed conditioning
156 on lead variants from credible sets identified via Bayesian fine-mapping (DAPG credible set 1 lead

157 variant: rs2369633, DAPG credible set 2 lead variant: rs12119098). For the conditional analyses, we
158 selected the following genomic window chr1: 203021577-206021577 covering both the DAP-G credible
159 set signals.

160 ***Variant effect prediction***

161 We annotated all fine-mapped variants using Variant Effect Predictor tool
162 (https://grch37.ensembl.org/Homo_sapiens/Tools/VEP)⁵⁹ based on human genome GENCODE version
163 19 protein coding transcripts. The rsID and “Consequence” columns were extracted and deduplicated to
164 obtain a list of rsIDs and their possible impacts. Additional descriptions of the predicted “consequences”
165 for a given variant can be found at
166 https://useast.ensembl.org/info/genome/variation/prediction/predicted_data.html.

167 ***Melanocyte cell culture***

168 We obtained frozen aliquots of melanocytes, isolated from discarded foreskin tissue of healthy newborn
169 males, from the SPORE in Skin Cancer Specimen Resource Core at Yale University as described
170 previously^{23,24}. For capture-C analysis we used three distinct cultures of European ancestry and two of
171 African ancestry (C24, C27, C56, C140, C205). Melanocytes were either grown in Dermal Cell Basal
172 Medium (ATCC PCS-200-030) supplemented with Melanocyte Growth Kit (ATCC PCS-200-041) and 1%
173 amphotericin B/penicillin/streptomycin (120-096-711, Quality Biological) for QTL^{23,24} and capture-HiC
174 analysis, or alternatively in M254 (Invitrogen, M254500) supplemented with HMGS-2 (Invitrogen, S0165)
175 for all other experiments. Cells were grown at 37°C with 5% CO₂. All cells tested negative for
176 mycoplasma contamination using MycoAlert PLUS mycoplasma detection kit (LT07-710, Lonza).

177 **Capture-HiC bait design and library preparation**

178 Capture-HiC baits were designed by Arima Genomics (San Diego, CA, 2x tiling, least stringent masking,
179 XTHSBoosting) to obtain an Agilent Sure Select library (Santa Clara, CA) targeting all restriction
180 fragments (recognition sequences: ^GATC, ^GANTC) covering entire regions of association for the 68
181 independent genome-wide significant signals (**Table S2**)¹⁰.

182 For most regions, we used the LLR- and LD-based fine-mapping to define regions of association. The
183 region of association was defined by the two outermost fine-mapped variants and was extended by at
184 least one restriction fragment. Exceptions to this were made for the following genomic loci: For the
185 *5p15.33* locus (locus 11, signals 11-13, **Table S1**), capture baits were designed to cover the entire region
186 spanning both the *TERT* and *CLPTM1L* genes (chr5: 1230000-1360000, ~130 kb). For the *7q31.11* locus
187 (locus 22, signal 25), we extended the region of association to encompass the complete LD block as
188 defined by LD link ⁵⁰⁻⁵² (chr7:124392512-124710858). For the *9p21.3* locus (locus 27, signals 30-35, **Table**
189 **S1**), we included the entire region spanning from the *MTAP* to *DMRTA1* genes (chr9:21790755-
190 22452478, ~660 kb). Finally, for the *21q22.3* locus (locus 51, signal 65, **Table S1**), we extended the
191 region to include the previously functionally fine-mapped variant rs398206⁶⁰.

192 Bait sequences are listed in **Table S2**. Hi-C libraries were generated using the Arima HiC kit (Arima
193 Genomics) and the KAPA HyperPrep kit (KAPA Biosystems) following the manufacturer's protocol.
194 Briefly, 2-4 million cells were crosslinked, enzyme digested, and ligated. The ligated DNA was reverse-
195 crosslinked, fragmented by sonication, and size-selected for adaptor ligation and library amplification.
196 The HiC library was then hybridized with the custom capture baits and captured by the SureSelect XT HS
197 and XT low input library preparation kit for ILM (Agilent). 15 capture-HiC libraries were made from five
198 human primary melanocyte cultures (C56, C140, C205, C24, and C27) with three technical replicates per
199 melanocyte culture. Barcoded capture-HiC libraries were pooled and sequenced using an Illumina

200 Novaseq, with one run on an SP and a second run on an S1 flowcell, generating ~5.7 billion paired-end
201 reads with 150bp read length, for a median coverage of ~350 million read pairs per technical replicate,
202 ~1.1 billion read pairs per culture.

203 ***Capture-HiC chromatin interaction analysis***

204 Paired-end sequencing reads were pre-processed using the HiCUP pipeline⁶¹ and aligned to the human
205 genome version 19 using Bowtie2⁶². The summary of quality-control (QC) of sequencing reads for each
206 replicate are shown in **Table S3**. For each melanocyte culture, the aligned reads were pooled across the
207 technical replicates. Chromatin interaction loops were detected at one and four restriction fragment
208 resolutions, separately, using CHiCAGO pipeline version 1.16.0⁶³, treating each of the five melanocyte
209 cultures as biological replicates. As described previously⁴², the four-fragment resolution was created
210 using the artificial “.baitmap” and “.rmap” files, where four consecutive restriction digestion fragments
211 were grouped into one fragment (baitmap files provided in **Tables S4 and S5**). We used the default
212 parameters for one- fragment analysis except for minFragLen, maxFragLen, binsize, maxLBrownEst
213 which were set to 75, 1200, 2000, and 150000 respectively (**Table S6**). Four-fragment analysis was
214 conducted using default parameters except for minFragLen, maxFragLen, binsize, maxLBrownEst which
215 were set to 150, 5000, 8000, and 600000 respectively (**Table S6**). Based on the literature^{39,42}, for four
216 fragment analysis the removeAdjacent parameter was set to FALSE. Following CHiCAGO tool
217 recommendations, chromatin interaction capture-HiC loops with CHiCAGO scores ≥ 5 were considered
218 high-confidence interactions and were further analyzed. The output file was generated using the long
219 range interaction format and big interact format for visualization on the WashU Epigenome Browser^{64,65}
220 and UCSC genome browser⁶⁶⁻⁷⁰ respectively.

221 Considering the wide possible range of cross-linking resolution, the potential for incomplete restriction
222 digestion, as well as the fact that some variants may be located at the edge of restriction fragment bins,

223 data from adjacent bins may also reflect physical interactions from fine-mapped variants to genes.
224 Therefore, to inclusively identify such interactions, we also considered interaction data from adjacent
225 restriction fragment bins. Specifically, for each fine-mapped variant, we defined a genomic window +/-
226 500 bases and assessed whether any adjacent restriction fragments overlap this window. In this case,
227 we assessed chromatin interactions from the restriction fragment bin harboring the variant itself as well
228 as any overlapping adjacent restriction fragment bin.

229 ***Capture-HiC target gene nomination for GWAS risk loci***

230 For each GWAS signal, we mapped chromatin interaction loops between baited restriction fragments
231 overlapping fine-mapped risk variants (see above) and the promoter regions of the target gene
232 transcripts as per the GENCODE version 19^{67,71}. We defined the promoter regions of the respective
233 target genes using three criteria, identifying genomic regions with histone marks consistent with active
234 promoters in melanocytes and melanoma cells, as well as using a broader definition for gene promoters
235 regardless of activity in melanocytic cells. Specifically:

236 1) **Melanocyte-specific active promoter regions:** We annotated genomic regions as melanocyte-
237 specific active promoters using the NIH Roadmap Epigenomics Mapping Consortium¹⁶
238 ChromHMM imputed state model annotations^{72,73} derived from human primary neonatal
239 melanocyte cultures (imputed ChromHMM state model data was downloaded from the UCSC
240 genome browser Roadmap Epigenomics Integrative Analysis Hub for melanocyte samples E059,
241 E061). Genomic regions were annotated as melanocyte-specific promoter if they overlapped
242 with the ChromHMM imputed model states annotated as: PromU (Promoter Upstream
243 transcriptional start site; TSS), PromD1 (Promoter Downstream TSS with DNase), PromD2, TssA
244 (Active TSS), PromP (Poised Promoter), PromBiv (Bivalent Promoter), Tx_Reg (transcription
245 regulator). Data were available from two melanocyte cultures (E059, E061) and were merged

246 such that regions were defined as promoter if they overlapped promoter annotated sequence
247 from either of the two cultures. Subsequently, target genes were assigned for these promoters
248 based on whether the promoter region overlapped +/-2.5 kb of a transcription start site (TSS)
249 for any GENCODE version 19 comprehensive protein coding and non-coding transcripts⁷¹.

250 2) **Melanoma-specific active promoter regions:** Similarly, we defined melanoma-specific active
251 gene promoters using publicly available ChromHMM data from two different engineered
252 melanoma cell-models, HMEL and PMEL⁷⁴. Both HMEL and PMEL cell lines were originally
253 derived from primary foreskin melanocytes, immortalized by overexpression of *TERT*, and
254 introduction of oncogenic *CDK4*^{R24C}, dominant negative *TP53*, and *BRAF*^{V600E}⁷⁵. We used the
255 ChromHMM data specifically from tumorigenic cell line variants with shRNA mediated *PTEN*
256 knockdown (HMEL-sh*PTEN* and PMEL-sh*PTEN*)⁷⁴. Melanoma-specific active promoter regions
257 were defined if annotated as the following ChromHMM states: 1_TssA, 2_PromWkD, 3_TssWkP.
258 Target genes were assigned where the promoter region overlapped +/-2.5 kb of a TSS for any
259 GENCODE version 19 comprehensive protein coding and non-coding transcripts⁷¹.

260 3) **Globally defined gene-promoter regions:** We utilized the ENCODE based⁷⁶ promoter boundary
261 criterion in order to more globally define promoter regions regardless of activity in melanocytic
262 cells. Global promoters were defined as regions +/-2.5 kb of a TSS for each of the GENCODE
263 version 19 comprehensive protein-coding transcripts⁷¹. 16,663 genes annotated with a
264 respective promoter using the globally defined promoter criteria also have melanocyte- and/or
265 melanoma-specific active promoters.

266 ***ATAC-seq library generation and data analysis***

267 As described previously⁴⁷, 30K-50K primary melanocytes were lysed with cold lysis buffer (10 mM Tris-
268 HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% IGEPAL CA-630) and centrifuged to obtain nuclei. The
269 nuclei were resuspended in the transposition reaction mix (2x TD Buffer, Illumina Cat #FC-121–1030,
270 Nextera), 2.5 µl Tn5 Transposase (Illumina Cat #FC-121–1030, Nextera) and Nuclease Free H₂O on ice
271 and then incubated for 30 min at 37C. The transposed DNA was then purified using the MinElute Kit
272 (Qiagen). PCR amplification was performed using Nextera primers for 12 cycles to generate each single
273 library and PCR reaction cleanup was performed using AMPureXP beads (Agencourt). ATAC libraries
274 were sequenced on an Illumina NovaSeq platform using paired-end sequencing. We sequenced 15 ATAC
275 libraries from five independent primary melanocyte cultures (C24, C27, C56, C140, C205), with three
276 technical replicates for each melanocyte culture. The ATAC-seq reads from the technical replicates were
277 merged for each melanocyte culture. We processed the ATAC sequencing (ATAC-seq) data using the
278 ENCODE ATAC-seq pipeline version 1.6.1 (<https://www.encodeproject.org/atac-seq/>), treating five
279 melanocyte cultures as biological replicates. Sequencing reads were aligned to hg19 using bowtie2^{62,77}.
280 The pipeline requires generating two pseudo replicates via random sampling of reads from pooled
281 biological replicates for peak calling. ATAC peaks were called using MACSv2 peak caller (2.1.0)⁷⁸ ($P <$
282 0.01) and regions overlapping ENCODE blacklisted regions were removed⁷⁹. ATAC peaks were called
283 from the five melanocyte biological replicates, pooled biological replicates, and the pseudo-replicates.
284 The ATAC overlap reproducibility peaks were identified via the optimal criteria assessing peak overlap
285 between individual biological replicates, pooled biological replicates, and the pseudo biological
286 replicates. ATAC-peaks were analyzed and visualized on the WashU Epigenome Browser^{64,65} and UCSC
287 genome browser⁶⁶⁻⁷⁰.

288 ***ATAC-seq data from melanoma cell lines***

289 We analyzed publicly available omni-ATAC-seq data⁸⁰ generated from nine melanoma cell lines derived
290 from melanoma patient biopsies⁸⁰. Melanoma cell lines were annotated as belonging to the following
291 melanoma tumor states: melanocyte-like melanoma cell line (data available from three individual cell
292 lines), the intermediate-like melanoma cell lines (data available from three individual cell lines), and
293 mesenchymal-like melanoma cell lines (data available from three individual cell lines). We assessed fine-
294 mapped variant overlap with ATAC-seq peaks in any of the nine melanoma cell lines using the Bedtools
295 package (intersect function)^{81,82}.

296 ***Melanocyte-specific and melanoma-specific enhancers***

297 To refine variant and candidate gene selection we utilized publicly available chromatin state data from
298 both melanocytes and melanoma cells to identify genomic regions with promoter or enhancer histone
299 marks. Promoter regions were defined as described above. We annotated genomic regions as
300 melanocyte-specific enhancers using Roadmap ChromHMM data from two primary human melanocyte
301 cultures^{16,73}. We utilized the following enhancer annotated states in the primary ChromHMM (Enh,
302 EnhG, EnhBiv), auxiliary ChromHMM (EnhG1, EnhG2, EnhA1, EnhA2, EnhWk, EnhBiv), or the imputed
303 ChromHMM (TxEnh5, TxEnh3, TxEnhW, EnhA1, EnhA2, EnhAF, EnhW1, EnhW2, EnhAc) model data to
304 define melanocyte specific enhancers. We classified a region as enhancer if marked as an enhancer in
305 either cell line for either model. Melanoma-specific active enhancers were defined as regions annotated
306 as any of the following ChromHMM states: 4_EnhA, 5_EnhM, 6_EnhW, 7_TxEnhM, 7_TxEnhW, and
307 9_TxWkEnhW, from the tumorigenic melanoma cell models described above⁷⁴; regions were classified
308 as an enhancer if marked as an enhancer in any cell line. Any region annotated as enhancer or promoter
309 was considered regulatory. Data were analyzed using the Bedtoolsr package^{81,82}.

310 ***Massively Parallel Reporter Assay (MPRA) data from melanocytes and melanoma cell***

311 ***lines***

312 We previously performed episomal Massively Parallel Reporter Assays (MPRA) in immortalized primary
313 melanocytes and the UACC903 melanoma cell line to assess variant allele-specific transcriptional
314 activity²⁵. The MPRA study design conducted fine-mapping using LLR- and LD-based criteria similar to the
315 fine-mapping strategies utilized in this manuscript (LLR<1:1000 to the primary lead variant at each locus,
316 LD $r^2>0.8$ 1000 Genomes Project Phase 3 EUR for secondary signals and variants not included in the
317 GWAS summary data). 1,701 out of the 1,948 fine-mapped candidate variants from this study were
318 successfully tested using MPRA. The remaining 247 variants could not be assessed for two reasons²⁵:
319 102 were not amenable to design or failed assay QC, and 145 were not fine-mapped in the MPRA study
320 due to slightly different fine-mapping criteria.

321 ***QTL colocalization and TWAS/MWAS in melanocytes and melanoma***

322 QTL Colocalization as well as TWAS/MWAS analyses of the GWAS summary data with the expression-
323 QTL or methylation-QTL datasets were available from our previous melanocyte QTL studies^{10,23,24}
324 (dbGaP: phs001500.v1.p1); melanocyte-eQTLs and meQTLs were generated using 106 primary
325 melanocyte cultures derived from individuals mainly of European descent; methylation probes were
326 assigned to genes as previously described (CpG probes located within 1.5 kb of the TSS, 5'-UTR, 1st exon,
327 gene body, or 3'-UTR of a gene)²³. Colocalization for the secondary marginal GWAS signal near locus 4
328 (signal 4) was performed using the ezQTL website (<https://analysistools.cancer.gov/ezqtl/#/home>)⁸³
329 (Parameters: LD- 1000 Genomes Project EUR population, window= +/-250 kb). ezQTL performs
330 colocalization using HyPrColoc⁸⁴ as well as eCAVIAR⁸⁵ (eCAVIAR results use a 100 kb window centered on
331 the lead GWAS variant). Consistent with prior studies of melanocyte QTLs^{23,24}, we considered

332 colocalization significant where the HyPrColoc posterior probability exceeded 0.5 or the eCAVIAR CLPP
333 exceeded 0.01.

