1 TITLE

2 Germline features associated with immune infiltration in solid tumors

3 AUTHORS

4 Sahar Shahamatdar^{1,2}
5 Meng Xiao He^{3,4,5}
6 Matthew Reyna^{6,7}
7 Alexander Gusev^{3,8}
8 Saud H. AlDubayan^{3,4,8}
9 Eliezer M. Van Allen^{3,4,9,*}
10 Sohini Ramachandran^{1,2,9,*}

11 AFFILIATIONS

12 ¹Center for Computational Molecular Biology, Brown University, Providence, RI 02912, USA
13 ²Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912,
14 USA
15 ³Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
16 ⁴Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
17 ⁵Harvard Graduate Program in Biophysics, Boston, MA 02115, USA
18 ⁶Department of Computer Science, Princeton University, Princeton, NJ 08544, USA
19 ⁷Department of Biomedical Informatics, Emory University, Atlanta, GA 30322, USA
20 ⁸Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA
21 ⁹Senior Authors
22 *Correspondence: eliezerm.vanallen@dfci.harvard.edu (E.M.V.A.), sramachandran@brown.edu
23 (S.R.)

24 ABSTRACT

25 Given the clinical success of immune checkpoint blockade (ICB) across a diverse set of solid
26 tumors, and the emerging role for different immune infiltrates in contributing to response to ICB,
27 a comprehensive assessment of the properties that dictate immune infiltrations may reveal new
28 biological insights and inform the development of new effective therapies. Multiple studies have
29 examined somatic and functional immune properties associated with different tumor infiltrates;
30 however, germline features that associate with specific immune infiltrates in cancers have been
31 incompletely characterized. Here, we analyzed over 7 million autosomal germline variants in
32 the TCGA cohort (5788 European-ancestry samples across 30 cancer types) and tested for pan-
33 cancer association with established immune-related phenotypes that describe the tumor immune
34 microenvironment. We identified: one SNP associated with the fraction of follicular helper T
35 cells in bulk tumor; 77 unique candidate genes, some of which are involved in cytokine-mediated
36 signaling (e.g. *CNTF* and *TRIM34*) and cancer pathogenesis (e.g. *ATR* and *AKAP9*); and
37 subnetworks with genes that are part of DNA repair (*RAD51* and *XPC*) and transcription

elongation (*CCNT2*) pathways. We found a positive association between polygenic risk for rheumatoid arthritis and absolute fraction of infiltrating CD8 T cells. Overall, we identified multiple germline genetic features associated with specific tumor-immune phenotypes across cancer, and developed a framework for probing inherited features that contribute to variation in immune infiltration.

INTRODUCTION

Immune checkpoint blockade (ICB) therapies have emerged as impactful treatments for a variety of cancers. The discovery of cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) and programmed cell death protein 1 (PD-1) as important modulators of the adaptive immune system (Tivol et al., 1995; Fife et al., 2009) led to the development of ICB therapies, which target these specific pathways. Antagonism of PD-1 and CTLA4, negative regulators of T cell activity, stimulates the host immune system to recognize and kill tumor cells. While these therapeutic strategies are effective in a wide variety of cancers, they elicit variable clinical response (Ribas and Wolchok, 2018; Keenan et al., 2019).

Tumor-intrinsic features correlated with ICB clinical activity, such as mutational load and microsatellite instability, have been characterized extensively (Snyder et al., 2014; Gentles et al., 2015; Rizvi et al., 2015; Rooney et al., 2015; Van Allen et al., 2015; Giannakis et al., 2016; Miao and Allen, 2016; Charoentong et al., 2017; Miao et al., 2018; Samstein et al., 2019). Numerous lines of evidence indicate that selective response to ICB is also driven by the composition of the tumor microenvironment (TME), particularly the immune infiltration patterns in the TME (Tumeh et al., 2014; Thorsson et al., 2018). Thorsson et al. (2018) conducted an immunogenomic analysis of over 10,000 tumor samples spanning 33 cancer types compiled by The Cancer Genome Atlas (TCGA), reported specific driver mutations (in genes such as *NRAS* and *CASP8*) correlated with leukocyte levels, and demonstrated the prognostic and therapeutic implications of the TME composition.

Germline determinants of immune infiltration in solid tumors remain incompletely characterized, although germline features associated with immune traits have been found (Orrù et al., 2013; Roederer et al., 2015; Astle et al., 2016). Astle et al. (2016) found that common autosomal genotypes explain up to 21% of variance in white blood cell indices in a GWA study of 170,000 participants. Recently, Lim et al. (2018) uncovered 103 germline SNPs associated with immune cell abundance in the TME in 12 different cancer types. However the study overlooked potential confounding due to population structure, and did not offer insight into how individuals variants interact through genes or pathways to affect immune infiltration patterns.

Here, we analyze germline variants and test for association with immune infiltration in solid tumors in a pan-cancer meta-analysis of 30 TCGA cancer cohorts across different genomic scales. We identified SNPs, genes, and networks that modulate immune infiltration, as well as an association between polygenic risk for autoimmune diseases and immune infiltration.

78 **RESULTS**

79 **Overview of Association Analyses**

80 In order to characterize how host genetics affect immune infiltration in solid tumors, we analyzed
81 the association between germline variants and 17 phenotypes describing the immune component
82 of the tumor microenvironment across 30 TCGA cancer cohorts (Figure 1A). We conducted
83 QTL studies of the 17 molecular phenotypes, and aggregated SNP-level signals across genes
84 and networks with gene and network-level tests of association. In addition, we asked whether
85 polygenic risk of autoimmune diseases are associated with immune infiltration measures.

Figure 1. Association study approach and GWAS results: (A) Schematic showing the type and size of dataset after quality control for association studies. Association studies are done at three genomic scales across all 17 phenotypes. (B) Manhattan plot for GWAS meta-analysis for the follicular helper T cell phenotype. Positions along the chromosomes are on the x axis, and $-\log_{10}$ -transformed p -values are on the y axis. Every autosome is represented, but for visualization some are unlabeled. The red line indicates genome-wide significance ($p < 5 \times 10^{-8}$).

