Integrative Genome-Wide Association Studies of COVID-19 Susceptibility and Hospitalization Reveal Risk Loci for Long COVID

Zhongshan Cheng[1*]

1. Center for Applied Bioinformatics, St. Jude Children's Research Hospital, 262 Danny Thomas Hospital, Memphis, TN 38105, USA

*Corresponding authors:

Zhongshan Cheng: Senior Bioinformatics Research Scientist, Center for applied bioinformatics, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, United States; email: cheng.zhong.shan@gmail.com

Keywords: COVID-19, SARS-CoV2, Long COVID, GWAS, TWAS, Hospitalization, Susceptibility

Abstract

Long COVID presents a significant public health challenge, characterized by over 200 reported symptoms spanning multiple organ systems, and genome-wide association studies (GWAS) have been hindered by the disorder's symptom heterogeneity and the limited power of available datasets. To overcome the challenges of limited sample size or heterogeneity of long COVID cohorts, a proxy-based, hypothesis-generating strategy is conducted to prioritize candidate risk loci that may contribute to long COVID by analyzing GWAS summary statistics for COVID-19 susceptibility and hospitalization, as well as long COVID from the COVID-19 Host Genetics Initiative (HGI, Release 7), resulting in 62 candidate loci represented by independent single nucleotide polymorphisms (SNPs). These SNPs are categorized into three groups based on their association with acute-phase COVID-19: (1) severe COVID-19-specific SNPs, such as SNPs mapped to two well-known loci involved in SARS-CoV2 entry (*ACE2* [rs190509934] and *TMPRSS2* [rs12329760]), as well as variants of *DPP9* (rs7251000), *FOXP4* (rs12660421) and *HLA-DQA1* (rs17219281), exhibiting associations with severe COVID-19 but displaying weaker signals in non-hospitalized COVID-19 cases; (2) SNPs associated with both severe and mild COVID-19, including the SNP close to *ABO* (rs505922); and (3) non-hospitalization-specific SNPs, such as variants in *KCTD16* (rs62401842) and *WASF3* (rs56143829), highlighting genetic contributors specific to mild COVID-19 cases. These candidate SNPs are investigated with recently published long COVID GWAS data from HGI, revealing most of these candidate SNPs display much weaker association with long COVID. Further transcriptome-wide association studies (TWASs) for GWASs of COVID-19 hospitalization/susceptibility/long COVID across 48 GTEx tissues demonstrate that genes adjacent to these candidate SNPs (43 out of 62 SNPs) exhibits relatively weaker association signals in long COVID across heart, brain, and muscle-related tissues, particular for notable TWAS hits, *DPP9* (rs7251000), *CCR1* (rs17078348), and *THBS3* (rs41264915) that display much weaker association in long COVID compared to acute-phase COVID-19. Phenome-wide TWASs also link *HLA-DQA1* (rs17219281), *HLA-A* (rs9260038), and *HLA-C* (rs1634761), *GSDMB* (rs9916158) and *FOXP4* (rs12660421) with other phenotypes closely relevant to long COVID. These results provide new insights to long COVID by leveraging acute-phase COVID-19.

Introduction

Long COVID remains a critical and complex public health issue, with a diverse array of over 200 reported symptoms affecting multiple organ systems[1]. These symptoms, which include fatigue, cognitive impairment, respiratory issues, and cardiovascular problems, can persist well beyond the acute phase of COVID-19 infection, revealing a unique challenge for researchers to identify genetic risk factors using genome-wide association studies (GWAS), as the wide variation in symptom clusters may arise from different biological mechanisms of long COVID and require extensive sample sizes to detect meaningful associations. To address this complexity, long COVID GWAS requires not only large and diverse patient datasets but also refined approaches to phenotype classification and statistical analyses.

Among the most significant efforts to understand the genetic basis of long COVID is the large GWAS from the COVID-19 Host Genetic Initiatives (HGI) conducted by Lama et al.[2], which analyzed approximately 3,500 long COVID cases but only identified a single gene, *FOXP4*, on chromosome 17, reaching genome-wide significance for association with long COVID. *FOXP4* variants were found to be strongly associated with both long COVID and severe COVID-19, in line with a shared genetic predisposition that may drive both conditions[3]. However, a stronger association of *FOXP4* with severe COVID-19 suggests that *FOXP4* may contribute primarily to risk factors influencing severe acute COVID-19 cases that later progress to long COVID[3]. Given that long COVID can also follow mild COVID-19, additional genetic contributors specifically associated with mild cases likely remain undiscovered.

Another long COVID GWAS conducted by Taylor et al.[4] provided further insights into genetic factors specific to a subtype of long COVID by investigating genetic overlap between long COVID and myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), a condition with substantial symptom overlap, including chronic fatigue and cognitive impairments. However, their GWAS faced limitations due to its relatively small sample size ($n < 1{,}000$ long COVID cases) and only reported SNPs derived by network analysis without identifying any SNPs that reached genome-wide significance for long COVID, emphasizing the need for larger, more statistically robust studies.

More recently, a multi-ancestry GWAS was conducted by Chaudhary et al.[5], which expanded the scope of long COVID GWAS by including participants from diverse ancestral backgrounds. Notably, several immune-related loci, such as *ABO*, *BPTF*, and *HLA-DQA1*, were reported to be linked with chronic fatigue syndrome (CFS), fibromyalgia, and depression, suggesting potential immune dysregulation in long COVID. Despite this, this multi-ancestry long COVID GWAS still has limited sample size in non-European samples,

leading to the lack of sufficient power to detect associations for less common symptom clusters or subtle genetic variants.

Compared to the limited long COVID GWAS hits reported by Lama et al. and others, large-scale GWAS from the HGI have robustly identified genetic variants associated with COVID-19 susceptibility (infection vs. population) and hospitalization (hospitalized vs. non-hospitalized or general population)[6]. Since both COVID-19 acute severity and susceptibility are known risk factors for long COVID[7], we hypothesize that genetic variants associated with these acute-phase COVID-19 may also influence the risk of developing long COVID given the strong phenotype-wide correlation between long COVID and acute-phase COVID-19[8]. To explore the rich data of acute-phase COVID-19 and search for potential long COVID risk loci, we performed an integrative analysis using COVID-19 susceptibility and hospitalization GWAS summary statistics as proxy phenotypes along with published long COVID GWAS from HGI. We first classified associated SNPs by their specificity/commonality to COVID-19 hospitalization or non-hospitalization and prioritized candidates SNPs for further evaluation of genes adjacent to them in published long COVID GWASs. Additionally, genes close to these candidate SNPs were evaluated in transcriptome-wide association studies (TWAS) of acute-phase COVID-19 and long COVID GWASs, as well as in phenome-wide association studies (PheWAS). Our objective was not to identify long COVID loci directly but to generate a set of candidate SNPs/genes that may contribute to both acute and chronic COVID-19 outcomes. This approach complements ongoing efforts in long COVID research and offers a framework for prioritizing long COVID loci to be validated in future studies.

Methods

Genome-wide Screening for SNPs Associated with COVID-19 Susceptibility, Hospitalization or Non-Hospitalization

To screen for unique/common genetic elements predisposing to COVID-19 susceptibility, hospitalization and non-hospitalization, COVID-19 GWAS datasets with large sample sizes were downloaded from the HGI (release 7), including HGI-C2 (COVID-19 [n=159,840] vs. the general population [n=2,782,977]), HGI-B1 (hospitalized [n=16,512] vs. non-hospitalized COVID-19 [n=71,321]), and HGI-B2 (hospitalized COVID-19 [n=44,986] vs. the general population [n=2,356,386]). The latter two GWASs, HGI-B1 and HGI-B2, were compared for effect sizes of each SNP genome-wide using differential z-score method, resulting in a new GWAS, called HGI-B1-vs-B2, that captures SNPs displaying differential effect sizes mainly because of the using of different controls by the two GWASs, i.e. HGI-B1 using non-hospitalized COVID-19 as controls but HGI-B2 employing general population as controls. HGI also conducted long COVID GWASs with 4 slightly different case-control designs that were published by Lammi et al[9] (see

supplementary Table S1), these long COVID GWASs were also downloaded but only used in evaluation of top hits derived from the comparison of the 4 acute-phase COVID-19 GWASs. Subsequently, by comparing these acute-phase COVID-19 GWASs, SNPs were separated into three categories, with the first categories labeled as "Severe COVID-19-Specific SNPs," which includes SNPs associated with both SARS-CoV-2 susceptibility and COVID-19 hospitalization but tends to be more prevalent within hospitalized than in non-hospitalized COVID-19, the second SNP category named as "SNPs Associated with Both Severe and Mild COVID-19," which influences both hospitalized and non-hospitalized COVID-19, and the last SNP category defined as "Mild COVID-19-Specific SNPs," which tends to be specifically associated with non-hospitalized COVID-19.

For the 1st SNP category, "Severe COVID-19-Specific SNPs," the following selection criteria were applied, including (a) $P \leq 0.05$ in HGI-C2, capturing SNPs linked to COVID-19 susceptibility; (b) $P > 0.05$ in the differential GWAS of HGI-B1 vs. HGI-B2, keeping SNPs without differential effect sizes between these comparisons; (c) $P \leq 0.05$ in HGI-B2, identifying SNPs distinguishing hospitalized COVID-19 individuals from the general population; (d) $P \leq 0.05$ in HGI-B1, identifying SNPs that predispose individuals to COVID-19 hospitalization rather than non-hospitalized COVID-19; (e) all SNPs required to display consistent effect size directions across HGI-C2 and HGI-B2. Only independent SNPs passing the $P < 1 \times 10^{-7}$ in any of the three GWASs (HGI-B1, HGI-B2, or HGI-C2) were selected as candidate SNPs.

In terms of the 2nd SNP category, "SNPs Associated with Both Severe and Mild COVID-19," representing SNPs not significant between hospitalized and non-hospitalized cases but exhibiting associations with SARS-CoV-2 susceptibility and COVID-19 hospitalization when compared to the general population. The selection criteria for this group were performed as follows: (a) $P \leq 0.05$ in HGI-C2, identifying COVID-19 susceptibility-related SNPs; (b) $P \leq 0.05$ in the differential GWAS of HGI-B1 vs. HGI-B2, capturing SNPs with distinct effects in the two GWASs; (b) $P \leq 0.05$ in HGI-B2, identifying SNPs associated with COVID-19 hospitalization; (c) $P > 0.05$ in HGI-B1, capturing SNPs with non-significant difference in effect sizes between severe and mild COVID-19; (e) all prioritized SNPs need to fulfill consistent effect size directions across HGI-C2 and HGI-B2. Similarly, only independent SNPs passing the $P < 1 \times 10^{-7}$ in any of these three GWASs (HGI-B1, HGI-B2, or HGI-C2) were selected as candidate SNPs.

The last SNP category is designated as "Mild COVID-19-Specific SNPs." The rationale is that current COVID-19 GWAS analyses from the HGI are predominantly biased toward acute-phase, hospitalized cases, potentially overlooking SNPs associated with mild COVID-19 that emerge specifically in non-hospitalized individuals who might later develop chronic COVID[7].

To address this gap, SNPs in the 3rd SNP category were obtained by using the following criteria: (a) $P \leq 0.05$ in the differential GWAS between HGI-B1 and HGI-B2, retaining SNPs with differential association signals between these datasets; (b) $P > 0.05$ in the GWAS comparing hospitalized cases with the general population (HGI-B2), which excludes SNPs associated with COVID-19 hospitalization; (c) $P \leq 0.05$ in the GWAS comparing hospitalized cases with non-hospitalized cases (HGI-B1), which captures SNPs significantly differing these two groups, obtaining SNPs specifically associated with mild COVID-19. To further refine the selection, independent SNPs with minor allele frequency (MAF) less than 0.01 in the largest cohort HGI-C2 and met a suggestive significance threshold ($P < 1 \times 10^{-5}$) were prioritized. These SNPs were also ensured that no other genome-wide significant SNPs to it within a 500-kb window in either HGI-B2 or HGI-C2. The use of a relaxed significance threshold and stricter MAF filter aimed to account for the higher heterogeneity among non-hospitalized cases in HGI-B1 and the inherent bias of HGI GWASs toward identifying SNPs associated with hospitalization. This approach enables the identification of SNPs that might represent unique genetic contributors to non-hospitalized COVID-19.

Finally, all these candidate SNPs obtained using the above criteria were subjected to evaluation among acute-phase COVID-19 GWASs of susceptibility and hospitalization from HGI, sex-stratification COVID-19 hospitalization GWASs from UK Biobank, and long COVID GWASs from HGI.

Comparing Two GWASs Using Differential Z-Score Method

To assess differences in effect sizes of SNPs between two GWASs, such as one GWAS comparing hospitalized cases with non-hospitalized cases (HGI-B1) and the other GWAS comparing hospitalized cases with the general population (HGI-B2), or two sex-stratified COVID-19 hospitalization GWASs, differential z-score method[10] was applied in current study. This approach quantifies the disparity in effect sizes (β) of each SNP between two GWASs by calculating a differential z-score, which incorporates adjustments for potential sample overlap between studies. The corresponding p-value was derived from the normalized differential z-scores. The differential z-score is defined as:

$$\Delta \mathrm{Z} - \mathrm{score} = \frac{\mathrm{gwas1.\beta} - \mathrm{gwas2.\beta}}{\sqrt{(\mathrm{gwas1.se^2} + \mathrm{gwas2.se^2})}}$$

$$\mathrm{diff.zscore} = \mathrm{Normalization}(\Delta \mathrm{Z} - \mathrm{score})$$

$$P = \mathrm{pnorm}(-|\mathrm{diff.zscore}|) * 2$$

Here, gwas1.β and gwas2.β represent the effect sizes for a given SNP from the two GWASs, while gwas1.se and gwas2.se denote the corresponding standard errors. To account for genome-wide correlation in effect sizes caused by overlapping samples, differential z-scores were normalized using the SAS procedure "proc stdize" with the "std" method. Subsequently, differential p-values were computed based on the normalized

differential z-scores. The cumulative distribution function of the standard normal distribution ("pnorm") was used to calculate p-values from the normalized scores.

Sex-Biased Associations with COVID-19 Hospitalization among Top Candidate SNPs

Epidemiological evidence indicates that women are more likely to experience long COVID but less likely to develop severe COVID-19[11,12], potentially due to stronger immune responses to SARS-CoV-2 infection in women[12]. To investigate whether these candidate SNPs exhibit sex-biased associations with COVID-19 hospitalization/non-hospitalization, we analyzed publicly available sex-stratified GWAS summary statistics from the GRASP COVID-19 database[13] generated based on UK Biobank samples. The selected sex-stratified COVID-19 hospitalization GWASs (cases=1,343 and controls=262,886 for females; cases=1,917 and controls=221,174 for males) includes mixed ancestry populations which are similar to the four HGI COVID-19 GWAS datasets used in our study. We used the aforementioned differential z-score method to compare male and female COVID-19 hospitalization GWAS results, and SNPs with differential effect sizes between sexes for COVID-19 hospitalization were identified. Genomic inflation factors for both sex-stratified GWASs and the sex-biased differential GWAS were also evaluated to ensure they don't have inflation.

Evaluate the Associations of Top Candidate SNPs and Its Adjacent SNPs with Long COVID

Four HGI long COVID GWASs, including LongCOVID-N1, LongCOVID-N2, LongCOVID-W1, and LongCOVID-W2 (see Table S1), with slightly different case and control designs were used to evaluate the association signals of these candidate SNPs. Due to the nature of phenotypic difference between acute-phase COVID-19 and long COVID, adjacent SNPs, located within a 1Mb-genomic window where each top candidate SNP residing in, were screened across these long COVID GWASs, and the one with the smallest long COVID association signals passed the significance threshold of $P<0.001$ were kept for further analyses, such as manual evaluation with local Manhattan plots and forest plots. These loci with nominal significant top SNPs or its adjacent SNPs passed the $P<0.001$ in any of these long COVID GWASs are treated as potential candidate loci involved in long COVID.

Evaluation of Previously Reported Long COVID SNPs in Current Acute-Phase COVID-19 GWASs

To assess whether these prioritized candidate SNPs identified in our study overlap with or are proximal to those reported in two independent long COVID GWASs by Taylor et al.[4] and Chaudhary et al.[5], a set of 83 SNPs derived from SNP network analyses and 3 SNPs identified via traditional GWAS methods specifically from the two studies were created and analyzed for their association signals among the three acute-phase GWASs from HGI.

204

205 **Transcriptome-Wide Association Analysis Using S-MetaXcan**

206 Transcriptome-wide association analysis (TWAS) was conducted using S-MetaXcan[14], which infers gene
207 expression association with a phenotype based on its corresponding GWAS summary statistics and GTEx
208 gene expression datasets across 48 tissues. Four COVID-19 GWAS datasets from the HGI were analyzed,
209 including HGI-B1, HGI-B2, HGI-C2, and HGI-B1-vs-B2, alongside four long COVID GWAS datasets
210 with subtly distinct case-control definitions from HGI[9] and an additional set of 222 GWASs from the UK
211 Biobank (UKB) (all sample sizes of cases >5000) shared by Benneal's lab (https://www.nealelab.is/uk-
212 biobank). Reference transcriptomic data from 48 tissues in GTEx v7[15] were utilized for gene expression
213 imputation in the analyses.

214

215 In detail, S-MetaXcan[14] employs multivariate regression to integrate evidence across different GTEx tissues
216 (n=48), enhancing the detection power. Default settings were applied, with the conditional number for
217 eigenvalue set at 30 during singular value decomposition (SVD) truncation and genes on autosomal
218 chromosomes were included in the analysis. First, TWAS results of COVID GWAS datasets (n=8,
219 collectively referred to as COVID TWASs, covering 4 long COVID TWASs as well as 4 acute-phase
220 COVID-19 TWASs) were evaluated for genes proximal to prioritized SNPs (±500kb to each SNP termed
221 as tag SNP). Genes showing nominal significance (P<0.05) in any of these COVID TWASs across the 48
222 GTEx tissues were prioritized for further investigation.

223

224 Also, to expand these findings, phenome-wide association analyses were conducted for these prioritized
225 genes. Associations were examined across a TWAS database aggregating 190 previously published TWAS
226 results covering a broad set of phenotypes[14] and 222 UKB TWASs created locally with the same settings
227 as that of the 190 published TWASs, as well as 8 COVID related TWASs generated in this study, yielding
228 a total of 420 TWAS datasets (see Table S2).

229

230 **Statistical Analysis Software and Availability of Codes and Data**

231 Unless otherwise specified, all analyses were performed using SAS OnDemand for Academics (freely
232 available at https://www.sas.com/en_us/software/on-demand-for-academics.html). Local Manhattan plots
233 for the prioritized candidate SNPs are publicly accessible at the GitHub link
234 (https://github.com/chengzhongshan/COVID19_GWAS_Analyzer/tree/main/LongCOVID_data_and_scri
235 pts/LongCOVID_SNP_LocalManhattanPlots). SAS scripts involved in post-GWAS analyses for these
236 acute-phase COVID-19 and long COVID GWASs are put into the same github folder. Acute-phase COVID-
237 19 and long COVID GWASs used in our study are accessible via Google Driver

238 (THREEGWASs4HGICOVID19.gz:
239 https://drive.google.com/file/d/1nZZsxbXJROjGmBGqq84WbCtPLlZWDT3C/view?usp=sharing;
240 CombineLongCOVIDGWAS.gz:
241 https://drive.google.com/file/d/1opd3BSC5JAHoPvDZPmqPB11Ox8YTrG-X/view?usp=sharing). A
242 readme document explaining details of the two compressed datasets is accessible here
243 (https://docs.google.com/document/d/1EVA-gUZyGeiyb4GPDuliUyT044WBya_I/edit).
244

245 **Results**

246 **Comparison of Three Acute-Phase COVID-19 GWASs from HGI**

247 To search for candidate genetic SNPs commonly or specifically associated with COVID-19 hospitalization
248 and non-hospitalization, an integrative workflow was developed by leveraging publicly available COVID-
249 19 GWAS (Fig. 1). Initially, acute COVID-19 phenotypes, such as SARS-CoV-2 susceptibility, COVID-19
250 hospitalization, and non-hospitalization, were used to prioritize genetic variants potentially, specifically
251 linked to COVID-19 hospitalization or non-hospitalization, as well as associated with both of the two
252 COVID-19 phenotypes. The analysis utilized GWAS data from the HGI release 7, including GWASs of
253 hospitalized vs. non-hospitalized cases (HGI-B1), hospitalized cases vs. the general population (HGI-B2),
254 and SARS-CoV-2 infection vs. the general population (HGI-C2). These GWASs comprise large sample
255 sizes (16,512, 44,986, and 159,840 COVID-19 cases for HGI-B1, HGI-B2, and HGI-C2, respectively; Fig.
256 1) and exhibit minimal genomic inflation (all λs close to 1), as shown in supplementary Fig. S1, and two
257 loci, including *SLC6A20* and *ABO*, displayed notable genome-wide significance, which were well-
258 established contributors to COVID-19 susceptibility and severity[16].

259 **Candidate SNPs Specifically Associated with Hospitalized COVID-19 or Commonly Associated with**
260 **Both Hospitalized and Non-hospitalized COVID-19**

261 Base on the filtering strategies illustrated in the workflow shown in Fig. 1, two SNP groups were prioritized
262 first, including "Severe COVID-19-Specific SNPs" (n=27) and "SNPs Associated with Both Severe and
263 Mild COVID-19" (n=15), resulting in 42 independent SNPs (Table 1 and Fig. 2). It is necessary to
264 emphasize that these two SNP categories are designated to capture SNPs either linked predominantly to
265 SARS-CoV-2 susceptibility and COVID-19 hospitalization coupled with higher prevalence in hospitalized
266 cases than in non-hospitalized cases, or associated with both severe and mild COVID-19 that demonstrate
267 consistent associations with SARS-CoV-2 susceptibility and hospitalization along with no frequency
268 difference between hospitalized and non-hospitalized cases. Among these 42 independent SNPs included
269 in the two SNP categories, all SNPs passed the conventional genome-wide significance threshold ($P<5\times10^{-}$
270 ^{8}) in at least one HGI COVID-19 GWAS, with the exception of two SNPs, rs150345524 (*SF3B1*, $P=5.6\times10^{-}$

^{8} in HGI-B2) and rs11766643 (*SLC22A5*, P=7.5x10^{-8} in HGI-C2), that are suggestively genome-wide significant.

To reveal biological insights, each of the above 42 COVID-19-associated SNPs was linked with its corresponding nearest gene to create a SNP-gene pair. It is necessary to note that proximity-based assignments may overlook distal regulatory effects, and this potential issue will be further addressed in subsequent transcriptome-wide association studies. In detail, notable SNPs along with its corresponding nearest gene in the 1st SNP category, "Severe COVID-19-Specific SNPs," include rs10774679 (*OAS3*), rs17169628 (*SLC22A31*), rs12329760 (*TMPRSS2*), rs12660421 (*FOXP4*), rs17078348 (*SLC6A20*), rs190509934 (*ACE2*), rs2834164 (*IFNAR2*), rs3785884 (*MAPT*), rs7251000 (*DPP9*), and rs76608815 (*SLC5A3*), among others (Table 1), and this SNP category contains a total of 27 SNPs showing stronger associations with severe COVID-19 than with mild cases (Fig. 2 and Fig. S2). Meanwhile, for the 2nd SNP category, "SNPs Associated with Both Severe and Mild COVID-19", five out 15 SNP-gene pairs, including rs2290859 (*NXPE3*), rs505922 (*ABO*), rs149533170 (*IFNA21*), rs34536443 (*TYK2*), and rs9916158 (*GSDMA/B*), demonstrates consistent associations across GWAS datasets without differential effects between hospitalized and non-hospitalized groups (Fig. 2 and Fig. S3). Notably, *ABO* (rs505922) has been reported as a genetic locus associated with long COVID[5]. Several other immune-related genes were found close to other candidate SNPs, including *IFNA21* (rs149533170), *TYK2* (rs34536443), *GSDMA/B* (rs9916158), *BCL11A* (rs1123573), *ELF5* (rs2924480), and *JAK1* (rs11208552), in accordance with the role of immune dysregulation in both hospitalized and non-hospitalized COVID-19[17-22]. Additionally, *FUT2* is found closely with rs492602 and encodes fucosyltransferase 2, a key enzyme in the ABO antigen synthesis pathway, further underscores the involvement of *ABO* in COVID-19 outcomes[23]. Additionally, four genes—*TAC4*[24], *NXPE3*[25], *NSF*[26], and *FOXG1*[27]—linked to neuronal functions were close to rs75432325, rs2290859, rs1378358, and rs150497803, respectively, while three genes—*ZKSCAN1*[28], *BCL11A*, and *FOXG1*—identified as transcription factors were linked to rs2897075, rs1123573, and rs150497803, respectively.