334 In addition, we assessed QTLs from TCGA melanoma tumors⁸⁶. eQTL colocalization was performed using
335 pre-analyzed QTL data on the ezQTL website. We performed melanoma TWAS using the pre-computed
336 weights from 103 TCGA melanoma samples ([http://gusevlab.org/projects/fusion/#the-cancer-genome-
337 atlas-tcga-tumornormal-expression](http://gusevlab.org/projects/fusion/#the-cancer-genome-atlas-tcga-tumornormal-expression)) using FUSION (<http://gusevlab.org/projects/fusion/>)⁸⁷. Melanoma
338 meQTL colocalization and MWAS were performed as described previously²³.

339 For assessing nominal eQTL support for 195 high-confidence candidate genes nominated via chromatin
340 interaction data, we assessed QTLs specifically between the interacting fine-mapped variant and its
341 putative target(s); where candidate genes were outside the +/- 1 Mb *cis*-window previously used for
342 melanocyte eQTL analysis²⁴, we specifically tested the association between fine-mapped variant
343 genotype and nominated target gene expression in the melanocyte eQTL data using a linear regression
344 model, where the input to the model included the interacting variant genotype and additional
345 covariates (3 genotyping principal components and 15 PEER factors) from the previous melanocyte
346 study²⁴. TCGA melanoma eQTLs were only assessed for genes within +/- 1Mb *cis* window.

347 ***Candidate gene expression in melanocytes and melanomas***

348 For use with an integrative scoring system (described below), we assessed gene expression in
349 melanocyte²⁴ and melanoma (The Cancer Genome Atlas project, SKCM Pan-Cancer build accessed via
350 cBioPortal)⁸⁶ gene expression datasets. For each gene expression dataset, we filtered out genes that are
351 not expressed by excluding those with an RSEM value <0.1 in more than 20% of samples criteria. For the
352 remaining genes, we calculated the median expression percentile across samples in the melanocyte and
353 melanoma datasets, respectively.

354 ***Pathway enrichment analyses***

355 Pathway and upstream-regulator enrichment analyses were performed using the Ingenuity Pathway
356 Analysis (IPA) tool (Qiagen)⁸⁸. Pathway enrichment *P*-values were calculated using the Ingenuity
357 knowledge base (genes only) as the reference set and using default parameters. The IPA tool parameter
358 “Relationships to consider” was set to “Direct relationships” for the upstream regulator analysis.

359 ***CRISPRi validation of regulation of target genes transcription***

360 CRISPR interference (CRISPRi) was performed in the immortalized human melanocyte cell line
361 C283T/dCas9-KRAB. The immortalized human melanocyte cell line C283T⁴⁷ was infected with a lentiviral
362 vector pLX_311-KRAB-dCas9 (gift from John Doench, William Hahn, and David Root; Addgene plasmid #
363 96918; <http://n2t.net/addgene:96918>; RRID:Addgene_96918)⁸⁹ followed by monoclonal cell selection.
364 We validated dCas9-KRAB expression and activity in the clone used for CRISPRi validation. For each
365 variant tested, three different guide RNAs (gRNAs) were designed to target the genomic regions
366 surrounding the variant, with the gRNA sequence located within/around +/- 50bps from the variant
367 (sequences for guides designed to target the region surrounding each variant are listed in **Table S7**). Two
368 non-targeting gRNAs were used (NTC1, NTC2). gRNAs were ligated into the lentiviral vector pXPR-050
369 (gift from John Doench and David Root, Addgene plasmid #96925; RRID: Addgene_96925)⁹⁰. Cells were
370 infected with lentiviral particles encoding gRNA and at 24h after infection, 1.5 µg/mL of puromycin was
371 added for selection. After two days of puromycin selection, puromycin was removed, and cells were
372 harvested for RNA collection on the same day or one day after puromycin removal. Total RNA was
373 isolated with RNeasy Mini Kit (Qiagen) and cDNA was generated with SuperScript IV VILO Master Mix
374 (Thermo Scientific). For each variant, at least three infections were performed, with two biological
375 replicates per infection for each of the three gRNAs against the variant. mRNA levels of the candidate
376 target genes being assessed were measured by Taqman assay (Thermo Scientific) and normalized to

377 *GAPDH* levels. qPCR triplicates (technical replicates) were averaged and subsequently considered as a
378 single data point. Data from NTC1 was used for statistical comparisons of other gRNAs, data in tables
379 and graphs were represented as fold-change relative to the average of NTC1. The statistical analysis was
380 performed using a paired two-tailed t-test comparing delta-Ct values.

381 ***Mutational cancer driver genes***

382 Mutational driver genes identified from analysis of melanoma cohorts and pan-cancer dataset were
383 available from the intOGen database⁹¹ (<https://www.intogen.org/search?cancer=MEL> and
384 <https://www.intogen.org/search#driver-genes:table>). We used this nominated target gene list to
385 identify if any of the candidate genes were also identified as cancer drivers in melanoma and pan-cancer
386 datasets.

387 ***Candidate gene prioritization via integrative scoring***

388 We created a scoring scheme (**Figure S1**) for candidate gene prioritization at each melanoma risk signal
389 using complementary information from fine-mapping, chromatin interaction, cell-type specific
390 epigenomic (chromatin state, accessibility), gene expression (eQTL/TWAS), DNA methylation
391 (meQTL/MWAS), MPRA, and mutational cancer driver datasets.

392 At each risk signal, if a candidate gene was nominated via melanocyte- or melanoma-specific expression
393 QTL colocalization analyses or TWAS, we considered it as a strong plausible candidate and added 6
394 points to the total score for the gene.

395 For the remaining genes, we cumulatively assigned scores if:

- 396 • The gene is nominated by a fine-mapped variant being linked to the gene promoter via
397 chromatin interaction or physical location within a promoter (+1)

- 398 • The gene is nominated by a fine-mapped variant within a melanocyte or melanoma
399 enhancer/promoter region and being linked to the gene promoter (+1)
- 400 • The gene is nominated by a fine-mapped variant that overlaps a melanocyte or melanoma
401 enhancer/promoter region, displays significant allelic transcriptional activity, and is linked to the
402 gene promoter (+1)
- 403 • The gene is nominated by a fine-mapped variant that overlaps a melanocyte or melanoma
404 enhancer/promoter region, displays significant allelic transcriptional activity, is a marginally
405 significant eQTL for the gene ($P < 0.05$), and is linked to the gene promoter (+1)
- 406 • The gene is nominated via methylation colocalization analysis or MWAS approaches (+2).

407 For all nominated candidate genes, we then added to the score if a candidate gene was identified as
408 driver in melanoma (+1) or pan-cancer (+1) analyses from the intOGen database⁹¹.

409 **RESULTS**

410 ***Fine-mapping 68 independent melanoma risk signals from GWAS***

411 We performed fine-mapping at 68 genome-wide significant melanoma GWAS risk signals¹⁰ using a
412 combination of complementary approaches in order to comprehensively and inclusively identify
413 potential causal variants (credible causal variants: CCVs). We first fine-mapped using GWAS summary
414 data, selecting variants with log-likelihood ratios (LLR) $< 1:100$ relative to the lead variant at each locus.
415 We also performed Bayesian fine-mapping using DAP-G^{53,54}, identifying credible sets that directly
416 correspond to each of the 68 signals. Lastly, to identify potential causal variants that are not present in
417 the summary data due to quality control filters and or imputation reference choice, we used an LD-
418 based fine-mapping strategy ($r^2 > 0.8$, 1000 Genomes Project, Phase 3, Version 5, EUR). We selected all

419 variants fine-mapped by at least one approach as CCVs, for a total of 1,948 variants. 1,477 variants were
420 fine-mapped by DAP-G, while 1,892 were fine-mapped using the LLR/LD approaches (**Table S8**). 1,421
421 were identified by both approaches, suggesting that distinct fine-mapping approaches largely prioritize
422 the same set of variants as credible causal variants (**Figure 1A, Figure S2**).

423 As expected, a large proportion of the fine-mapped CCVs were in non-protein coding regions and only a
424 few variants were identified as directly impacting the protein coding sequence (**Tables S9 and S10**).

425 Among the 1,948 fine-mapped variants, variants altering protein coding sequence (i.e., variant
426 annotated as missense variant or frame shift variant) were observed at the 20 of the 68 GWAS risk
427 signals (20 genes; **Tables S9 and S10**); we observed no variants in consensus splice donors or acceptors.
428 Therefore, at these risk signals we considered the affected gene as a potential target gene based on the
429 variant impacting protein coding sequence; amongst these are well-established pigmentation genes
430 (*MC1R*⁹²⁻⁹⁴, *SLC45A2*^{92,95-97}, *TYR*^{2,10,92,98,99}), well-characterized protein-coding melanoma risk variants
431 (*MITF*^{100,101}), and variants in well-established cancer genes (*OBFC1*^{8,102,103}, *ATM*^{104,105}). Most melanoma
432 GWAS risk signals lack protein-altering CCVs, and even still, the presence of such variants within a
433 credible set for a locus does not exclude the possibility of *cis*-regulation being the causal mechanism.

434 ***Mapping chromatin interactions at 68 GWAS signals using a custom region-focused*** 435 ***capture-HiC assay***

436 We performed a custom capture-HiC assay to resolve chromatin interaction patterns at melanoma
437 GWAS risk signals. We designed custom capture baits tagging all restriction digestion fragments tiled
438 across the entire region of association for each locus (**Table S5**). 88% of fine-mapped variants (1,717 out
439 of the total 1,948 fine-mapped variants) were located within a baited restriction fragment (**Table S8**,
440 **Figure S3**). 94% of variants were either directly baited or were located adjacent to at least one baited
441 restriction fragment (**Table S8**). As binning groups of adjacent restriction fragments has been shown to

442 increase sensitivity to detect long range interactions^{39,40,47,106}, we also assessed baiting coverage for bins
443 of four adjacent restriction fragments (**Table S5**). We observed slightly better coverage of fine-mapped
444 CCVs in four-fragment analysis, as 95% of variants (1858 out of 1948) were located within a four-
445 fragment bin with at least one restriction fragment that was baited (**Table S8, Figure S3**). The capture-
446 HiC assay was performed using five primary cultures of human melanocytes from unrelated individuals
447 drawn from the melanocyte collection we previously used for eQTL and meQTL studies^{23,24}. To maximize
448 sensitivity to detect long-range interactions as others have done, as well as to assess regions harboring
449 variants that were not directly baited, we assessed collective groupings of four adjacent restriction
450 fragments (e.g. four-fragment analysis, 4F; **Table S5**) and compared to the analysis of individual
451 fragments (one-fragment analysis, 1F; settings for both analyses summarized in **Table S6**).

452 Consistent with previous studies^{39,42}, 4F analysis identified chromatin interactions spanning longer
453 distances (**Table S11**) in comparison to the 1F analysis. As expected given the larger number of bins
454 analyzed, the 1F analysis overall identified a higher number of unique chromatin interactions with better
455 resolution (>2.3 fold). We also observed similar distributions of CHiCAGO scores between the 1F
456 (median score=8.38) and 4F (median score=7.07) analyses (**Table S12**). Given the better coverage of
457 variants and increased sensitivity to detect long-range interactions, we used 4F data in the subsequent
458 analyses.

459 ***Capture-HiC links fine-mapped risk variants to candidate target genes at most loci***

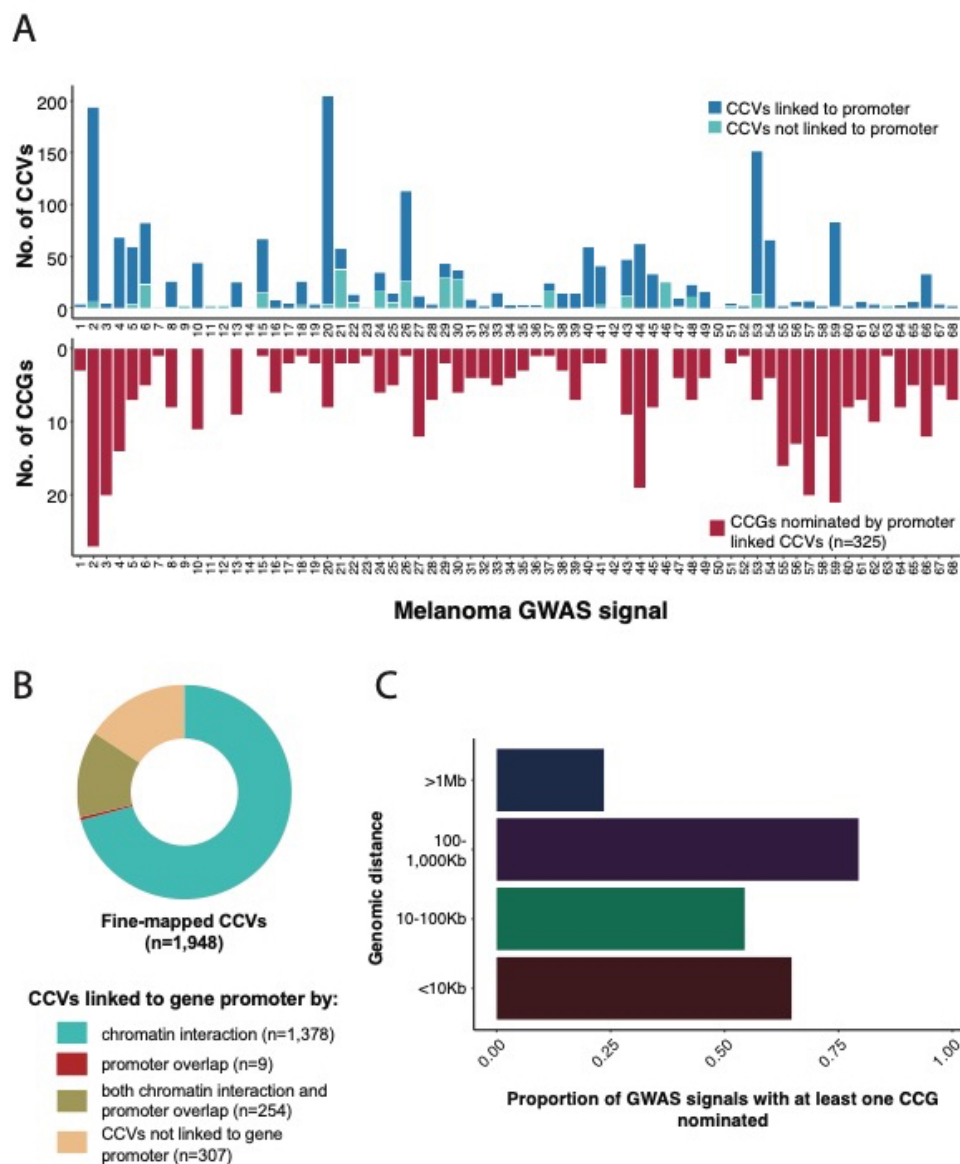
460 Next, we analyzed the capture-HiC chromatin interaction data to identify physical interactions between
461 fine-mapped CCVs at the 68 melanoma GWAS signals and gene promoter(s). We defined promoters in
462 three different ways. Firstly, in order to identify promoters in melanoma relevant cell-types where
463 target gene is more likely to be causal, we separately defined melanocyte- and melanoma-specific
464 promoters using ChromHMM state model data from two primary melanocyte cultures (ROADMAP

465 epigenome project¹⁶), as well as two engineered melanoma cell-models⁷⁴, respectively. We also more
466 broadly defined promoters regardless of activity in melanocytic cells using a general promoter definition
467 from the ENCODE consortium^{69,76,107}.

468 The capture-HiC data identified chromatin interaction loops from 84% of fine-mapped risk variants
469 (n=1,632) to at least one annotated promoter region, nominating 323 genes as candidate causal genes
470 (CCGs) for 61 melanoma GWAS signals (**Figures 2A-B, Figures S4, S5, and S6**). A small proportion of fine-
471 mapped variants were located directly within annotated gene promoters (13%, 263 out of 1,948
472 variants, 56 unique genes; **Figure 2B**). Most of these not surprisingly showed physical interactions within
473 the promoter itself, however there were nine additional CCGs without such an interaction nominating
474 an additional two genes as potential CCGs (**Figure S7**). 122 promoter-overlapping variants (6%) showed
475 chromatin interaction loops to an alternative promoter for the same gene (located at least 10 kb away).
476 234 promoter overlapping variants (12% out of 1,948 variants) showed chromatin interactions with
477 gene(s) other than the gene nominated by promoter overlap. Therefore, we considered both chromatin
478 interaction or direct promoter overlap criteria for linking variant to genes and collectively identified a
479 total set of 1,641 unique variants linked to 325 genes (**Figures 2A-B**).