SNP-level Association with Follicular Helper T Cell Phenotype

Genome-wide association (GWA) studies of 5788 patients across 17 immune infiltration phenotypes reveal two associations at genome-wide significance ($p < 5 \times 10^{-8}$). rs3366, a variant in

the 3' UTR of *SIK1* (effect size = 0.1550, minor allele frequency = 18.42%, $p = 2.99 \times 10^{-9}$), is associated with the absolute fraction of follicular helper T (T_{FH}) cells in bulk tumor (Figure 1B). This SNP currently has no published associations in the GWAS catalog (McMahon et al., 2018). Although the biological role of *SIK1* in T_{FH} cells is unknown, there is evidence of differential expression of *SIK1* in this cell type (Newman et al., 2015).

rs4819959 is associated with the T helper 17 cells (T_H17) signature (effect size = -0.1682, $p = 1.71 \times 10^{-16}$). However, this variant is a known eQTL of *IL17RA* in 31 tissues in GTEx (Carithers and Moore, 2015), meaning the observed association is likely a byproduct of the T_H17 signature phenotype definition (gene expression of three genes including *IL17RA* (Thorsson et al., 2018; Bindea et al., 2013)).

Gene-level Association Studies Reveal 77 Candidate Genes

We then performed gene-level tests of association with immune infiltration patterns using PEGASUS (Nakka et al., 2016). We found 87 candidate gene-phenotype relationships ($p < 2.9 \times 10^{-5}$ after Bonferroni correction for 1703 independent haplotype blocks in the autosomes (Berisa and Pickrell, 2016)), compromising 77 unique genes across 17 phenotypes. We annotated these candidate genes based on: (1) expressed at mean transcripts per million (TPM) > 1 in either bulk tumor or immune cell populations from the Database of Immune Cell Expression, Expression quantitative trait loci, and Epigenomics (Schmiedel et al., 2018), (2) previously published GWAS hits in the GWAS catalog (McMahon et al., 2018), focusing on traits related to cancer, immunity, or autoimmunity, and (3) evidence for promoting oncogenic transformation via mutations according to the Cancer Gene Census (Futreal et al., 2004). The results are summarized in Figure 2A; full results can be found in Table S1.

We focused on candidate genes expressed in bulk tumor or immune cells, with 65 out of the 77 unique genes satisfying these criteria. Out of the 65 genes, we observed 10 unique gene candidates that contain reported GWAS hits in a related trait. Six out of ten genes (*AKAP9*, *CDK14*, *COL21A1*, *GPATCH1*, *MASTL*, *SBF2*) contain SNPs associated with different cancers, such as breast carcinoma, small cell lung carcinoma, and colorectal cancer. Five out of 10 genes (*COL21A1*, *IL17RA*, *KIAA1109*, *PXK*, *SIK1*) contain SNPs associated with immune or autoimmune traits, such as allergies, Crohn's disease, and systemic lupus erythematosus. We refer to genes with no published GWAS hits in traits related to cancer, immunity, or autoimmunity as novel candidate genes.

Six novel candidate genes were associated ($p < 2.28 \times 10^{-5}$) with the CD8 T cell phenotype, an established effector cell in the antitumor activity of the immune system (Figure 2B). *TCF12* is one of the candidate genes associated with the CD8 T cell phenotype, and it codes for a transcription factor called HEB. HEB regulates lineage-specific transcriptional profiles of $CD4^+CD8^+$ thymocytes (Futreal et al., 2004).

Two of the novel candidate genes (*ATR* and *EML4*), both associated with the leukocyte fraction phenotype, have been previously implicated in cancer pathogenesis according to the Cancer

Figure 2. Summary of candidate genes results: (A) Gene-level association testing identified 77 unique candidate genes, after Bonferroni correction for number of haplotype blocks in the autosomes. Out of the 77 genes, 65 genes are expressed at mean TPM > 1 in either bulk tumor samples or immune cells from healthy donors. Ten of the genes had published GWAS hits in traits related to cancer, immunity, or autoimmunity. The other 55 genes are designated as novel candidate genes. For the novel genes, the two boxes contain candidate genes that represent the two Gene Ontology (GO) terms with the most members. For the 10 previously-associated candidates, the two boxes contain candidate genes with published GWAS hits in cancer or immune/autoimmune traits. Genes are colored according to the phenotype category for which they are most significant. Genes significant for multiple phenotypes are denoted with a colored asterisk. (B) Manhattan plot for gene-level association analysis for the CD8 T cell phenotype. Each point represents a gene. Positions along the chromosomes are on the x axis, and $-\log_{10}$ -transformed p -values are on the y axis. The red line indicates Bonferroni-corrected significance ($p < 2.9 \times 10^{-5}$).

Gene Census (Futreal et al., 2004). *ATR* is inactivated via somatic missense mutations, and reported germline mutations predispose an individual to cancer (Tanaka et al., 2012).

Finally, we find that several immune-related phenotypes share candidate genes. For example, *ZFP91* is associated with T_H17 cells, lymphocytes, and macrophages phenotypes. This gene activates the NF- κ B pathway by stabilizing the NF- κ B inducing kinase, and therefore plays an important role in mounting an immune response (Jin et al., 2010).