We also evaluated sex-biased associations for these 42 candidate SNPs in the first 2 SNP categories (Fig. S4). Due to the smaller sample sizes in the sex-stratified COVID-19 hospitalization GWASs from UK Biobank, only 15 SNPs reached nominal significance (P<0.05) in at least one of these three GWASs, including COVID-19 hospitalization for males and females, and differential COVID-19 hospitalization between males and females. Of these, rs17078348 (*SLC6A20*), a well-established severe COVID-19 risk SNP[29], showed nominal significance in both male and female cohorts of COVID-19 hospitalization. Five SNPs displayed nominal significance exclusively in females, including rs75432325 (*TAC4*), rs505922 (*ABO*), rs12660421 (*FOXP4*), rs7251000 (*DPP9*), and rs11766643 (*TFR2*). Eight SNPs were nominally

304 significant only in males, including rs897075 (*ZKSCAN1*), rs117169628 (*SLC22A31*), rs7515509 (*AK5*),
305 rs76608815 (*SLC5A3*), rs12585036 (*ATP11A*), rs5023077 (*FBRSL1*), rs10774679 (*IFNAR21*), and
306 rs12329760 (*HLA-C*). None of the SNPs passed the stringent multiple testing correction threshold ($P<6x10^{-}$
307 ^{4}), and local Manhattan plots for 12 of the 42 SNPs revealed relatively weaker association signals in the
308 sex-stratified GWAS in contrast with the stronger signals observed in HGI-B2 (Fig. S5). This reduction in
309 signal strength likely reflects the smaller sample sizes in the sex-stratified datasets, implicating reduced
310 statistical power to detect sex-biased associations. In summary, while the sex-biased GWAS analysis did
311 not identify SNPs with statistically significant sex-biased associations with COVID-19 hospitalization, it
312 highlighted 15 SNPs with nominal significance, including rs505922 (*ABO*) and rs12660421 (*FOXP4*),
313 which have been previously implicated as risk variants for both severe COVID-19 and long COVID. These
314 findings warrant further investigation in larger sex-stratified COVID-19 GWAS as well as long COVID
315 GWAS.

316 **Candidate SNPs Specifically Associated with Non-Hospitalized COVID-19**

317 To identify potential candidate SNPs specific to non-hospitalized COVID-19 cases, we screened for SNPs
318 that showed suggestive significant association signals in HGI-B1 (hospitalized vs. non-hospitalized, $P<$
319 $1x10^{-5}$) and HGI-B1-vs-B2 (differential GWAS between HGI-B1 and HGI-B2, $P<0.05$) but were non-
320 significant in HGI-B2 (hospitalized vs. general population, $P>0.05$) and were not in proximity to any
321 genome-wide significant SNPs in HGI-B2. We also only focused on top hits with MAF<0.01 in the HGI-
322 C2 with the largest sample size compared to other COVID-19 GWASs from HGI. This selection prioritized
323 20 independent SNPs potentially specifically associated with mild (non-hospitalized) COVID-19 cases
324 (Table 2; Fig. S6). All 20 SNPs showed suggestive genome-wide significant differences in effect sizes
325 between HGI-B1 and HGI-B2 ($P<1x10^{-5}$). Although none met the genome-wide significance threshold
326 ($P<5 x10^{-8}$), notable findings were observed by linking these candidate SNPs with their adjacent genes. For
327 example, rs62401842 (*KCTD16*) exhibited the strongest association in HGI-B1 ($P=8.53x10^{-7}$), while
328 showing no association in HGI-B2 ($P=0.38$), resulting in a significant difference between HGI-B1 and
329 HGI-B2 ($P=3.18x10^{-7}$). Gene *KCTD16*, encoding potassium channel tetramerization domain-containing
330 16[30], is highly expressed in brain tissues according to the GTEx database[15]. In addition, eight SNPs
331 illustrated in Fig. S6 were mapped to genes with brain-related functions, including rs9799354 (*NLGN1*[4]),
332 rs112842080 (*CPLX2*[31]), rs61858037 (*NRG3*[32]), rs61939166 (*KIF21A*[33]), rs367777 (*NAV3*[34]), rs56143829
333 (*WASF3*[35]), rs11454577 (*AKAP6*[36]), and rs6049828 (*SYNDIG1*[37]). Among these, *WASF3* has been
334 implicated in mitochondrial dysfunction and may mediate exercise intolerance in myalgic
335 encephalomyelitis/chronic fatigue syndrome[35]. Additionally, among these 20 candidate SNPs identified,
336 only three SNPs (rs62401842 [*KCTD16*], rs9799354 [*NLGN1*], and rs112842080 [*CPLX2*]) exhibit male-

biased associations with COVID-19 hospitalization at nominal significance levels (Fig. S7). Furthermore, six candidate SNPs, including rs62401842 (*KCTD16*), rs9799354 (*NLGN1*), rs112842080 (*CPLX2*), rs56143829 (*WASF3*), rs4737438 (*PENK*), and rs6049828 (*SYNDIG1*), demonstrate sex-biased association patterns within a genomic window ranging from 500kb to 1000kb (Fig. S7). However, the limited sample sizes in sex-stratified COVID-19 GWAS preclude any candidate SNPs from achieving statistical significance after correction for multiple testing in analyses of sex-biased associations with COVID-19 hospitalization. In summary, while no genome-wide significant SNPs specific to non-hospitalized COVID-19 cases were identified, these 20 SNPs along with their adjacent genes provide suggestive evidence of association with mild COVID-19.

Association of Prioritized Candidates SNPs with Long COVID

To investigate the associations of these 62 candidates SNPs with long COVID, the published long COVID GWASs from HGI (see Table S1) were used to examine association signals of them. Notably, only one candidate SNP, rs12660421 of *FOXP4*, passed the traditional genome-wide significance threshold and significantly associated with long COVID (P=1x10^{-9}) in one of four long COVID GWASs, i.e. LongCOVID-N2 GWAS (strict cases of long COVID after test-verified SARS-CoV-2 infection [n = 3,018] vs. general population controls [n = 994,582)) (Table S3, Fig. S8-S10). Among other 3 alternative long COVID GWASs with slightly different case-control designs for long COVID, including LongCOVID-W2, LongCOVID-N1, and LongCOVID-W1, rs12660421 (*FOXP4*) is also normally significantly associated with long COVID, with p-values of which equal to 9.7x10^{-7}, 4.3x10^{-3}, and 4.95x10^{-2}, respectively. However, rs12660421 (*FOXP4*) tends to be more strongly associated with COVID susceptibility (P=7.4x10^{-15}) and hospitalization (P=2.9x10^{-10}), supporting the SNP's association with multiple COVID-19 and long COVID phenotypes. For another 61 remaining candidate SNPs, there are 8 SNPs, comprising rs505922 (*ABO*), rs7660881 (*SLC5A3*), rs6186040 (*SFTPD*), rs2834164 (*IFNAR2*), rs1077467 (*OAS3*), rs2290859 (*NXPE3*), rs1707834 (*SLC6A20*) and rs1128420 (*CPLX2*), showing normally significance in at least one long COVID GWAS, all of which displaying much weaker association signals compared to that from acute-phase COVID-19 GWASs (Table 1 and 2, Table S3; Fig. S8-S10). Additionally, 43 out of 62 candidate SNPs, including loci *ELF5* (rs141160516), *TAC4* (rs28704677), *FUT2* (rs557806), *RAB2A* (rs113059174), *ABO* (rs544873), *GSDMA/B* (rs72832927), *HLA*-C (rs3218825), *TYK2* (rs8112028), *OAS3* (rs72487538), *KCTD16* (rs114210747), *WASF3* (rs75947067), *ATP11A* (rs3832907), *PCDH20* (rs34409482), *SLC27A6* (rs704234), and others (see Table S4), have at least one adjacent SNP displaying long COVID association signals with P<0.001, indicating that these loci are potentially associated with long COVID due to different genetic variants at the same locus (Table S4 and Fig. S10). The observation of much more weaker associations of these SNPs as well as its adjacent SNPs with long COVID is in line with the winners' curse,

since these candidate SNPs were initially selected from COVID-19 GWASs by its strong association signals that are likely to be overestimated; meanwhile, it is also possible that the current long COVID GWASs are biased to risk factors associated with severe COVID-19, mainly due to the overlapping samples used by both long COVID cohorts and COVID-19 cohorts, leading to the observation of relatively weaker association signals for these candidate SNPs in long COVID.

Evaluating Other Published Long COVID SNPs in Current COVID-19 GWASs

Previously published long COVID risk SNPs reported by Taylor et al.[4] and Chaudhary et al.[5]. were investigated among three acute-phase COVID-19 GWASs of HGI (Table S5). The rationale is based on the hypothesis that larger sample sizes and higher statistical power of the HGI datasets have enough power to detect associations with traits closely related to long COVID, such as severe COVID-19 and susceptibility to SARS-CoV-2 infection. Moreover, overlap between our prioritized long COVID SNPs and those reported in previous studies would strengthen the hypothesis that the loci represented by these 62 candidate SNPs in question might be implicated in long COVID pathogenesis.

Among the three SNPs (rs643434, rs2080090, and rs9273363) reported by Chaudhary et al.[5] (Fig. S11 and Table S5), rs643434 (*ABO*) and rs2080090 (*BPTF*) demonstrated significant associations with both COVID-19 hospitalization (HGI-B2) and susceptibility to COVID-19 (HGI-C2). In detail, rs643434 (*ABO*) surpassed genome-wide significance thresholds in both HGI-B2 (P= 6.7x10^{-24}) and HGI-C2 (P=8.7x10^{-85}) and showed non-significant association in HGI-B1 (P=0.6) but demonstrated significantly differential effect sizes (P=6x10^{-5}) between HGI-B1 (hospitalized vs. non-hospitalized COVID-19) and HGI-B2. This supported its classification as a "category 2" COVID-19 SNP (associated with both severe and mild COVID-19 phenotypes). Similarly, rs2080090 (*BPTF*) showed stronger associations in HGI-B2 (P=1.9x10^{-4}) than in HGI-C2 (P=2.5x10^{-2}), with no significant association in HGI-B1. This SNP was also categorized as a suggestive "category 2" COVID-19 SNP. In contrast, rs9273363 (*HLA-DQB1*) did not exhibit significant associations across any of the HGI GWAS datasets analyzed (all Ps>0.2). Nevertheless, among these 62 candidate SNPs, rs17219281 (*HLA-DQA1*), a “category 1” COVID-19 SNP that tends to be more likely associated with hospitalized rather than mild COVID-19 (see Table 1), is ~50kb downstream of rs9273363 (*HLA-DQB1*), strongly implicating that rs17219281 may be the real SNP associated with COVID-19 hospitalization or long COVID. Taken together, the three recently published long COVID loci, including *ABO*, *BPTF*, and *HLA-DQB1*, were completely covered by these 62 prioritized SNPs and display association with either COVID-19 hospitalization or both COVID-19 hospitalization and non-hospitalization.

Meanwhile, of the 80 SNPs reported by Taylor et al.[4], 12 displayed nominally significant associations in at

402 least one HGI GWAS dataset (Table S5 and Fig. S11). These included four SNPs associated with HGI-B2,
403 five with HGI-B1, and seven with HGI-C2. Notably, rs4303401 (*FRMD6*) surpassed the multiple-testing
404 correction threshold ($P<1.5x10^{-4}$), representing a strong candidate SNP. Expanding the analysis to include
405 a ±500kb genomic window around the reported SNPs identified additional suggestive genome-wide
406 significant associations in the HGI datasets. For example, in Fig. S11, rs62401842 (proximal to *KCTD16*
407 that shows brain-specific expression) and rs7671107 and rs1017716 (a GTEx eQTL of *ANAPC4*, which is
408 a gene implicated in aging) emerged as promising candidates. Other noteworthy SNPs in Fig. S11 include
409 rs12602210 (*BPTF*) and rs17412601 and rs2290859 (both mapped to *NXPE3*), which demonstrate
410 suggestive or genome-wide significance in at least one HGI GWAS dataset. In summary, four loci, including
411 *KCTD16*, *ANAPC4*, *BPTF*, and *NXPE3*, revealed by Taylor et al[4]. derived from SNP network analysis are
412 independently confirmed to be nominally associated with COVID-19 related phenotypes in the current
413 study; among them *BPTF* is a locus confirmed to be associated with long COVID by Chaudhary et al.,
414 Taylor et al., and associated with acute-phase COVID-19 in current study. Therefore, these findings
415 underscore the potential of integrating data from diverse GWAS to elucidate the genetic architecture of
416 acute-phase COVID-19 and long COVID to highlight candidate loci for further investigation.

417 **Transcriptome-Wide Association Study Identifies Candidate Genes Adjacent to Prioritized SNPs for**
418 **COVID-19 and Long COVID Phenotypes**

419 To elucidate genetic mechanisms underlying these COVID-19 candidate SNPs and potential causal genes
420 represented by these SNPs, we performed a transcriptome-wide association study (TWAS) using S-
421 MetaXcan across 48 GTEx tissues by correlating genotype-imputed gene expression with acute-phase
422 COVID-19 and long COVID. This approach first utilized COVID-19 GWAS summary statistics from HGI:
423 HGI-B1, HGI-B2, HGI-C2, and HGI-B1-vs-B2. The analysis targeted 62 candidate SNPs previously
424 prioritized and aimed to identify their adjacent genes whose expression levels may be associated with
425 COVID-19 hospitalization or susceptibility phenotypes, and then 4 long COVID GWASs with slightly
426 different definitions in terms of cases and controls were performed for TWAS with the same method.

427 We first evaluated whether these candidate SNPs and its adjacent genes were covered by TWASs of acute-
428 phase COVID-19. Of the 62 candidate SNPs, two SNPs (rs190509934 near *ACE2* and rs9570861 near
429 *OR7E156P*) were not included in the TWAS analysis. The former resides on chromosome X, whom S-
430 MetaXcan excludes due to the focus of S-MetaXcan is only on autosomal genes, while the latter SNP has
431 no protein-coding genes within a 500-kb window. Although there are 879 adjacent genes to these 62
432 candidate SNPs, only 776 out of them are corresponding to the remaining 60 SNPs that were tested by these
433 COVID TWASs (Fig. S12) with detectable expression across the 48 GTEx tissues.

434 Next, to simplify the interpretation for TWASs, we retained only the strongest association (smallest p-value)
435 for each gene across 48 GTEx tissues. This approach was justified by the hypothesis that specific tissues,
436 rather than all tissues, are likely to drive the pathogenicity of COVID-19. Applying varying significance
437 thresholds (P<0.05, 0.01, 0.001, 0.0001, 0.00001), the analysis revealed 422, 152, 43, 19, and 11 adjacent
438 genes meeting these criteria in at least one of the four HGI datasets (Fig. S12, Table S6). Using a genome-
439 wide significance threshold ($P<3.5\times10^{-7}$), we found that top hits are originated exclusively from COVID-
440 19 hospitalization and susceptibility GWAS, with no genes passed this threshold in long COVID TWASs
441 (Fig. 3). The top three loci displayed significant associations are identified as: (1) *DPP9*, showing the
442 strongest association signal ($P=4.6\times10^{-18}$) in adipose visceral omentum tissue; (2) *CXCR6* gene cluster,
443 which included multiple genes (*CXCR6*, *CCR3*, *CCR1*, and *LZTFL1*), displaying the most significant
444 association ($P=1.3\times10^{-10}$) in nerve tibial tissue; (3) *MUC1* gene cluster, comprising *MUC1*, *THBS3* and
445 *FAM189B*, highlighted in multiple GTEx tissues, including lung and colon. Additional loci with suggestive
446 TWAS significance ($P<3.5\times10^{-6}$) included *FYCO1*, *SLC6A20*, and *XCR1*, specifically showing associations
447 in HGI GWASs B1, B2, and C2 in distinct GTEx tissues.

448 Further determination revealed that these candidate genes proximal to 60 prioritized SNPs accounted for
449 all top TWAS signals across the eight HGI TWASs, with the exception of *CCR5*, *EXOSC7*, *FLT1P1*, *RLN1*,
450 *GBAP1*, *ARHGEF2*, *ZSCAN31*, *ZKSCAN3*, *PGMSP2*, *PAQR5*, and *TMIE* (Fig. S13A). Further in-depth
451 evaluation of these outlier genes in long COVID TWASs revealed most of them lack of significance (Fig.
452 S13B and S14), reinforcing the utility of prioritizing top candidate SNPs and their adjacent genes that might
453 be associated with long COVID.

454 We then focused on the associations of these adjacent genes in long COVID TWASs across all 48 GTEx
455 tissues using a relatively weak association threshold, i.e., the nominal significance threshold of P<0.05,
456 among these top COVID-19 TWAS genes (Fig. 3, Table S7 and Fig. S14). In details, *THBS3* is the gene
457 showing ubiquitous association with long COVID among multiple GTEx tissues, such as brain cerebellum,
458 nerve tibial, speen, adrenal gland, colon sigmoid, testis and esophagus mucosa; *CCR3* is observed to be
459 correlated with long COVID in the GTEx tissue heart atrial appendage; *MUC1* is predicted to be associated
460 with long COVID in the GTEx tissue colon transverse, and *DPP9* expression associates with long COVID
461 in both GTEx tissues colon transverse and testis. Furthermore, using a relaxed threshold (P<0.05), we
462 identified additional adjacent genes associated with long COVID (Fig. 4 & Fig. S15-16). These included
463 *OAS3* (rs10774679), *TMPRSS2* (rs12329760), *HLA-A* (rs4s9260038), *HLA-C* (rs1634761), *HLA-DQA1*
464 (rs17219281), *WASF3* (rs65143829), *IFNAR2* (rs2834164), *SFTPD* (rs61860402), and *GSDMA/B*
465 (rs9916158), the associations of which enriched in tissues such as heart, brain, colon, adipose, and muscle.
466 Notably, the well-established *ABO* locus, a COVID-19 susceptibility locus, did not show significant TWAS

467 signals, suggesting that its effects may stem from non-expression-related mechanisms, such as genetic
468 mutations influencing blood types[38].

469 In summary, TWAS analyses identified 52 SNP-gene pairs nominally associated with long COVID across
470 tissues. While genes such as *DPP9, CCR1*, *THBS3*, and *MUC1* emerged as potential shared contributors to
471 acute and long COVID phenotypes, their tissue-specific associations suggest distinct mechanisms may
472 drive long COVID. These findings highlight the importance of integrative approaches in uncovering the
473 molecular underpinnings of COVID-19 susceptibility/hospitalization/non-hospitalization and long COVID.

474 **Phenome-Wide Analysis of Gene Expression Associations with Candidate Genes**

475 To investigate gene expression associations with acute-phase COVID-19 and long COVID-related
476 phenotypes, we performed a phenome-wide analysis of gene expression associations (Pheno-TWAS)
477 analysis targeting adjacent genes located within a 1Mb-window of 62 candidate SNPs previously linked to
478 COVID-19. Of the 918 tested TWAS genes, 400 genes (43.6%), including 87 associated with *HLA* loci (e.g.,
479 *HLA-DQA2*, *HLA-A*, and *HLA-C*; Fig. S17-19), passed a stringent Bonferroni-adjusted significance
480 threshold ($P<2.7\text{x}10^{-9}$). Genes located near *HLA* loci emerged as top hits, demonstrating significant
481 associations with immune-related disorders such as asthma, rheumatoid arthritis, psoriasis, coeliac disease,
482 and systemic lupus erythematosus—conditions that may contribute to acute COVID-19 or long COVID. In
483 addition to *HLA*-associated loci, other 313 genes across 44 non-*HLA* loci surpassed the genome-wide
484 significance threshold in TWASs. Key findings include: (1) *IFNAR2*: strongly associated with COVID-19
485 hospitalization ($P=2.6\text{x}10^{-10}$; Fig. S20); (2) *GSDMB*: linked to asthma ($P=2.7\text{x}10^{-57}$; Fig. S21); (3) *BPTF*:
486 associated with trunk fat percentage ($P=1\text{x}10^{-18}$; Fig. S22); (4) *CCDC171*: associated with body fat
487 percentage ($P=2\text{x}10^{-35}$; Fig. S23). Interestingly, *FOXP4* expression, which did not exhibit significant
488 associations with long COVID phenotypes in prior TWASs, demonstrated a suggestive association with
489 sleep duration ($P=1.1\text{x}10^{-5}$; Fig. S24). This finding highlights the potential for phenome-wide approaches
490 to uncover previously uncharacterized associations relevant to acute COVID-19 or long COVID. In total,
491 Pheno-TWAS identified 45 candidate SNPs associated with genome-wide significant expression-linked
492 phenotypes. These include three major *HLA* loci (*HLA-DQA2*, *HLA-A*, and *HLA-C*) as well as non-*HLA*
493 loci such as *IFNAR2*, *BPTF*, *CCDC171*, and *GSDMB*. The implicated phenotypes—ranging from COVID-
494 19 hospitalization and immune disorders to body composition traits and sleep duration—provide a broad
495 spectrum of potential biological pathways linking to acute COVID-19 or long COVID pathophysiology.
496 Full results, including nominally significant associations, are provided in Table S8.

497 **Discussion**

498 Long COVID, characterized by over 200 symptoms spanning multiple organ systems, presents significant

challenges for GWAS due to difficulties in assembling large, representative cohorts. To date, only four genome-wide significant loci—*FOXP4*, *ABO*, *HLA-DQA1*, and *BPTF*—have been identified for long COVID[5,9]. Our integrative analysis identifies 62 candidate loci associated with acute-phase COVID-19 phenotypes, such as COVID-19 susceptibility, hospitalization, and non-hospitalization. These candidate SNPs were obtained by leveraging large-scale COVID-19 GWAS datasets from the HGI. This approach provide potential candidates for validation in future long COVID studies, as currently published long COVID GWASs have the inherent limitations of limited sample sizes and phenotype heterogeneity, and SARS-CoV-2 susceptibility and COVID-19 hospitalization can be used as proxy phenotypes to study long COVID in the future. By categorizing genetic variants into three groups—those associated with hospitalized COVID-19, those linked to both hospitalized and non-hospitalized cases, and those specific to non-hospitalized cases—our findings reveal new insights about potential SNPs that might be associated with long COVID.

There are 42 SNPs prioritized to be associated with either COVID-19 hospitalization or both COVID-19 hospitalization and non-hospitalizaiton, including well-documented variants such as rs12660421 (*FOXP4*) and rs505922 (*ABO*), which have been implicated in severe COVID-19 outcomes and long COVID susceptibility. Notably, additional loci like rs17219281 (*HLA-DQA1*), rs9260038 (*HLA-A*), rs1634761 (*HLA-C*), rs7251000 (*DPP9*), rs9916158 (*GSDMB*), rs12602210 (*BPTF*), rs17078348 (*CCR1*), and rs41264915 (*THBS3*), rs190509934 (*ACE2*) and rs12329760 (*TMPRSS2*) were identified. Furthermore, a unique category comprising 20 SNPs emerged specifically in non-hospitalized COVID-19 cases. These include rs62401842 (*KCTD16*) and rs56143829 (*WASF3*), mapped to genes enriched in brain-related functions, strongly suggesting their potential involvement in long COVID. However, future investigations are warranted for these candidate SNP-gene pairs in long COVID.