480 A median of five candidate target genes were nominated per risk signal; nine GWAS risk signals had only
481 one target gene nominated. Eight fine-mapped CCGs per risk signal (median) were linked to at least one
482 target gene promoter; three risk signals had only one fine-mapped CCG linked to a gene promoter. For
483 seven signals (~10%), no candidate target could be nominated due to the lack of observed capture-HiC
484 chromatin interaction from the fine-mapped CCGs to any target gene promoter. Of these, three have
485 well characterized pathogenic protein coding changes (signal 9, *MITF*^{100,101}; signal 14, *SLC45A2*^{92,95-97};
486 and signal 42, *TYR*^{2,10,92,98,99}), while three others harbor known melanoma driver or pigmentation genes
487 (signal 11, *TERT*; signal 12, *TERT*; signal 50, *OCA2*);^{5,108-110}. We note that we did not observe the

488 previously reported interaction between rs12913832 within an enhancer in the gene body of *HERC2* and
 489 the promoter of *OCA2*¹⁰⁹.



490
 491 **Figure 2. Summary of fine-mapped candidate causal variants (CCVs) linked to potential target**
 492 **candidate causal genes (CCGs) at 68 melanoma GWAS risk signals.** (A) Stacked bar plot summary of
 493 fine-mapped CCVs and nominated target CCGs. The top bar plot (dark blue color) shows the number of
 494 CCVs linked by chromatin interaction or overlap with at least one gene promoter, while the light blue
 495 color shows the number of CCVs not linked to a promoter. The bottom plot shows the total number of
 496 nominated CCGs per locus. (B) Pie chart showing the proportion of all fine-mapped CCVs that are linked
 497 to target CCGs via distant promoter interactions, direct overlap with gene promoter regions, or both. (C)
 498 Bar plot summarizing the proportion of GWAS risk signals with at least one gene nominated through
 499 chromatin interactions over varying distances.

500 Notably, we nominated several distant target CCGs via long-range chromatin interaction between fine-
501 mapped CCVs and target gene promoter(s) (**Figure 2C, Figure S8**). A large proportion of the GWAS risk
502 signals (79%, or 54/68) had at least one CCG nominated by a chromatin interaction loop spanning
503 between 100-1000 kb distance from the fine-mapped CCV to the target gene promoter (**Figure 2C**). For
504 23% (16/68) of GWAS risk signals, at least one nominated CCG was located >1 Mb away from the fine-
505 mapped CCV (**Figure S8**).

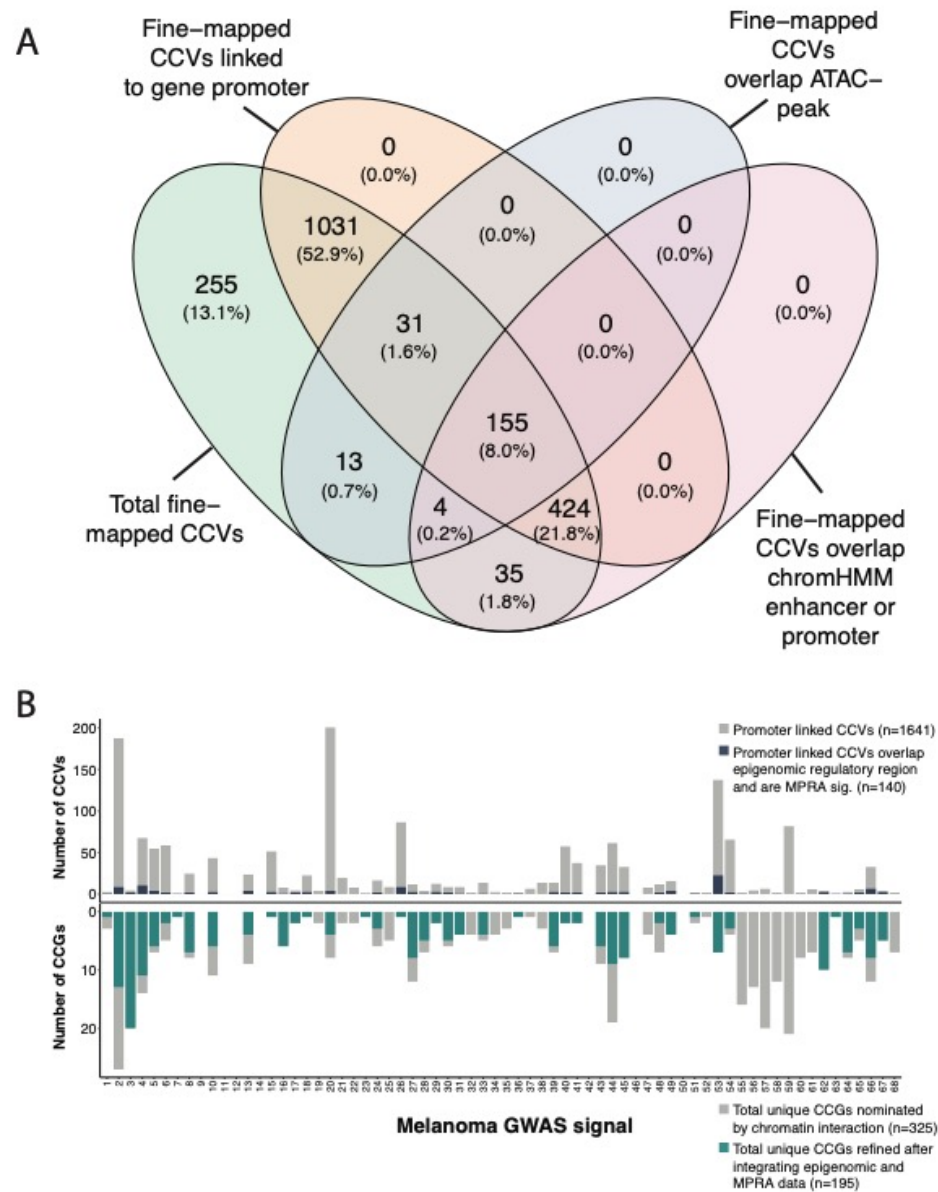
506 ***Refining variant and gene nomination via integration with cell-type specific***

507 ***epigenomic and massively parallel reporter assay data***

508 Not all interactions between fine-mapped risk variant and gene promoters are necessarily functional *cis*-
509 regulatory interactions, thus we sought to further refine candidate gene nomination using melanocyte-
510 and melanoma- specific epigenomic datasets as well as data from cell-type specific massively parallel
511 reporter assays (MPRA). Firstly, we utilized chromatin accessibility data (ATAC-seq) and chromatin state
512 annotations (ChromHMM) from human melanocytes and melanoma cell lines to identify those
513 interactions between gene promoter(s) and risk variants in potential regulatory elements. We generated
514 ATAC-seq data for the same five melanocyte cultures used in the capture-HiC assay (three replicates per
515 culture) and analyzed treating the five cultures as biological replicates. Additionally, we also utilized
516 publicly available ATAC-seq data from nine melanoma cell cultures⁸⁰. 203 fine-mapped variants (10%)
517 were located within annotated accessible chromatin regions in melanocyte or melanoma cells (**Figure**
518 **3A**); of these, 186 variants (9% of 1,948) were linked to a gene promoter via chromatin looping or direct
519 promoter overlap (**Figure S9, Table S13**). These variants were linked to the promoter(s) of 223 unique
520 genes nominated as potential candidates at 46 GWAS risk signals. In addition, we performed a similar
521 analysis using melanocyte- and melanoma-derived ChromHMM data, which provides broader regulatory
522 region definitions using multiple histone marks. Specifically, we utilized ChromHMM enhancer and

523 promoter annotations from two human melanocytes cultures from the Roadmap Epigenome Project^{16,73}
524 as well as published data from two engineered melanoma cell models⁷⁴. In contrast to the ATAC-seq
525 data, analysis using ChromHMM data identified considerably more fine-mapped variants located within
526 potentially regulatory regions (n=618, 32%) (**Figure 3A**), with 579 linked via looping or direct promoter
527 overlap to 275 total candidate genes at 56 GWAS risk signals (**Figure S10, Table S13**). In total, 610
528 potentially *cis*-regulatory variant-promoter interactions were identified using either ATAC-seq or
529 ChromHMM data (linked to 282 unique genes, 57 risk signals (**Figure S11, Table S13**), while 155 (209
530 genes) were identified using both datasets. The latter potentially represent stronger functional
531 evidence, however, given functional *cis*-regulatory variants may be found outside of ATAC-seq peaks, we
532 moved forward with a more inclusive approach to identify potential *cis*-regulatory variants (n=610).

533 We then sought to further refine our variant and candidate gene nomination using additional evidence
534 from melanocyte- and melanoma-specific episomal massively parallel reporter assays. We previously
535 assessed allele-specific *cis*-regulatory activity of fine-mapped melanoma risk variants²⁵. 1,701 out of the
536 1,948 fine-mapped variants were assessed in both melanocyte cultures and melanoma cells, of which
537 349 were FDR-significant in either melanocytes or melanoma cells. We subsequently assessed those
538 MPRA-significant variants located within potentially *cis*-regulatory regions as described above,
539 identifying 140 variants linked to 195 genes at 42 risk signals (“high-confidence gene set”; **Figure 3B**,
540 **Table S13**). While we cannot exclude the possibility that variants and genes filtered out using this
541 strategy may be functional under specific cellular contexts not evaluated here, this set of 140 variants
542 and 195 genes represent strong functional leads.



543

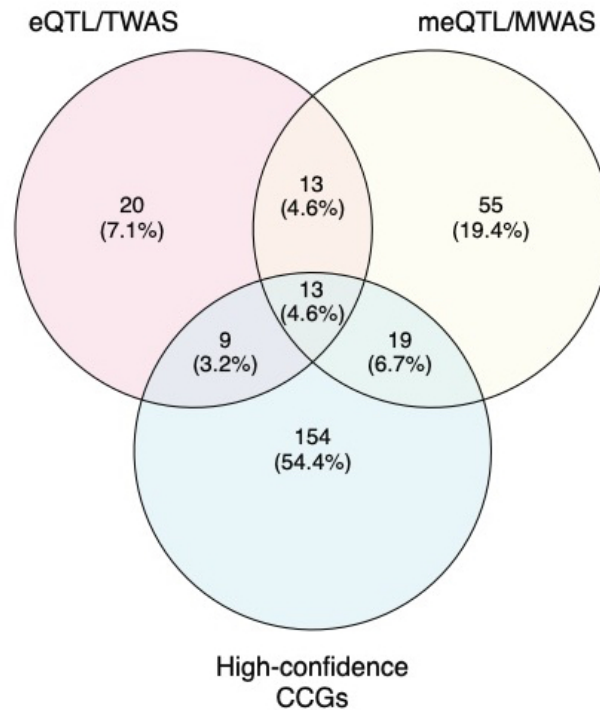
544 **Figure 3.** (A) Summary of fine-mapped variants overlap with chromatin interaction *cis*-regulatory regions
 545 in the ATAC-seq and ChromHMM datasets. (B) Stacked bar plot summary of fine-mapped variants (CCVs)
 546 and nominated target genes (CCGs) after integrating the chromatin interaction dataset with
 547 melanocyte- and melanoma-specific ATAC-seq, ChromHMM, and MPRA datasets for each of 68
 548 melanoma risk signals. The top bar plot shows the total number of fine-mapped variants that are linked
 549 to at least one target gene using the chromatin interaction dataset, while blue color shows the number
 550 of interacting variants overlapping a potential regulatory region in any of the ATAC-seq or ChromHMM
 551 datasets and the variant is also FDR significant in MPRA dataset. The bottom plot shows the number of
 552 unique genes nominated as potential candidates using chromatin interaction data only, while the green
 553 color shows the number of candidate genes following integration with epigenomic (ATAC-seq and
 554 ChromHMM) and MPRA datasets.

555 We looked into cell-type specificity of the high-confidence gene set. In general, roughly half of these
556 high-confidence capture-HiC-nominated genes were identified via interactions with variants that were
557 MPRA-significant and/or located within regulatory regions of both melanocytes and melanomas. Slightly
558 less than half of these high-confidence genes were nominated by analysis of only melanocyte
559 epigenomic/MPRA data but not a similar analysis using melanoma data, e.g. melanocyte-specific
560 candidates (n=88 genes, **Table S14**). Very few gene candidates were identified solely by analysis using
561 only melanoma epigenomic/MPRA data (n=17 genes, **Table S15**). Perhaps unsurprisingly given the role
562 of normal melanocytes in regulating pigmentation, analyses using only melanocyte data nominated
563 candidates for far more loci found to be implicated in pigmentation phenotypes by Landi and
564 colleagues¹⁰ (n=13 loci) versus analyses using only melanoma data (n=1 locus). Thus, in all, these data
565 suggest that a substantial number of loci may retain function in both melanocytes and melanoma cells
566 and that analyses using melanocyte data appear to better annotate melanoma GWAS loci.

567 ***An integrative scoring system to prioritize GWAS candidate causal genes.***

568 We previously used cell-type specific expression and methylation QTL datasets (eQTL, meQTL) to
569 nominate candidate genes for melanoma GWAS risk loci^{10,23,24}. Specifically, we performed colocalization
570 for FDR-significant QTLs, as well as identifying FDR-significant transcriptome- and methylome-wide
571 association study (TWAS, MWAS) genes using data from both a panel of primary human melanocyte
572 cultures as well as melanoma tumors from The Cancer Genome Atlas Project^{10,23,24}. 40% of candidate
573 genes (22 out of 55) nominated via eQTL colocalization and/or TWAS from melanocytes and melanoma
574 were also nominated as high-confidence genes from capture-HiC data (**Figure 4, Table S16**). For meQTL-
575 colocalizing and MWAS-significant genes, only 32% (32 out of 100) were identified in the high-
576 confidence gene set. Of the meQTL/MWAS-nominated candidate genes that lack significant eQTL/TWAS

577 support, 26% (19 out of 74) were nominated as high-confidence candidates in chromatin interaction
578 analyses.



579

580 **Figure 4. Summary of overlapping CCGs between QTL datasets and capture-HiC chromatin interaction**
581 **analyses.** eQTL/TWAS CCGs were nominated when colocalization of eQTL and GWAS data was observed,
582 or alternatively when the gene was identified as FDR-significant via Transcriptome Wide Association
583 Study (TWAS), using either primary melanocyte or melanoma tumor eQTL reference datasets. Likewise,
584 meQTL/MWAS CCGs were nominated via meQTL colocalization or an FDR-significant Methylome-Wide
585 Association Study finding, where the significant CpG probe was located within a gene promoter or gene
586 body, and meQTL reference datasets from melanocytes and melanoma tumors were tested separately.
587 High confidence CCGs were nominated via integration analyses of fine-mapping, chromatin interactions
588 datasets with epigenomic (ATAC-seq and ChromHMM) and MPRA data derived from melanocytes and
589 melanoma cells.

590 Given that a large proportion of the 195 genes nominated via chromatin interaction analyses lack

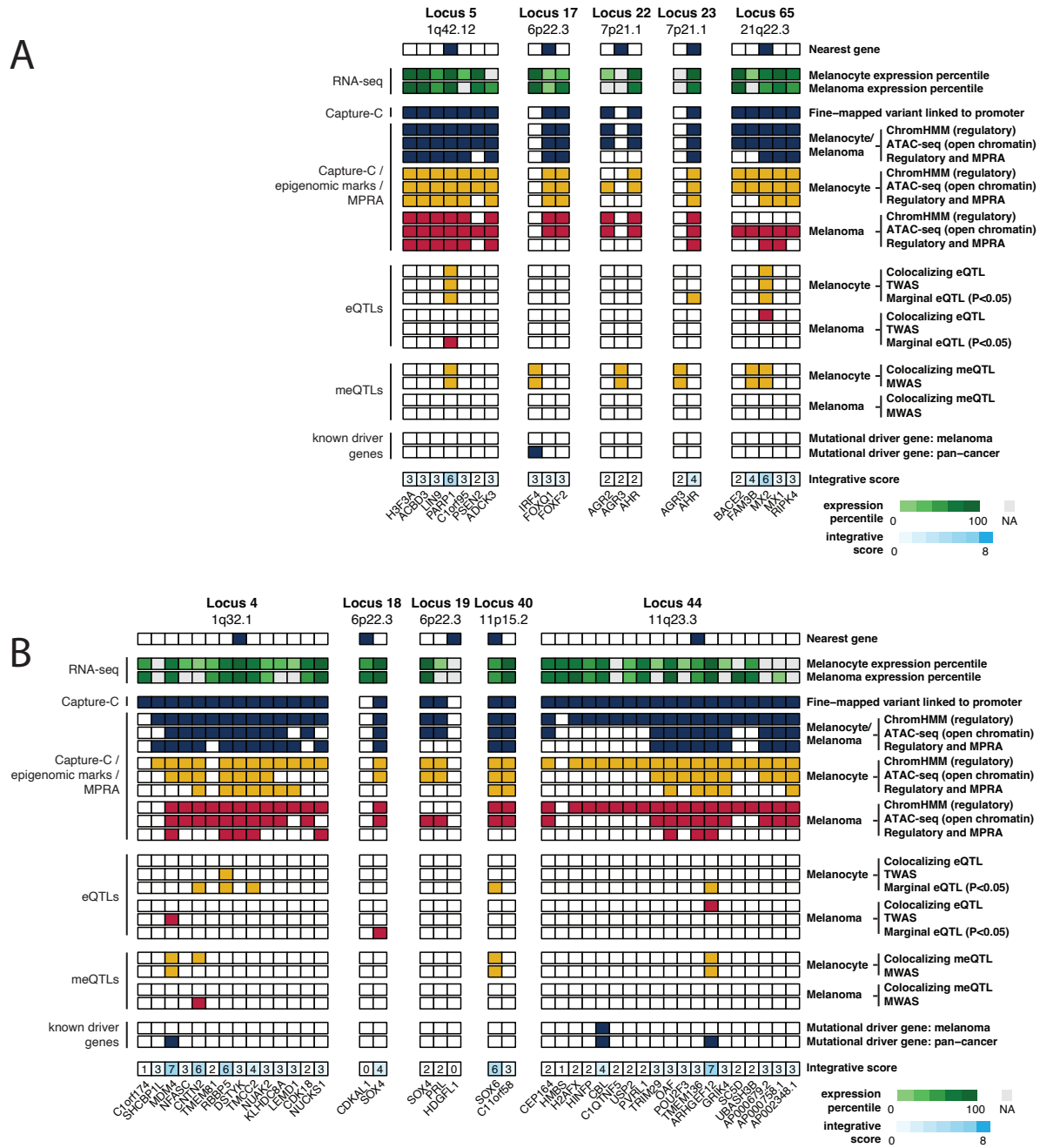
591 significant eQTL/meQTL/TWAS/MWAS support, we assessed whether any of these genes may be

592 supported via marginal eQTLs between the promoter-interacting variant and its putative target(s). In

593 total, 36 such genes displayed at least nominal QTL support ($P < 0.05$) in melanocyte and/or TCGA eQTL

594 datasets (**Table S13**).