Genes in DNA Repair and Transcription Elongation Pathways Correlated with Leukocyte Fraction

We then used network propagation to identify gene subnetworks enriched for genes with low gene-level p -values whose protein products are topologically connected on a protein-protein interaction

A

B

GO terms: * DNA Repair, § Metabolism, # Nucleotide Excision Repair, † Transcription Elongation

Figure 3. Altered subnetworks in leukocyte fraction phenotype: Two statistically significant ($p < 0.05$) altered subnetworks associated with the leukocyte fraction phenotype in the iRefIndex 15 interaction network. Each rectangle represents a gene, and is colored according to the $-\log_{10}$ -transformed PEGASUS gene p-value. Two genes are connected if their protein products interact in the iRefIndex 15 interaction network. Underlined genes are significantly associated genes from gene-level analysis. (A) Two candidate genes, *ATR* and *HSPA2*, are part of a larger subnetwork involved in DNA repair. Genes involved in DNA repair are indicated by *. In addition, genes involved in metabolism are indicated by §. (B) A subnetwork containing important members of the nucleotide excision repair and transcription elongation pathway, indicated by # and † respectively. *CCNT2* and *GTF2H4* are marginally significant ($p < 0.00018$).

network. Network analysis using Hierarchical HotNet (Reyna et al., 2018) was applied to each of the 17 phenotypes. We found statistically significant subnetworks for the leukocyte fraction phenotype ($p < 10^{-3}$) with the iRefIndex 15 interaction network; two of these subnetworks are highlighted in Figure 3. The second largest connected subgraph includes two candidate genes: *ATR* and *HSPA2* ($p < 2.8 \times 10^{-5}$). These genes are connected via *SYCP2*, which is involved in meiosis (Yang et al., 2006). Although not significant in our gene-level analysis, somatic mutations in *SYCP2* were previously reported to lower regulatory T cell to CD8 T cell ratios in head and neck cancers (Siemers et al., 2017). Other biologically relevant genes in this subnetwork include *FANCM*, *RAD51*, *PRIM1*, and *TOPBP1*, which participate in DNA repair pathways. Components of the subnetwork shown in Figure 3B are involved in the transcription elongation pathway (*CCNT2*, *LEO1*, *CD3EAP*, *GRF2H4*, and *IWS1*) and nucleotide excision repair pathway (*XPC*, *GTF2H4*, *COPS4*, and *COPS5*). None of the genes in this subnetwork had significant gene-level p-values, and were only discovered through network analysis.

Autoimmune Disease Polygenic Risk Associated With Immune Infiltration Patterns

Lastly, we explored how the pre-existing state of an individual's immune system may impact phenotypes of interest by investigating if common variants that affect the risk for autoimmune diseases are also correlated with immune infiltration (Figure 4A). We calculated polygenic risk scores (PRS) for five autoimmune disorders: rheumatoid arthritis, inflammatory bowel disease, celiac disease, systemic lupus erythematosus, and multiple sclerosis. These diseases were chosen based on availability of summary statistics in large, well-powered published GWA studies (Dubois et al., 2010; Sawcer et al., 2011; Anderson et al., 2011; Okada et al., 2013; Bentham et al., 2015). When computing PRS, we used the pruning and thresholding technique (Purcell et al., 2009), and based our scores on SNPs with GWA p-values of 0.001 or smaller (see Method Details: Polygenic Risk Score Analysis).

We identified statistically significant associations ($p < 0.0029$, Bonferroni corrected for number of immune infiltration phenotypes, 17) between PRS for rheumatoid arthritis and three immune infiltration phenotypes: lymphocytes, CD8 T cells, and macrophages (Figure 4B). The effect sizes are: CD8 T cells effect size = 0.0088, lymphocytes effect size = 0.0091, and macrophages effect size = -0.0073. It is important to note that the lymphocytes phenotype is defined as the sum of 12 cell types, one of which is amount of CD8 T cells (Thorsson et al., 2018). To test whether the lymphocyte and CD8 T cell hits were independent, we subtracted the amount of CD8 T cells from lymphocytes and repeated the analysis. In this reanalysis, we no longer observed a significant association between PRS of rheumatoid arthritis and this phenotype ($p = 0.0092$), demonstrating that the association signal of the lymphocytes phenotype is driven by the CD8 T cells phenotype.

Figure 4. Polygenic risk score associations with immune infiltration(A) Workflow for calculating polygenic risk scores of autoimmune disorders based on published GWAS summary statistics, followed by regression of the 17 immune infiltration phenotypes of interest onto the polygenic risk scores. (B) Bar plot showing the strength of association between 17 immune infiltration phenotypes and the polygenic risk score for rheumatoid arthritis. The phenotypes are on the x axis, and $-\log_{10}$ -transformed p -values are on the y axis. Each bar is colored according to the phenotype category. The red line indicates the Bonferroni-corrected significance value ($p = 0.0029$)

DISCUSSION

The abundance and composition of immune cell populations in the tumor microenvironment are known to affect response to immune checkpoint blockade. Here, we presented the first pan-cancer germline analysis of immune infiltration in solid tumors, demonstrating that host genetics are associated with phenotypes describing the immune component of the tumor microenvironment. Through integrative analysis of DNA-seq, RNA-seq, and DNA methylation data, we identified features at multiple genomic scales (SNP-level, gene-level, and pathway-level) that are correlated with amount of infiltrating follicular helper T cells (T_{FH}) and fraction of leukocytes in bulk tumor, among other phenotypes.

We found evidence for only one SNP-level association; rs3366 is associated with the amount of T_{FH} cells. The associated locus is in the 3' UTR of SIK1, a gene that is differentially expressed

191 in T_{FH} cells, among others, compared to other immune cells (Newman et al., 2015). The sparsity
192 of results from our GWA analysis is not surprising as the GWA framework is underpowered to
193 detect SNP-level associations in complex traits (McClellan and King, 2010; Stranger et al., 2011).

194
195 By aggregating SNP-level signals and testing for phenotype associations at the gene and pathway
196 levels, we uncovered multiple genes and pathways that are associated with immune infiltration
197 patterns. Out of 77 unique candidate genes, six were previously identified in GWA studies on
198 autoimmune disorders or immune-related traits; these results suggest host genomic factors that
199 cause variation or disease in the immune system also affect immune infiltration of tumors. We
200 found an additional five genes containing SNPs significant in cancer GWA studies; these genes
201 may be affecting cancer risk by altering the innate anti-tumor immune response. There is evi-
202 dence that non-genic cancer-risk SNPs are enriched in immune response processes, and therefore
203 may affect immune function (Fagny et al., 2018).