While this approach provides critical insights, certain limitations warrant consideration. First, the reliance on proxy phenotypes may obscure SNPs specific to mild COVID-19 cases, given the skewed representation of hospitalized patients in HGI datasets. Indeed, non-hospitalized-specific SNPs did not reach genome-wide significance, likely reflecting this bias. Nevertheless, 20 suggestive SNPs, enriched near genes implicated in neurocognitive and age-related pathways, merit further validation in independent cohorts. Second, the TWAS conducted using S-MetaXcan are constrained by the limited tissue specificity of available GTEx datasets and COVID related GWASs. Despite identifying 7 loci—*HLA-DQA1*, *HLA-A*, *HLA-C*, *DPP9*, *CCR1*, *THBS3*, *GSDMB*—that emerged as promising TWAS hits for long COVID, the potential absence of COVID-19-specific tissues, such as nasopharyngeal or immune-relevant sites, may hinder the detection of tissue-specific effects pertinent to acute-phase COVID-19 and long COVID. Finally, while our approach

offers new insights into long COVID, we acknowledge that proxy-based inference carries inherent uncertainty. The overlaps observed with long COVID loci are suggestive rather than confirmatory. Nonetheless, the concordance of TWAS-implicated genes and Pheno-TWAS-trait enrichment supports the biological plausibility association between acute-phase COVID-19 and long COVID. Future efforts should validate these findings in larger, dedicated long COVID cohorts with more refined phenotypes. Until then, our results serve as a resource for prioritizing loci for further investigation in post-acute COVID research.

In conclusion, our findings underscore the utility of integrative approaches to prioritize genetic risk loci for future investigation of long COVID. The discovery of loci such as *FOXP*4, *ABO*, and *HLA-DQA1*, as well as *BPTF*, suggests potentially shared biological pathways between acute COVID-19 severity and its chronic manifestations, such as long COVID. Additionally, the identification of brain-enriched loci in non-hospitalized cases aligns with emerging evidence linking long COVID to neuroinflammatory and neurodegenerative processes. Future studies, particularly those incorporating diverse populations and long-term follow-up data, are essential to refine these associations and unravel the intricate mechanisms underlying long COVID.

Ethics approval and consent to participate

All data used in this study were freely shared by other researchers. Since no patient-identifiable information was involved, ethical approval and informed consent were not applicable.

Consent for publication

This study is a secondary analysis of publicly available COVID-19 and long COVID-related GWAS data. No individual-identifiable information, such as personal identifiers or images, were used; therefore, consent for publication is not applicable.

Availability of data and material

All data are freely accessible upon request by contacting the corresponding author, Z.C., at cheng.zhong.shan@gmail.com.

Declaration of interests

The authors declare no competing interests.

Funding

No funding is available for supporting current study.

Authors' contributions

Z.C. conceived the study, analyzed the data, and wrote the manuscript.

Acknowledgements

I am thankful to all researchers who participated and contributed to the HGI and GRASP project, as their generosities in freely sharing these COVID-19 hospitalization GWASs and other COVID-19/long COVID related GWASs are essential for the current investigation.

571 **TABLES**

572 **Table 1. Candidate SNPs likely associated with either COVID-19 hospitalization or both COVID-19**
573 **hospitalization and non-hospitalization**

Variant information						HGI-B1			HGI-B2			HGI-C2			HGI-B1-vs-B2	COVID-19
chr	pos	rsid	allele	AF	gene	beta	se	p	beta	se	p	beta	se	p	DiffPval	SNP groups
1	47275770	rs10890422	C	0.402	*CYP4B1*	-0.05	0.02	3.28E-03	-0.05	0.01	4.43E-08	-0.02	0.00	1.10E-04	6.63E-01	Hosp
1	65412830	rs11208552	T	0.647	*JAK1*	-0.01	0.02	6.95E-01	-0.05	0.01	2.19E-09	-0.02	0.00	4.21E-04	5.14E-03	Both
1	77949123	rs7515509	A	0.366	*AK5*	0.08	0.02	5.71E-06	0.06	0.01	2.94E-12	0.02	0.00	2.16E-05	1.82E-01	Hosp
1	155167786	rs41264915	G	0.103	*THBS3*	-0.10	0.03	3.59E-04	-0.13	0.01	9.33E-23	-0.02	0.01	1.13E-03	2.64E-01	Hosp
2	26833309	rs1275969	T	0.560	*CIB4*	0.04	0.02	2.37E-02	0.05	0.01	2.03E-08	0.01	0.00	9.18E-03	8.15E-01	Hosp
2	60707588	rs1123573	G	0.367	*BCL11A*	-0.03	0.02	1.86E-01	-0.07	0.01	1.32E-14	-0.03	0.00	4.62E-09	1.84E-02	Both
2	198245297	rs150345524	A	0.413	*SF3B1*	0.05	0.02	1.69E-02	0.05	0.01	5.56E-08	0.01	0.00	4.66E-02	9.52E-01	Hosp
3	45847241	rs17078348	G	0.101	*SLC6A20*	0.29	0.03	1.15E-19	0.34	0.01	2.53E-130	0.08	0.01	2.01E-36	6.34E-02	Hosp
3	101525625	rs2290859	T	0.341	*NXPE3*	-0.03	0.02	1.55E-01	-0.07	0.01	2.34E-13	-0.05	0.00	4.43E-31	2.02E-02	Both
4	25435159	rs7684660	C	0.829	*ANAPC4*	-0.06	0.03	1.98E-02	-0.08	0.01	8.71E-12	-0.02	0.01	2.86E-05	4.57E-01	Hosp
5	131776967	rs10066378	C	0.131	*SLC22A5*	0.06	0.02	1.80E-02	0.07	0.01	1.63E-10	0.02	0.01	1.00E-03	4.94E-01	Hosp
6	29907544	rs9260038	G	0.622	*HLA-A*	-0.04	0.02	2.72E-02	-0.06	0.01	1.94E-10	-0.01	0.00	1.30E-02	3.40E-01	Hosp
6	31274027	rs1634761	T	0.500	*HLA-C*	-0.04	0.02	1.81E-02	-0.06	0.01	4.50E-12	-0.02	0.00	9.99E-07	2.95E-01	Hosp
6	32675645	rs17219281	A	0.072	*HLA-DQA1*	-0.10	0.04	1.01E-02	-0.12	0.02	5.53E-11	-0.03	0.01	1.81E-03	5.30E-01	Hosp
6	41488378	rs12660421	A	0.041	*FOXP4*	0.26	0.04	2.88E-10	0.24	0.02	2.82E-31	0.08	0.01	7.36E-15	6.55E-01	Hosp
7	99630342	rs2897075	T	0.364	*ZKSCAN1*	0.01	0.02	7.21E-01	0.05	0.01	8.92E-09	0.02	0.00	6.17E-06	7.17E-03	Both
7	100220850	rs11766643	G	0.204	*TFR2*	-0.05	0.02	2.45E-02	-0.04	0.01	4.71E-04	-0.03	0.01	7.46E-08	4.92E-01	Hosp
8	61537523	rs1120591	C	0.393	*RAB2A*	0.02	0.02	2.75E-01	0.05	0.01	8.64E-10	0.02	0.00	2.48E-05	4.32E-02	Both
9	15795833	rs79611697	T	0.068	*CCDC171*	0.08	0.04	3.75E-02	0.10	0.02	3.57E-08	0.03	0.01	3.26E-03	4.72E-01	Hosp
9	21172825	rs149533170	A	0.009	*IFNA21*	0.06	0.11	6.03E-01	0.28	0.05	1.38E-08	0.08	0.03	7.64E-04	2.21E-02	Both
9	136149229	rs505922	T	0.348	*ABO*	0.01	0.02	6.16E-01	0.09	0.01	4.14E-24	0.08	0.00	7.93E-90	2.33E-06	Both
10	81723695	rs61860402	T	0.112	*SFTPD*	0.08	0.03	2.65E-03	0.08	0.01	1.38E-10	0.03	0.01	2.76E-08	9.61E-01	Hosp
11	1241221	rs35705950	T	0.100	*MUC5B*	-0.09	0.03	3.84E-03	-0.10	0.01	3.41E-12	-0.02	0.01	3.31E-03	6.00E-01	Hosp
11	34529831	rs2924480	C	0.399	*ELF5*	-0.03	0.02	1.67E-01	-0.08	0.01	1.63E-16	-0.02	0.00	1.59E-06	1.31E-02	Both
12	113374748	rs10774679	T	0.653	*OAS3*	0.04	0.02	2.31E-02	0.07	0.01	3.25E-16	0.03	0.00	4.75E-10	7.42E-02	Hosp
12	133141973	rs5023077	C	0.513	*FBRSL1*	-0.05	0.02	1.03E-02	-0.07	0.01	2.84E-15	-0.02	0.00	3.14E-07	1.79E-01	Hosp
13	113535741	rs12585036	T	0.209	*ATP11A*	0.07	0.02	4.58E-04	0.10	0.01	3.15E-21	0.02	0.01	2.85E-06	2.22E-01	Hosp
14	29497258	rs150497803	T	0.043	*FOXG1*	0.00	0.05	9.79E-01	-0.09	0.02	6.68E-05	-0.07	0.01	2.79E-10	3.11E-02	Both
16	89262657	rs117169628	A	0.135	*SLC22A31*	0.09	0.03	4.68E-04	0.10	0.01	1.37E-14	0.03	0.01	2.78E-05	7.97E-01	Hosp
17	38182229	rs9916158	T	0.365	*GSDMA/B*	-0.01	0.02	4.85E-01	0.03	0.01	2.05E-04	0.02	0.00	1.11E-08	5.88E-03	Both
17	44057595	rs3785884	A	0.187	*MAPT*	-0.05	0.02	2.22E-02	-0.09	0.01	3.84E-16	-0.02	0.01	3.51E-04	6.69E-02	Hosp
17	44787312	rs1378358	T	0.167	*NSF*	-0.04	0.02	9.81E-02	-0.08	0.01	1.18E-13	-0.02	0.01	2.58E-03	2.75E-02	Both
17	47950329	rs75432325	C	0.042	*TAC4*	0.08	0.04	5.82E-02	0.18	0.02	2.24E-19	0.05	0.01	7.00E-07	2.04E-02	Both
19	4714468	rs7251000	T	0.432	*DPP9*	-0.07	0.02	3.98E-04	-0.08	0.01	1.17E-19	-0.03	0.00	1.18E-11	4.61E-01	Hosp
19	9009161	rs12976386	T	0.552	*MUC16*	0.04	0.03	1.14E-01	-0.03	0.01	9.37E-03	-0.03	0.01	3.75E-08	2.43E-03	Both
19	10463118	rs34536443	C	0.043	*TYK2*	0.05	0.06	3.58E-01	0.23	0.02	1.74E-21	0.05	0.01	1.29E-05	4.66E-04	Both
19	49206417	rs492602	G	0.460	*FUT2*	-0.02	0.02	1.69E-01	-0.06	0.01	1.71E-11	-0.04	0.00	3.44E-18	4.00E-02	Both
19	50882619	rs1405655	C	0.332	*NR1H2*	0.07	0.02	6.03E-05	0.08	0.01	5.92E-19	0.03	0.00	1.20E-09	7.95E-01	Hosp
21	34621948	rs2834164	C	0.480	*IFNAR2*	-0.10	0.02	1.51E-08	-0.11	0.01	4.14E-41	-0.03	0.00	6.54E-14	4.83E-01	Hosp
21	35353264	rs76608815	T	0.090	*SLC5A3*	0.11	0.03	9.64E-05	0.13	0.01	6.60E-21	0.04	0.01	1.49E-07	4.83E-01	Hosp
21	42852497	rs12329760	T	0.258	*TMPRSS2*	-0.07	0.02	8.60E-04	-0.06	0.01	2.16E-08	-0.01	0.00	5.32E-03	4.55E-01	Hosp
23	15620340	rs190509934	C	0.003	*ACE2*	-0.46	0.20	2.12E-02	-0.44	0.10	1.08E-05	-0.36	0.04	1.70E-18	9.38E-01	Hosp

574

575 Note: HGI: the COVID-19 host genetics initiative; chr: chromosome; pos: hg38 genomic position; allele:
576 alternative allele; AF: allele frequency of alternative allele; diffPval: statistical significance *P* for
577 differential effect size of each common SNP by comparing the two HGI GWASs, including HGI-B1 and
578 HGI-B2; Hosp: SNPs tend to be more specifically emerged in hospitalized COVID-19 cases; Both: SNPs
579 tend to be emerged in both hospitalized and non-hospitalized COVID-19 cases. Association *P*s passed the
580 suggestive genome-wide association significance threshold of $P < 1\text{x}10^{-7}$ are highlighted in light red color,

581 and differential effect size significance (DiffPval) $P < 0.05$ is colored with light green. Note: HGI-C2
582 (SASR-CoV2 susceptibility vs. general population), HGI-B1 (hospitalized vs. non-hospitalized COVID-
583 19), and HGI-B2 (hospitalized vs. general population).

584 **Table 2. Candidate SNPs specifically associated with non-hospitalized COVID-19 cases**

Variant information						HGI-B1			HGI-B2			HGI-C2			HGI-B1-vs-B2
chr	pos	rsid	allele	AF	gene	beta	se	p	beta	se	p	beta	se	p	DiffPval
3	174147743	rs9799354	C	0.556	*NLGN1*	0.08	0.02	5.38E-06	0.01	0.01	2.71E-01	-0.01	0.00	1.79E-02	6.21E-06
5	128241667	rs201484359	G	0.304	*SLC27A6*	0.13	0.03	8.55E-06	0.02	0.01	1.25E-01	-0.01	0.01	8.91E-02	2.58E-05
5	143904835	rs62401842	A	0.042	*KCTD16*	0.27	0.06	8.53E-07	0.02	0.03	3.82E-01	0.00	0.01	8.72E-01	3.18E-07
5	175248251	rs112842080	T	0.021	*CPLX2*	-0.36	0.07	1.58E-06	-0.04	0.03	2.53E-01	0.03	0.02	6.69E-02	1.25E-06
6	37086436	rs147171940	A	0.038	*PIM1*	-0.19	0.04	5.96E-06	-0.04	0.02	1.33E-01	0.01	0.01	5.74E-01	5.92E-05
6	151444801	rs76392050	T	0.010	*MTHFD1L*	0.70	0.16	6.10E-06	0.10	0.06	6.16E-02	0.03	0.03	3.84E-01	5.88E-06
8	57438048	rs4737438	A	0.378	*PENK*	0.08	0.02	2.97E-06	0.01	0.01	2.79E-01	0.00	0.00	5.30E-01	2.98E-06
8	68035003	rs148340257	T	0.065	*CSPP1*	-0.22	0.05	8.05E-06	-0.03	0.02	1.77E-01	0.00	0.01	8.82E-01	7.21E-06
9	123344199	rs10760104	T	0.819	*CDK5RAP2*	-0.13	0.03	2.37E-06	-0.01	0.01	4.57E-01	0.00	0.01	8.77E-01	5.36E-07
10	54976003	rs55654117	C	0.272	*MBL2*	0.10	0.02	1.85E-06	0.01	0.01	2.61E-01	0.00	0.00	4.50E-01	1.13E-06
10	84981211	rs61858037	T	0.091	*NRG3*	0.15	0.03	9.21E-06	0.02	0.02	1.62E-01	0.00	0.01	9.58E-01	2.31E-05
12	39502966	rs61939166	G	0.011	*KIF21A*	-0.40	0.09	3.60E-06	-0.03	0.04	5.36E-01	-0.02	0.02	3.63E-01	1.10E-06
12	78552292	rs367777	T	0.358	*NAV3*	-0.10	0.02	3.13E-06	-0.01	0.01	2.77E-01	0.01	0.00	2.39E-01	1.89E-06
13	27062460	rs56143829	T	0.025	*WASF3*	0.24	0.05	1.26E-06	0.01	0.03	6.03E-01	0.00	0.01	7.56E-01	3.23E-07
13	63988259	rs9570861	A	0.552	*PCDH20*	0.08	0.02	5.51E-06	0.02	0.01	1.07E-01	-0.01	0.00	4.06E-02	4.83E-05
14	32937383	rs11454577	G	0.589	*AKAP6*	-0.09	0.02	5.54E-06	-0.01	0.01	3.22E-01	0.00	0.01	6.24E-01	7.71E-06
16	84887921	rs113067468	G	0.030	*CRISPLD2*	-0.30	0.07	4.48E-06	-0.05	0.03	1.20E-01	0.00	0.02	8.21E-01	2.18E-05
17	3454928	rs115744301	G	0.035	*TRPV3*	0.66	0.15	5.21E-06	0.02	0.05	6.89E-01	-0.03	0.03	2.52E-01	2.53E-07
17	13775336	rs11869231	T	0.310	*COX10*	-0.11	0.02	2.38E-06	0.00	0.01	6.52E-01	0.00	0.00	9.85E-01	2.26E-07
20	24627964	rs6049828	C	0.567	*SYNDIG1*	0.09	0.02	7.13E-06	0.01	0.01	2.65E-01	0.00	0.00	3.51E-01	5.43E-06

585

586 Note: abbreviations in the table headers are similar to that of Table 1.

587

588 **SUPPLEMENTARY TABLES**

589 **Table S1. COVID-19 GWASs and long COVID GWASs included in current study for prioritizing**
590 **candidate SNPs**

591 **Table S2. Transcriptome-wide association studies (TWAS) curated and used in current study**

592 **Table S3. Association signals for these 62 prioritized candidate SNPs in Long COVID GWASs from**
593 **HGI**

594 **Table S4. Association signals for adjacent SNPs of these 62 prioritized candidate SNPs in Long**
595 **COVID GWASs from HGI**

596 **Table S5. Association signals for published long COVID SNPs across 4 HGI COVID-19 GWASs**

597 **Table S6 Top genes associated with acute-phase COVID-19 susceptibility and hospitalization across**
598 **48 GTEx tissues in transcriptome-wide association studies (TWAS)**

599 **Table S7. Long COVID association signals for the adjacent genes of these 62 candidate SNPs that**
600 **pass the nominal significance in the transcriptome-wide association studies (TWAS) of COVID-19**
601 **hospitalization and susceptibility**

602 **Table S8. Adjacent genes of these 62 candidate SNPs that pass the nominal significance threshold in**
603 **the phenome-wide transcriptome association studies (Pheno-TWASs)**

604

FIGURES

Figure 1. Workflow for prioritizing genetic single nucleotide polymorphisms (SNPs) associated with either COVID-19 hospitalization or non-hospitalization or both as well as its association with sex-stratified COVID-19 hospitalization and long COVID

An integrative analysis workflow is developed to analyze acute-phase COVID-19 GWAS, including HGI-C2 (SARS-CoV2 infected cases vs. general population), HGI-B1 (hospitalized COVID-19 vs. non-hospitalized COVID-19), and HGI-B2 (hospitalized COVID-19 vs. general population), as well as the differential GWAS (HGI-B1-vs-B2) generated in current study by comparing effect sizes of each SNPs between HGI-B1 and HGI-B2. Three groups of top candidate SNPs, comprising "Severe COVID-19-Specific SNPs," "SNPs Associated with Both Severe and Mild COVID-19," and "Mild COVID-19-Specific SNPs," are obtained by applying different filters of association p-values as demonstrated in the right part of the figure. These filters result in 62 candidate SNPs that are subsequently subjected to the determination of their associations with long COVID or sex-biased association with COVID-19 hospitalization. Further transcriptome-wide association studies and phenotype-wide transcriptome-wide association studies are conducted to investigate genes adjacent to these candidate SNPs and their association with acute-phase COVID-19 and long COVID.

Figure 2. Candidate SNPs associated with COVID-19 hospitalization or both hospitalization and non-hospitalization

(A) 42 top prioritized candidate SNPs, with 27 SNPs specifically associated with hospitalization risk and 15 showing common risk patterns across both hospitalized and non-hospitalized cases. These SNPs were selected based on the analysis workflow illustrated in Fig. 1, and each SNP is centered within a 1Mb window in a local Manhattan plot to highlight the genomic context for its association with acute-phase COVID-19. Key genes near these SNPs, including *FOXP4*—a locus previously implicated as a risk locus for both acute-phase COVID-19 and long COVID—and other notable COVID-19-associated genes (*ACE2*, *OAS3*, *JAK1*, *IFNA2*, *DDP9*, and *ABO*), are annotated to emphasize their potential relevance to SARS-CoV-2 pathology. (B) Forest plot of 42 SNPs identified from COVID-19 GWAS on susceptibility (HGI-C2) and hospitalization (HGI-B1 and HGI-B2), showing suggestive ($P<1x10^{-7}$ or genome-wide significant associations ($P<5x10^{-8}$). Allele distributions are compared between hospitalized and non-hospitalized cases (HGI-B1). SNPs are grouped as follows: (1) hospitalization-specific, genome-wide significant (Hosp-specific, $P<5x10^{-8}$); (2) common across cases, genome-wide significant (Common, $P<5x10^{-8}$); (3) hospitalization-specific, suggestive or non-significant (hosp-specific, P = ($5x10^{-8}$, 1]); (4) common across cases, suggestive or non-significant (Common, P = ($5x10^{-8}$, 1]). Colored dots represent odds ratios (ORs)

with 95% confidence intervals (CIs) as error bars. ORs with CIs overlapping the vertical line (OR = 1) indicate no significance.

Figure 3. Evaluation of top genes adjacent to prioritized candidate SNPs (n=62) in transcriptome-wide association studies (TWAS) of COVID-19 susceptibility, hospitalization, and long COVID

(A) Bar plot showing the top genes adjacent to prioritized candidate SNPs based on transcriptome-wide association studies (TWAS) for COVID-19 susceptibility and hospitalization, as well as long COVID. Genes are categorized by the GTEx tissue in which they display the most significant association. Besides rs17078348, other two SNPs have multiple genes with the smallest TWAS *P* are highlighted. (B) Heatmap depicting the significance of gene-tissue associations for prioritized SNP-adjacent genes across 48 GTEx tissues. Each cell represents the *P* for a specific gene in a given tissue, with a color gradient indicating statistical significance: gray ($P>0.05$), yellow ($P<0.05$), and red ($P<1\text{x}10^{-6}$). Genes are sorted by TWAS names, and GTEx tissues are clustered to highlight patterns of tissue specificity.

Figure 4. Top genes adjacent to prioritized candidate SNPs (n=23) in long COVID related transcriptome-wide association studies (TWASs)

Bar plots illustrating the best *P* across 48 GTEx tissues for genes adjacent to prioritized candidate SNPs (n=23) in 4 long COVID TWASs.

SUPPLEMENTARY FIGURES

Figure S1. QQ plots for four COVID-19 GWASs from HGI

Figure S2. Local Manhattan plots for candidate SNPs more likely associated with hospitalized than non-hospitalized COVID-19

Figure S3. Local Manhattan plots for candidate SNPs likely associated with both hospitalized and non-hospitalized COVID-19

Figure S4. Evaluation of prioritized candidate SNPs that tend to be either more likely associated with COVID-19 hospitalization or commonly associated with both hospitalization and non-hospitalization of COVID-19 in sex-biased COVID-19 hospitalization GWAS from UK Biobank

(A-B) Manhattan plots and QQ plots for sex-stratified COVID-19 hospitalization GWASs, as well as the differential effect size GWAS between females and males. (C) Local Manhattan plot for these 42 SNPs separated into two groups, including candidate SNPs that tend to be specific in hospitalized cases and that are more frequent in both hospitalized and non-hospitalized cases. All association signals shown in the figure are within a 100Mb genomic window for each candidate SNP residing in the middle. Genes close to these 42 SNPs are added into the upper track. (C) Forest plots for these 42 candidate SNPs in sex-stratified COVID-19 hospitalization GWASs, as well as the differential effect size GWAS between females and males. Odd Ratios (ORs) and the corresponding 95% confidence intervals (CI) of these SNPs are illustrated, with SNPs showing nominal significant associations among female or male COVID-19 patients are colored in dark red and indicated by '*'. Note that there are 8 and 6 SNPs showing sex-biased associations with females and males, respectively, meanwhile, only one SNP rs17078348 mapped to *SLC6A20* displaying nominally significant association signals in both sexes. The genomic position of rs190509934 is used to collect all other SNPs around it, as it is not included in the two UKB sex-stratified COVID-19 hospitalization GWASs.