595 Finally, to better prioritize *cis*-regulatory functional leads at each risk signal, we created a candidate
596 gene prioritization score integrating melanocyte and melanoma functional datasets (**Figure S1**). At each
597 signal, we considered the presence of significant colocalizing eQTL or TWAS findings, which suggest
598 potential shared causal variants between gene expression and melanoma risk, to be strong evidence for
599 candidate genes, contributing a total of six points to the gene score. For the remaining gene candidates,
600 we combined fine-mapping, chromatin interaction, cell-type specific epigenomic (chromatin state,
601 accessibility), DNA methylation (meQTL/MWAS), and MPRA evidence to assign a gene score up to six
602 points. For all candidates, we further added one point each where genes have been identified as
603 melanoma or pan-cancer driver genes in the intOGen database⁹¹, allowing for a maximum gene score of
604 8 (**Table S17, Figures 5A-B, Figures S12A-H**). In total, six risk signals had at least one candidate gene
605 score ≥ 7 , 37 with a gene scoring ≥ 6 , and 49 with a score ≥ 4 . At previously characterized loci, the
606 scoring system ranked the likely casual as the highest scoring gene, including *PARP1* at 1q42¹¹¹ (locus 5,
607 score = 6), *AHR* at 7p21⁴⁷ (locus 23, score = 4), and *MX2* at 21q22⁶⁰ (locus 65, score = 6) (**Figure 5A**). For
608 novel loci, the scoring system ranks *MDM4* as the best scoring and *CBL* as the second-best candidate for
609 loci on chromosome bands 1q32 (locus 4, score =7) and 11q23 (locus 44, score =4) respectively. Both
610 genes are located more than 500 kb and 1 Mb from their respective lead variants (**Figure 5B**). In
611 addition, the scoring system nominates two SRY-related HMG-box genes, *SOX4* and *SOX6*, as the best
612 candidates at 6p22 (loci 18-19, scores =4,2) and 11p15 (locus 40, score=6), respectively (**Figure 5B**).
613 Collectively, our integrative scoring system-based gene prioritization approach re-identified previously
614 characterized susceptibility genes and provide additional support for functional investigation of novel
615 candidates.



616

617 **Figure 5. Integrative evidence for candidate causal genes at select melanoma risk signals.** (A) Loci with
 618 previously characterized candidate causal genes, and (B) select novel loci. For each locus, the figure
 619 indicates the nearest gene to the lead variant, summarizes candidate gene expression in primary
 620 melanocytes and melanoma tumors, indicates genes implicated by interaction of fine-mapped variants
 621 to the gene's promoter, along with further refined evidence for these interacting variants integrated
 622 with melanocyte and melanoma epigenomic and MPRA data. Also summarized are melanocyte
 623 eQTL/TWAS evidence, meQTL/MWAS evidence, and whether the candidate gene has been implicated as
 624 a melanoma or pan-cancer driver gene. Finally, the figures show an overall integrative score for each
 625 candidate scored from 0-8 with 8 being the highest score.

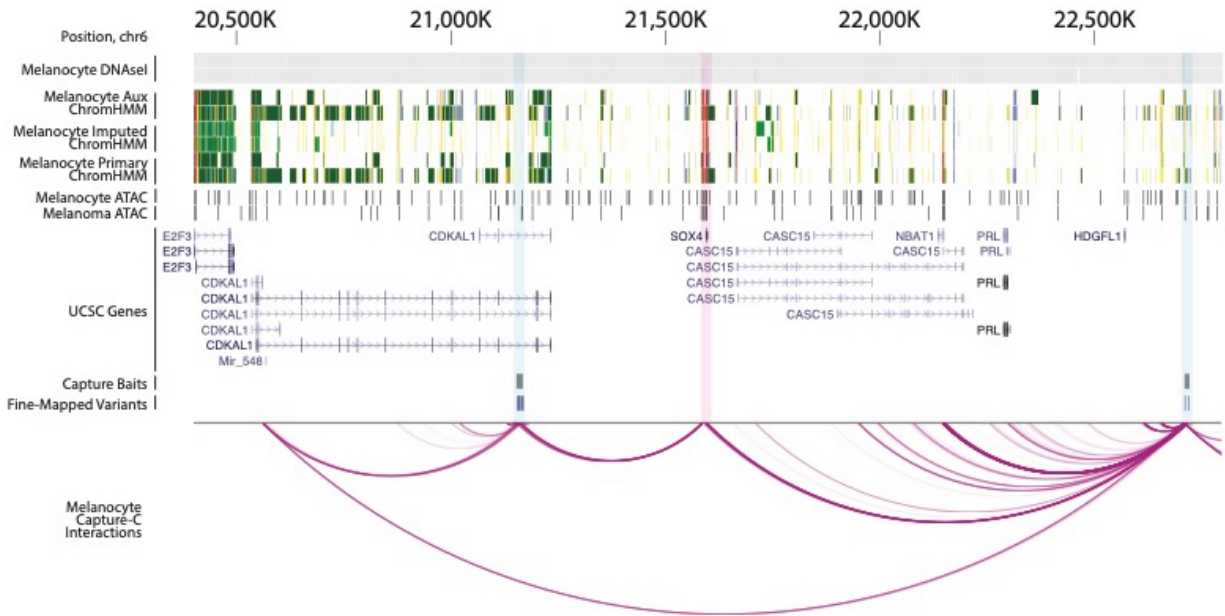
626 **Biological pathway enrichment analysis of capture-HiC nominated gene candidates**

627 To identify biological pathways associated with melanoma risk, we performed pathway enrichment
628 analysis of the collective set of genes nominated by (1) identification of protein-coding variants amongst
629 the full set of fine-mapped variants, (2) colocalizing eQTLs/meQTLs or FDR-significant TWAS/MWAS
630 findings, and (3) the set of high-confidence gene candidates (n=195 genes) nominated via capture-HiC
631 analyses (**Table S18**). We compared this to a pathway analysis of only these genes nominated by prior
632 QTL (eQTL/meQTL colocalization, TWAS/MWAS) analyses or protein-coding fine-mapped variants (**Table**
633 **S19**). Overall, pathway analyses including the high-confidence capture-C nominated genes had a greater
634 number of pathways enriched in comparison to that of protein-coding/QTL nominated genes. Novel
635 pathways identified only by the former analysis include p53 signaling ($P = 0.0005$), WNT/beta-catenin
636 signaling ($P = 0.002$), and interferon gamma signaling ($P = 0.003$) (**Table S18**); the capture-HiC gene set
637 further strengthened the evidence for enrichment in numerous pathways, including MITF-M-dependent
638 gene expression ($-\log_{10}P = 4.26$ vs 2.87), telomere maintenance ($-\log_{10}P = 3.52$ vs 1.9) and aryl
639 hydrocarbon receptor signaling ($-\log_{10}P = 3.22$ vs 1.87 ; **Tables S18-S19**) along with pigmentation
640 related pathways (melanin biosynthesis, $-\log_{10}P = 7.28$ vs 5.92 ; L-dopachrome biosynthesis, $-\log_{10}P =$
641 3.99 vs 2.0 ; melanocyte development and pigmentation signaling 3.29 vs 1.88 ; **Tables S18-S19**) which is
642 consistent with skin pigmentation-related phenotypes as critical risk factors for melanoma. We also
643 performed enrichment analysis for upstream transcriptional regulators of genes collectively nominated
644 by capture-HiC, eQTL/TWAS, meQTL/MWAS, and protein-coding fine-mapped variants (**Table S20**). We
645 found MITF to be the most enriched upstream regulator, consistent with its well-established role in
646 pigmentation and melanoma risk and progression. In addition, multiple upstream regulators related to
647 MITF were identified as enriched with this gene set, including ZEB2, which has been itself shown to
648 regulate *MITF* levels¹¹², as well as SMARCA4 (BRG1) which has previously been shown to be required for
649 MITF activation of melanocyte-specific target genes¹¹³.

650 ***CRISPRi validation of long-range cis-regulatory interactions***

651 To further assess and validate the regulation of potential target genes by fine-mapped functional
652 variants nominated by integrative analysis, we performed CRISPRi experiments to test the
653 transcriptional regulation of four nominated high-confidence candidate genes at five independent loci.
654 Firstly, we assessed two SOX family transcription factors given a well-established role for SOX proteins in
655 neural crest and melanocyte development, as well as the fact that capture-HiC data identify chromatin
656 interactions with multiple SOX genes. Specifically, *SOX4* was nominated via long-range interactions from
657 two independent loci located more than 1.5 Mb apart (locus 16, signal 18, ~400 kb; locus 17, signal 19,
658 ~1.1 Mb, **Figure 6**). *SOX6* was nominated via ~130 kb chromatin interactions (locus 32, signal 40) and
659 was also identified in melanocyte methylation QTL colocalization and MWAS analyses (**Figure S13A**). In
660 addition, we assessed loci interacting with known cancer driver genes. Specifically, *MDM4* was
661 nominated by long-range chromatin interaction (locus 4, signal 4, ~550 kb) and was further nominated
662 by melanocyte MWAS and TCGA melanoma TWAS analyses (**Figure S13B**). Finally, *CBL* was nominated by
663 a long-range ~1.1 Mb interaction (locus 36, signal 44; **Figure S13C**). As eQTLs for *CBL* were not previously
664 been tested¹⁰ given the 1.1 Mb distance, we evaluated multiple fine-mapped variants and observed a
665 marginal correlation between risk allele and *CBL* expression in primary melanocytes (including
666 rs2120430, $P = 0.008$; rs61898347, $P = 0.009$; rs11217853, and $P = 0.02$; rs61900794, $P = 0.02$; **Figure**
667 **S14; Table S21**).

668 For each locus, we chose fine-mapped variants using the collective evidence from capture-HiC,
669 epigenomic, and MPRA datasets, designed three guides targeting the region surrounding each variant,
670 and tested these guides for effect on target gene expression relative to a non-targeting control gRNA
671 (NTC1) in an immortalized human melanocyte cell clone stably expressing dCas9-KRAB.



672

673 **Figure 6. Chromatin looping from two independent loci on chromosome 6 to the promoter of *SOX4*.**

674 Figure shows data from melanocyte DNase I hypersensitivity sequencing (Roadmap, n=2 melanocyte
675 cultures), melanocyte ChromHMM (Roadmap, n=2 melanocyte cultures), melanocyte ATAC-seq (n=5
676 cultures), and melanoma cell ATAC-seq relative to genes in the region. Fine-mapped variants for both
677 loci and location of capture-HiC baits is shown along with chromatin looping. Fine-mapped variants from
678 both loci located within the *CDKAL1* gene and near *HDGFL1*, respectively, directly interact with the *SOX4*
679 promoter region.

680 At locus 16 (signal 18) within the *CDKAL1* gene, we tested four regions harboring a set of five fine-

681 mapped variants within a ~3 kb stretch (**Figure 7A**); at least one gRNA targeting each of the four CCVs

682 showed significant reduction of *SOX4* without affecting *CDKAL1*. All three gRNAs targeting both

683 rs7776158 or rs2125570 showed significant reductions of *SOX4* (0.63-0.68-fold and 0.67-fold expression

684 relative to non-targeting guide 1, NTC1), as did two out of three gRNAs simultaneously targeting both

685 rs6935117 and rs6935124 (0.72-0.84-fold expression relative to NTC1; **Figure 7B, Table S22**). Only one of

686 three gRNAs targeting rs6914598 showed inhibition of *SOX4* (0.72-fold expression relative to NTC1;

687 **Figure 7B, Table S22**). None of the guides targeting any of these variants significantly influenced *CDKAL1*

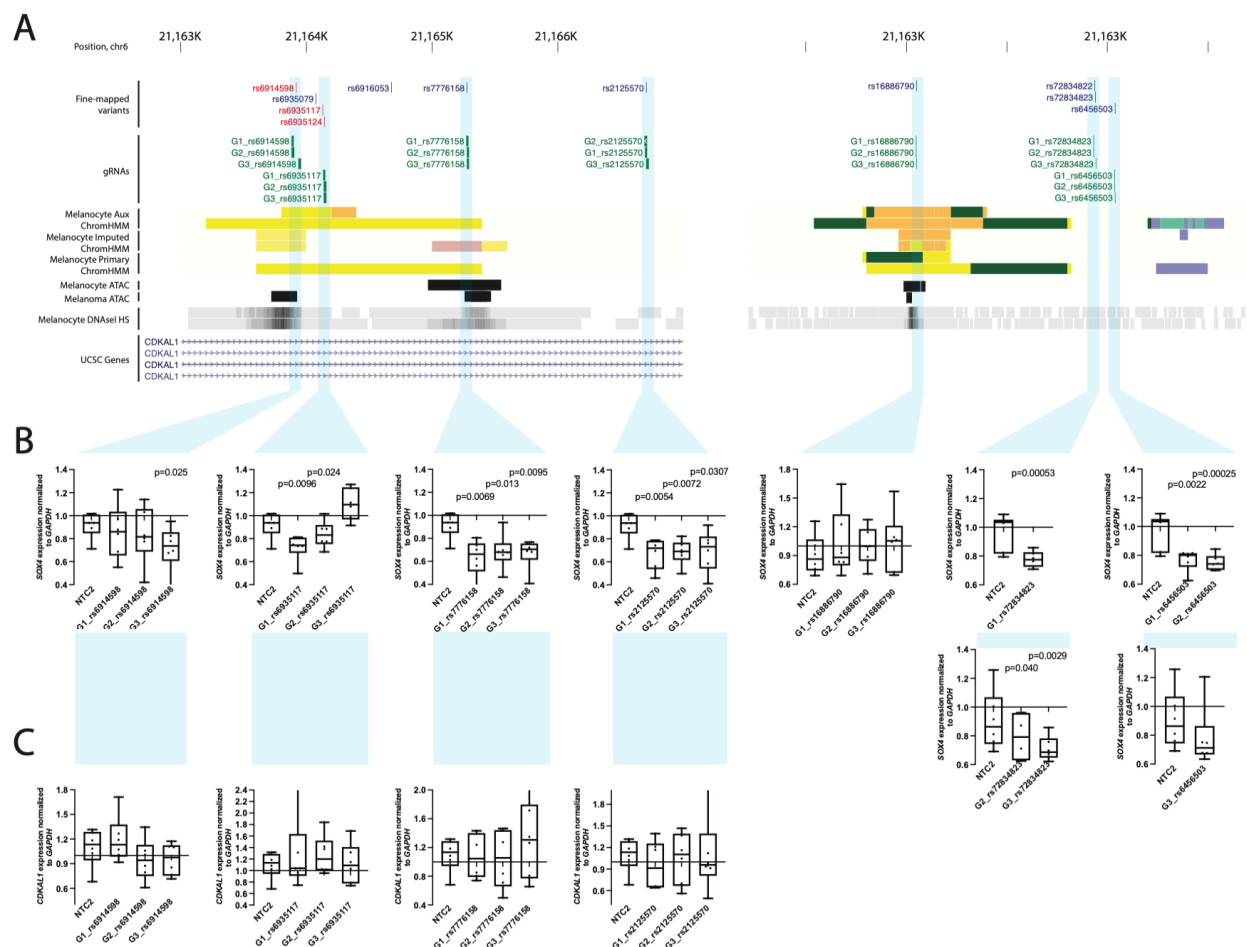
688 (**Figure 7C, Table S22**). While further characterization is necessary to disentangle the exact functional

689 variant or combination thereof, these data clearly indicate specific transcriptional regulation of *SOX4* via

690 this risk locus. At a second locus (locus 17, signal 19) closest to the *HDGFL1* gene, we targeted three

691 regions harboring four fine-mapped variants (**Figure 7A, Table S22**). While guides targeting rs16886790
 692 showed no significant reduction of *SOX4* levels, all three gRNAs simultaneously targeting both
 693 rs72834823 and rs72834822, as well as those targeting rs6456503 significantly reduced *SOX4* levels
 694 (0.72-0.80 and 0.75-0.78-fold expression relative to NTC1, respectively; **Figure 7B, Table S22**), consistent
 695 with the potential regulation of the *SOX4* gene by the region. In contrast to rs16886790, rs72834823,
 696 rs72834822, and rs6456503 do not directly overlap with a restriction fragment bin interacting with *SOX4*
 697 nor are located within ATAC open or ChromHMM enhancer regions but are in close proximity to
 698 interacting enhancer regions. These CRISPRi data nonetheless suggest the regions harboring these SNPs
 699 directly regulate *SOX4* expression.