204
205 Our gene-level analysis also identified *ATR* as a novel candidate gene associated with leukocyte
206 fraction. Germline and somatic mutations in *ATR* have been reported to play a role in tumori-
207 genesis (Tanaka et al., 2012; Harsha et al., 2016). Somatic *ATR* mutations have also been shown
208 to modulate the tumor microenvironment in melanomas, recruiting macrophages and blocking
209 T cell recruitment (Chen et al., 2017).

210
211 *ATR* and interacting genes are found to be associated with the leukocyte fraction phenotype in
212 our network propagation analysis. Significantly associated subnetworks contain genes involved
213 in important pathways such as DNA repair, nucleotide excision repair, and transcription elonga-
214 tion. Somatic mutations in genes involved in DNA repair, such as *ATR* and *RAD51* associated
215 with leukocyte fraction in our network analyses, can increase the neoantigen load in the TME
216 and affect response to immunotherapy (Mouw et al., 2017; Knijnenburg et al., 2018). In addi-
217 tion, defective transcription elongation is known to confer resistance to immunotherapy despite
218 increased levels of infiltrating T cells (Modur et al., 2018).

219
220 Finally, we showed that the polygenic risk score for rheumatoid arthritis is correlated with
221 amount of CD8 T cells, suggesting a shared genetic etiology between rheumatoid arthritis and
222 cytotoxic immune response to solid tumors. In the synovial compartment of rheumatic joints,
223 40% of T cells are CD8 T cells (McInnes, 2003). Past studies have found associations between
224 rheumatoid arthritis and MHC class I polymorphisms (Raychaudhuri et al., 2012) as well as be-
225 tween amount of CD8 T cells in synovial fluid and disease activity (Cho et al., 2012), suggesting
226 a potential role for CD8 T cells in the development and progression of rheumatoid arthritis.

227
228 While we implemented many quality control filters of genotype and phenotype data to remove
229 confounders in our analyses, replication is necessary. We were unable to conduct a replication
230 analysis; replication studies are currently not feasible due to a lack of a large, independent, pan-
231 cancer cohort with matched germline and gene expression data. The TCGA dataset provided
232 a unique opportunity to conduct integrative association analyses that leverage germline data,
233 which have largely been under-appreciated (besides investigation of predisposition germline vari-
234 ants in cancer (Kim et al., 2013; Palles et al., 2012; Huang et al., 2018)).

235

236 We note that 16 out of 17 phenotypes we studied here were based on bulk RNA-seq data, and
 237 six of those 16 were derived using a deconvolution method CIBERSORT (Newman et al., 2015).
 238 CIBERSORT has several limitations, including reliance on the fidelity of a reference expression
 239 panel for deconvolution, and not being explicitly tested on RNA-seq data during development
 240 (Newman et al., 2015). Ideally, future studies will integrate germline and somatic variation
 241 with orthogonal measures of immune infiltration patterns (such as flow cytometry based mea-
 242 surements), but such study design does not currently exist to validate the results presented here.

243

244 Future studies incorporating other immune cell populations known to affect response to im-
 245 munotheapy (such as amount of neutrophils or CD4 T cells) and joint analysis of germline
 246 variants and somatic mutations will further understanding of predictors of response to immune
 247 checkpoint blockade. And ultimately experimental investigations are needed to determine the
 248 biological mechanisms driving the reported associations.

249

250 In conclusion, we reported germline variation in SNPs, genes, and pathways associated with
 251 immune infiltration patterns. These results highlight the important yet previously overlooked
 252 role that inherited variants play in determining the immune composition of the TME, a crucial
 253 step towards understanding predictors of response to immune checkpoint blockade therapies.

254 METHODS

255 Sample Inclusion Criteria

256 The Cancer Genome Atlas (TCGA) dataset consists of tumor and matched normal samples from
 257 over 11,000 patients. The Genomic Data Commons (GDC) legacy archive contains germline
 258 data for 11,440 samples from 10,776 unique participants. Samples with the following TCGA
 259 project IDs: DLBC, LAML, LCML, MISC, and THYM were excluded as they represent uniden-
 260 tified cancer or cancers derived from immune cells. Samples indicated as problematic by either
 261 GDC-issued or TCGA-issued annotations were removed. The reasons for exclusion ranged from
 262 mismatched genotypes in tumor and normal samples to incorrect barcodes on aliquots. Strict
 263 genetic ancestry filtering was applied to account for population structure.

264 Raw Germline Variant Data

265 Germline variants were derived from the Affymetrix SNP6.0 microarray. Raw CEL files for the
 266 TCGA cohort were downloaded from FireCloud (<https://software.broadinstitute.org/firecloud/>)
 267 and the Genomic Data Commons (GDC) legacy archive ([https://portal.gdc.cancer.gov/legacy-](https://portal.gdc.cancer.gov/legacy-archive)
 268 archive). Probesets with non-unique mapping in the genome or not mapping to the location
 269 provided by Affymetrix (NetAffx Annotation Release 35) were removed.

270 Germline Variant Calling

271 Genotypes calls from the CEL files were made using Birdseed (Korn et al., 2008) in batches;
 272 samples from the same TCGA batch were included in the same run. Because Birdseed recom-
 273 mends more than 50 samples in each run, batches with less than 50 samples were combined with
 274 samples from temporally adjacent batches. Genotype calls with Birdseed confidence scores more
 275 than 0.1 were removed.

276
 277 Samples with autosomal SNP missingness $> 2\%$ or unexpected sex chromosome genotypes (males
 278 with missing Y chromosome calls or females with Y chromosome calls) were removed. Partici-
 279 pants with more than two replicate samples were removed. Participants with replicate samples
 280 with $> 1\%$ discordance among genotype calls were removed. Among these samples, SNPs with
 281 missingness $> 5\%$, sex effect (Fisher's exact $p < 10^{-20}$), or batch effect (each batch versus all
 282 others, Fisher's exact $p < 10^{-12}$) were removed. Several participants had two replicate samples
 283 remaining after the filtering process. SNPs with $> 2\%$ replicate discordance were removed. For
 284 each participant, the sample with the higher genotype missingness was removed and discordant
 285 genotypes were excluded.