Figure S5. Local Manhattan plots for candidate SNPs either more likely associated with COVID-19 hospitalization or commonly associated with both hospitalization and non-hospitalization of COVID-19 in sex-biased COVID-19 hospitalization GWAS from UK Biobank

Figure S6. Local Manhattan plots for candidate SNPs specifically associated with non-hospitalized COVID-19

Figure S7. Evaluation of candidate SNPs associated with non-hospitalized COVID-19 in sex-biased COVID-19 hospitalization GWAS from UK Biobank

(A) Local Manhattan plots for candidate SNPs identified exclusively in the non-hospitalized COVID-19 group in sex-stratified COVID-19 hospitalization GWAS (the differential effect size GWAS between females and males put on the top, female GWAS added in the middle, and male GWAS shown at the bottom). (B) Forest plots displaying the odds ratios (ORs) and its corresponding 95% confidence intervals (CIs) of these SNPs in non-hospitalized COVID-19 patients from the sex-stratified GWAS (male results on the top and female results at the bottom). SNPs with nominally significant associations in either female or male groups are highlighted in dark red and marked with an asterisk (*). Notably, three SNPs exhibit male-biased associations with COVID-19 hospitalization. (C-H) Detailed local Manhattan plots for six prioritized SNPs specifically identified in the non-hospitalized COVID-19 group within the sex-stratified GWAS.

Figure S8. Local Manhattan plots for SNPs in the SNP category "Severe COVID-19-Specific SNPs" across 4 long COVID GWASs

Each target SNP is located at the center of each local Manhattan plot, and its linked gene is put along SNP rsid at the bottom of the figure.

Figure S9. Local Manhattan plots for SNPs in the two SNP categories, including "Mild COVID-19-Specific SNPs" and "SNPs Associated with Both Severe and Mild COVID-19" across 4 long COVID GWASs

Figure annotation is similar to that of Figure S8.

Figure S10. Forest plots for 62 candidate SNPs and its adjacent SNPs passed the association threshold of *P*<0.001 in at least one of four long COVID GWASs

SNPs adjacent to candidate SNPs that are prioritized from 4 HGI COVID-19 GWASs (colored according to its membership to 3 SNP categories in Figure 1) are indicated by adding "+" after the corresponding gene. "*" represents the SNP's association with long COVID is nominally significant. To the sake of clarity, the OR is restricted to be less than 2. SNPs that are not include in specific long COVID GWASs show no OR and 95% CI in the forest plots.

Figure S11. Evaluation of other published long COVID SNPs in the current HGI COVID-19 GWASs

(A) Forest plots for 15 published long COVID SNPs in currently acute-phase COVID-19 GWASs (from left to right: HGI-B1, HGI-B2, and HGI-C2), with genes closed to each of them listed on the right-y-axis. Note: beside 3 long COVID risk SNPs reported by Chaudhary et al., other 12 long COVID SNPs were

obtained by screening 83 previously reported long COVID SNPs by Taylor et al. (see Table S5) that show nominally significant association in at least one of the 3 acute-phase COVID-19 GWASs. (B-G) Local Manhattan plots illustrate published long COVID SNPs approximate to these top candidate SNPs identified in current study with each candidate SNP is located in the center of each genomic window; GWAS labels from bottom to top in each panel are HGI-B1, HGI-B2, HGI-B1-vs-B2, and HGI-C2, respectively.

Figure S12. Total number of genes adjacent to prioritized candidate SNPs (n=62) as well as total number of genes that were tested and passed different association thresholds in COVID-19 transcriptome-wide association studies (TWASs)

Bar plots illustrating the number of genes adjacent to prioritized candidate SNPs (n=62) as well as number of adjacent genes tested and passed different association *P* thresholds in TWASs. Note that the 62 SNPs are colored according to its corresponding SNP category as depicted in Figure 1. There are 4 acute-phase COVID-19 related TWASs and 4 long COVID TWASs performed across 48 GTEx tissues.

Figure S13. Transcriptome-wide association studies (TWAS) for COVID-19 susceptibility, hospitalization, and long COVID

(A) Top genome-wide significant TWAS-identified genes associated with COVID-19 susceptibility and hospitalization, annotated by their best significant tissue expression association (based on 48 GTEx tissues). No genes reached genome-wide significance in TWAS analyses of long COVID-related GWASs. (B) Manhattan plots depicting TWAS results for COVID-19 susceptibility, and hospitalization (HGI datasets), as well as long COVID GWASs (LongCOVID-N1, LongCOVID-N2, LongCOVID-W1, and LongCOVID-W2) derived from four distinct case-control designs (see Table S1). Each dot represents a gene, with its genomic position along the x-axis and -log10(*P*) on the y-axis.

Figure S14. Evaluation of top genes emerged from transcriptome-wide association studies (TWAS) of COVID-19 susceptibility and hospitalization in the corresponding long COVID-related TWASs

Heatmap displaying the significance (*P*) of top genes identified in TWAS analyses of COVID-19 susceptibility and hospitalization across multiple tissues in long COVID-related TWASs. Each cell represents the *P* for the association of a gene with long COVID in a specific tissue from GTEx database. The color gradient denotes statistical significance: blue ($P>0.05$), yellow ($P<0.05$), and red ($P<1\text{x}10^{-6}$). Rows are sorted by genes, and columns represent tissues, allowing visualization of tissue-specific patterns of gene expression relevance.

Figure S15. Other top genes adjacent to prioritized candidate SNPs (n=29) in long COVID-19 transcriptome-wide association studies (TWASs)

Bar plots illustrating the best *P* for genes adjacent to prioritized candidate SNPs (n=29) in 4 long COVID TWASs across 48 GTEx tissues, which is a supplement to Figure 4.

Figure S16. Evaluation of top genes adjacent to prioritized candidate SNPs (n=62) in transcriptome-wide association studies (TWASs) of long COVID across 48 GTEx tissues

(A) Bar plot showing the number of SNP-gene pairs passed the nominal TWAS association threshold in GTEx tissues. Note that SNP-gene pairs are defined by combining prioritized candidate SNPs with its adjacent genes. (B) Heatmap depicting the significance (*P*) of SNP-gene pairs across 48 GTEx tissues in long COVID TWASs. Each cell represents the *P* for a specific gene in a given tissue, with a color gradient indicating statistical significance: white ($P>0.05$), yellow ($P<0.05$), and red ($P<1E-6$). Genes are sorted by SNP-gene pairs and TWAS names.

Figure S17. Evaluation of genes adjacent to prioritized candidate SNP rs17219281 (*HLA-DQA1*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

To the sake of clarity, heatmap only depicting the significance of SNP-gene pairs with $P<1x10^{-60}$ across 48 GTEx tissues in TWASs. Each cell represents the *P* for a specific gene in a given tissue, with blank cell indicating the *P* not passed the threshold of $P<1x10^{-60}$.

Figure S18. Evaluation of genes adjacent to prioritized candidate SNP rs9260038 (*HLA-A*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

Annotation is similar to heatmap as that in Figure S17.

Figure S19. Evaluation of genes adjacent to prioritized candidate SNP rs1634761 (*HLA-C*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

Annotation is similar to heatmap as that in Figure S17.

Figure S20. Evaluation of genes adjacent to prioritized candidate SNP rs2834164 (*IFNAR2*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

To the sake of clarity, the color gradient of heatmap denotes statistical significance: blue (P>0.05), yellow (P<0.05), and red (P<11x10^{-6}). Rows are sorted by genes, and columns represent TWAS names, allowing visualization of tissue-specific patterns of gene expression relevance across 48 GTEx tissues in TWASs.

Figure S21. Evaluation of genes adjacent to prioritized candidate SNP rs9916158 (*GSDMA/B*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S20

Figure S22. Evaluation of genes adjacent to prioritized candidate SNP rs12602210 (*BPTF*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S20

Figure S23. Evaluation of genes adjacent to prioritized candidate SNP rs79611697 (*CCDC171*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S20

Figure S24. Evaluation of genes adjacent to prioritized candidate SNP rs12660421 (*FOXP4*) in phenome-transcriptome-wide association studies (Pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S20

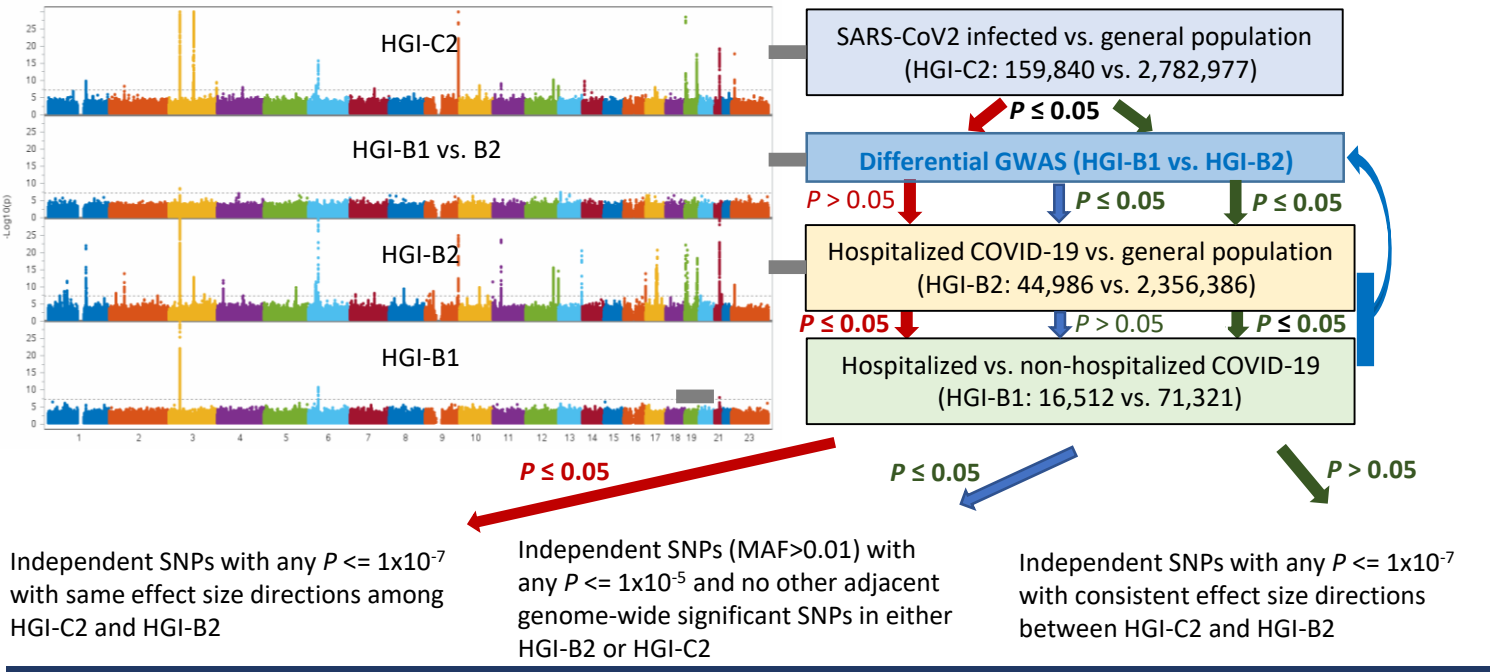
References

1. Ely, E.W., Brown, L.M., Fineberg, H.V., National Academies of Sciences, E., and Medicine Committee on Examining the Working Definition for Long, C. (2024). Long Covid Defined. N Engl J Med *391*, 1746-1753. 10.1056/NEJMsb2408466.
2. Lammi, V., Nakanishi, T., Jones, S.E., Andrews, S.J., Karjalainen, J., Cortes, B., O'Brien, H.E., Ochoa-Guzman, A., Fulton-Howard, B.E., Broberg, M., et al. (2025). Genome-wide association study of long COVID. Nat Genet. 10.1038/s41588-025-02100-w.
3. Luo, Y.-S., Zhang, K., and Cheng, Z.-S. (2023). Absence of Association between a Long COVID and Severe COVID-19 Risk Variant of FOXP4 and Lung Cancer. Frontiers in genetics *14*. 10.3389/fgene.2023.1258829.
4. Taylor, K., Pearson, M., Das, S., Sardell, J., Chocian, K., and Gardner, S. (2023). Genetic risk factors for severe and fatigue dominant long COVID and commonalities with ME/CFS identified by combinatorial analysis. Journal of translational medicine *21*, 775. 10.1186/s12967-023-04588-4.
5. Chaudhary, N.S., Weldon, C.H., Nandakumar, P., Team, a.R., Holmes, M.V., and Aslibekyan, S. (2024). Multi-ancestry GWAS of Long COVID identifies immune-related loci and etiological links to chronic fatigue

802 syndrome, fibromyalgia and depression. medRxiv, 2024.2010.2007.24315052.
803 10.1101/2024.10.07.24315052.
804 6. Covid-Host-Genetics-Initiative (2023). A second update on mapping the human genetic architecture of
805 COVID-19. Nature *621*, E7-E26. 10.1038/s41586-023-06355-3.
806 7. Davis, H.E., McCorkell, L., Vogel, J.M., and Topol, E.J. (2023). Long COVID: major findings, mechanisms and
807 recommendations. Nature reviews. Microbiology *21*, 133-146. 10.1038/s41579-022-00846-2.
808 8. Cai, M., Xie, Y., Topol, E.J., and Al-Aly, Z. (2024). Three-year outcomes of post-acute sequelae of COVID-19.
809 Nat Med *30*, 1564-1573. 10.1038/s41591-024-02987-8.
810 9. Lammi, V., Nakanishi, T., Jones, S.E., Andrews, S.J., Karjalainen, J., Cortés, B., O'Brien, H.E., Fulton-Howard,
811 B.E., Haapaniemi, H.H., Schmidt, A., et al. (2023). Genome-wide Association Study of Long COVID.
812 medRxiv, 2023.2006.2029.23292056. 10.1101/2023.06.29.23292056.
813 10. Cheng, Z., Cai, Y., Zhang, K., Zhang, J., Gui, H., Luo, Y.S., Zhou, J., and DeVeale, B. (2023). MAP3K19
814 regulatory variation in populations with African ancestry may increase COVID-19 severity. iScience *26*,
815 107555. 10.1016/j.isci.2023.107555.
816 11. Bai, F., Tomasoni, D., Falcinella, C., Barbanotti, D., Castoldi, R., Mulè, G., Augello, M., Mondatore, D.,
817 Allegrini, M., Cona, A., et al. (2022). Female gender is associated with long COVID syndrome: a prospective
818 cohort study. Clinical microbiology and infection : the official publication of the European Society of
819 Clinical Microbiology and Infectious Diseases *28*, 611.e619-611.e616. 10.1016/j.cmi.2021.11.002.
820 12. Neuhouser, M.L., Butt, H.I., Hu, C., Shadyab, A.H., Garcia, L., Follis, S., Mouton, C., Harris, H.R., Wactawski-
821 Wende, J., Gower, E.W., et al. (2024). Risk factors for long COVID syndrome in postmenopausal women
822 with previously reported diagnosis of COVID-19. Annals of epidemiology *98*, 36-43.
823 10.1016/j.annepidem.2024.08.003.
824 13. Thibord, F., Chan, M.V., Chen, M.H., and Johnson, A.D. (2022). A year of Covid-19 GWAS results from the
825 GRASP portal reveals potential genetic risk factors. HGG Adv, 100095. 10.1016/j.xhgg.2022.100095.
826 14. Barbeira, A.N., Dickinson, S.P., Bonazzola, R., Zheng, J., Wheeler, H.E., Torres, J.M., Torstenson, E.S., Shah,
827 K.P., Garcia, T., Edwards, T.L., et al. (2018). Exploring the phenotypic consequences of tissue specific gene
828 expression variation inferred from GWAS summary statistics. Nat Commun *9*, 1825. 10.1038/s41467-018-
829 03621-1.
830 15. GTEx-Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues.
831 Science *369*, 1318-1330. 10.1126/science.aaz1776.
832 16. Severe-COVID-GWAS-Group, Ellinghaus, D., Degenhardt, F., Bujanda, L., Buti, M., Albillos, A., Invernizzi, P.,
833 Fernandez, J., Prati, D., Baselli, G., et al. (2020). Genomewide Association Study of Severe Covid-19 with
834 Respiratory Failure. N Engl J Med *383*, 1522-1534. 10.1056/NEJMoa2020283.
835 17. Galbraith, M.D., Kinning, K.T., Sullivan, K.D., Araya, P., Smith, K.P., Granrath, R.E., Shaw, J.R., Baxter, R.,
836 Jordan, K.R., Russell, S., et al. (2022). Specialized interferon action in COVID-19. Proc Natl Acad Sci U S A
837 *119*. 10.1073/pnas.2116730119.
838 18. Zabihi Rizi, F., Ghorbani, A., Zahtab, P., Darbaghshahi, N.N., Ataee, N., Pourhamzeh, P., Hamzei, B.,
839 Dolatabadi, N.F., Zamani, A., and Hooshmand, M. (2023). TYK2 single-nucleotide variants associated with
840 the severity of COVID-19 disease. Archives of virology *168*, 119. 10.1007/s00705-023-05729-2.
841 19. Bittner, Z.A., Schrader, M., George, S.E., and Amann, R. (2022). Pyroptosis and Its Role in SARS-CoV-2
842 Infection. Cells *11*. 10.3390/cells11101717.
843 20. Choudhri, Y., Maslove, D.M., and Rauh, M.J. (2023). COVID-19 and the Genetics of Inflammation. Critical
844 care medicine *51*, 817-825. 10.1097/CCM.0000000000005843.
845 21. Pietzner, M., Chua, R.L., Wheeler, E., Jechow, K., Willett, J.D.S., Radbruch, H., Trump, S., Heidecker, B.,
846 Zeberg, H., Heppner, F.L., et al. (2022). ELF5 is a potential respiratory epithelial cell-specific risk gene for
847 severe COVID-19. Nature Communications *13*, 4484. 10.1038/s41467-022-31999-6.
848 22. Hammersen, J., Birndt, S., Dohner, K., Reuken, P., Stallmach, A., Sauerbrey, P., La Rosee, F., Pfirrmann, M.,
849 Fabisch, C., Weiss, M., et al. (2023). The JAK1/2 inhibitor ruxolitinib in patients with COVID-19 triggered
850 hyperinflammation: the RuxCoFlam trial. Leukemia *37*, 1879-1886. 10.1038/s41375-023-01979-w.
851 23. Galeev, A., Suwandi, A., Cepic, A., Basu, M., Baines, J.F., and Grassl, G.A. (2021). The role of the blood
852 group-related glycosyltransferases FUT2 and B4GALNT2 in susceptibility to infectious disease.
853 International journal of medical microbiology : IJMM *311*, 151487. 10.1016/j.ijmm.2021.151487.

24. Rao, S., Baranova, A., Cao, H., Chen, J., Zhang, X., and Zhang, F. (2021). Genetic mechanisms of COVID-19 and its association with smoking and alcohol consumption. Briefings in bioinformatics *22*. 10.1093/bib/bbab284.
25. Matveeva, N., Kiselev, I., Baulina, N., Semina, E., Kakotkin, V., Agapov, M., Kulakova, O., and Favorova, O. (2023). Shared genetic architecture of COVID-19 and Alzheimer's disease. Frontiers in aging neuroscience *15*, 1287322. 10.3389/fnagi.2023.1287322.
26. Simon-Sanchez, J., Schulte, C., Bras, J.M., Sharma, M., Gibbs, J.R., Berg, D., Paisan-Ruiz, C., Lichtner, P., Scholz, S.W., Hernandez, D.G., et al. (2009). Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet *41*, 1308-1312. 10.1038/ng.487.
27. Lottini, G., Baggiani, M., Chesi, G., D'Orsi, B., Quaranta, P., Lai, M., Pancrazi, L., Onorati, M., Pistello, M., Freer, G., and Costa, M. (2022). Zika virus induces FOXG1 nuclear displacement and downregulation in human neural progenitors. Stem cell reports *17*, 1683-1698. 10.1016/j.stemcr.2022.05.008.
28. Kovanda, A., Lukezic, T., Maver, A., Vokac Krizaj, H., Cizek Sajko, M., Selb, J., Rijavec, M., Bidovec-Stojkovic, U., Biteznik, B., Rituper, B., et al. (2024). Genomic Landscape of Susceptibility to Severe COVID-19 in the Slovenian Population. Int J Mol Sci *25*. 10.3390/ijms25147674.
29. Kasela, S., Daniloski, Z., Bollepalli, S., Jordan, T.X., tenOever, B.R., Sanjana, N.E., and Lappalainen, T. (2021). Integrative approach identifies SLC6A20 and CXCR6 as putative causal genes for the COVID-19 GWAS signal in the 3p21.31 locus. Genome biology *22*, 242. 10.1186/s13059-021-02454-4.
30. Zuo, H., Glaaser, I., Zhao, Y., Kurinov, I., Mosyak, L., Wang, H., Liu, J., Park, J., Frangaj, A., Sturchler, E., et al. (2019). Structural basis for auxiliary subunit KCTD16 regulation of the GABA(B) receptor. Proc Natl Acad Sci U S A *116*, 8370-8379. 10.1073/pnas.1903024116.
31. Zhou, S., and Zhang, J. (2021). An epigenome-wide DNA methylation study of patients with COVID-19. *85*, 221-234. 10.1111/ahg.12440.
32. Song, E., Bartley, C.M., Chow, R.D., Ngo, T.T., Jiang, R., Zamecnik, C.R., Dandekar, R., Loudermilk, R.P., Dai, Y., Liu, F., et al. (2021). Divergent and self-reactive immune responses in the CNS of COVID-19 patients with neurological symptoms. Cell Reports Medicine *2*, 100288. https://doi.org/10.1016/j.xcrm.2021.100288.
33. Bartley, C.M., Johns, C., Ngo, T.T., Dandekar, R., Loudermilk, R.L., Alvarenga, B.D., Hawes, I.A., Zamecnik, C.R., Zorn, K.C., Alexander, J.R., et al. (2021). Anti–SARS-CoV-2 and Autoantibody Profiles in the Cerebrospinal Fluid of 3 Teenaged Patients With COVID-19 and Subacute Neuropsychiatric Symptoms. JAMA Neurology *78*, 1503-1509. 10.1001/jamaneurol.2021.3821.
34. Yang, X., Zhu, J., Wang, Q., Tang, B., Shen, Y., Wang, B., Ji, L., Liu, H., and Wuchty, S. (2024). Comparative analysis of dynamic transcriptomes reveals specific COVID-19 features and pathogenesis of immunocompromised populations. *9*, e0138523. 10.1128/msystems.01385-23.
35. Wang, P.Y., Ma, J., Kim, Y.C., Son, A.Y., Syed, A.M., Liu, C., Mori, M.P., Huffstutler, R.D., Stolinski, J.L., Talagala, S.L., et al. (2023). WASF3 disrupts mitochondrial respiration and may mediate exercise intolerance in myalgic encephalomyelitis/chronic fatigue syndrome. Proc Natl Acad Sci U S A *120*, e2302738120. 10.1073/pnas.2302738120.
36. Guruparan, D., and Kumar, S. (2024). Long COVID: G Protein-Coupled Receptors (GPCRs) Associated Genes and Pathways as a Promising Therapeutic Potential Through Computational Analysis. Preprints. Preprints.
37. Deuis, J.R., Klasfauseweh, T., Walker, L., and Vetter, I. (2024). The ‘dispanins’ and related proteins in physiology and neurological disease. Trends in Neurosciences *47*, 622-634. 10.1016/j.tins.2024.06.004.
38. Chester, M.A., and Olsson, M.L. (2001). The ABO blood group gene: a locus of considerable genetic diversity. Transfusion medicine reviews *15*, 177-200. 10.1053/tmrv.2001.24591.