700



701

702 **Figure 7. CRISPR-inhibition validation of *SOX4* as a target of regulatory regions harboring fine-mapped**
703 **variants at two independent melanoma risk loci on chromosome 6.** (A) Guide RNAs were designed to
704 target four regions collectively harboring five fine-mapped sequence variants in a risk locus located
705 within an intron of the *CDKAL1* gene (left), as well as three regions harboring four fine-mapped variants
706 for an independent locus nearest the *HDGFL1* gene (right); three guides were designed per region and
707 tested along with two non-targeting guides (NTC1 and NTC2). (B) Each guide was individually tested for
708 effects on *SOX4* expression relative to NTC1 in immortalized melanocytes stably expressing dCas9-KRAB
709 via a TaqMan quantitative RT-PCR assay. Expression values from six replicate experiments are shown as
710 fold change relative to NTC1. Where SNP-targeting guides were tested in separate experiments, they are
711 shown grouped with respective values for NTC2 from the same experiments. Whiskers show minimum
712 and maximum values. *P*-values were calculated using a two-sample two-sided paired t-test comparing
713 delta-Ct values from individual guides to those from NTC1.

714 For *SOX6* (locus 32, signal 40), we tested five regions harboring six SNPs (rs1455114/rs1455115,
715 rs2953060, rs4617548, rs7108091 and rs7941496). Only one of the five regions (targeting both
716 rs1455114 and rs1455115; interact with *SOX6*; located within melanocyte enhancer; within ATAC-open
717 region in melanocytes and melanoma) showed a reduction of *SOX6* by a single guide (0.71 fold
718 expression relative to NTC1; **Table S22**). A second region targeting rs7108091 (interacts with *SOX6*,
719 MPRA-significant in melanocytes and melanoma) showed marginal reductions for two gRNAs (**Table**
720 **S22**). On the other hand, at least one guide targeting each of the five regions showed a reduction in
721 *C11orf58*, another interacting candidate gene ~650 kb from the GWAS signal. All three guides each
722 respectively targeting rs2953060 or rs4617548 showed significant reductions in *C11orf58* (0.34-0.56 and
723 0.72-0.81 fold-expression relative to NTC1, respectively; **Table S22**), suggesting this gene as a strong
724 candidate causal gene. For *CBL* (locus 36, signal 44), two out of three gRNAs for one of four tested SNPs
725 (rs61900794; melanoma and melanocyte enhancer, ATAC open in melanoma) showed small but
726 significant effects on *CBL* transcription (0.81-0.84 fold expression relative to NTC1). We note that while
727 the restriction fragment bin harboring rs61900794 was not found to interact directly with *CBL* promoter,
728 it is located within a contiguous enhancer region that does, consistent with the region surrounding
729 rs61900794 functioning as an enhancer for *CBL*. Finally, for *MDM4* (locus 4, signal 4), all three gRNAs for
730 one of the four SNPs tested (rs6700182; interacts with *MDM4*; within melanocyte and melanoma

731 ChromHMM regulatory region; MPRA-significant in melanoma) showed significantly reduced *MDM4*
732 levels (0.82-0.85-fold expression relative to NTC1; **Table S22**); only one of these variants reduced
733 expression of another candidate in this region, *RBBP5*, with no effect on a third candidate, *TMEM81*.

734 ***Additional evidence for MDM4 as a melanoma predisposition gene***

735 Bayesian fine-mapping of the locus for which *MDM4* was nominated as a potential causal gene (locus 4,
736 signal 4, lead SNP rs2369633) identified a second set of credible causal variants (**Table S23**; variant with
737 highest posterior inclusion probability is rs12119098). This second credible set appears to be an
738 independent GWAS signal of marginal significance marked by rs12119098 (GWAS $P = 1.30 \times 10^{-7}$,
739 rs12119098-G OR = 0.95; LD to rs2369633, $r^2 = 0.0016$, $D' = 0.18$) which remains strongly significant after
740 conditioning on the lead SNP for locus 4 (signal 4, rs2369633; $P_{\text{conditional}} = 2.17 \times 10^{-7}$). In addition,
741 rs12119098 is a significant colocalizing melanocyte eQTL for *MDM4* ($P = 5.83 \times 10^{-6}$; HyPrColoc posterior
742 probability = 0.78; eCAVIAR CLPP = 0.073; **Figure S15**) but no other gene, where the risk-associated allele
743 is associated with lower *MDM4* levels. rs12119098 is not an eQTL for *MDM4* ($P = 0.10$) or any other
744 gene in melanoma tumors. Taken together with the capture-HiC and CRISPRi already linking locus 4 to
745 regulation of *MDM4*, these reinforce a potential role for *MDM4* in melanoma risk.

746 **DISCUSSION**

747 With ever growing GWAS sample-sizes for GWAS resulting in increasing numbers of genome-wide
748 significant loci, high-throughput analyses are critically needed to link risk variants with their respective
749 target genes. Most common trait-associated loci identified by GWAS harbor risk variants located
750 primarily in non-coding regions, with the underlying causal variants largely hypothesized to function via
751 altering expression of causal genes. Here, we report a post-GWAS follow-up study aimed at identifying
752 potential causal genes underlying common melanoma risk loci by evaluating cell-type specific chromatin

753 interactions using a custom, GWAS region-focused chromatin capture-HiC assay in human primary
754 melanocytes.

755 Where eQTL colocalization and TWAS using a primary melanocyte expression reference dataset
756 previously nominated candidate genes for only roughly 25% of the loci from the most recent (2020)
757 melanoma risk GWAS¹⁰, our GWAS region-focused capture-HiC data identified fine-mapped risk variants
758 either interacting with or overlapping with gene promoters for 61 out of 68 risk signals, with a median
759 number of five candidate genes per locus. These data suggest that region-focused capture-HiC assays
760 are highly sensitive for identifying variant-to-gene promoter associations, but alone are not likely
761 sufficiently specific to pinpoint the individual causal gene or genes at many loci. We applied multiple
762 additional filters to narrow down the number of interactions to retain (1) only those involving fine-
763 mapped variants located within cell-type specific (melanocyte or melanoma cell) *cis*-regulatory regions
764 (nominates 282 genes at 57 risk signals), and (2) subsequently requiring these interacting variants to
765 have been significant in a large-scale parallel reporter assay screen conducted in both melanocytes and
766 melanoma cells²⁵. This resulted in prioritization of 140 *cis*-regulatory variants interacting with 195 genes
767 at 42 signals, with at least one gene nominated at 62% of risk signals.

768 Our analysis appeared to be highly sensitive for identifying long-distance interactions. Most risk signals
769 (54/68) had at least one gene nominated by an interaction ranging from 100 kb to 1 Mb. Notably,
770 roughly a quarter of GWAS signals (16/68) had a fine-mapped variant interacting with a gene more than
771 1Mb away, distances conventionally ignored in *cis*-eQTL analyses. Indeed, as an example we identified
772 multiple fine-mapped SNPs interacting with the promoter of *CBL* that were in fact previously-untested
773 marginal eQTLs for *CBL* (locus 36, signal 44, rs2120430, $P = 0.007$; rs11217853, $P = 0.02$), consistent with
774 potential long-range allelic *cis*-regulation of *CBL* by risk-associated variants. We focused our CRISPRi-
775 based validation efforts primarily on validating longer-range *cis*-regulatory interactions with strong

776 causal candidates, validating such regulatory interactions for two independent loci and *SOX4* (loci 16 and
777 17, signals 18 and 19, located ~400 kb and 1.1 Mb away from *SOX4*, respectively), one with *MDM4*
778 (locus 4, signal 4, ~550 kb distance), and one with *CBL* (locus 36, signal 44; 1.1 Mb distance). Of the loci
779 we tested, we only failed to validate *cis*-regulation of *SOX6* (locus 32, signal 40), but instead showed
780 strong regulation of a distant candidate (*C11orf58*) located more than 500 kb from the risk signal.

781 *SOX4* was nominated as a potential causal gene via variant-to-promoter looping for two loci originally
782 considered independent based on distance (~1.6 Mb) and lack of LD, highlighting the limitations of
783 assigning loci in this manner. For the first locus, located within the gene body of *CDKAL1*, we identified
784 three variants each of which physically interact with *SOX4*, show allelic *cis*-regulatory potential via MPRA
785 where the risk allele is associated with higher reporter expression, and are located within CRISPRi-
786 validated *SOX4* regulatory regions (locus 16, signal 18, rs6935117, rs6935124, and rs2125570). In
787 addition, all three of these variants are marginal eQTLs for *SOX4* in TCGA melanomas (rs6935117, $P =$
788 0.01; rs6935124, $P = 0.01$; and rs2125570, $P = 0.01$; **Table S21**) with direction of effect matching reporter
789 assay data, i.e. risk alleles for these variants are associated with higher *SOX4*. We likewise see physical
790 interaction between fine-mapped variants for the locus near *HDGFL1* and *SOX4* (locus 17, signal 19);
791 while none showed clear allelic regulatory potential via MPRA or were marginal QTLs (**Table S21**),
792 CRISPRi validated regulation of *SOX4* from regions harboring these fine-mapped variants. Collectively,
793 these data provide strong evidence establishing *SOX4* as a potential melanoma risk gene. *SOX4* plays a
794 pivotal role in regulating stemness, promoting cell survival, and epithelial to mesenchymal transition¹¹⁴⁻
795 ¹¹⁸. Single-cell sequencing studies have identified *SOX4* as a marker of multiple melanoma cell states
796 including a melanocytic state in human tumors⁸⁰, and an invasive-like melanoma program in patient
797 derived xenografts under RAF/MEK inhibition¹¹⁹. Intriguingly, analysis of a zebrafish model found *SOX4*
798 to be a marker of stress-like cell population that more efficiently seeded new tumors; induction of a
799 stress-like program conferred resistance to both BRAF and MEK inhibition in zebrafish melanoma

800 cells¹²⁰. We do not observe a correlation between risk variant genotype and expression of *SOX4* in
801 melanocytes. We did observe at least a marginal correlation with *SOX4* expression in melanomas for one
802 of the two signals, suggesting that the function of causal variants in these two regions may be context-
803 dependent, e.g. dependent on cell state, differentiation, or requiring oncogenic mutations. Beyond
804 *SOX4*, we provide weaker evidence for a second SRY-box transcription factor, *SOX6*. Specifically, while
805 we observe interaction between multiple fine-mapped variants and *SOX6* promoter, CRISPRi validation
806 for the few variant regions we tested did not firmly establish a regulatory link between risk variants and
807 *SOX6*. *SOX6* was recently found to be a marker of a hypothesized intermediate melanoma cell state⁸⁰
808 between melanocytic cells and a mesenchymal-like state associated with increased migration and
809 resistance to therapies.

810 These data also provide strong evidence linking two risk loci to distant established melanoma driver
811 genes. We find physical associations between multiple fine-mapped variants (locus 4, signal 4) and
812 *MDM4* amongst other candidate genes. On one hand, rs6700182 is located within a CRISPRi-validated
813 *MDM4* regulatory region and shows allelic regulatory potential via MPRA with the risk allele associated
814 with higher reporter expression (**Table S21**). In contrast a second variant with two CRISPRi guides
815 showing significant or marginal knockdown of *MDM4* (**Table S22**), showed significant allelic regulation in
816 the opposite direction via MPRA (**Table S21**). Fine-mapping of this locus identified a second
817 independent-but-marginal melanoma GWAS signal over the *MDM4* gene itself (rs12119098, $P = 1.30 \times$
818 10^{-7} ; **Figure S15**), a signal that colocalizes with a significant melanocyte *MDM4* eQTL ($P = 5.83 \times 10^{-6}$;
819 **Figure S15**), considerably strengthening the evidence for a role for *MDM4* in melanoma susceptibility
820 and potentially resolving a role for *MDM4* in promoting or alternatively protecting against melanoma.
821 Here, we observe a positive correlation between the rs12119098-protective allele and *MDM4*
822 expression in melanocytes (as well as many other tissue types assayed by The Genotype-Tissue
823 Expression (GTEx) project¹²¹), suggesting higher *MDM4* expression is protective. In contrast to

824 melanoma risk, *MDM4* is found to be amplified in ~5% of melanoma tumors^{86,122}, is over-expressed in
825 much larger proportion (65%) of melanomas¹²³, and antagonizes TP53 function^{123,124}. These data suggest
826 potentially pleiotropic roles for *MDM4* across different stages of melanomagenesis; on one hand higher
827 expression of *MDM4* protects against melanoma development while on the other overexpression is
828 selected for during tumor progression and promotes melanoma cell survival. We note that a previously-
829 published small pooled CRISPR knockout screen in melanocytes⁴⁷ found that *MDM4* knockout
830 significantly reduced melanocyte viability and/or growth (FDR = 0.000102), and both RNAi and CRISPR
831 screen data from the Cancer Dependency Map project¹²⁵ show *MDM4* knockout to be strongly selective
832 in the same direction. This locus appears to be potentially pleiotropic in terms of cancer risk;
833 rs12133735 near *MDM4* has been reported as a suggestive association for aerodigestive squamous cell
834 cancers (rs12133735 LD to rs12119098 $r^2 = 0.82$), where the rs12133735-G risk allele is on a shared
835 haplotype with the melanoma rs12119098-G protective allele¹²⁶. Further work will be required to
836 understand this pleiotropy and the mechanistic role of *MDM4* in risk. More broadly, pathway analysis of
837 high-confidence candidate genes suggest enrichment for genes involved in p53 signaling including
838 *MDM4* including *TP53* itself (loci 60 and 61, integrative scores of 3 and 4, respectively).

839 For locus 36 (signal 44), we observe multiple interactions between fine-mapped risk variants and the
840 *CBL* promoter and verified *cis*-regulation of *CBL* by a region harboring at one such variant (rs61900794)
841 via CRISPRi. While this variant did not show significant allelic *cis*-regulatory potential via MPRA, fine-
842 mapped variants at this locus are indeed eQTLs for *CBL* in melanocytes ($P = 0.02$ to 0.008 for the four
843 variants targeted by CRISPRi; **Figure S14; Table S21**), suggesting the potential for shared causal variants
844 between melanoma risk and germline regulation of *CBL* expression at this locus. *CBL* plays a role in
845 downregulation of receptor tyrosine kinase signaling including through the RAS-MAPK pathway.
846 Germline *CBL* mutations, primarily missense mutations confined to the linker and RING domains, have
847 been found to be associated with cancer including juvenile myelomonocytic leukemia¹²⁷ as well as a

848 variable syndrome overlapping with Noonan syndrome^{128,129}. Somatically, *CBL* has been identified as a
849 potential melanoma driver⁹¹, with loss of function alterations specifically enriched in desmoplastic
850 melanomas (11%)¹³⁰. While our analysis does not identify one or more clear-cut causal sequence variant
851 candidates, the physical and gene-regulatory connection from this risk locus as well as melanocyte eQTL
852 for *CBL* nonetheless establishes regulation of *CBL* from this locus and suggests that a potential role for
853 *CBL* and common variation underlying RAS-MAPK signaling in melanoma risk should be further explored.