286
 287 We imputed genotypes with the Michigan Imputation Server (Das et al., 2016), using data from
 288 the Haplotype Reference Consortium (McCarthy et al., 2016) as the reference panel. Loci with
 289 imputation quality $R^2 < 0.8$ were excluded.

290
 291 To prepare the genotype data for association studies, the following additional quality control
 292 steps were taken using plink (Chang et al., 2015):

1. SNPs with minor allele frequency $< 1\%$ were removed.
 2. SNPs not in Hardy Weinberg equilibrium ($p < 10^{-6}$) were removed.
 3. Related individuals (IBD $\hat{\pi} > 0.185$) were removed.
 4. Samples with missing GDC demographic data (sex and birth year) were removed.
- The final genotype data consists of 7,070,031 variants and 5788 samples.

Genetic Ancestry Calculation and Inclusion Criteria

Strict ancestry filtering was applied to samples using two techniques: (1) project TCGA samples onto a ten-dimensional principal component (PC)-space derived from PCA all individuals in the 1000 Genomes Project (Auton et al., 2015), and retain only TCGA samples whose five nearest 1000 Genome neighbors were labelled as "European" and whose mean distance to those neighbors was < 0.1 . (2) Run supervised Admixture (Alexander et al., 2009) with K set to 3 — using the Utah Residents with Northern and Western European Ancestry (CEU), Yoruba in Ibadan, Nigeria (YRI), and Han Chinese in Beijing, China (CHB) + Japanese in Tokyo, Japan (JPT) populations as reference data — and keep TCGA samples with greater than 90% membership in the CEU cluster.

Phenotype Data

CIBERSORT-derived fraction of 22 types of immune cells, immune gene expression signatures, and leukocyte fraction from methylation analysis were downloaded from Thorsson et al. (2018). Cytolytic activity immune signature was added from Rooney et al. (2015). Twenty phenotypes with more than 10% zero-values were excluded, with 17 phenotype remaining. Within each cancer cohort, a rank-based inverse normal transformation was applied to each phenotype. The transformed value of phenotype j for the i th subject in cohort k is:

$$Y_{ijk} = \phi^{-1}\left(\frac{r_{ijk} - 0.5}{N_{jk}}\right), \quad (1)$$

where r_{ijk} is the rank of the i th case in non-null observations in phenotype j in cohort k , N_{jk} is the number of non-null observations of phenotype j in cohort k , and ϕ^{-1} is the probit function.

SNP-level and Gene-level Association Studies

Genome-wide association (GWA) studies were conducted for 17 phenotypes within each cancer-specific cohort using plink (Chang et al., 2015). The first ten genetic PCs, age, and sex were included in the regression analysis as covariates. We then used METAL (Willer et al., 2010) with a sample size weighting scheme to perform a pan-cancer meta-analysis for each phenotype. The effect sizes of significant SNPs ($p < 5 \times 10^{-8}$) were calculated using an inverse-variance weighting scheme.

These GWA SNP-level summary statistics were then used as input to the gene-level association test method PEGASUS (Nakka et al., 2016). Gene-level p-values are reported for genes with at least one SNP in the gene boundary $\pm 50\text{kb}$ window (17,563 autosomal genes). Genes with p-values less than 2.9×10^{-5} (Bonferroni corrected for number of independent haplotype blocks in the autosomes, 1703 (Berisa and Pickrell, 2016)) were reported as significant.

Network Analysis

We performed network analysis with Hierarchical HotNet Reyna et al. (2018), on the log transformed p-values ($-\log_{10}(p)$) from gene-level association testing to identify significantly altered subnetworks. For our analysis, we used the following interaction networks, which were the most recent versions available as of February 23, 2018.

- HINT+HI (Das and Yu, 2012; Rolland et al., 2014): HINT binary + HINT co-complex + HuRI HI
- iRefIndex 15.0 (Razick et al., 2008)
- ReactomeFI 2016 (Fabregat et al., 2018)

For the ReactomeFI network, we considered the set of interactions with a confidence score of 0.75 (out of 1) or larger. For each network, we restricted our attention to the largest connected subgraph of the network.

To reduce the influence of genes for which we have low confidence of association with a phenotype, we assigned p-values of 1 to genes with p-values of $p > 0.1$ and ran Hierarchical HotNet (10^3 permutations) on these thresholded gene scores. This provides sparser, more interpretable, and higher confidence networks. Similar p-value thresholds were applied in similar network analyses (Nakka et al., 2016).

Polygenic Risk Score Analysis

We downloaded the summary statistics from GWA studies of five autoimmune traits: celiac disease (Dubois et al., 2010); multiple sclerosis (Sawcer et al., 2011); ulcerative colitis (Anderson et al., 2011); rheumatoid arthritis (Okada et al., 2013); systemic lupus erythematosus (Bentham et al., 2015). Records with missing odds ratio, p-values, and risk alleles were excluded from analysis. For each autoimmune disease, we extracted SNPs at various p-value thresholds ($p = 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, 5 \times 10^{-8}$) that overlapped with our genotype data, excluding ambiguous and mismatched variants. At each threshold, the SNPs were further filtered via LD-clumping, with a 250kb window and an r^2 threshold of 0.1 (Table S2). PRSice (Lewis et al., 2014) was used to calculate the polygenic risk score for each autoimmune trait for each sample by summing over the log odds ratio of the selected SNPs, weighted by allele dosage of risk alleles.

361 The polygenic risk score for each disease was first regressed against each of the 17 immune
362 infiltration phenotypes within each cancer cohort, using the first 10 PCs, age, and sex as covari-
363 ates. The reported results are from a sample size based meta-analysis of all cancer cohorts.
364 Effect sizes of significant associations (Bonferroni corrected for number of immune infiltration
365 phenotypes tested) were calculated using an inverse-variance weighted analysis.