Figure 1



Independent SNPs with any $P <= 1x10^{-7}$ with same effect size directions among HGI-C2 and HGI-B2

Independent SNPs (MAF>0.01) with any $P <= 1x10^{-5}$ and no other adjacent genome-wide significant SNPs in either HGI-B2 or HGI-C2

Independent SNPs with any $P <= 1x10^{-7}$ with consistent effect size directions between HGI-C2 and HGI-B2

Prioritize each top hit among 4 HGI-GWASs by manually evaluating local Manhattan plots

Category 1: Severe COVID-19-Specific SNPs (n=27)

Category 3: Mild COVID-19-Specific SNPs (n=20)

Category 2: SNPs Associated with Both Severe and Mild COVID-19 (n=15)

Determine association signals for these 62 candidate SNPs in sex-stratified COVID-19 GWASs and long COVID GWASs

Check genes close to these candidate SNPs in transcriptome-wide association studies (TWASs) of COVID-19 hospitalization/susceptibility and long COVID GWASs across 48 GTEx tissues

Evaluate genes adjacent to these candidate SNPs in phenotype-wide transcription association studies (Pheno-TWASs) across 48 GTEx tissues, covering a broad set of 420 phenotypes

Figure 2

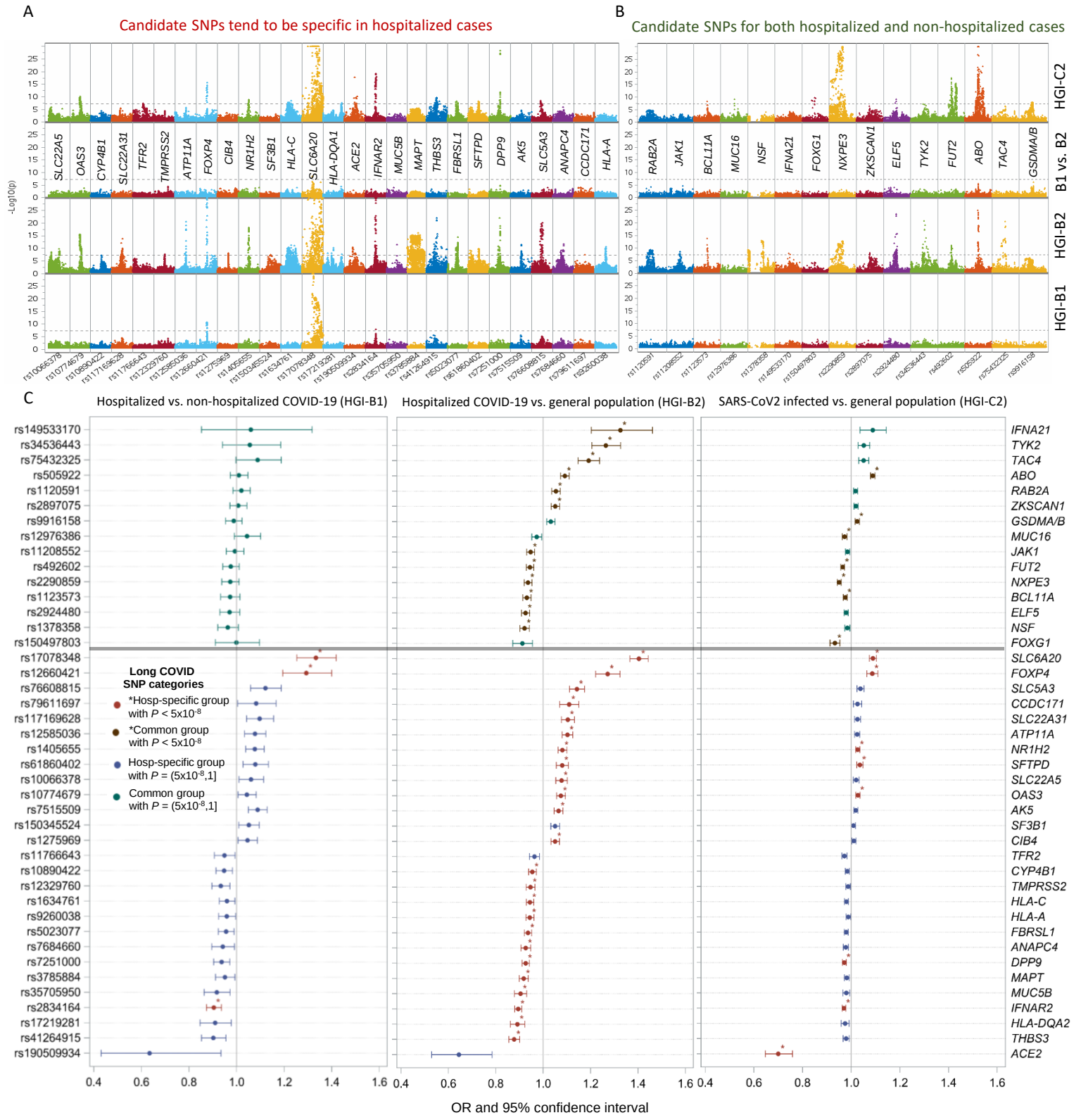

A
Candidate SNPs tend to be specific in hospitalized cases
B
Candidate SNPs for both hospitalized and non-hospitalized cases
-Log10(p)
HGI-C2
B1 vs. B2
HGI-B2
HGI-B1
SLC22A5
OAS3
CYP4B1
SLC22A31
TFR2
TMPRSS2
ATP11A
FOXP4
CIB4
NR1H2
SF3B1
HLA-C
SLC6A20
HLA-DQA1
ACE2
IFNAR2
MUC5B
MAPT
THBS3
FBRSL1
SFTPD
DPP9
AK5
SLC5A3
ANAPC4
CCDC171
HLA-A
RAB2A
JAK1
BCL11A
MUC16
NSF
IFNA21
FOXG1
NXPE3
ZKSCAN1
ELF5
TYK2
FUT2
ABO
TAC4
GSDMA/B
C
Hospitalized vs. non-hospitalized COVID-19 (HGI-B1)
Hospitalized COVID-19 vs. general population (HGI-B2)
SARS-CoV2 infected vs. general population (HGI-C2)
Long COVID
SNP categories
*Hosp-specific group with P < 5x10-8
*Common group with P < 5x10-8
Hosp-specific group with P = (5x10-8,1]
Common group with P = (5x10-8,1]
OR and 95% confidence interval

Figure 3

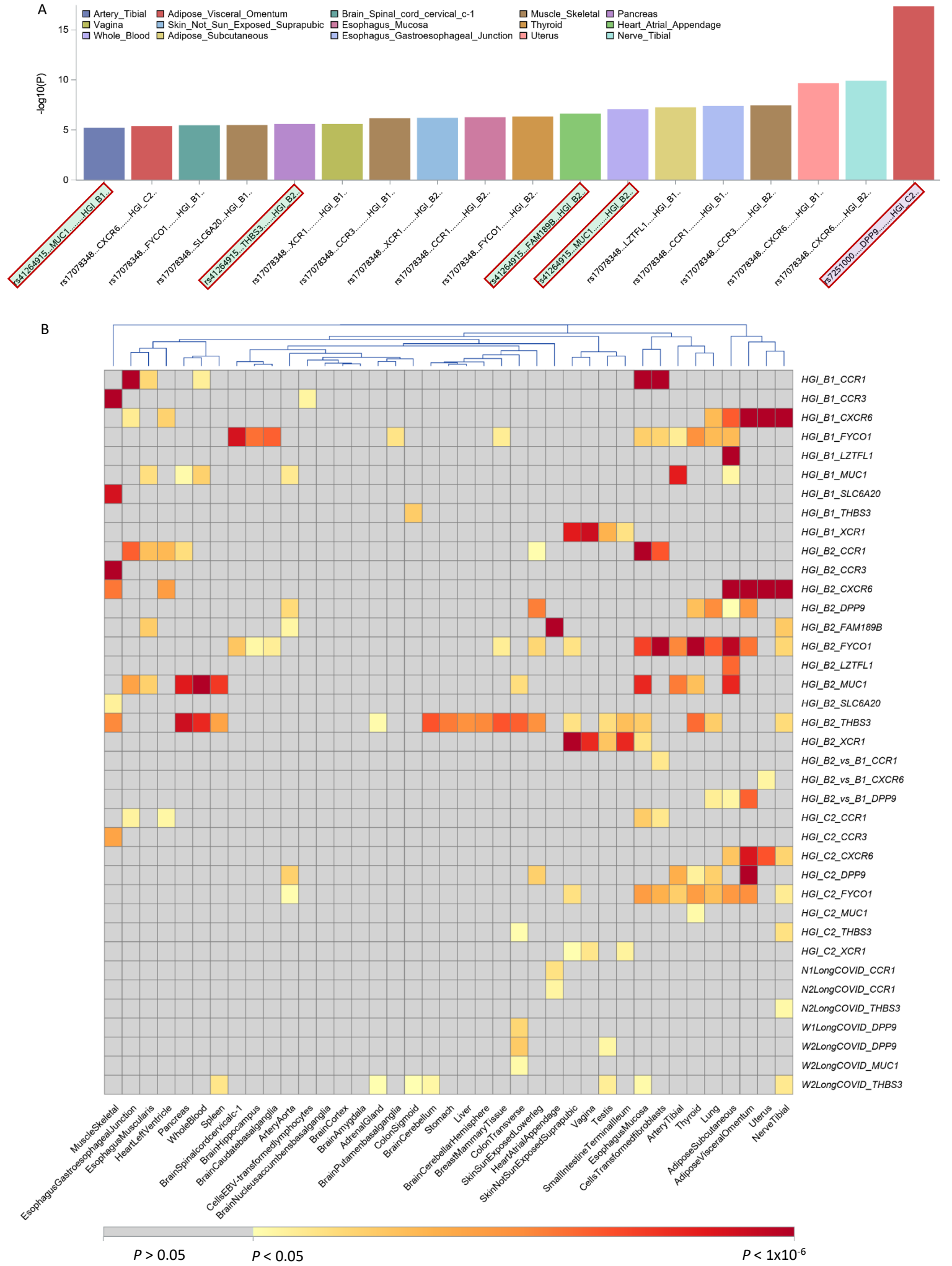

A
Artery_Tibial
Vagina
Whole_Blood
Adipose_Visceral_Omentum
Skin_Not_Sun_Exposed_Suprapubic
Adipose_Subcutaneous
Brain_Spinal_cord_cervical_c-1
Esophagus_Mucosa
Esophagus_Gastroesophageal_Junction
Muscle_Skeletal
Thyroid
Uterus
Pancreas
Heart_Atrial_Appendage
Nerve_Tibial
-log10(P)
0
5
10
15
rs41264915...MUC1.....HGI_B1..
rs17078348...CXCR6.....HGI_C2..
rs17078348...FYCO1.....HGI_B1..
rs17078348...SLC6A20...HGI_B1..
rs41264915...THBS3.....HGI_B2..
rs17078348...XCR1.....HGI_B1..
rs17078348...CCR3.....HGI_B1..
rs17078348...XCR1.....HGI_B2..
rs17078348...CCR1.....HGI_B2..
rs17078348...FYCO1.....HGI_B2..
rs41264915...FAM189B...HGI_B2..
rs41264915...MUC1.....HGI_B2..
rs17078348...LZTFL1...HGI_B1..
rs17078348...CCR1.....HGI_B1..
rs17078348...CCR3.....HGI_B2..
rs17078348...CXCR6.....HGI_B1..
rs17078348...CXCR6.....HGI_B2..
rs7251000....DPP9.....HGI_C2..
B
HGI_B1_CCR1
HGI_B1_CCR3
HGI_B1_CXCR6
HGI_B1_FYCO1
HGI_B1_LZTFL1
HGI_B1_MUC1
HGI_B1_SLC6A20
HGI_B1_THBS3
HGI_B1_XCR1
HGI_B2_CCR1
HGI_B2_CCR3
HGI_B2_CXCR6
HGI_B2_DPP9
HGI_B2_FAM189B
HGI_B2_FYCO1
HGI_B2_LZTFL1
HGI_B2_MUC1
HGI_B2_SLC6A20
HGI_B2_THBS3
HGI_B2_XCR1
HGI_B2_vs_B1_CCR1
HGI_B2_vs_B1_CXCR6
HGI_B2_vs_B1_DPP9
HGI_C2_CCR1
HGI_C2_CCR3
HGI_C2_CXCR6
HGI_C2_DPP9
HGI_C2_FYCO1
HGI_C2_MUC1
HGI_C2_THBS3
HGI_C2_XCR1
N1LongCOVID_CCR1
N2LongCOVID_CCR1
N2LongCOVID_THBS3
W1LongCOVID_DPP9
W2LongCOVID_DPP9
W2LongCOVID_MUC1
W2LongCOVID_THBS3
MuscleSkeletal
EsophagusGastroesophagealJunction
EsophagusMuscularis
HeartLeftVentricle
Pancreas
WholeBlood
Spleen
BrainSpinalcordcervicalc-1
BrainHippocampus
BrainCaudatebasalganglia
ArteryAorta
CellsEBV-transformedlymphocytes
BrainNucleusaccumbensbasalganglia
BrainCortex
BrainAmygdala
AdrenalGland
BrainPutamenbasalganglia
ColonSigmoid
BrainCerebellum
Stomach
Liver
BrainCerebellarHemisphere
BreastMammaryTissue
ColonTransverse
SkinSunExposedLowerleg
HeartAtrialAppendage
SkinNotSunExposedSuprapubic
Vagina
Testis
SmallIntestineTerminalIleum
EsophagusMucosa
CellsTransformedfibroblasts
ArteryTibial
Thyroid
Lung
AdiposeSubcutaneous
AdiposeVisceralOmentum
Uterus
NerveTibial
P > 0.05
P < 0.05
P < 1x10-6

Figure 4

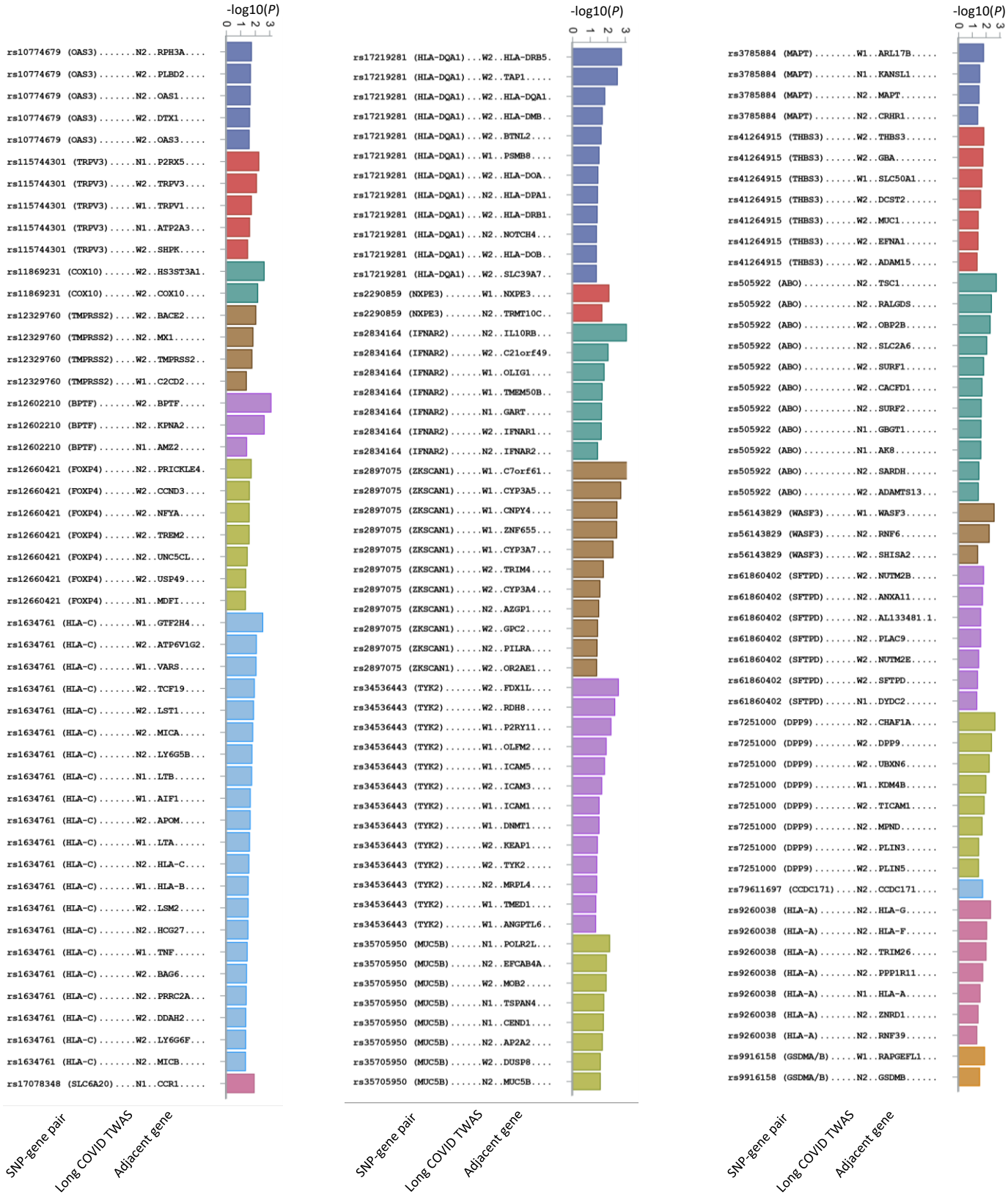

-log10(P)
0 1 2 3
rs10774679 (OAS3).....N2..RPH3A....
rs10774679 (OAS3).....W2..PLBD2....
rs10774679 (OAS3).....N2..OAS1.....
rs10774679 (OAS3).....W2..DTX1.....
rs10774679 (OAS3).....W2..OAS3.....
rs115744301 (TRPV3).....N1..P2RX5....
rs115744301 (TRPV3).....W2..TRPV3....
rs115744301 (TRPV3).....W1..TRPV1....
rs115744301 (TRPV3).....N1..ATP2A3...
rs115744301 (TRPV3).....W2..SHPK.....
rs11869231 (COX10).....W2..HS3ST3A1.
rs11869231 (COX10).....W2..COX10....
rs12329760 (TMPRSS2)....W2..BACE2....
rs12329760 (TMPRSS2)....N2..MX1.....
rs12329760 (TMPRSS2)....W2..TMPRSS2..
rs12329760 (TMPRSS2)....W1..C2CD2....
rs12602210 (BPTF).....W2..BPTF.....
rs12602210 (BPTF).....N2..KPNA2....
rs12602210 (BPTF).....N1..AMZ2.....
rs12660421 (FOXP4).....N2..PRICKLE4.
rs12660421 (FOXP4).....W2..CCND3....
rs12660421 (FOXP4).....W2..NFYA.....
rs12660421 (FOXP4).....W2..TREM2....
rs12660421 (FOXP4).....N2..UNC5CL...
rs12660421 (FOXP4).....W2..USP49....
rs12660421 (FOXP4).....N1..MDFI.....
rs1634761 (HLA-C).....W1..GTF2H4...
rs1634761 (HLA-C).....W2..ATP6V1G2.
rs1634761 (HLA-C).....W1..VARS.....
rs1634761 (HLA-C).....W2..TCF19....
rs1634761 (HLA-C).....W2..LST1.....
rs1634761 (HLA-C).....W2..MICA.....
rs1634761 (HLA-C).....N2..LY6G5B...
rs1634761 (HLA-C).....N1..LTB......
rs1634761 (HLA-C).....W1..AIF1.....
rs1634761 (HLA-C).....W2..APOM.....
rs1634761 (HLA-C).....W1..LTA......
rs1634761 (HLA-C).....N2..HLA-C....
rs1634761 (HLA-C).....W1..HLA-B....
rs1634761 (HLA-C).....W2..LSM2.....
rs1634761 (HLA-C).....N2..HCG27....
rs1634761 (HLA-C).....W1..TNF......
rs1634761 (HLA-C).....W2..BAG6.....
rs1634761 (HLA-C).....N2..PRRC2A...
rs1634761 (HLA-C).....W2..DDAH2....
rs1634761 (HLA-C).....W2..LY6G6F...
rs1634761 (HLA-C).....N2..MICB.....
rs17078348 (SLC6A20)....N1..CCR1.....
SNP-gene pair
Long COVID TWAS
Adjacent gene
-log10(P)
0 1 2 3
rs17219281 (HLA-DQA1)...W2..HLA-DRB5.
rs17219281 (HLA-DQA1)...W2..TAP1.....
rs17219281 (HLA-DQA1)...W2..HLA-DQA1.
rs17219281 (HLA-DQA1)...W2..HLA-DMB..
rs17219281 (HLA-DQA1)...W2..BTNL2....
rs17219281 (HLA-DQA1)...W1..PSMB8....
rs17219281 (HLA-DQA1)...W2..HLA-DOA..
rs17219281 (HLA-DQA1)...N2..HLA-DPA1.
rs17219281 (HLA-DQA1)...W2..HLA-DRB1.
rs17219281 (HLA-DQA1)...N2..NOTCH4...
rs17219281 (HLA-DQA1)...W2..HLA-DOB..
rs17219281 (HLA-DQA1)...W2..SLC39A7..
rs2290859 (NXPE3).....W1..NXPE3....
rs2290859 (NXPE3).....N2..TRMT10C..
rs2834164 (IFNAR2).....N2..IL10RB...
rs2834164 (IFNAR2).....W2..C21orf49.
rs2834164 (IFNAR2).....W1..OLIG1....
rs2834164 (IFNAR2).....W1..TMEM50B..
rs2834164 (IFNAR2).....N1..GART.....
rs2834164 (IFNAR2).....W2..IFNAR1...
rs2834164 (IFNAR2).....N2..IFNAR2...
rs2897075 (ZKSCAN1).....W1..C7orf61..
rs2897075 (ZKSCAN1).....W1..CYP3A5...
rs2897075 (ZKSCAN1).....W1..CNPY4....
rs2897075 (ZKSCAN1).....W1..ZNF655...
rs2897075 (ZKSCAN1).....W1..CYP3A7...
rs2897075 (ZKSCAN1).....W2..TRIM4....
rs2897075 (ZKSCAN1).....W2..CYP3A4...
rs2897075 (ZKSCAN1).....N2..AZGP1....
rs2897075 (ZKSCAN1).....W2..GPC2.....
rs2897075 (ZKSCAN1).....N2..PILRA....
rs2897075 (ZKSCAN1).....W2..OR2AE1...
rs34536443 (TYK2).....W2..FDX1L....
rs34536443 (TYK2).....W2..RDH8.....
rs34536443 (TYK2).....W1..P2RY11...
rs34536443 (TYK2).....W1..OLFM2....
rs34536443 (TYK2).....W1..ICAM5....
rs34536443 (TYK2).....W2..ICAM3....
rs34536443 (TYK2).....W1..ICAM1....
rs34536443 (TYK2).....W1..DNMT1....
rs34536443 (TYK2).....W2..KEAP1....
rs34536443 (TYK2).....W2..TYK2.....
rs34536443 (TYK2).....N2..MRPL4....
rs34536443 (TYK2).....W1..TMED1....
rs34536443 (TYK2).....W1..ANGPTL6..
rs35705950 (MUC5B).....N1..POLR2L...
rs35705950 (MUC5B).....N2..EFCAB4A..
rs35705950 (MUC5B).....W2..MOB2.....
rs35705950 (MUC5B).....N1..TSPAN4...
rs35705950 (MUC5B).....N1..CEND1....
rs35705950 (MUC5B).....N2..AP2A2....
rs35705950 (MUC5B).....W2..DUSP8....
rs35705950 (MUC5B).....N2..MUC5B....
SNP-gene pair
Long COVID TWAS
Adjacent gene
-log10(P)
0 1 2 3
rs3785884 (MAPT).....W1..ARL17B.....
rs3785884 (MAPT).....N1..KANSL1.....
rs3785884 (MAPT).....N2..MAPT.......
rs3785884 (MAPT).....N2..CRHR1......
rs41264915 (THBS3).....W2..THBS3.....
rs41264915 (THBS3).....W2..GBA.......
rs41264915 (THBS3).....W1..SLC50A1....
rs41264915 (THBS3).....W2..DCST2.....
rs41264915 (THBS3).....W2..MUC1.....
rs41264915 (THBS3).....W2..EFNA1.....
rs41264915 (THBS3).....W2..ADAM15....
rs505922 (ABO).....N2..TSC1.....
rs505922 (ABO).....N2..RALGDS.....
rs505922 (ABO).....W2..OBP2B.....
rs505922 (ABO).....N2..SLC2A6.....
rs505922 (ABO).....W2..SURF1.....
rs505922 (ABO).....W2..CACFD1.....
rs505922 (ABO).....N2..SURF2.....
rs505922 (ABO).....N1..GBGT1.....
rs505922 (ABO).....N1..AK8.....
rs505922 (ABO).....N2..SARDH.....
rs505922 (ABO).....W2..ADAMTS13...
rs56143829 (WASF3).....W1..WASF3.....
rs56143829 (WASF3).....N2..RNF6.....
rs56143829 (WASF3).....W2..SHISA2....
rs61860402 (SFTPD).....W2..NUTM2B....
rs61860402 (SFTPD).....N2..ANXA11....
rs61860402 (SFTPD).....N2..AL133481.1.
rs61860402 (SFTPD).....N2..PLAC9.....
rs61860402 (SFTPD).....W2..NUTM2E....
rs61860402 (SFTPD).....W2..SFTPD.....
rs61860402 (SFTPD).....N1..DYDC2.....
rs7251000 (DPP9).....N2..CHAF1A....
rs7251000 (DPP9).....W2..DPP9.....
rs7251000 (DPP9).....W2..UBXN6.....
rs7251000 (DPP9).....W1..KDM4B.....
rs7251000 (DPP9).....W2..TICAM1....
rs7251000 (DPP9).....N2..MPND.....
rs7251000 (DPP9).....W2..PLIN3.....
rs7251000 (DPP9).....W2..PLIN5.....
rs79611697 (CCDC171)....N2..CCDC171....
rs9260038 (HLA-A).....N2..HLA-G.....
rs9260038 (HLA-A).....N2..HLA-F.....
rs9260038 (HLA-A).....N2..TRIM26....
rs9260038 (HLA-A).....N2..PPP1R11....
rs9260038 (HLA-A).....N1..HLA-A.....
rs9260038 (HLA-A).....N2..ZNRD1.....
rs9260038 (HLA-A).....N2..RNF39.....
rs9916158 (GSDMA/B).....W1..RAPGEFL1...
rs9916158 (GSDMA/B).....N2..GSDMB.....
SNP-gene pair
Long COVID TWAS
Adjacent gene