854 Recently Pudjihartono and colleagues performed an integrative analyses of the 2020 melanoma GWAS
855 risk signals to nominate likely causal variants-target gene pairs using keratinocyte and melanoma
856 genome-scale HiC data, melanocyte and melanoma-specific epigenomic (promoter, enhancer histone
857 marks, and DNA accessibility), and melanocyte and GTEx skin tissue gene expression datasets¹³¹. While
858 these studies share similarity in approach, there are several key differences. Pudjihartono and
859 colleagues relied on genome-scale Hi-C data generated in keratinocytes and melanoma cells, while we
860 applied a higher-resolution capture-HiC approach to primary melanocytes coupled with deep library
861 sequencing, potentially allowing for more sensitive assessment of interaction in the cell-type of origin of
862 melanoma. We further sought to take advantage of the resolution of this approach by requiring
863 interactions between fine-mapped variants and annotated gene promoters, rather than including
864 interactions with the gene body. Finally, we focused on integrating data specifically from melanocytic
865 cells (melanocytes or melanomas/melanoma cells) and were able to further refine our candidate gene
866 nomination by taking advantage of a larger massively parallel reporter assay which assessed the vast
867 majority of fine-mapped variants from this study for allelic cis-regulatory potential in melanocytes and
868 melanomas. Comparing the high-confidence set of candidate genes identified here, 44 of the 151 genes
869 identified by Pudjihartono were likewise found amongst our 195 gene high-confidence gene set (**Table**
870 **S23**). These approaches are complementary; inclusion of keratinocyte data and skin eQTLs has the
871 advantage of potentially identifying causal variants and genes that function via gene regulation in a cell

872 type other than melanocytes, while our approach is highly focused on identifying genes that function in
873 a cell-type intrinsic manner. We anticipate both approaches to be of considerable utility as applied to
874 future melanoma GWAS.

875 We also acknowledge several limitations of our current study. We performed the capture-HiC assay in
876 primary melanocytes, the cell-of-origin for melanoma, and our downstream analyses integrating
877 epigenomic and MPRA data focused on integrating data from melanocytes or melanomas. This approach
878 could fail to identify appropriate gene candidates for loci where the causal variant(s) function in a non-
879 melanocytic cell type, e.g. keratinocytes or immune cells for example. Further, cultured melanocytes
880 and melanomas may fail to replicate conditions or cellular contexts under which regulatory elements
881 and causal variants within them may function, and thus may miss interactions with some causal genes.
882 As the degree to which chromatin interactions are stable across such contexts is not entirely clear,
883 interaction data alone may still identify potential candidate genes that only become functional under
884 specific contexts. Lastly, capture-HiC sensitivity and precision are dependent on several factors,
885 including whether a capture bait was designable to any given variant, bait efficiency, the size of the
886 restriction fragment harboring a variant, and bin size used for analyses. Despite considerable apparent
887 sensitivity of this method, some variant to gene interactions could be missed.

888

889 **DECLARATION OF INTERESTS**

890 The authors declare no competing interests.

891 **ACKNOWLEDGEMENTS**

892 This work has been supported by the Intramural Research Program (IRP) of the Division of Cancer
893 Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health. This work utilized
894 the Biowulf cluster computing system at the NIH. The results appearing here are in part based on data
895 generated by the TCGA Research Network. We would like to thank members at the National Cancer
896 Institute Cancer Genomics Research Laboratory (CGR) for help with sequencing efforts. We also thank all
897 the cohorts, funders, and investigators who contributed to the melanoma GWAS, as originally
898 acknowledged by Landi and colleagues. We would like to thank the research participants and employees
899 of 23andMe. The content of this publication does not necessarily reflect the views or policies of the US
900 Department of Health and Human Services, nor does the mention of trade names, commercial products,
901 or organizations imply endorsement by the US Government. Mark Iles is supported in part by the
902 National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre. The views
903 expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department
904 of Health and Social Care.

905 AUTHOR CONTRIBUTIONS

906 R.T., M.X., H.S., J.Yon, L.J., T.M., T.Z., R.C., E.L., T.R., R.H., K.F., J.Yin, M.E.J., A.D.W., A.C., S.F.A.G,
907 M.M.I., M.T.L., M.H.L., J.C., and K.M.B. contributed to the research activities described in this
908 manuscript. K.M.B. led and supervised the research described. M.M. provided additional supervision
909 and mentorship. R.T., M.X., and K.M.B. wrote the manuscript. J. Yon, R.H., M.M., M.M.I., M.H.L., and
910 J.C. provided significant contributions to writing, review, and editing.

911 WEB RESOURCES

- 912 • CHiCAGO: <https://www.functionalgenecontrol.group/chicago>
- 913 • HiCUP: <https://www.bioinformatics.babraham.ac.uk/projects/hicup/>
- 914 • LDlink: <https://ldlink.nih.gov/?tab=home>
- 915 • Ensembl variant Effect Predictor: <https://useast.ensembl.org/info/docs/tools/vep/index.html>
- 916 • ATAC-seq pipeline: <https://github.com/ENCODE-DCC/atac-seq-pipeline>
- 917 • UCSC browser: <https://genome.ucsc.edu>
- 918 • WashU Epigenome Browser: <http://epigenomegateway.wustl.edu>
- 919 • DAP-G: <https://github.com/xqwen/dap>
- 920 • Bedtools: <https://bedtools.readthedocs.io/en/latest/>
- 921 • Bedtoolsr: <http://phanstiel-lab.med.unc.edu/bedtoolsr.html>
- 922 • ROADMAP epigenomic project:
923 https://egg2.wustl.edu/roadmap/web_portal/chr_state_learning.html
- 924 • Intogen: <https://www.intogen.org/search>
- 925 • NIH Biowulf Cluster, <http://hpc.nih.gov>
- 926 • PLINK, <https://www.cog-genomics.org/plink/>

- 927 • The Cancer Genome Atlas (TCGA) Research Network, <http://cancergenome.nih.gov/>
928 • TWAS FUSION, <http://gusevlab.org/projects/fusion/>

929 **DATA AND CODE AVAILABILITY**

930 Capture-HiC data are being deposited in ArrayExpress, including both called interactions as well as raw
931 sequencing data.

932 Melanocyte ATAC-seq data are being deposited in ArrayExpress, including called peaks and raw
933 sequencing data.

934 Data from the 2020 melanoma GWAS meta-analysis performed by Landi and colleagues¹⁰ were obtained
935 from dbGaP (dbGaP: phs001868.v1.p1), with the exclusion of self-reported data from 23andMe, Inc. and
936 UK Biobank. The full GWAS summary statistics for the 23andMe discovery dataset will be made available
937 through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy
938 of the 23andMe participants. Please visit [https://research.23andme.com/collaborate/#dataset-](https://research.23andme.com/collaborate/#dataset-access/)
939 [access/](https://research.23andme.com/collaborate/#dataset-access/) for more information and to apply to access the data. Summary data from the remaining self-
940 reported cases are available from the corresponding authors of that manuscript (Matthew
941 Law, matthew.law@qimrberghofer.edu.au; Mark Iles, m.m.iles@leeds.ac.uk; and Maria Teresa
942 Landi, landim@mail.nih.gov).

943 Melanocyte genotype data, RNA-seq expression data, and all eQTL/meQTL association results^{23,24} are
944 accessible through Genotypes and Phenotypes (dbGaP) under accession dbGaP: phs001500.v2.p1. The
945 MPRA sequencing²⁵ and associated MPRA sequencing data are accessible through Gene Expression
946 Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under the accession GEO: [GSE210356](https://www.ncbi.nlm.nih.gov/geo/).

947 REFERENCES

- 948 1. Brown, K.M., Macgregor, S., Montgomery, G.W., Craig, D.W., Zhao, Z.Z., Iyadurai, K., Henders,
949 A.K., Homer, N., Campbell, M.J., Stark, M., et al. (2008). Common sequence variants on 20q11.22
950 confer melanoma susceptibility. *Nature Genetics* 40, 838-840. 10.1038/ng.163.
- 951 2. Bishop, D.T., Demenais, F., Iles, M.M., Harland, M., Taylor, J.C., Corda, E., Randerson-Moor, J.,
952 Aitken, J.F., Avril, M.F., Azizi, E., et al. (2009). Genome-wide association study identifies three loci
953 associated with melanoma risk. *Nature Genetics* 41, 920-925. 10.1038/ng.411.
- 954 3. Amos, C.I., Wang, L.E., Lee, J.E., Gershenwald, J.E., Chen, W.V., Fang, S., Kosoy, R., Zhang, M.,
955 Qureshi, A.A., Vattathil, S., et al. (2011). Genome-wide association study identifies novel loci
956 predisposing to cutaneous melanoma. *Hum Mol Genet* 20, 5012-5023. 10.1093/hmg/ddr415.
- 957 4. Macgregor, S., Montgomery, G.W., Liu, J.Z., Zhao, Z.Z., Henders, A.K., Stark, M., Schmid, H.,
958 Holland, E.A., Duffy, D.L., Zhang, M., et al. (2011). Genome-wide association study identifies a new
959 melanoma susceptibility locus at 1q21.3. *Nature Genetics* 43, 1114-1118. 10.1038/ng.958.
- 960 5. Law, M.H., Montgomery, G.W., Brown, K.M., Martin, N.G., Mann, G.J., Hayward, N.K., and
961 MacGregor, S. (2012). Meta-analysis combining new and existing data sets confirms that the TERT-
962 CLPTM1L locus influences melanoma risk. *J Invest Dermatol* 132, 485-487. 10.1038/jid.2011.322.
- 963 6. Iles, M.M., Law, M.H., Stacey, S.N., Han, J., Fang, S., Pfeiffer, R., Harland, M., Macgregor, S., Taylor,
964 J.C., Aben, K.K., et al. (2013). A variant in FTO shows association with melanoma risk not due to
965 BMI. *Nature Genetics* 45, 428-432, 432e421. 10.1038/ng.2571.
- 966 7. Barrett, J.H., Taylor, J.C., Bright, C., Harland, M., Dunning, A.M., Akslen, L.A., Andresen, P.A., Avril,
967 M.F., Azizi, E., Bianchi Scarra, G., et al. (2015). Fine mapping of genetic susceptibility loci for
968 melanoma reveals a mixture of single variant and multiple variant regions. *Int J Cancer* 136, 1351-
969 1360. 10.1002/ijc.29099.
- 970 8. Law, M.H., Bishop, D.T., Lee, J.E., Brossard, M., Martin, N.G., Moses, E.K., Song, F., Barrett, J.H.,
971 Kumar, R., Easton, D.F., et al. (2015). Genome-wide meta-analysis identifies five new susceptibility
972 loci for cutaneous malignant melanoma. *Nature Genetics* 47, 987-995. 10.1038/ng.3373.
- 973 9. Ransohoff, K.J., Wu, W., Cho, H.G., Chahal, H.C., Lin, Y., Dai, H.J., Amos, C.I., Lee, J.E., Tang, J.Y.,
974 Hinds, D.A., et al. (2017). Two-stage genome-wide association study identifies a novel
975 susceptibility locus associated with melanoma. *Oncotarget* 8, 17586-17592.
976 10.18632/oncotarget.15230.
- 977 10. Landi, M.T., Bishop, D.T., MacGregor, S., Machiela, M.J., Stratigos, A.J., Ghiorzo, P., Brossard, M.,
978 Calista, D., Choi, J., Fargnoli, M.C., et al. (2020). Genome-wide association meta-analyses
979 combining multiple risk phenotypes provide insights into the genetic architecture of cutaneous
980 melanoma susceptibility. *Nat Genet* 52, 494-504. 10.1038/s41588-020-0611-8.
- 981 11. Cano-Gamez, E., and Trynka, G. (2020). From GWAS to Function: Using Functional Genomics to
982 Identify the Mechanisms Underlying Complex Diseases. *Frontiers in Genetics* 11.
- 983 12. Uffelmann, E., Huang, Q.Q., Munung, N.S., de Vries, J., Okada, Y., Martin, A.R., Martin, H.C.,
984 Lappalainen, T., and Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews*
985 *Methods Primers* 1, 59. 10.1038/s43586-021-00056-9.
- 986 13. Gallagher, M.D., and Chen-Plotkin, A.S. (2018). The Post-GWAS Era: From Association to Function.
987 *Am J Hum Genet* 102, 717-730. 10.1016/j.ajhg.2018.04.002.

- 988 14. Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P.,
989 Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic localization of common disease-
990 associated variation in regulatory DNA. *Science* 337, 1190-1195. 10.1126/science.1222794.
- 991 15. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A.,
992 Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The NIH Roadmap Epigenomics
993 Mapping Consortium. *Nat Biotechnol* 28, 1045-1048. 10.1038/nbt1010-1045.
- 994 16. Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-
995 Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111
996 reference human epigenomes. *Nature* 518, 317-330. 10.1038/nature14248.
- 997 17. Satterlee, J.S., Chadwick, L.H., Tyson, F.L., McAllister, K., Beaver, J., Birnbaum, L., Volkow, N.D.,
998 Wilder, E.L., Anderson, J.M., and Roy, A.L. (2019). The NIH Common Fund/Roadmap Epigenomics
999 Program: Successes of a comprehensive consortium. *Science Advances* 5, eaaw6507.
1000 doi:10.1126/sciadv.aaw6507.
- 1001 18. G. TEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human
1002 tissues. *Science* 369, 1318-1330. 10.1126/science.aaz1776.
- 1003 19. Flynn, E.D., and Lappalainen, T. (2022). Functional Characterization of Genetic Variant Effects on
1004 Expression. *Annu Rev Biomed Data Sci* 5, 119-139. 10.1146/annurev-biodatasci-122120-010010.
- 1005 20. Zuber, V., Grinberg, N.F., Gill, D., Manipur, I., Slob, E.A.W., Patel, A., Wallace, C., and Burgess, S.
1006 (2022). Combining evidence from Mendelian randomization and colocalization: Review and
1007 comparison of approaches. *Am J Hum Genet* 109, 767-782. 10.1016/j.ajhg.2022.04.001.
- 1008 21. Liu, B., Gloudemans, M.J., Rao, A.S., Ingelsson, E., and Montgomery, S.B. (2019). Abundant
1009 associations with gene expression complicate GWAS follow-up. *Nat Genet* 51, 768-769.
1010 10.1038/s41588-019-0404-0.
- 1011 22. Barbeira, A.N., Bonazzola, R., Gamazon, E.R., Liang, Y., Park, Y., Kim-Hellmuth, S., Wang, G., Jiang,
1012 Z., Zhou, D., Hormozdiari, F., et al. (2021). Exploiting the GTEx resources to decipher the
1013 mechanisms at GWAS loci. *Genome Biol* 22, 49. 10.1186/s13059-020-02252-4.
- 1014 23. Zhang, T., Choi, J., Dilshat, R., Einarsdóttir, B., Kovacs, M.A., Xu, M., Malasky, M., Chowdhury, S.,
1015 Jones, K., Bishop, D.T., et al. (2021). Cell-type-specific meQTLs extend melanoma GWAS
1016 annotation beyond eQTLs and inform melanocyte gene-regulatory mechanisms. *Am J Hum Genet*
1017 108, 1631-1646. 10.1016/j.ajhg.2021.06.018.
- 1018 24. Zhang, T., Choi, J., Kovacs, M.A., Shi, J., Xu, M., Program, N.C.S., Melanoma Meta-Analysis, C.,
1019 Goldstein, A.M., Trower, A.J., Bishop, D.T., et al. (2018). Cell-type-specific eQTL of primary
1020 melanocytes facilitates identification of melanoma susceptibility genes. *Genome Res* 28, 1621-
1021 1635. 10.1101/gr.233304.117.
- 1022 25. Long, E., Yin, J., Funderburk, K.M., Xu, M., Feng, J., Kane, A., Zhang, T., Myers, T., Golden, A.,
1023 Thakur, R., et al. (2022). Massively parallel reporter assays and variant scoring identified
1024 functional variants and target genes for melanoma loci and highlighted cell-type specificity. *Am J*
1025 *Hum Genet* 109, 2210-2229. 10.1016/j.ajhg.2022.11.006.
- 1026 26. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I.,
1027 Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex
1028 diseases. *Nature* 461, 747-753. 10.1038/nature08494.