366 References

- 367 Alexander, D. H., Novembre, J. and Lange, K. (2009). Fast model-based estimation of ancestry
368 in unrelated individuals. *Genome research* 19, 1655–1664.
- 369 Anderson, C. A., Boucher, G., Lees, C. W., Franke, A., D’Amato, M., Taylor, K. D., Lee,
370 J. C., Goyette, P., Imielinski, M., Latiano, A. and et al. (2011). Meta-analysis identifies 29
371 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47.
372 *Nature Genetics* 43, 246.
- 373 Astle, W. J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A. L., Mead, D., Bouman,
374 H., Riveros-Mckay, F., Kostadima, M. A. and et al. (2016). The Allelic Landscape of Human
375 Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* 167, 1415 – 1429.e19.
- 376 Auton, A., Abecasis, G. R., Altshuler (Co-Chair), D. M., Durbin (Co-Chair), R. M., Abecasis,
377 G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E. and et al.
378 (2015). A global reference for human genetic variation. *Nature* 526, 68.
- 379 Bentham, J., Morris, D. L., Cunninghame Graham, D. S., Pinder, C. L., Tomblinson, P., Behrens,
380 T. W., Martín, J., Fairfax, B. P., Knight, J. C. and et al. (2015). Genetic association analyses
381 implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of
382 systemic lupus erythematosus. *Nature Genetics* 47, 1457.
- 383 Berisa, T. and Pickrell, J. K. (2016). Approximately independent linkage disequilibrium blocks
384 in human populations. *Bioinformatics (Oxford, England)* 32, 283–285.
- 385 Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A., Angell, H.,
386 Fredriksen, T., Lafontaine, L., Berger, A. and et al. (2013). Spatiotemporal Dynamics of
387 Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* 39,
388 782–795.
- 389 Carithers, L. J. and Moore, H. M. (2015). The Genotype-Tissue Expression (GTEx) Project.
390 *Biopreservation and Biobanking* 13, 307–308.
- 391 Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M. and Lee, J. J. (2015).
392 Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*
393 4, 7.
- 394 Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl,
395 H. and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-
396 Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell*
397 *Reports* 18, 248–262.

- 398 Chen, C.-F., Ruiz-Vega, R., Vasudeva, P., Espitia, F., Krasieva, T. B., de Feraudy, S., Tromberg,
399 B. J., Huang, S., Garner, C. P., Wu, J., Hoon, D. S. and Ganesan, A. K. (2017). ATR Muta-
400 tions Promote the Growth of Melanoma Tumors by Modulating the Immune Microenviron-
401 ment. *Cell Reports* 18, 2331–2342.
- 402 Cho, B.-A., Sim, J. H., Park, J. A., Kim, H. W., Yoo, W.-H., Lee, S.-H., Lee, D.-S., Kang,
403 J. S., Hwang, Y.-I., Lee, W. J. and et al. (2012). Characterization of Effector Memory CD8+
404 T Cells in the Synovial Fluid of Rheumatoid Arthritis. *Journal of Clinical Immunology* 32,
405 709–720.
- 406 Das, J. and Yu, H. (2012). HINT: High-quality protein interactomes and their applications in
407 understanding human disease. *BMC systems biology* 6, 92.
- 408 Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A. E., Kwong, A., Vrieze, S. I., Chew, E. Y.,
409 Levy, S., McGue, M. and et al. (2016). Next-generation genotype imputation service and
410 methods. *Nature Genetics* 48, 1284.
- 411 Dubois, P. C. A., Trynka, G., Franke, L., Hunt, K. A., Romanos, J., Curtotti, A., Zhernakova,
412 A., Heap, G. A. R., Ádány, R., Aromaa, A. and et al. (2010). Multiple common variants for
413 celiac disease influencing immune gene expression. *Nature Genetics* 42, 295.
- 414 Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R.,
415 Jassal, B., Korkmeyer, F., May, B. and et al. (2018). The Reactome Pathway Knowledgebase.
416 *Nucleic acids research* 46, D649–D655.
- 417 Fagny, M., Platig, J., Kuijjer, M. L., Lin, X. and Quackenbush, J. (2018). Nongenetic cancer-risk
418 SNPs affect oncogenes, tumor suppressor genes, and immune function. *bioRxiv* 1.
- 419 Fife, B. T., Pauken, K. E., Eagar, T. N., Obu, T., Wu, J., Tang, Q., Azuma, M., Krummel,
420 M. F. and Bluestone, J. A. (2009). Interactions between PD-1 and PD-L1 promote tolerance
421 by blocking the TCR-induced stop signal. *Nature immunology* 10, 1185.
- 422 Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N. and
423 Stratton, M. R. (2004). A census of human cancer genes. *Nature Reviews Cancer* 4, 177.
- 424 Gentles, A. J., Newman, A. M., Liu, C. L., Bratman, S. V., Feng, W., Kim, D., Nair, V. S.,
425 Xu, Y., Khuong, A., Hoang, C. D., Diehn, M., West, R. B., Plevritis, S. K. and Alizadeh,
426 A. A. (2015). The prognostic landscape of genes and infiltrating immune cells across human
427 cancers. *Nature Medicine* 21, 938.
- 428 Giannakis, M., Mu, X., Shukla, S., Qian, Z., Cohen, O., Nishihara, R., Bahl, S., Cao, Y., Amin-
429 Mansour, A., Yamauchi, M. and et al. (2016). Genomic Correlates of Immune-Cell Infiltrates
430 in Colorectal Carcinoma. *Cell Reports* 15, 857 – 865.
- 431 Harsha, B., Kok, C. Y., Cole, C. G., Beare, D., Dawson, E., Boutselakis, H., Jubb, H., Tate, J.,
432 Ponting, L., Jia, M. and et al. (2016). COSMIC: somatic cancer genetics at high-resolution.
433 *Nucleic Acids Research* 45, D777–D783.