Figure S1

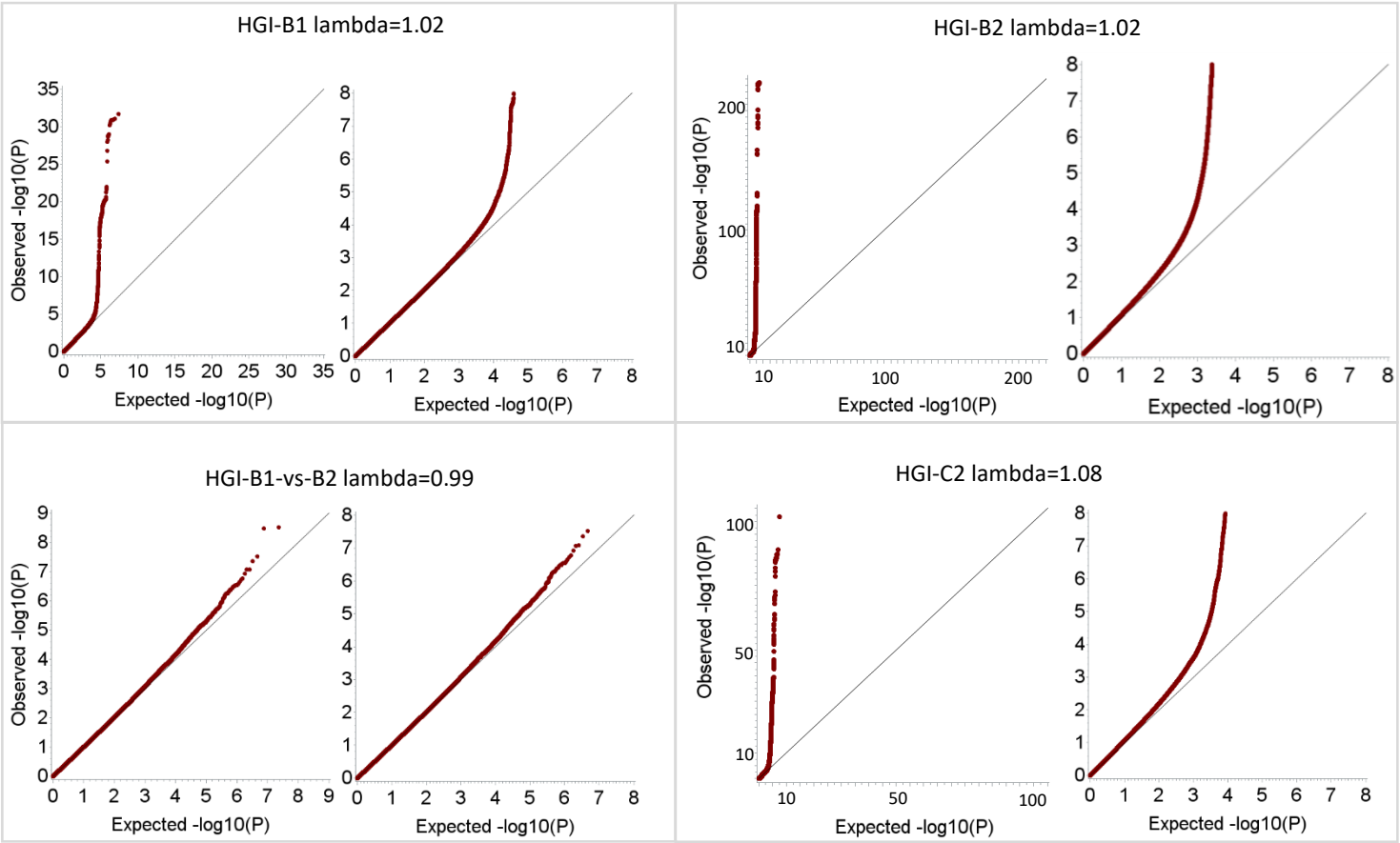

HGI-B1 lambda=1.02
Observed -log10(P)
Expected -log10(P)
Expected -log10(P)
HGI-B2 lambda=1.02
Observed -log10(P)
Expected -log10(P)
Expected -log10(P)
HGI-B1-vs-B2 lambda=0.99
Observed -log10(P)
Expected -log10(P)
Expected -log10(P)
HGI-C2 lambda=1.08
Observed -log10(P)
Expected -log10(P)
Expected -log10(P)

Figure S2

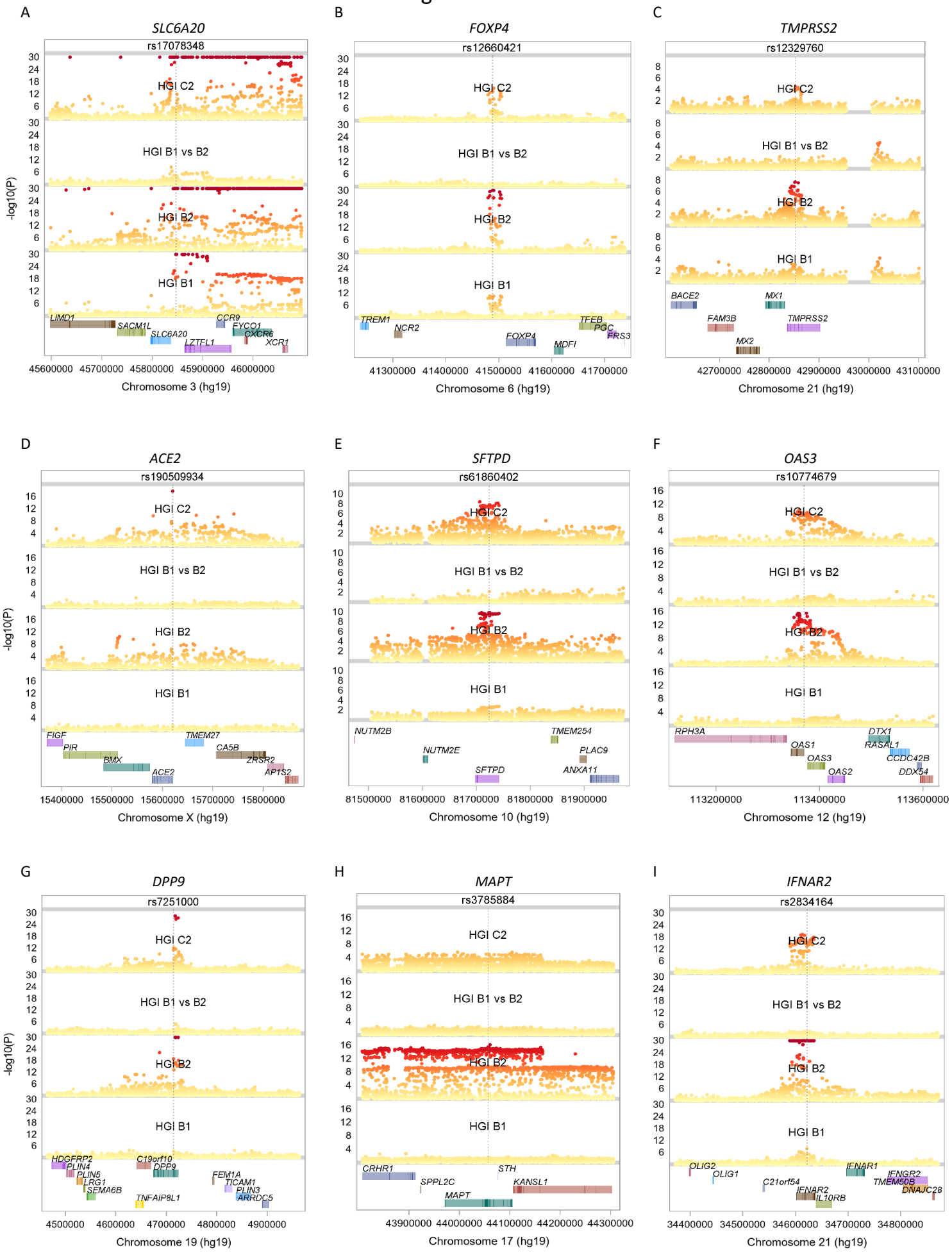

A
SLC6A20
rs17078348
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
LIMD1
SACM1L
SLC6A20
LZTFL1
CCR9
FYCO1
CXCR6
XCR1
45600000 45700000 45800000 45900000 46000000
Chromosome 3 (hg19)
B
FOXP4
rs12660421
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
TREM1
NCR2
FOXP4
MDFI
TFEB
PGC
FRS3
41300000 41400000 41500000 41600000 41700000
Chromosome 6 (hg19)
C
TMPRSS2
rs12329760
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
BACE2
MX1
FAM3B
TMPRSS2
MX2
42700000 42800000 42900000 43000000 43100000
Chromosome 21 (hg19)
D
ACE2
rs190509934
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
FIGF
PIR
BMX
ACE2
TMEM27
CA5B
ZRSR2
AP1S2
15400000 15500000 15600000 15700000 15800000
Chromosome X (hg19)
E
SFTPD
rs61860402
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
NUTM2B
NUTM2E
SFTPD
TMEM254
PLAC9
ANXA11
81500000 81600000 81700000 81800000 81900000
Chromosome 10 (hg19)
F
OAS3
rs10774679
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
RPH3A
OAS1
OAS3
OAS2
DTX1
RASAL1
CCDC42B
DDX54
113200000 113400000 113600000
Chromosome 12 (hg19)
G
DPP9
rs7251000
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
HDGFRP2
PLIN4
PLIN5
LRG1
SEMA6B
C19orf10
DPP9
TNFAIP8L1
FEM1A
TICAM1
PLIN3
ARRDC5
4500000 4600000 4700000 4800000 4900000
Chromosome 19 (hg19)
H
MAPT
rs3785884
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
CRHR1
SPPL2C
MAPT
STH
KANSL1
43900000 44000000 44100000 44200000 44300000
Chromosome 17 (hg19)
I
IFNAR2
rs2834164
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
OLIG2
OLIG1
C21orf54
IFNAR2
IL10RB
IFNAR1
IFNGR2
TMEM50B
DNAJC28
34400000 34500000 34600000 34700000 34800000
Chromosome 21 (hg19)

Figure S3

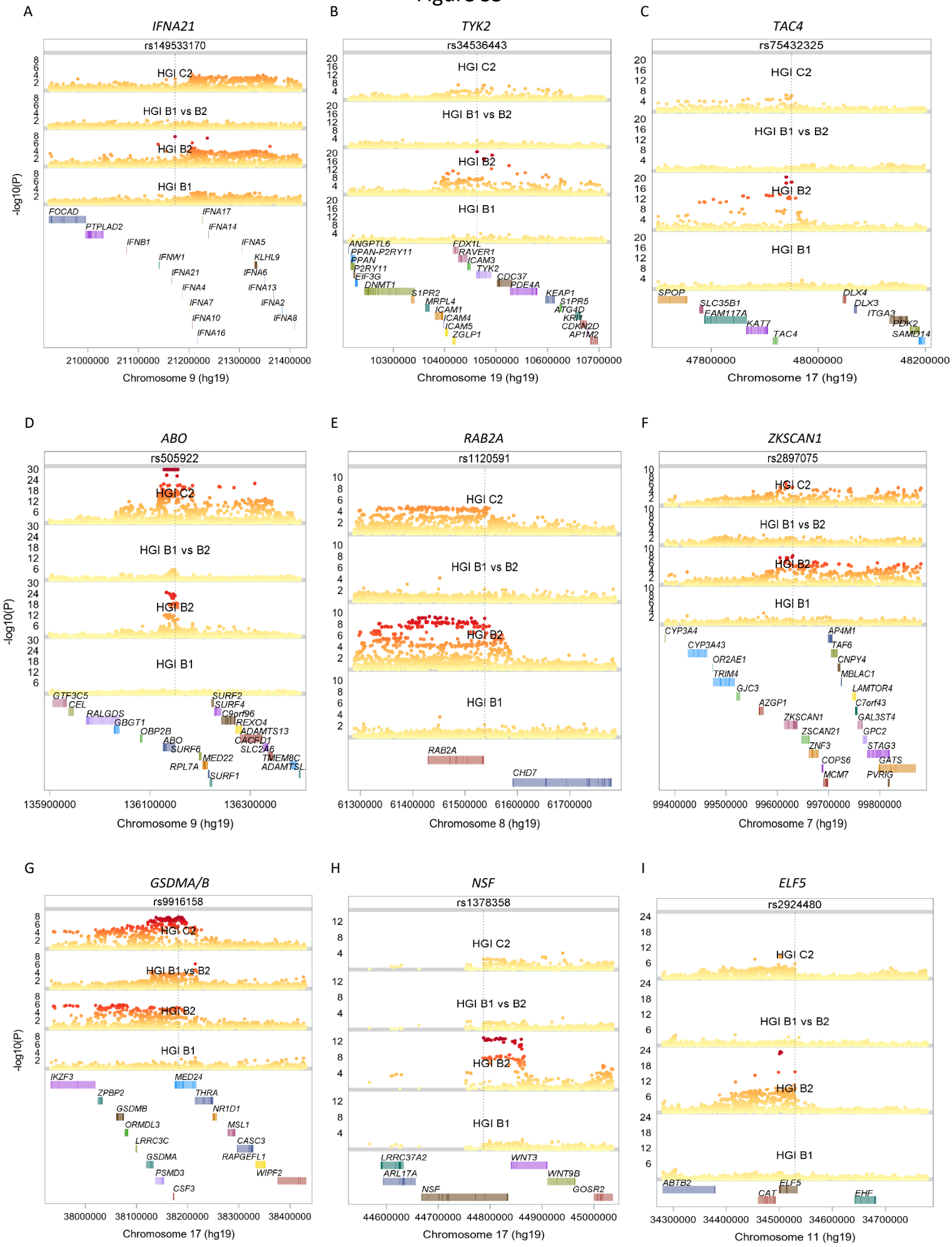

A
IFNA21
rs149533170
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
FOCAD
PTPLAD2
IFNB1
IFNW1
IFNA21
IFNA4
IFNA7
IFNA10
IFNA16
IFNA17
IFNA14
IFNA5
KLHL9
IFNA6
IFNA13
IFNA2
IFNA8
21000000 21100000 21200000 21300000 21400000
Chromosome 9 (hg19)
B
TYK2
rs34536443
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
ANGPTL6
PPAN-P2RY11
PPAN
P2RY11
EIF3G
DNMT1
S1PR2
MRPL4
ICAM1
ICAM4
ICAM5
ZGLP1
FDX1L
RAVER1
ICAM3
TYK2
CDC37
PDE4A
KEAP1
S1PR5
ATG4D
KRR1
CDKN2D
AP1M2
10300000 10400000 10500000 10600000 10700000
Chromosome 19 (hg19)
C
TAC4
rs75432325
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
SPOP
SLC35B1
FAM117A
KAT7
TAC4
DLX4
DLX3
ITGA3
PDK2
SAMD14
47800000 48000000 48200000
Chromosome 17 (hg19)
D
ABO
rs505922
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
GTF3C5
CEL
RALGDS
GBGT1
OBP2B
ABO
SURF6
MED22
RPL7A
SURF1
SURF2
SURF4
C9orf96
REXO4
ADAMTS13
CACFD1
SLC2A6
TMEM8C
ADAMTSL
135900000 136100000 136300000
Chromosome 9 (hg19)
E
RAB2A
rs1120591
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
RAB2A
CHD7
61300000 61400000 61500000 61600000 61700000
Chromosome 8 (hg19)
F
ZKSCAN1
rs2897075
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
CYP3A4
CYP3A43
OR2AE1
TRIM4
GJC3
AZGP1
ZKSCAN1
ZSCAN21
ZNF3
COPS6
MCM7
AP4M1
TAF6
CNPY4
MBLAC1
LAMTOR4
C7orf43
GAL3ST4
GPC2
STAG3
GATS
PVRIG
99400000 99500000 99600000 99700000 99800000
Chromosome 7 (hg19)
G
GSDMA/B
rs9916158
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
IKZF3
ZPBP2
GSDMB
ORMDL3
LRRC3C
GSDMA
PSMD3
CSF3
MED24
THRA
NR1D1
MSL1
CASC3
RAPGEFL1
WIPF2
38000000 38100000 38200000 38300000 38400000
Chromosome 17 (hg19)
H
NSF
rs1378358
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
LRRC37A2
ARL17A
NSF
WNT3
WNT9B
GOSR2
44600000 44700000 44800000 44900000 45000000
Chromosome 17 (hg19)
I
ELF5
rs2924480
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
ABTB2
CAT
ELF5
EHF
34300000 34400000 34500000 34600000 34700000
Chromosome 11 (hg19)

Figure S4

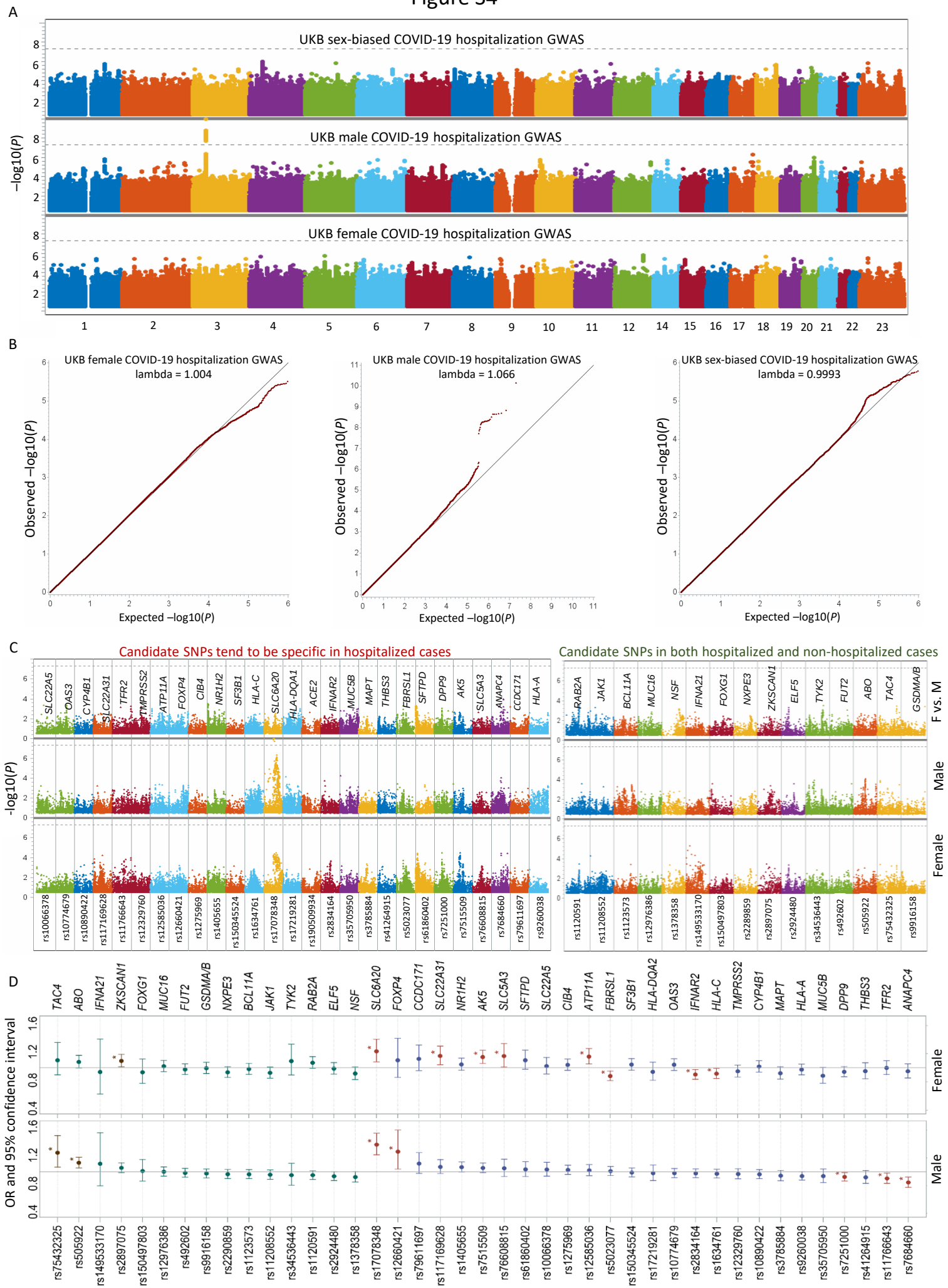

A
UKB sex-biased COVID-19 hospitalization GWAS
UKB male COVID-19 hospitalization GWAS
UKB female COVID-19 hospitalization GWAS
−log10(P)
B
UKB female COVID-19 hospitalization GWAS
lambda = 1.004
UKB male COVID-19 hospitalization GWAS
lambda = 1.066
UKB sex-biased COVID-19 hospitalization GWAS
lambda = 0.9993
Observed −log10(P)
Expected −log10(P)
C
Candidate SNPs tend to be specific in hospitalized cases
Candidate SNPs in both hospitalized and non-hospitalized cases
-log10(P)
F vs. M
Male
Female
D
OR and 95% confidence interval
Female
Male