- 1029 27. Farashi, S., Kryza, T., Clements, J., and Batra, J. (2019). Post-GWAS in prostate cancer: from genetic
1030 association to biological contribution. *Nat Rev Cancer* 19, 46-59. 10.1038/s41568-018-0087-3.
- 1031 28. Liu, X., Li, Y.I., and Pritchard, J.K. (2019). Trans Effects on Gene Expression Can Drive Omnigenic
1032 Inheritance. *Cell* 177, 1022-1034.e1026. <https://doi.org/10.1016/j.cell.2019.04.014>.
- 1033 29. Strober, B.J., Elorbany, R., Rhodes, K., Krishnan, N., Tayeb, K., Battle, A., and Gilad, Y. (2019).
1034 Dynamic genetic regulation of gene expression during cellular differentiation. *Science* 364, 1287-
1035 1290. 10.1126/science.aaw0040.
- 1036 30. Ward, M.C., Banovich, N.E., Sarkar, A., Stephens, M., and Gilad, Y. (2020). Dynamic effects of
1037 genetic variation on gene expression revealed following hypoxic stress in cardiomyocytes. *BioRxiv*.
- 1038 31. Moyerbrailean, G.A., Richards, A.L., Kurtz, D., Kalita, C.A., Davis, G.O., Harvey, C.T., Alazizi, A.,
1039 Watzka, D., Sorokin, Y., Hauff, N., et al. (2016). High-throughput allele-specific expression across
1040 250 environmental conditions. *Genome Res* 26, 1627-1638. 10.1101/gr.209759.116.
- 1041 32. Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional
1042 organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 14, 390-403.
1043 10.1038/nrg3454.
- 1044 33. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012).
1045 Topological domains in mammalian genomes identified by analysis of chromatin interactions.
1046 *Nature* 485, 376-380. 10.1038/nature11082.
- 1047 34. Sati, S., and Cavalli, G. (2017). Chromosome conformation capture technologies and their impact
1048 in understanding genome function. *Chromosoma* 126, 33-44. 10.1007/s00412-016-0593-6.
- 1049 35. Akgol Oksuz, B., Yang, L., Abraham, S., Venev, S.V., Krietenstein, N., Parsi, K.M., Ozadam, H.,
1050 Oomen, M.E., Nand, A., Mao, H., et al. (2021). Systematic evaluation of chromosome
1051 conformation capture assays. *Nat Methods* 18, 1046-1055. 10.1038/s41592-021-01248-7.
- 1052 36. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozcy, T., Telling, A., Amit, I.,
1053 Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range
1054 interactions reveals folding principles of the human genome. *Science* 326, 289-293.
1055 10.1126/science.1181369.
- 1056 37. Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn,
1057 A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome
1058 at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665-1680.
1059 10.1016/j.cell.2014.11.021.
- 1060 38. Davidson, C., Wordsworth, B.P., Cohen, C.J., Knight, J.C., and Vecellio, M. (2023). Chromosome
1061 conformation capture approaches to investigate 3D genome architecture in Ankylosing
1062 Spondylitis. *Front Genet* 14, 1129207. 10.3389/fgene.2023.1129207.
- 1063 39. Su, C., Johnson, M.E., Torres, A., Thomas, R.M., Manduchi, E., Sharma, P., Mehra, P., Le Coz, C.,
1064 Leonard, M.E., Lu, S., et al. (2020). Mapping effector genes at lupus GWAS loci using promoter
1065 Capture-C in follicular helper T cells. *Nat Commun* 11, 3294. 10.1038/s41467-020-17089-5.
- 1066 40. Palermo, J., Chesi, A., Zimmerman, A., Sonti, S., Pahl, M.C., Lasconi, C., Brown, E.B., Pippin, J.A.,
1067 Wells, A.D., Doldur-Balli, F., et al. (2023). Variant-to-gene mapping followed by cross-species
1068 genetic screening identifies GPI-anchor biosynthesis as a regulator of sleep. *Sci Adv* 9, eabq0844.
1069 10.1126/sciadv.abq0844.

- 1070 41. Beesley, J., Sivakumaran, H., Moradi Marjaneh, M., Lima, L.G., Hillman, K.M., Kaufmann, S., Tuano,
1071 N., Hussein, N., Ham, S., Mukhopadhyay, P., et al. (2020). Chromatin interactome mapping at 139
1072 independent breast cancer risk signals. *Genome Biol* 21, 8. 10.1186/s13059-019-1877-y.
- 1073 42. Chesi, A., Wagley, Y., Johnson, M.E., Manduchi, E., Su, C., Lu, S., Leonard, M.E., Hodge, K.M.,
1074 Pippin, J.A., Hankenson, K.D., et al. (2019). Genome-scale Capture C promoter interactions
1075 implicate effector genes at GWAS loci for bone mineral density. *Nat Commun* 10, 1260.
1076 10.1038/s41467-019-09302-x.
- 1077 43. Jager, R., Migliorini, G., Henrion, M., Kandaswamy, R., Speedy, H.E., Heindl, A., Whiffin, N.,
1078 Carnicer, M.J., Broome, L., Dryden, N., et al. (2015). Capture Hi-C identifies the chromatin
1079 interactome of colorectal cancer risk loci. *Nat Commun* 6, 6178. 10.1038/ncomms7178.
- 1080 44. Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W.,
1081 Andrews, S., Grey, W., Ewels, P.A., et al. (2015). Mapping long-range promoter contacts in human
1082 cells with high-resolution capture Hi-C. *Nat Genet* 47, 598-606. 10.1038/ng.3286.
- 1083 45. Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B.M.,
1084 Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S.W., et al. (2015). The pluripotent regulatory
1085 circuitry connecting promoters to their long-range interacting elements. *Genome Res* 25, 582-
1086 597. 10.1101/gr.185272.114.
- 1087 46. Dryden, N.H., Broome, L.R., Dudbridge, F., Johnson, N., Orr, N., Schoenfelder, S., Nagano, T.,
1088 Andrews, S., Wingett, S., Kozarewa, I., et al. (2014). Unbiased analysis of potential targets of breast
1089 cancer susceptibility loci by Capture Hi-C. *Genome Res* 24, 1854-1868. 10.1101/gr.175034.114.
- 1090 47. Xu, M., Mehl, L., Zhang, T., Thakur, R., Sowards, H., Myers, T., Jessop, L., Chesi, A., Johnson, M.E.,
1091 Wells, A.D., et al. (2021). A UVB-responsive common variant at chromosome band 7p21.1 confers
1092 tanning response and melanoma risk via regulation of the aryl hydrocarbon receptor, AHR. *Am J*
1093 *Hum Genet* 108, 1611-1630. 10.1016/j.ajhg.2021.07.002.
- 1094 48. Fang, S., Lu, J., Zhou, X., Wang, Y., Ross, M.I., Gershenwald, J.E., Cormier, J.N., Wargo, J., Sui, D.,
1095 Amos, C.I., and Lee, J.E. (2020). Functional annotation of melanoma risk loci identifies novel
1096 susceptibility genes. *Carcinogenesis* 41, 452-457. 10.1093/carcin/bgz173.
- 1097 49. Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korb, J.O.,
1098 Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for
1099 human genetic variation. *Nature* 526, 68-74. 10.1038/nature15393.
- 1100 50. Machiela, M.J., and Chanock, S.J. (2018). LDassoc: an online tool for interactively exploring
1101 genome-wide association study results and prioritizing variants for functional investigation.
1102 *Bioinformatics* 34, 887-889. 10.1093/bioinformatics/btx561.
- 1103 51. Machiela, M.J., and Chanock, S.J. (2015). LDlink: a web-based application for exploring population-
1104 specific haplotype structure and linking correlated alleles of possible functional variants.
1105 *Bioinformatics* 31, 3555-3557. 10.1093/bioinformatics/btv402.
- 1106 52. Myers, T.A., Chanock, S.J., and Machiela, M.J. (2020). LDlinkR: An R Package for Rapidly Calculating
1107 Linkage Disequilibrium Statistics in Diverse Populations. *Front Genet* 11, 157.
1108 10.3389/fgene.2020.00157.
- 1109 53. Wen, X., Lee, Y., Luca, F., and Pique-Regi, R. (2016). Efficient Integrative Multi-SNP Association
1110 Analysis via Deterministic Approximation of Posteriors. *Am J Hum Genet* 98, 1114-1129.
1111 10.1016/j.ajhg.2016.03.029.

- 1112 54. Lee, Y., Luca, F., Pique-Regi, R., and Wen, X. (2018). Bayesian Multi-SNP Genetic Association
1113 Analysis: Control of FDR and Use of Summary Statistics. *bioRxiv*, 316471. 10.1101/316471.
- 1114 55. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D.,
1115 Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and
1116 genomic data. *Nature* 562, 203-209. 10.1038/s41586-018-0579-z.
- 1117 56. Weissbrod, O., Hormozdiari, F., Benner, C., Cui, R., Ulirsch, J., Gazal, S., Schoech, A.P., van de Geijn,
1118 B., Reshef, Y., Márquez-Luna, C., et al. (2020). Functionally informed fine-mapping and polygenic
1119 localization of complex trait heritability. *Nat Genet* 52, 1355-1363. 10.1038/s41588-020-00735-5.
- 1120 57. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide
1121 complex trait analysis. *Am J Hum Genet* 88, 76-82. 10.1016/j.ajhg.2010.11.011.
- 1122 58. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of, A.T.C., Replication,
1123 D.I.G., Meta-analysis, C., Madden, P.A., Heath, A.C., Martin, N.G., et al. (2012). Conditional and
1124 joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing
1125 complex traits. *Nat Genet* 44, 369-375, S361-363. 10.1038/ng.2213.
- 1126 59. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P., and Cunningham,
1127 F. (2016). The Ensembl Variant Effect Predictor. *Genome Biol* 17, 122. 10.1186/s13059-016-0974-
1128 4.
- 1129 60. Choi, J., Zhang, T., Vu, A., Ablain, J., Makowski, M.M., Colli, L.M., Xu, M., Hennessey, R.C., Yin, J.,
1130 Rothschild, H., et al. (2020). Massively parallel reporter assays of melanoma risk variants identify
1131 MX2 as a gene promoting melanoma. *Nat Commun* 11, 2718. 10.1038/s41467-020-16590-1.
- 1132 61. Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and Andrews,
1133 S. (2015). HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res* 4, 1310.
1134 10.12688/f1000research.7334.1.
- 1135 62. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*
1136 9, 357-359. 10.1038/nmeth.1923.
- 1137 63. Cairns, J., Freire-Pritchett, P., Wingett, S.W., Varnai, C., Dimond, A., Plagnol, V., Zerbino, D.,
1138 Schoenfelder, S., Javierre, B.M., Osborne, C., et al. (2016). CHiCAGO: robust detection of DNA
1139 looping interactions in Capture Hi-C data. *Genome Biol* 17, 127. 10.1186/s13059-016-0992-2.
- 1140 64. Li, D., Hsu, S., Purushotham, D., Sears, R.L., and Wang, T. (2019). WashU Epigenome Browser
1141 update 2019. *Nucleic Acids Res* 47, W158-W165. 10.1093/nar/gkz348.
- 1142 65. Li, D., Purushotham, D., Harrison, J.K., Hsu, S., Zhuo, X., Fan, C., Liu, S., Xu, V., Chen, S., Xu, J., et
1143 al. (2022). WashU Epigenome Browser update 2022. *Nucleic Acids Res* 50, W774-W781.
1144 10.1093/nar/gkac238.
- 1145 66. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D.
1146 (2002). The human genome browser at UCSC. *Genome Res* 12, 996-1006. 10.1101/gr.229102.
- 1147 67. Navarro Gonzalez, J., Zweig, A.S., Speir, M.L., Schmelter, D., Rosenbloom, K.R., Raney, B.J., Powell,
1148 C.C., Nassar, L.R., Maulding, N.D., Lee, C.M., et al. (2021). The UCSC Genome Browser database:
1149 2021 update. *Nucleic Acids Res* 49, D1046-D1057. 10.1093/nar/gkaa1070.
- 1150 68. Raney, B.J., Dreszer, T.R., Barber, G.P., Clawson, H., Fujita, P.A., Wang, T., Nguyen, N., Paten, B.,
1151 Zweig, A.S., Karolchik, D., and Kent, W.J. (2014). Track data hubs enable visualization of user-
1152 defined genome-wide annotations on the UCSC Genome Browser. *Bioinformatics* 30, 1003-1005.
1153 10.1093/bioinformatics/btt637.

- 1154 69. Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C.,
1155 Maddren, M., Fang, R., Heitner, S.G., et al. (2013). ENCODE data in the UCSC Genome Browser:
1156 year 5 update. *Nucleic Acids Res* 41, D56-63. 10.1093/nar/gks1172.
- 1157 70. Lee, B.T., Barber, G.P., Benet-Pages, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez,
1158 J.N., Hinrichs, A.S., Lee, C.M., et al. (2022). The UCSC Genome Browser database: 2022 update.
1159 *Nucleic Acids Res* 50, D1115-D1122. 10.1093/nar/gkab959.
- 1160 71. Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu,
1161 C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and
1162 mouse genomes. *Nucleic Acids Res* 47, D766-d773. 10.1093/nar/gky955.
- 1163 72. Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and
1164 characterization. *Nat Methods* 9, 215-216. 10.1038/nmeth.1906.
- 1165 73. Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with
1166 ChromHMM. *Nat Protoc* 12, 2478-2492. 10.1038/nprot.2017.124.
- 1167 74. Fiziev, P., Akdemir, K.C., Miller, J.P., Keung, E.Z., Samant, N.S., Sharma, S., Natale, C.A., Terranova,
1168 C.J., Maitituohti, M., Amin, S.B., et al. (2017). Systematic Epigenomic Analysis Reveals Chromatin
1169 States Associated with Melanoma Progression. *Cell Rep* 19, 875-889.
1170 10.1016/j.celrep.2017.03.078.
- 1171 75. Garraway, L.A., Widlund, H.R., Rubin, M.A., Getz, G., Berger, A.J., Ramaswamy, S., Beroukhim, R.,
1172 Milner, D.A., Granter, S.R., Du, J., et al. (2005). Integrative genomic analyses identify MITF as a
1173 lineage survival oncogene amplified in malignant melanoma. *Nature* 436, 117-122.
1174 10.1038/nature03664.
- 1175 76. Encode Project Consortium (2012). An integrated encyclopedia of DNA elements in the human
1176 genome. *Nature* 489, 57-74. 10.1038/nature11247.
- 1177 77. Langmead, B., Wilks, C., Antonescu, V., and Charles, R. (2019). Scaling read aligners to hundreds
1178 of threads on general-purpose processors. *Bioinformatics* 35, 421-432.
1179 10.1093/bioinformatics/bty648.
- 1180 78. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers,
1181 R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome*
1182 *Biol* 9, R137. 10.1186/gb-2008-9-9-r137.
- 1183 79. Amemiya, H.M., Kundaje, A., and Boyle, A.P. (2019). The ENCODE Blacklist: Identification of
1184 Problematic Regions of the Genome. *Sci Rep* 9, 9354. 10.1038/s41598-019-45839-z.
- 1185 80. Wouters, J., Kalender-Atak, Z., Minnoye, L., Spanier, K.I., De Waegeneer, M., Bravo Gonzalez-Blas,
1186 C., Mauduit, D., Davie, K., Hulselmans, G., Najem, A., et al. (2020). Robust gene expression
1187 programs underlie recurrent cell states and phenotype switching in melanoma. *Nat Cell Biol* 22,
1188 986-998. 10.1038/s41556-020-0547-3.
- 1189 81. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic
1190 features. *Bioinformatics* 26, 841-842. 10.1093/bioinformatics/btq033.
- 1191 82. Patwardhan, M.N., Wenger, C.D., Davis, E.S., and Phanstiel, D.H. (2019). Bedtoolsr: An R package
1192 for genomic data analysis and manipulation. *J Open Source Softw* 4. 10.21105/joss.01742.
- 1193 83. Zhang, T., Klein, A., Sang, J., Choi, J., and Brown, K.M. (2022). ezQTL: A Web Platform for
1194 Interactive Visualization and Colocalization of QTLs and GWAS Loci. *Genomics Proteomics*
1195 *Bioinformatics* 20, 541-548. 10.1016/j.gpb.2022.05.004.