- 434 Huang, K.-l., Mashl, R. J., Wu, Y., Ritter, D. I., Wang, J., Oh, C., Paczkowska, M., Reynolds,
435 S., Wyczalkowski, M. A., Oak, N. and et al. (2018). Pathogenic Germline Variants in 10,389
436 Adult Cancers. *Cell* 173, 355–370.e14.
- 437 Jin, X., Jin, H. R., Jung, H. S., Lee, S. J., Lee, J.-H. and Lee, J. J. (2010). An Atypical E3
438 Ligase Zinc Finger Protein 91 Stabilizes and Activates NF- κ B-inducing Kinase via Lys63-
439 linked Ubiquitination. *Journal of Biological Chemistry* 285, 30539–30547.
- 440 Keenan, T. E., Burke, K. P. and Van Allen, E. M. (2019). Genomic correlates of response to
441 immune checkpoint blockade. *Nature Medicine* 25, 389–402.
- 442 Kim, H., Minna, J. and White, M. (2013). GWAS Meets TCGA to Illuminate Mechanisms of
443 Cancer Predisposition. *Cell* 152, 387–389.
- 444 Knijnenburg, T. A., Wang, L., Zimmermann, M. T., Chambwe, N., Gao, G. F., Cherniack,
445 A. D., Fan, H., Shen, H., Way, G. P., Greene, C. S. and et al. (2018). Genomic and Molecular
446 Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Reports*
447 23, 239–254.e6.
- 448 Korn, J. M., Kuruvilla, F. G., McCarroll, S. A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell,
449 E., Veitch, J., Collins, P. J., Darvishi, K., Lee, C., Nizzari, M. M., Gabriel, S. B., Purcell, S.,
450 Daly, M. J. and Altshuler, D. (2008). Integrated genotype calling and association analysis of
451 SNPs, common copy number polymorphisms and rare CNVs. *Nature genetics* 40, 1253–1260.
- 452 Lewis, C. M., Euesden, J. and O'Reilly, P. F. (2014). PRSice: Polygenic Risk Score software.
453 *Bioinformatics* 31, 1466–1468.
- 454 Lim, Y. W., Chen-Harris, H., Mayba, O., Lianoglou, S., Wuster, A., Bhangale, T., Khan,
455 Z., Mariathasan, S., Daemen, A., Reeder, J., Haverty, P. M., Forrest, W. F., Brauer, M.,
456 Mellman, I. and Albert, M. L. (2018). Germline genetic polymorphisms influence tumor gene
457 expression and immune cell infiltration. *Proceedings of the National Academy of Sciences*
458 115, E11701–E11710.
- 459 McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang, H. M.,
460 Fuchsberger, C., Danecek, P., Sharp, K. and et al. (2016). A reference panel of 64,976 haplo-
461 types for genotype imputation. *Nature Genetics* 48, 1279.
- 462 McClellan, J. and King, M.-C. (2010). Genetic Heterogeneity in Human Disease. *Cell* 141,
463 210–217.
- 464 McInnes, I. B. (2003). Leukotrienes, mast cells, and T cells. *Arthritis research & therapy* 5,
465 288–289.
- 466 McMahon, A., Malangone, C., Suveges, D., Sollis, E., Cunningham, F., Riat, H. S., MacArthur,
467 J. A., Hayhurst, J., Morales, J., Guillen, J. A. and et al. (2018). The NHGRI-EBI GWAS
468 Catalog of published genome-wide association studies, targeted arrays and summary statistics
469 2019. *Nucleic Acids Research* 47, D1005–D1012.

- 470 Miao, D. and Allen, E. M. V. (2016). Genomic determinants of cancer immunotherapy. *Current*
471 *Opinion in Immunology* 41, 32 – 38.
- 472 Miao, D., Margolis, C. A., Gao, W., Voss, M. H., Li, W., Martini, D. J., Norton, C., Bossé, D.,
473 Wankowicz, S. M., Cullen, D. and et al. (2018). Genomic correlates of response to immune
474 checkpoint therapies in clear cell renal cell carcinoma. *Science* 359, 801–806.
- 475 Modur, V., Singh, N., Mohanty, V., Chung, E., Muhammad, B., Choi, K., Chen, X., Chetal, K.,
476 Ratner, N., Salomonis, N. and et al. (2018). Defective transcription elongation in a subset of
477 cancers confers immunotherapy resistance. *Nature Communications* 9, 4410.
- 478 Mouw, K. W., Goldberg, M. S., Konstantinopoulos, P. A. and D’Andrea, A. D. (2017). DNA
479 Damage and Repair Biomarkers of Immunotherapy Response. *Cancer Discovery* 7, 675–693.
- 480 Nakka, P., Raphael, B. J. and Ramachandran, S. (2016). Gene and Network Analysis of Common
481 Variants Reveals Novel Associations in Multiple Complex Diseases. *Genetics* 204, 783–798.
- 482 Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., Hoang, C. D., Diehn,
483 M. and Alizadeh, A. A. (2015). Robust enumeration of cell subsets from tissue expression
484 profiles. *Nature Methods* 12, 453.
- 485 Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki,
486 A., Yoshida, S. and et al. (2013). Genetics of rheumatoid arthritis contributes to biology and
487 drug discovery. *Nature* 506, 376.
- 488 Orrù, V., Steri, M., Sole, G., Sidore, C., Virdis, F., Dei, M., Lai, S., Zoledziowska, M., Busonero,
489 F., Mulas, A. and et al. (2013). Genetic Variants Regulating Immune Cell Levels in Health
490 and Disease. *Cell* 155, 242 – 256.
- 491 Palles, C., Cazier, J.-B., Howarth, K. M., Domingo, E., Jones, A. M., Broderick, P., Kemp, Z.,
492 Spain, S. L., Guarino, E., Salguero, I. and et al. (2012). Germline mutations affecting the
493 proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcino-
494 mas. *Nature Genetics* 45, 136.
- 495 Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O’Donovan, M. C., Sullivan, P. F.,
496 Sklar, P., Purcell (Leader), S. M., Stone, J. L., Sullivan, P. F. and et al. (2009). Common
497 polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460,
498 748.
- 499 Raychaudhuri, S., Sandor, C., Stahl, E. A., Freudenberg, J., Lee, H.-S., Jia, X., Alfredsson, L.,
500 Padyukov, L., Klareskog, L., Worthington, J. and et al. (2012). Five amino acids in three HLA
501 proteins explain most of the association between MHC and seropositive rheumatoid arthritis.
502 *Nature Genetics* 44, 291.
- 503 Razick, S., Magklaras, G. and Donaldson, I. M. (2008). iRefIndex: a consolidated protein
504 interaction database with provenance. *BMC bioinformatics* 9, 405.
- 505 Reyna, M. A., Leiserson, M. D. M. and Raphael, B. J. (2018). Hierarchical HotNet: identifying
506 hierarchies of altered subnetworks. *Bioinformatics (Oxford, England)* 34, i972–i980.