Figure S5

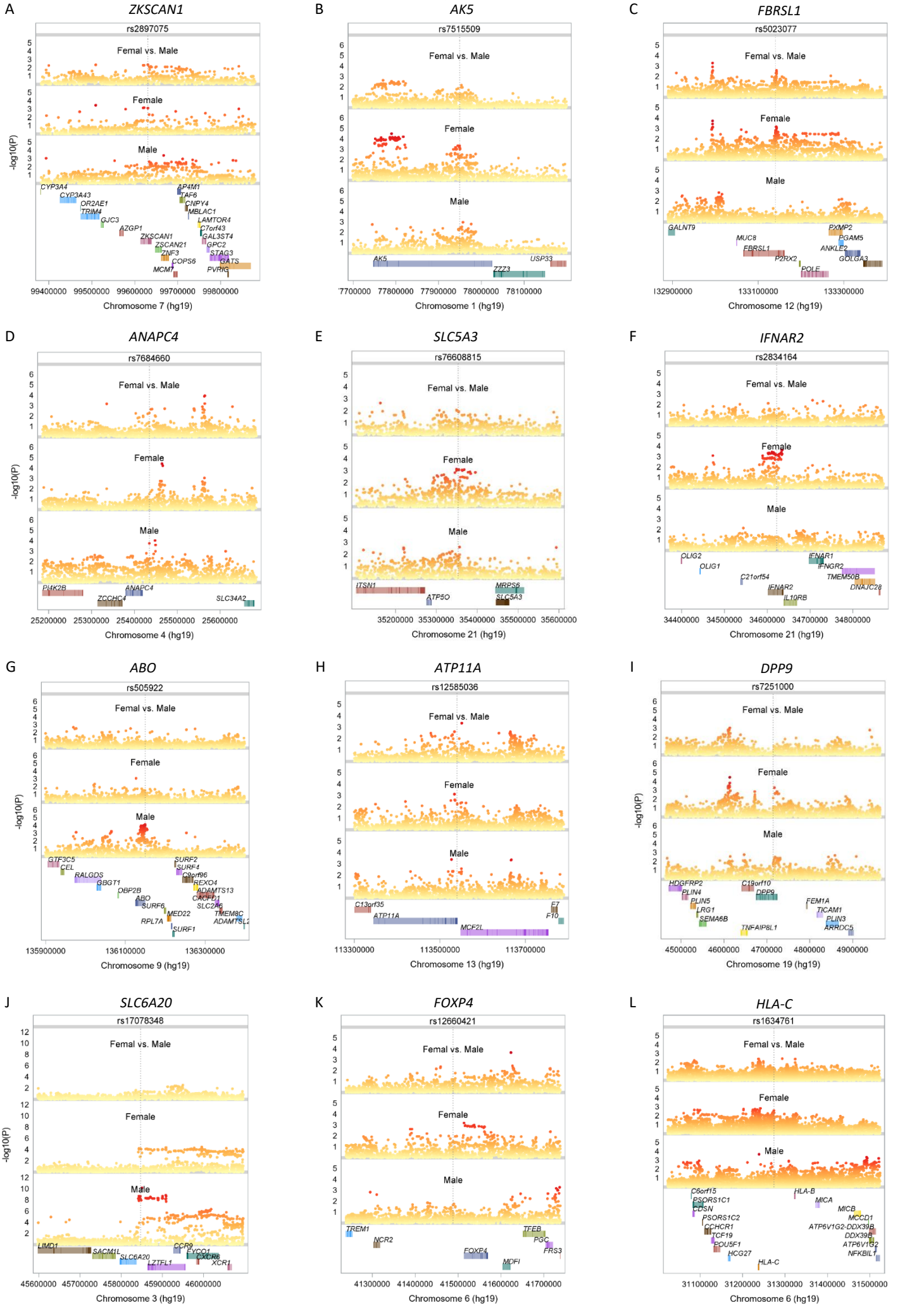

A
ZKSCAN1
rs2897075
Femal vs. Male
Female
Male
-log10(P)
CYP3A4
CYP3A43
OR2AE1
TRIM4
GJC3
AZGP1
ZKSCAN1
ZSCAN21
ZNF3
COPS6
MCM7
AP4M1
TAF6
CNPY4
MBLAC1
LAMTOR4
C7orf43
GAL3ST4
GPC2
STAG3
GATS
PVRIG
99400000 99500000 99600000 99700000 99800000
Chromosome 7 (hg19)
B
AK5
rs7515509
Femal vs. Male
Female
Male
AK5
USP33
ZZZ3
77700000 77800000 77900000 78000000 78100000
Chromosome 1 (hg19)
C
FBRSL1
rs5023077
Femal vs. Male
Female
Male
GALNT9
MUC8
FBRSL1
P2RX2
PXMP2
PGAM5
ANKLE2
GOLGA3
POLE
132900000 133100000 133300000
Chromosome 12 (hg19)
D
ANAPC4
rs7684660
Femal vs. Male
Female
Male
-log10(P)
PI4K2B
ZCCHC4
ANAPC4
SLC34A2
25200000 25300000 25400000 25500000 25600000
Chromosome 4 (hg19)
E
SLC5A3
rs76608815
Femal vs. Male
Female
Male
ITSN1
ATP5O
MRPS6
SLC5A3
35200000 35300000 35400000 35500000 35600000
Chromosome 21 (hg19)
F
IFNAR2
rs2834164
Femal vs. Male
Female
Male
OLIG2
OLIG1
C21orf54
IFNAR2
IFNAR1
IFNGR2
TMEM50B
DNAJC28
IL10RB
34400000 34500000 34600000 34700000 34800000
Chromosome 21 (hg19)
G
ABO
rs505922
Femal vs. Male
Female
Male
-log10(P)
GTF3C5
CEL
RALGDS
GBGT1
DBP2B
ABO
SURF6
RPL7A
MED22
SURF1
SURF2
SURF4
C9orf96
REXO4
ADAMTS13
CACFD1
SLC2A6
TMEM8C
ADAMTSL2
135900000 136100000 136300000
Chromosome 9 (hg19)
H
ATP11A
rs12585036
Femal vs. Male
Female
Male
C13orf35
ATP11A
MCF2L
F7
F10
113300000 113500000 113700000
Chromosome 13 (hg19)
I
DPP9
rs7251000
Femal vs. Male
Female
Male
-log10(P)
HDGFRP2
PLIN4
PLIN5
LRG1
SEMA6B
C19orf10
DPP9
TNFAIP8L1
FEM1A
TICAM1
PLIN3
ARRDC5
4500000 4600000 4700000 4800000 4900000
Chromosome 19 (hg19)
J
SLC6A20
rs17078348
Femal vs. Male
Female
Male
-log10(P)
LIMD1
SACM1L
SLC6A20
LZTFL1
CCR9
CXCR6
XCR1
45600000 45700000 45800000 45900000 46000000
Chromosome 3 (hg19)
K
FOXP4
rs12660421
Femal vs. Male
Female
Male
TREM1
NCR2
FOXP4
MDFI
TFEB
PGC
FRS3
41300000 41400000 41500000 41600000 41700000
Chromosome 6 (hg19)
L
HLA-C
rs1634761
Femal vs. Male
Female
Male
C6orf15
PSORS1C1
CDSN
PSORS1C2
CCHCR1
TCF19
POU5F1
HCG27
HLA-C
HLA-B
MICA
MICB
MCCD1
ATP6V1G2-DDX39B
DDX39B
ATP6V1G2
NFKBIL1
31100000 31200000 31300000 31400000 31500000
Chromosome 6 (hg19)

Figure S6

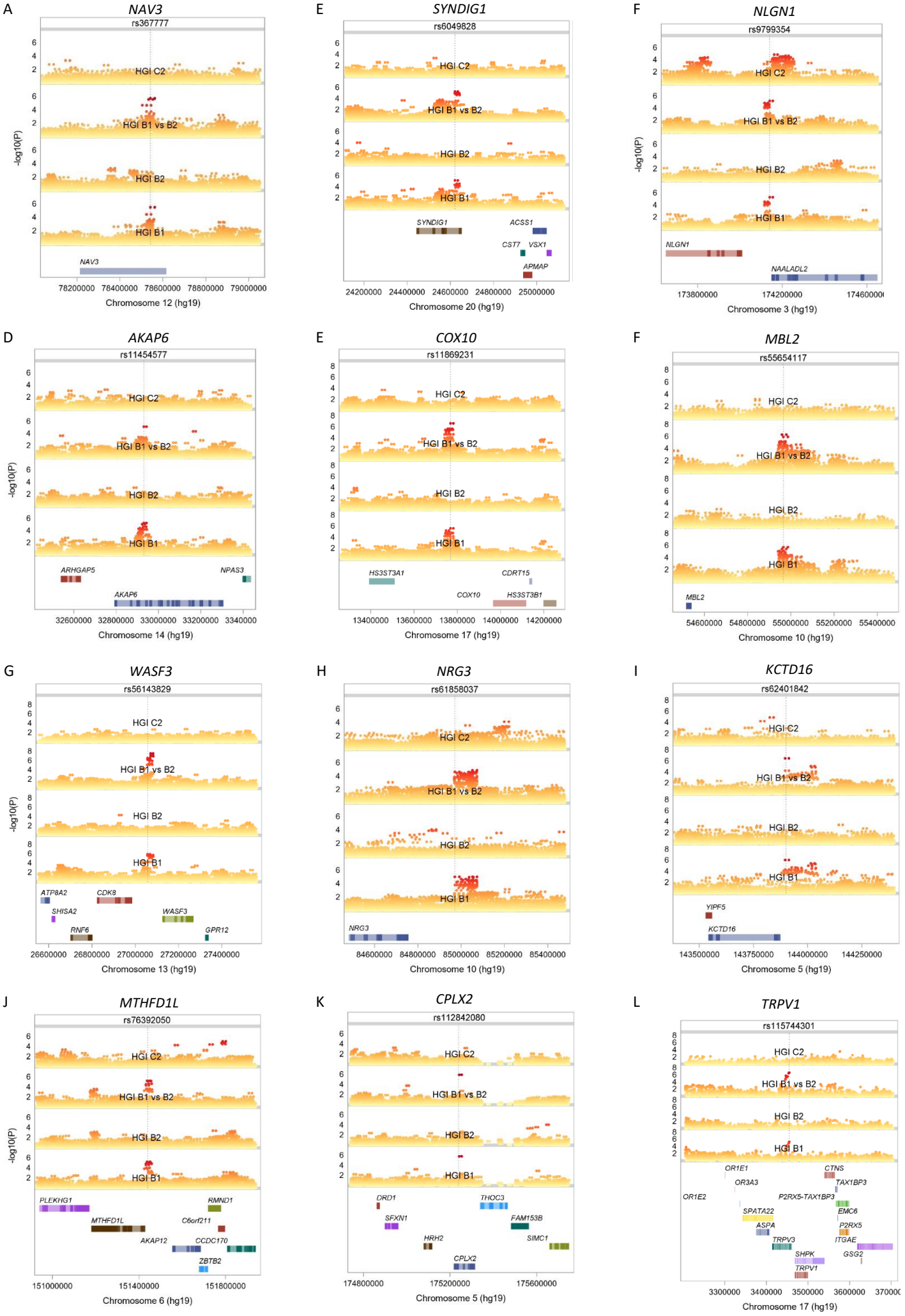

A
NAV3
rs367777
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
NAV3
78200000 78400000 78600000 78800000 79000000
Chromosome 12 (hg19)
E
SYNDIG1
rs6049828
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
SYNDIG1
ACSS1
CST7
VSX1
APMAP
24200000 24400000 24600000 24800000 25000000
Chromosome 20 (hg19)
F
NLGN1
rs9799354
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
NLGN1
NAALADL2
173800000 174200000
Chromosome 3 (hg19)
D
AKAP6
rs11454577
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
ARHGAP5
NPAS3
AKAP6
32600000 32800000 33000000 33200000 3340000
Chromosome 14 (hg19)
E
COX10
rs11869231
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
HS3ST3A1
CDRT15
COX10
HS3ST3B1
13400000 13600000 13800000 14000000 14200000
Chromosome 17 (hg19)
F
MBL2
rs55654117
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
MBL2
54600000 54800000 55000000 55200000 55400000
Chromosome 10 (hg19)
G
WASF3
rs56143829
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
ATP8A2
CDK8
SHISA2
WASF3
RNF6
GPR12
26600000 26800000 27000000 27200000 27400000
Chromosome 13 (hg19)
H
NRG3
rs61858037
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
NRG3
84600000 84800000 85000000 85200000 85400000
Chromosome 10 (hg19)
I
KCTD16
rs62401842
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
YIPF5
KCTD16
143500000 143750000 144000000 144250000
Chromosome 5 (hg19)
J
MTHFD1L
rs76392050
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
PLEKHG1
RMND1
MTHFD1L
C6orf211
AKAP12
CCDC170
ZBTB2
151000000 151400000 151800000
Chromosome 6 (hg19)
K
CPLX2
rs112842080
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
DRD1
THOC3
SFXN1
FAM153B
HRH2
SIMC1
CPLX2
174600000 175200000 175600000
Chromosome 5 (hg19)
L
TRPV1
rs115744301
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
OR1E1
CTNS
OR3A3
TAX1BP3
OR1E2
P2RX5-TAX1BP3
SPATA22
EMC6
ASPA
P2RX5
TRPV3
ITGAE
SHPK
GSG2
TRPV1
3300000 3400000 3500000 3600000
Chromosome 17 (hg19)

Figure S7

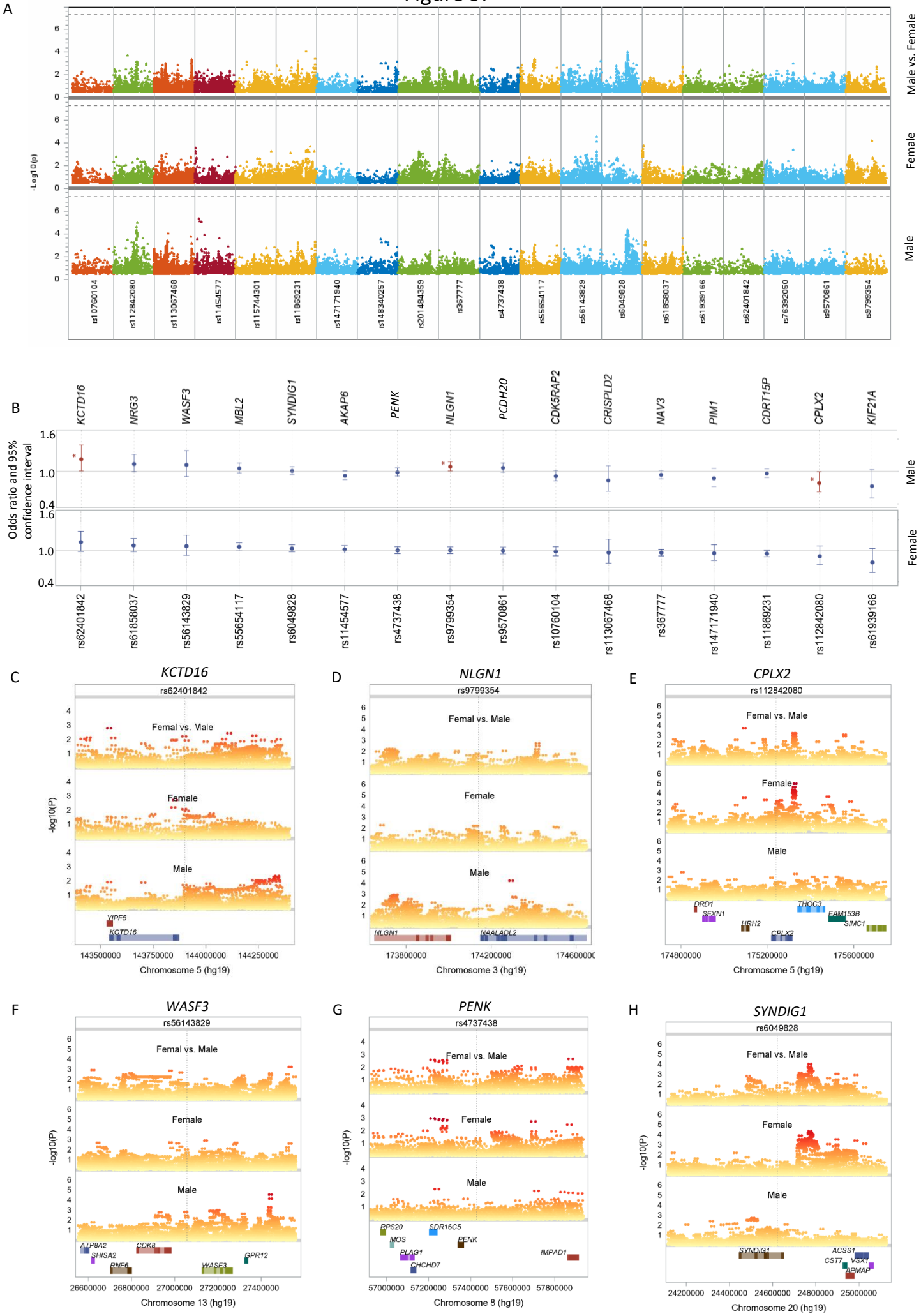
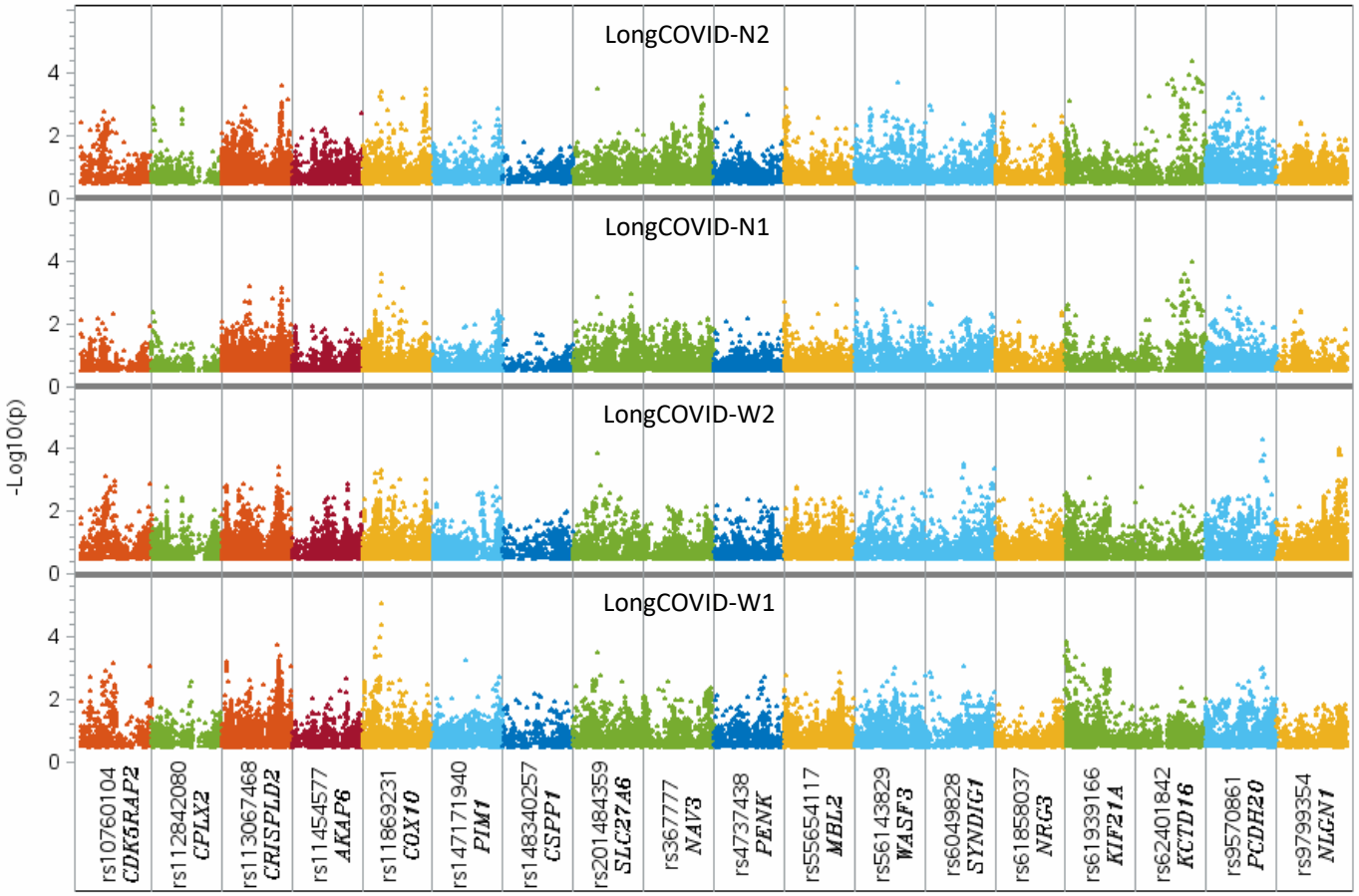

A
-Log10(p)
Male vs. Female
Female
Male
rs10760104
rs112842080
rs113067468
rs11454577
rs115744301
rs11869231
rs147171940
rs148340257
rs201484359
rs367777
rs4737438
rs55654117
rs56143829
rs6049828
rs61858037
rs61939166
rs62401842
rs76392050
rs9570861
rs9799354
B
KCTD16
NRG3
WASF3
MBL2
SYNDIG1
AKAP6
PENK
NLGN1
PCDH20
CDK5RAP2
CRISPLD2
NAV3
PIM1
CDRT15P
CPLX2
KIF21A
Odds ratio and 95% confidence interval
Male
Female
rs62401842
rs61858037
rs56143829
rs55654117
rs6049828
rs11454577
rs4737438
rs9799354
rs9570861
rs10760104
rs113067468
rs367777
rs147171940
rs11869231
rs112842080
rs61939166
C
KCTD16
rs62401842
Femal vs. Male
Female
Male
-log10(P)
YIPF5
KCTD16
Chromosome 5 (hg19)
D
NLGN1
rs9799354
Femal vs. Male
Female
Male
NLGN1
NAALADL2
Chromosome 3 (hg19)
E
CPLX2
rs112842080
Femal vs. Male
Female
Male
DRD1
SFXN1
HRH2
THOC3
FAM153B
SIMC1
CPLX2
Chromosome 5 (hg19)
F
WASF3
rs56143829
Femal vs. Male
Female
Male
-log10(P)
ATP8A2
SHISA2
CDK8
RNF6
WASF3
GPR12
Chromosome 13 (hg19)
G
PENK
rs4737438
Femal vs. Male
Female
Male
-log10(P)
RPS20
SDR16C5
MOS
PENK
PLAG1
CHCHD7
IMPAD1
Chromosome 8 (hg19)
H
SYNDIG1
rs6049828
Femal vs. Male
Female
Male
-log10(P)
SYNDIG1
ACSS1
CST7
VSX1
APMAP
Chromosome 20 (hg19)

Figure S8

Long COVID association signals for SNPs in the SNP category “Severe COVID-19-Specific SNPs”

Figure S9

Long COVID association signals for SNPs in the SNP category "Mild COVID-19-Specific SNPs"



Long COVID association signals for SNPs in the SNP category "SNPs Associated with Both Severe and Mild COVID-19"