- 1196 84. Foley, C.N., Staley, J.R., Breen, P.G., Sun, B.B., Kirk, P.D.W., Burgess, S., and Howson, J.M.M.
1197 (2021). A fast and efficient colocalization algorithm for identifying shared genetic risk factors
1198 across multiple traits. *Nat Commun* 12, 764. 10.1038/s41467-020-20885-8.
- 1199 85. Hormozdiari, F., van de Bunt, M., Segre, A.V., Li, X., Joo, J.W.J., Bilow, M., Sul, J.H., Sankararaman,
1200 S., Pasaniuc, B., and Eskin, E. (2016). Colocalization of GWAS and eQTL Signals Detects Target
1201 Genes. *Am J Hum Genet* 99, 1245-1260. 10.1016/j.ajhg.2016.10.003.
- 1202 86. TCGA (2015). Genomic Classification of Cutaneous Melanoma. *Cell* 161, 1681-1696.
1203 10.1016/j.cell.2015.05.044.
- 1204 87. Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B.W., Jansen, R., de Geus, E.J., Boomsma,
1205 D.I., Wright, F.A., et al. (2016). Integrative approaches for large-scale transcriptome-wide
1206 association studies. *Nat Genet* 48, 245-252. 10.1038/ng.3506.
- 1207 88. Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in
1208 Ingenuity Pathway Analysis. *Bioinformatics* 30, 523-530. 10.1093/bioinformatics/btt703.
- 1209 89. Rosenbluh, J., Xu, H., Harrington, W., Gill, S., Wang, X., Vazquez, F., Root, D.E., Tsherniak, A., and
1210 Hahn, W.C. (2017). Complementary information derived from CRISPR Cas9 mediated gene
1211 deletion and suppression. *Nat Commun* 8, 15403. 10.1038/ncomms15403.
- 1212 90. Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E., Vaimberg, E.W.,
1213 Goodale, A., Root, D.E., Piccioni, F., and Doench, J.G. (2018). Optimized libraries for CRISPR-Cas9
1214 genetic screens with multiple modalities. *Nat Commun* 9, 5416. 10.1038/s41467-018-07901-8.
- 1215 91. Martinez-Jimenez, F., Muinos, F., Sentis, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C.,
1216 Mularoni, L., Pich, O., Bonet, J., Kranas, H., et al. (2020). A compendium of mutational cancer
1217 driver genes. *Nat Rev Cancer* 20, 555-572. 10.1038/s41568-020-0290-x.
- 1218 92. Ibarrola-Villava, M., Hu, H.H., Guedj, M., Fernandez, L.P., Descamps, V., Basset-Seguín, N., Bagot,
1219 M., Bensussan, A., Saiag, P., Fargnoli, M.C., et al. (2012). MC1R, SLC45A2 and TYR genetic variants
1220 involved in melanoma susceptibility in southern European populations: results from a meta-
1221 analysis. *Eur J Cancer* 48, 2183-2191. 10.1016/j.ejca.2012.03.006.
- 1222 93. Landi, M.T., Kanetsky, P.A., Tsang, S., Gold, B., Munroe, D., Rebbeck, T., Swoyer, J., Ter-Minassian,
1223 M., Hedayati, M., Grossman, L., et al. (2005). MC1R, ASIP, and DNA repair in sporadic and familial
1224 melanoma in a Mediterranean population. *J Natl Cancer Inst* 97, 998-1007. 10.1093/jnci/dji176.
- 1225 94. Palmer, J.S., Duffy, D.L., Box, N.F., Aitken, J.F., O'Gorman, L.E., Green, A.C., Hayward, N.K., Martin,
1226 N.G., and Sturm, R.A. (2000). Melanocortin-1 receptor polymorphisms and risk of melanoma: is
1227 the association explained solely by pigmentation phenotype? *Am J Hum Genet* 66, 176-186.
1228 10.1086/302711.
- 1229 95. Guedj, M., Bourillon, A., Combadières, C., Rodero, M., Dieudé, P., Descamps, V., Dupin, N.,
1230 Wolkenstein, P., Aegerter, P., Lebbe, C., et al. (2008). Variants of the MATP/SLC45A2 gene are
1231 protective for melanoma in the French population. *Hum Mutat* 29, 1154-1160.
1232 10.1002/humu.20823.
- 1233 96. Fernandez, L.P., Milne, R.L., Pita, G., Avilés, J.A., Lázaro, P., Benítez, J., and Ribas, G. (2008).
1234 SLC45A2: a novel malignant melanoma-associated gene. *Hum Mutat* 29, 1161-1167.
1235 10.1002/humu.20804.
- 1236 97. Le, L., Escobar, I.E., Ho, T., Lefkovith, A.J., Latteri, E., Haltaufderhyde, K.D., Dennis, M.K., Plowright,
1237 L., Sviderskaya, E.V., Bennett, D.C., et al. (2020). SLC45A2 protein stability and regulation of

- 1238 melanosome pH determine melanocyte pigmentation. *Mol Biol Cell* *31*, 2687-2702.
1239 10.1091/mbc.E20-03-0200.
- 1240 98. Gudbjartsson, D.F., Sulem, P., Stacey, S.N., Goldstein, A.M., Rafnar, T., Sigurgeirsson, B.,
1241 Benediktsdottir, K.R., Thorisdottir, K., Ragnarsson, R., Sveinsdottir, S.G., et al. (2008). ASIP and TYR
1242 pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet*
1243 *40*, 886-891. 10.1038/ng.161.
- 1244 99. Sulem, P., Gudbjartsson, D.F., Stacey, S.N., Helgason, A., Rafnar, T., Jakobsdottir, M., Steinberg,
1245 S., Gudjonsson, S.A., Palsson, A., Thorleifsson, G., et al. (2008). Two newly identified genetic
1246 determinants of pigmentation in Europeans. *Nat Genet* *40*, 835-837. 10.1038/ng.160.
- 1247 100. Yokoyama, S., Woods, S.L., Boyle, G.M., Aoude, L.G., MacGregor, S., Zismann, V., Gartside, M.,
1248 Cust, A.E., Haq, R., Harland, M., et al. (2011). A novel recurrent mutation in MITF predisposes to
1249 familial and sporadic melanoma. *Nature* *480*, 99-103. 10.1038/nature10630.
- 1250 101. Bertolotto, C., Lesueur, F., Giuliano, S., Strub, T., de Lichy, M., Bille, K., Dessen, P., d'Hayer, B.,
1251 Mohamdi, H., Remenieras, A., et al. (2011). A SUMOylation-defective MITF germline mutation
1252 predisposes to melanoma and renal carcinoma. *Nature* *480*, 94-98. 10.1038/nature10539.
- 1253 102. Cardinale, A., Cantalupo, S., Lasorsa, V.A., Montella, A., Cimmino, F., Succoio, M., Vermeulen, M.,
1254 Baltissen, M.P., Esposito, M., Avitabile, M., et al. (2022). Functional annotation and investigation
1255 of the 10q24.33 melanoma risk locus identifies a common variant that influences transcriptional
1256 regulation of OBFC1. *Hum Mol Genet* *31*, 863-874. 10.1093/hmg/ddab293.
- 1257 103. Duffy, D.L., Zhu, G., Li, X., Sanna, M., Iles, M.M., Jacobs, L.C., Evans, D.M., Yazar, S., Beesley, J.,
1258 Law, M.H., et al. (2018). Novel pleiotropic risk loci for melanoma and nevus density implicate
1259 multiple biological pathways. *Nat Commun* *9*, 4774. 10.1038/s41467-018-06649-5.
- 1260 104. Dalmaso, B., Pastorino, L., Nathan, V., Shah, N.N., Palmer, J.M., Howlie, M., Johansson, P.A.,
1261 Freedman, N.D., Carter, B.D., Beane-Freeman, L., et al. (2021). Germline ATM variants predispose
1262 to melanoma: a joint analysis across the GenoMEL and MelaNostrum consortia. *Genet Med* *23*,
1263 2087-2095. 10.1038/s41436-021-01240-8.
- 1264 105. Derheimer, F.A., and Kastan, M.B. (2010). Multiple roles of ATM in monitoring and maintaining
1265 DNA integrity. *FEBS Lett* *584*, 3675-3681. 10.1016/j.febslet.2010.05.031.
- 1266 106. Pahl, M.C., Doege, C.A., Hodge, K.M., Littleton, S.H., Leonard, M.E., Lu, S., Rausch, R., Pippin, J.A.,
1267 De Rosa, M.C., Basak, A., et al. (2021). Cis-regulatory architecture of human ESC-derived
1268 hypothalamic neuron differentiation aids in variant-to-gene mapping of relevant complex traits.
1269 *Nat Commun* *12*, 6749. 10.1038/s41467-021-27001-4.
- 1270 107. Consortium, E.P., Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shores, N., Adrian, J., Kawli,
1271 T., Davis, C.A., Dobin, A., et al. (2020). Expanded encyclopaedias of DNA elements in the human
1272 and mouse genomes. *Nature* *583*, 699-710. 10.1038/s41586-020-2493-4.
- 1273 108. Codd, V., Nelson, C.P., Albrecht, E., Mangino, M., Deelen, J., Buxton, J.L., Hottenga, J.J., Fischer,
1274 K., Esko, T., Surakka, I., et al. (2013). Identification of seven loci affecting mean telomere length
1275 and their association with disease. *Nat Genet* *45*, 422-427, 427e421-422. 10.1038/ng.2528.
- 1276 109. Visser, M., Kayser, M., and Palstra, R.J. (2012). HERC2 rs12913832 modulates human
1277 pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the
1278 OCA2 promoter. *Genome Res* *22*, 446-455. 10.1101/gr.128652.111.

- 1279 110. Zocchi, L., Lontano, A., Merli, M., Dika, E., Nagore, E., Quaglino, P., Puig, S., and Ribero, S. (2021).
1280 Familial Melanoma and Susceptibility Genes: A Review of the Most Common Clinical and
1281 Dermoscopic Phenotypic Aspect, Associated Malignancies and Practical Tips for Management. *J*
1282 *Clin Med* *10*. 10.3390/jcm10163760.
- 1283 111. Choi, J., Xu, M., Makowski, M.M., Zhang, T., Law, M.H., Kovacs, M.A., Granzhan, A., Kim, W.J.,
1284 Parikh, H., Gartside, M., et al. (2017). A common intronic variant of PARP1 confers melanoma risk
1285 and mediates melanocyte growth via regulation of MITF. *Nat Genet* *49*, 1326-1335.
1286 10.1038/ng.3927.
- 1287 112. Denecker, G., Vandamme, N., Akay, O., Koludrovic, D., Taminau, J., Lemeire, K., Gheldof, A., De
1288 Craene, B., Van Gele, M., Brochez, L., et al. (2014). Identification of a ZEB2-MITF-ZEB1
1289 transcriptional network that controls melanogenesis and melanoma progression. *Cell Death Differ*
1290 *21*, 1250-1261. 10.1038/cdd.2014.44.
- 1291 113. de la Serna, I.L., Ohkawa, Y., Higashi, C., Dutta, C., Osias, J., Kommajosyula, N., Tachibana, T., and
1292 Imbalzano, A.N. (2006). The microphthalmia-associated transcription factor requires SWI/SNF
1293 enzymes to activate melanocyte-specific genes. *J Biol Chem* *281*, 20233-20241.
1294 10.1074/jbc.M512052200.
- 1295 114. Grimm, D., Bauer, J., Wise, P., Krüger, M., Simonsen, U., Wehland, M., Infanger, M., and Corydon,
1296 T.J. (2020). The role of SOX family members in solid tumours and metastasis. *Semin Cancer Biol*
1297 *67*, 122-153. 10.1016/j.semcancer.2019.03.004.
- 1298 115. Moreno, C.S. (2020). SOX4: The unappreciated oncogene. *Semin Cancer Biol* *67*, 57-64.
1299 10.1016/j.semcancer.2019.08.027.
- 1300 116. Roukens, M.G., Frederiks, C.L., Seinstra, D., Braccioli, L., Khalil, A.A., Pals, C., De Neck, S., Bornes,
1301 L., Beerling, E., Mokry, M., et al. (2021). Regulation of a progenitor gene program by SOX4 is
1302 essential for mammary tumor proliferation. *Oncogene* *40*, 6343-6353. 10.1038/s41388-021-
1303 02004-z.
- 1304 117. Jafarnejad, S.M., Ardekani, G.S., Ghaffari, M., Martinka, M., and Li, G. (2013). Sox4-mediated Dicer
1305 expression is critical for suppression of melanoma cell invasion. *Oncogene* *32*, 2131-2139.
1306 10.1038/onc.2012.239.
- 1307 118. Dai, W., Xu, X., Li, S., Ma, J., Shi, Q., Guo, S., Liu, L., Guo, W., Xu, P., He, Y., et al. (2017). SOX4
1308 Promotes Proliferative Signals by Regulating Glycolysis through AKT Activation in Melanoma Cells.
1309 *J Invest Dermatol* *137*, 2407-2416. 10.1016/j.jid.2017.06.026.
- 1310 119. Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D.,
1311 Chang, Y.H., Debiec-Rychter, M., et al. (2018). Toward Minimal Residual Disease-Directed Therapy
1312 in Melanoma. *Cell* *174*, 843-855.e819. 10.1016/j.cell.2018.06.025.
- 1313 120. Baron, M., Tagore, M., Hunter, M.V., Kim, I.S., Moncada, R., Yan, Y., Campbell, N.R., White, R.M.,
1314 and Yanai, I. (2020). The Stress-Like Cancer Cell State Is a Consistent Component of Tumorigenesis.
1315 *Cell Syst* *11*, 536-546.e537. 10.1016/j.cels.2020.08.018.
- 1316 121. Consortium, G.T. (2020). The GTEx Consortium atlas of genetic regulatory effects across human
1317 tissues. *Science* *369*, 1318-1330. 10.1126/science.aaz1776.
- 1318 122. Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico,
1319 A., Wendl, M.C., Kim, J., Reardon, B., et al. (2018). Comprehensive Characterization of Cancer
1320 Driver Genes and Mutations. *Cell* *173*, 371-385.e318. 10.1016/j.cell.2018.02.060.

- 1321 123. Gembarska, A., Luciani, F., Fedele, C., Russell, E.A., Dewaele, M., Villar, S., Zwolinska, A., Haupt,
1322 S., de Lange, J., Yip, D., et al. (2012). MDM4 is a key therapeutic target in cutaneous melanoma.
1323 *Nat Med* 18, 1239-1247. 10.1038/nm.2863.
- 1324 124. Migliorini, D., Lazzarini Denchi, E., Danovi, D., Jochemsen, A., Capillo, M., Gobbi, A., Helin, K.,
1325 Pelicci, P.G., and Marine, J.C. (2002). Mdm4 (Mdmx) regulates p53-induced growth arrest and
1326 neuronal cell death during early embryonic mouse development. *Mol Cell Biol* 22, 5527-5538.
1327 10.1128/mcb.22.15.5527-5538.2002.
- 1328 125. Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S.,
1329 Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a Cancer Dependency Map.
1330 *Cell* 170, 564-576 e516. 10.1016/j.cell.2017.06.010.
- 1331 126. Lesseur, C., Ferreiro-Iglesias, A., McKay, J.D., Bosse, Y., Johansson, M., Gaborieau, V., Landi, M.T.,
1332 Christiani, D.C., Caporaso, N.C., Bojesen, S.E., et al. (2021). Genome-wide association meta-
1333 analysis identifies pleiotropic risk loci for aerodigestive squamous cell cancers. *PLoS Genet* 17,
1334 e1009254. 10.1371/journal.pgen.1009254.
- 1335 127. Loh, M.L., Sakai, D.S., Flotho, C., Kang, M., Fliegau, M., Archambeault, S., Mullighan, C.G., Chen,
1336 L., Bergstraesser, E., Bueso-Ramos, C.E., et al. (2009). Mutations in CBL occur frequently in juvenile
1337 myelomonocytic leukemia. *Blood* 114, 1859-1863. 10.1182/blood-2009-01-198416.
- 1338 128. Martinelli, S., Stellacci, E., Pannone, L., D'Agostino, D., Consoli, F., Lissewski, C., Silvano, M.,
1339 Cencelli, G., Lepri, F., Maitz, S., et al. (2015). Molecular Diversity and Associated Phenotypic
1340 Spectrum of Germline CBL Mutations. *Hum Mutat* 36, 787-796. 10.1002/humu.22809.
- 1341 129. Perez, B., Mechinaud, F., Galambrun, C., Ben Romdhane, N., Isidor, B., Philip, N., Derain-Court, J.,
1342 Cassinat, B., Lachenaud, J., Kaltenbach, S., et al. (2010). Germline mutations of the CBL gene define
1343 a new genetic syndrome with predisposition to juvenile myelomonocytic leukaemia. *J Med Genet*
1344 47, 686-691. 10.1136/jmg.2010.076836.
- 1345 130. Shain, A.H., Garrido, M., Botton, T., Talevich, E., Yeh, I., Sanborn, J.Z., Chung, J., Wang, N.J.,
1346 Kakavand, H., Mann, G.J., et al. (2015). Exome sequencing of desmoplastic melanoma identifies
1347 recurrent NFKBIE promoter mutations and diverse activating mutations in the MAPK pathway.
1348 *Nat Genet* 47, 1194-1199. 10.1038/ng.3382.
- 1349 131. Pudjihartono, M., Golovina, E., Fadason, T., O'Sullivan, J.M., and Schierding, W. (2024). Links
1350 between melanoma germline risk loci, driver genes and comorbidities: insight from a tissue-
1351 specific multi-omic analysis. *Mol Oncol* 18, 1031-1048. 10.1002/1878-0261.13599.
- 1352