- 507 Ribas, A. and Wolchok, J. D. (2018). Cancer immunotherapy using checkpoint blockade. *Science*
508 359, 1350–1355.
- 509 Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., Lee, W.,
510 Yuan, J., Wong, P., Ho, T. S. and et al. (2015). Mutational landscape determines sensitivity
511 to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128.
- 512 Roederer, M., Quaye, L., Mangino, M., Beddall, M., Mahnke, Y., Chattopadhyay, P., Tosi, I.,
513 Napolitano, L., TerranovaBarberio, M., Menni, C. and et al. (2015). The Genetic Architecture
514 of the Human Immune System: A Bioresource for Autoimmunity and Disease Pathogenesis.
515 *Cell* 161, 387 – 403.
- 516 Rolland, T., Taşan, M., Charleaux, B., Pevzner, S., Zhong, Q., Sahni, N., Yi, S., Lemmens, I.,
517 Fontanillo, C., Mosca, R. and et al. (2014). A Proteome-Scale Map of the Human Interactome
518 Network. *Cell* 159, 1212–1226.
- 519 Rooney, M., Shukla, S., Wu, C., Getz, G. and Hacohen, N. (2015). Molecular and Genetic
520 Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* 160, 48 – 61.
- 521 Samstein, R. M., Lee, C.-H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y.,
522 Barron, D. A., Zehir, A., Jordan, E. J., Omuro, A. and et al. (2019). Tumor mutational load
523 predicts survival after immunotherapy across multiple cancer types. *Nature Genetics* 51,
524 202–206.
- 525 Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C. A., Patsopoulos, N. A., Moutsianas, L.,
526 Dilthey, A., Su, Z., Freeman, C., Hunt, S. E. and et al. (2011). Genetic risk and a primary
527 role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 476, 214.
- 528 Schmiedel, B. J., Singh, D., Madrigal, A., Valdovino-Gonzalez, A. G., White, B. M., Zapardiel-
529 Gonzalo, J., Ha, B., Altay, G., Greenbaum, J. A. and et al. (2018). Impact of Genetic
530 Polymorphisms on Human Immune Cell Gene Expression. *Cell* 175, 1701–1715.e16.
- 531 Siemers, N. O., Holloway, J. L., Chang, H., Chasalow, S. D., Ross-MacDonald, P. B., Voliva,
532 C. F. and Szustakowski, J. D. (2017). Genome-wide association analysis identifies genetic
533 correlates of immune infiltrates in solid tumors. *PloS one* 12, e0179726–e0179726.
- 534 Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J. M., Desrichard, A., Walsh, L. A.,
535 Postow, M. A., Wong, P., Ho, T. S., Hollmann, T. J., Bruggeman, C., Kannan, K., Li, Y.,
536 Elipenahli, C., Liu, C., Harbison, C. T., Wang, L., Ribas, A., Wolchok, J. D. and Chan, T. A.
537 (2014). Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *New England*
538 *Journal of Medicine* 371, 2189–2199.
- 539 Stranger, B. E., Stahl, E. A. and Raj, T. (2011). Progress and Promise of Genome-Wide
540 Association Studies for Human Complex Trait Genetics. *Genetics* 187, 367–383.
- 541 Tanaka, A., Weinell, S., Nagy, N., O'Driscoll, M., Lai-Cheong, J., Kulp-Shorten, C., Knable,
542 A., Carpenter, G., Fisher, S., Hiragun, M., Yanase, Y., Hide, M., Callen, J. and McGrath,
543 J. (2012). Germline Mutation in *ATR* in Autosomal-Dominant Oropharyngeal
544 Cancer Syndrome. *The American Journal of Human Genetics* 90, 511–517.

- 545 Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Yang, T.-H. O., Porta-
546 Pardo, E., Gao, G. F., Plaisier, C. L., Eddy, J. A. and et al. (2018). The Immune Landscape
547 of Cancer. *Immunity* 48, 812 – 830.e14.
- 548 Tivol, E. A., Borriello, F., Schweitzer, A. N., Lynch, W. P., Bluestone, J. A. and Sharpe, A. H.
549 (1995). Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue
550 destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* 3, 541–547.
- 551 Tumeh, P. C., Harview, C. L., Yearley, J. H., Shintaku, I. P., Taylor, E. J. M., Robert, L.,
552 Chmielowski, B., Spasic, M., Henry, G., Ciobanu, V. and et al. (2014). PD-1 blockade induces
553 responses by inhibiting adaptive immune resistance. *Nature* 515, 568.
- 554 Van Allen, E. M., Miao, D., Schilling, B., Shukla, S. A., Blank, C., Zimmer, L., Sucker, A.,
555 Hillen, U., Geukes Foppen, M. H. and et al. (2015). Genomic correlates of response to CTLA-
556 4 blockade in metastatic melanoma. *Science* 350, 207–211.
- 557 Willer, C. J., Abecasis, G. R. and Li, Y. (2010). METAL: fast and efficient meta-analysis of
558 genomewide association scans. *Bioinformatics* 26, 2190–2191.
- 559 Yang, F., Fuente, R. D. L., Leu, N. A., Baumann, C., McLaughlin, K. J. and Wang, P. J.
560 (2006). Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal
561 synapsis during male meiosis. *The Journal of Cell Biology* 173, 497–507.