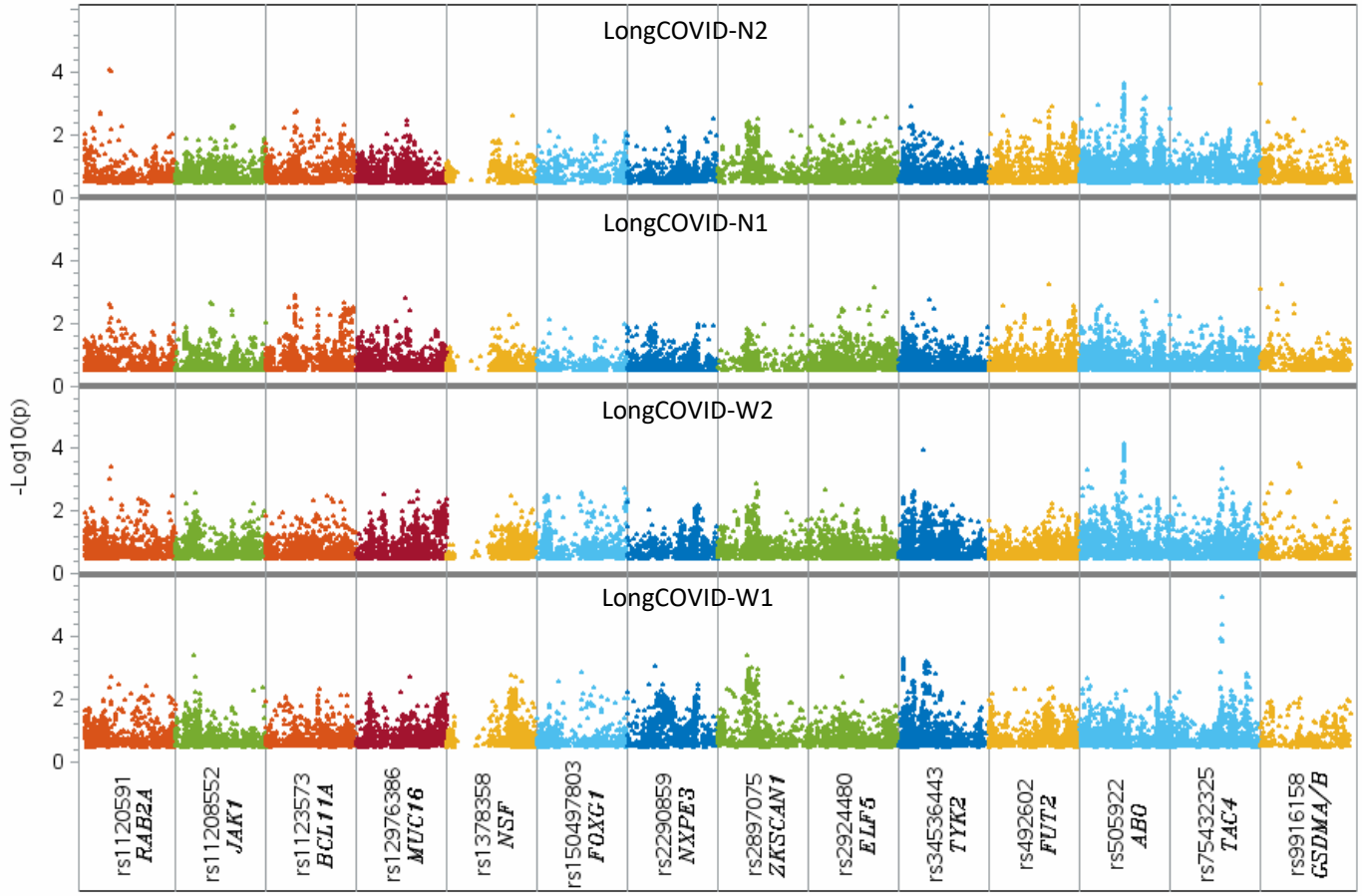


Figure S10

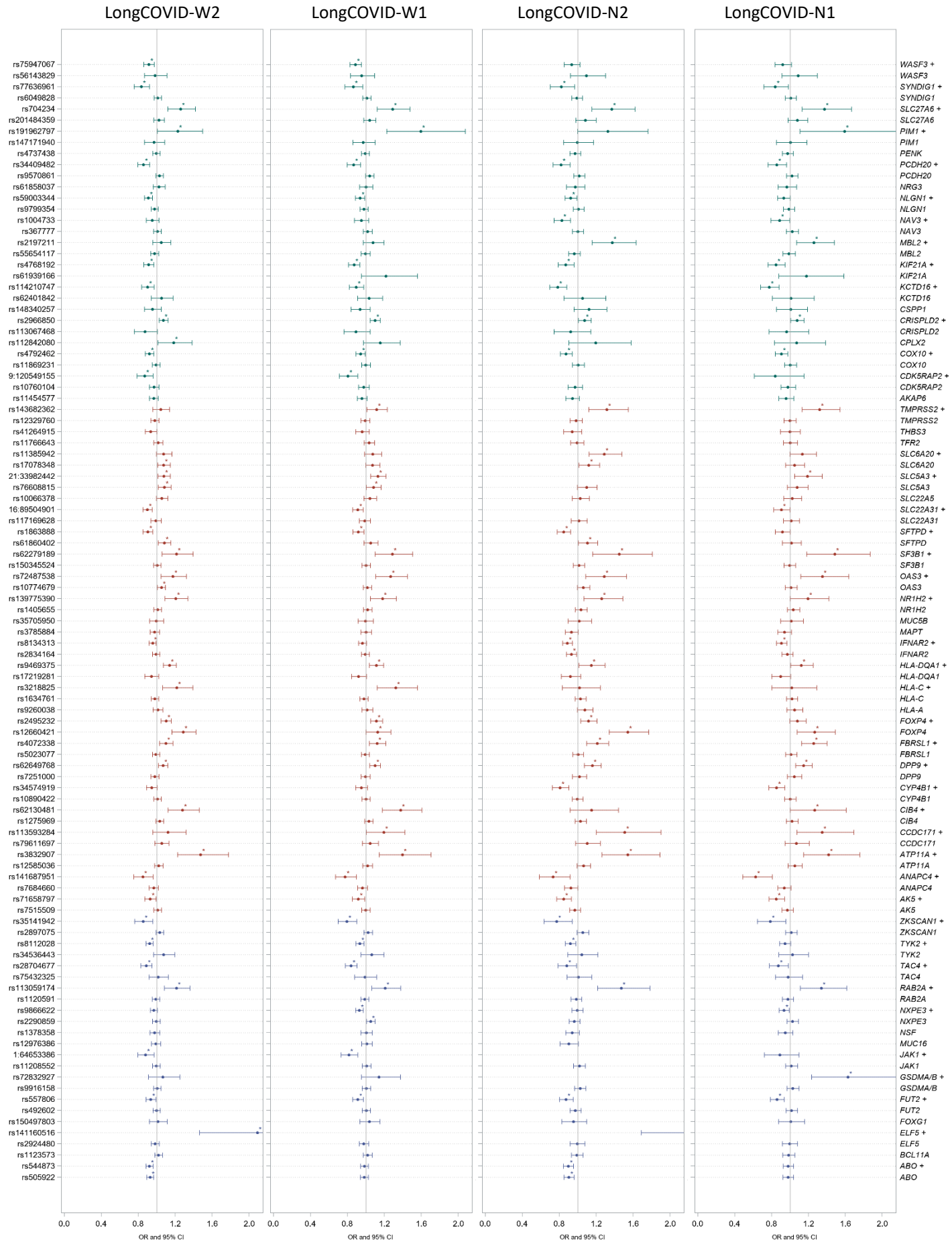

LongCOVID-W2
LongCOVID-W1
LongCOVID-N2
LongCOVID-N1
rs75947067 WASF3 +
rs56143829 WASF3
rs77636961 SYNDIG1 +
rs6049828 SYNDIG1
rs704234 SLC27A6 +
rs201484359 SLC27A6
rs191962797 PIM1 +
rs147171940 PIM1
rs4737438 PENK
rs34409482 PCDH20 +
rs9570861 PCDH20
rs61858037 NRG3
rs59003344 NLGN1 +
rs9799354 NLGN1
rs1004733 NAV3 +
rs367777 NAV3
rs2197211 MBL2 +
rs55654117 MBL2
rs4768192 KIF21A +
rs61939166 KIF21A
rs114210747 KCTD16 +
rs62401842 KCTD16
rs148340257 CSPP1
rs2966850 CRISPLD2 +
rs113067468 CRISPLD2
rs112842080 CPLX2
rs4792462 COX10 +
rs11869231 COX10
9:120549155 CDK5RAP2 +
rs10760104 CDK5RAP2
rs11454577 AKAP6
rs143682362 TMPRSS2 +
rs12329760 TMPRSS2
rs41264915 THBS3
rs11766643 TFR2
rs11385942 SLC6A20 +
rs17078348 SLC6A20
21:33982442 SLC5A3 +
rs76608815 SLC5A3
rs10066378 SLC22A5
16:89504901 SLC22A31 +
rs117169628 SLC22A31
rs1863888 SFTPD +
rs61860402 SFTPD
rs62279189 SF3B1 +
rs150345524 SF3B1
rs72487538 OAS3 +
rs10774679 OAS3
rs139775390 NR1H2 +
rs1405655 NR1H2
rs35705950 MUC5B
rs3785884 MAPT
rs8134313 IFNAR2 +
rs2834164 IFNAR2
rs9469375 HLA-DQA1 +
rs17219281 HLA-DQA1
rs3218825 HLA-C +
rs1634761 HLA-C
rs9260038 HLA-A
rs2495232 FOXP4 +
rs12660421 FOXP4
rs4072338 FBRSL1 +
rs5023077 FBRSL1
rs62649768 DPP9 +
rs7251000 DPP9
rs34574919 CYP4B1 +
rs10890422 CYP4B1
rs62130481 CIB4 +
rs1275969 CIB4
rs113593284 CCDC171 +
rs79611697 CCDC171
rs3832907 ATP11A +
rs12585036 ATP11A
rs141687951 ANAPC4 +
rs7684660 ANAPC4
rs71658797 AK5 +
rs7515509 AK5
rs35141942 ZKSCAN1 +
rs2897075 ZKSCAN1
rs8112028 TYK2 +
rs34536443 TYK2
rs28704677 TAC4 +
rs75432325 TAC4
rs113059174 RAB2A +
rs1120591 RAB2A
rs9866622 NXPE3 +
rs2290859 NXPE3
rs1378358 NSF
rs12976386 MUC16
1:64653386 JAK1 +
rs11208552 JAK1
rs72832927 GSDMA/B +
rs9916158 GSDMA/B
rs557806 FUT2 +
rs492602 FUT2
rs150497803 FOXG1
rs141160516 ELF5 +
rs2924480 ELF5
rs1123573 BCL11A
rs544873 ABO +
rs505922 ABO
0.0 0.4 0.8 1.2 1.6 2.0
OR and 95% CI

Figure S11

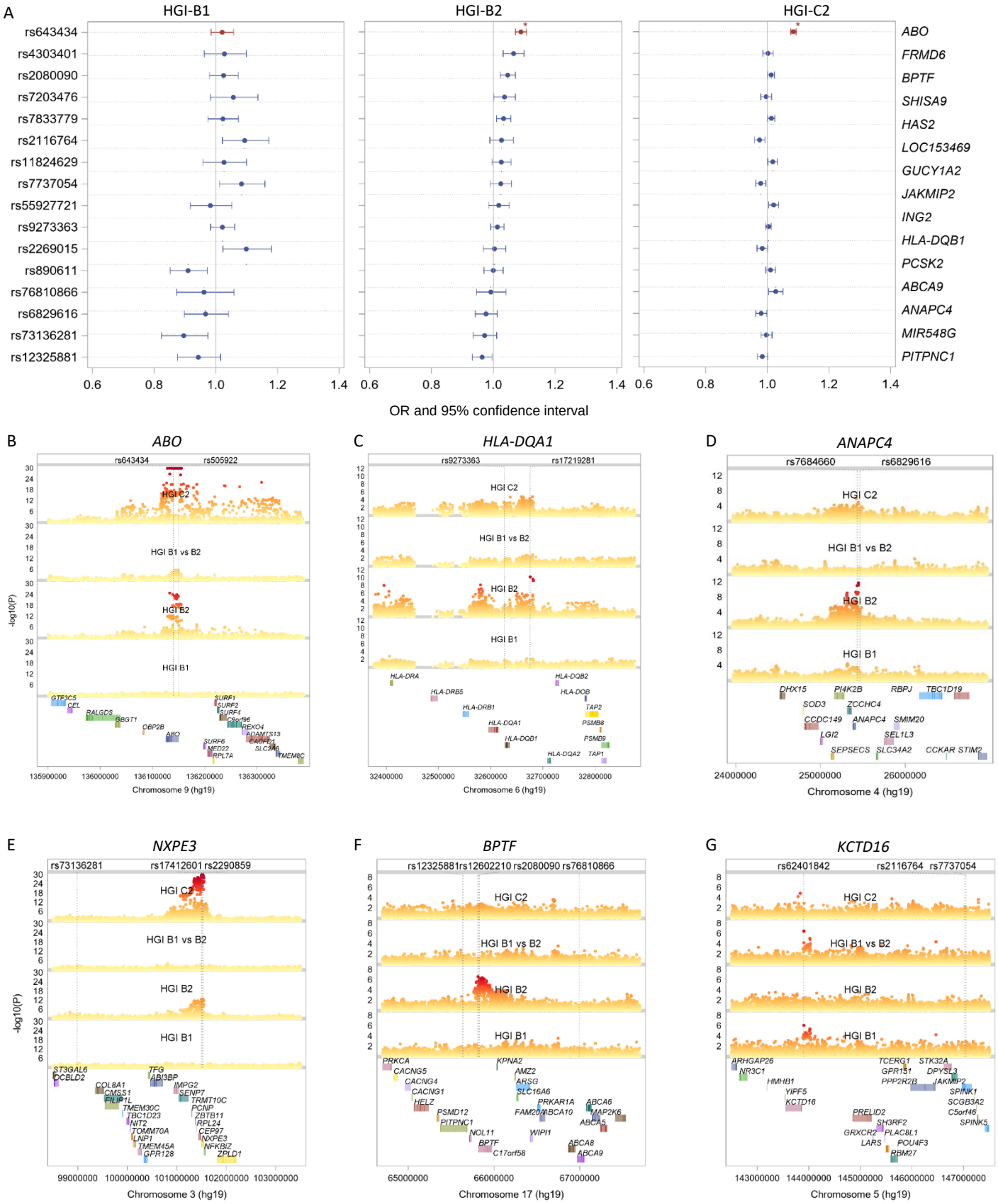

A
HGI-B1
HGI-B2
HGI-C2
rs643434
rs4303401
rs2080090
rs7203476
rs7833779
rs2116764
rs11824629
rs7737054
rs55927721
rs9273363
rs2269015
rs890611
rs76810866
rs6829616
rs73136281
rs12325881
ABO
FRMD6
BPTF
SHISA9
HAS2
LOC153469
GUCY1A2
JAKMIP2
ING2
HLA-DQB1
PCSK2
ABCA9
ANAPC4
MIR548G
PITPNC1
0.6 0.8 1.0 1.2 1.4
OR and 95% confidence interval
B
ABO
rs643434 rs505922
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
Chromosome 9 (hg19)
C
HLA-DQA1
rs9273363 rs17219281
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
Chromosome 6 (hg19)
D
ANAPC4
rs7684660 rs6829616
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
Chromosome 4 (hg19)
E
NXPE3
rs73136281 rs17412601 rs2290859
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
Chromosome 3 (hg19)
F
BPTF
rs12325881 rs12602210 rs2080090 rs76810866
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
Chromosome 17 (hg19)
G
KCTD16
rs62401842 rs2116764 rs7737054
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
Chromosome 5 (hg19)

Figure S12

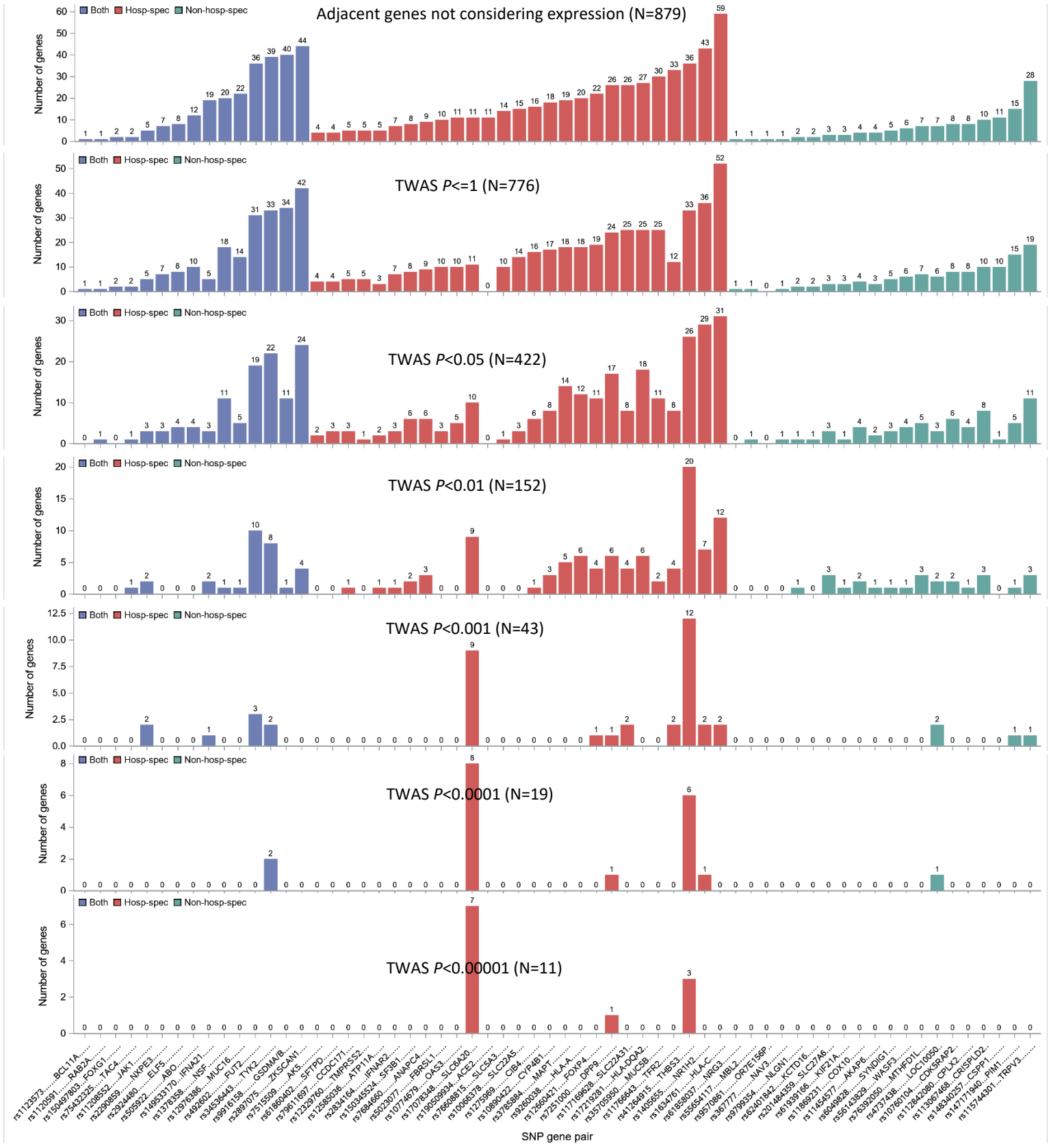

Both
Hosp-spec
Non-hosp-spec
Adjacent genes not considering expression (N=879)
TWAS P<=1 (N=776)
TWAS P<0.05 (N=422)
TWAS P<0.01 (N=152)
TWAS P<0.001 (N=43)
TWAS P<0.0001 (N=19)
TWAS P<0.00001 (N=11)
Number of genes
SNP gene pair

Figure S13

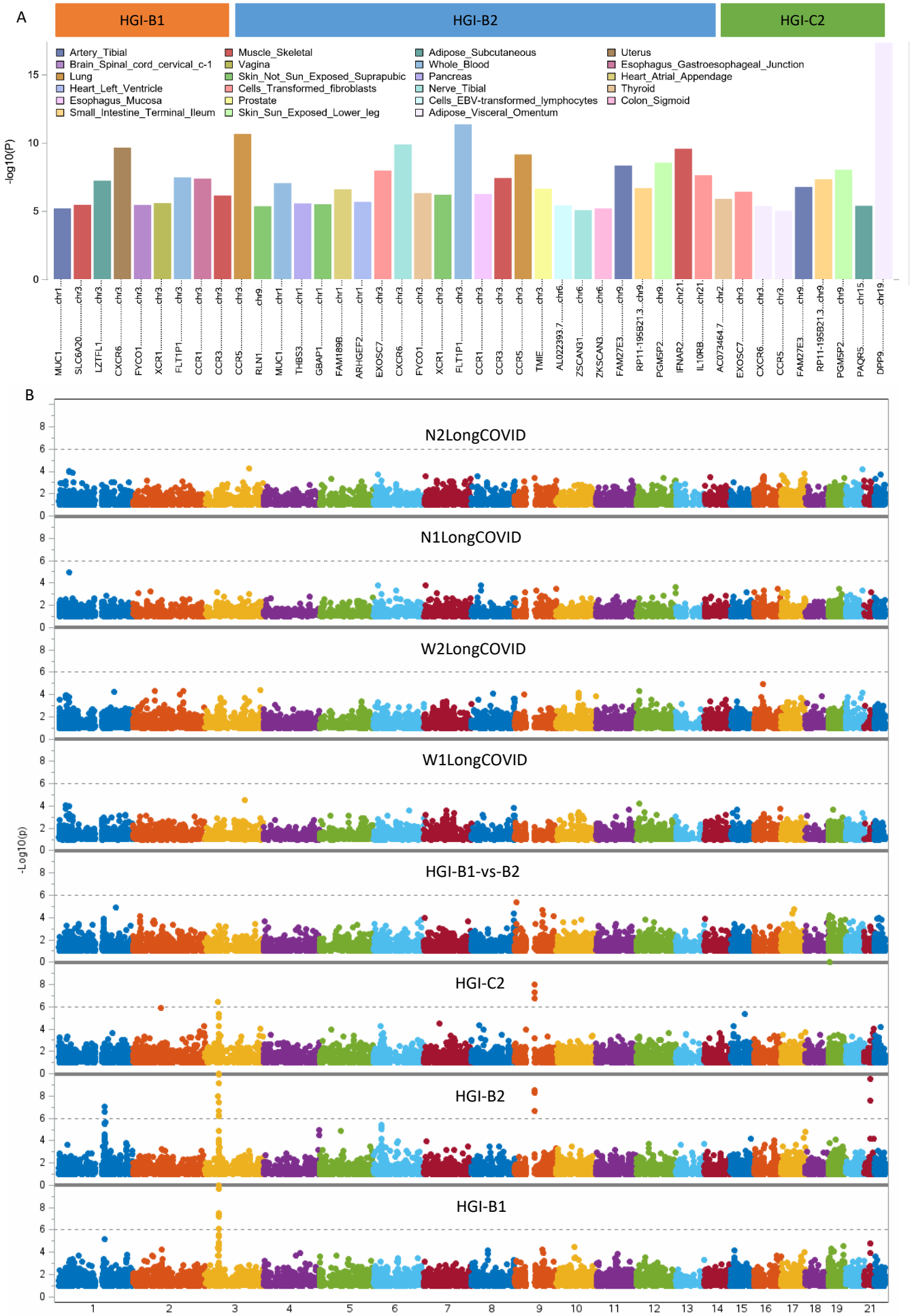

A
HGI-B1
HGI-B2
HGI-C2
Artery_Tibial
Brain_Spinal_cord_cervical_c-1
Lung
Heart_Left_Ventricle
Esophagus_Mucosa
Small_Intestine_Terminal_Ileum
Muscle_Skeletal
Vagina
Skin_Not_Sun_Exposed_Suprapubic
Cells_Transformed_fibroblasts
Prostate
Skin_Sun_Exposed_Lower_leg
Adipose_Subcutaneous
Whole_Blood
Pancreas
Nerve_Tibial
Cells_EBV-transformed_lymphocytes
Adipose_Visceral_Omentum
Uterus
Esophagus_Gastroesophageal_Junction
Heart_Atrial_Appendage
Thyroid
Colon_Sigmoid
-log10(P)
MUC1
SLC6A20
LZTFL1
CXCR6
FYCO1
XCR1
FLT1P1
CCR1
CCR3
CCR5
RLN1
MUC1
THBS3
GBAP1
FAM189B
ARHGEF2
EXOSC7
CXCR6
FYCO1
XCR1
FLT1P1
CCR1
CCR3
CCR5
TMIE
AL022393.7
ZSCAN31
ZKSCAN3
FAM27E3
RP11-195B21.3
PGM5P2
IFNAR2
IL10RB
AC073464.7
EXOSC7
CXCR6
CCR5
FAM27E3
RP11-195B21.3
PGM5P2
PAQR5
DPP9
B
N2LongCOVID
N1LongCOVID
W2LongCOVID
W1LongCOVID
HGI-B1-vs-B2
HGI-C2
HGI-B2
HGI-B1
-Log10(p)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
21

Figure S14

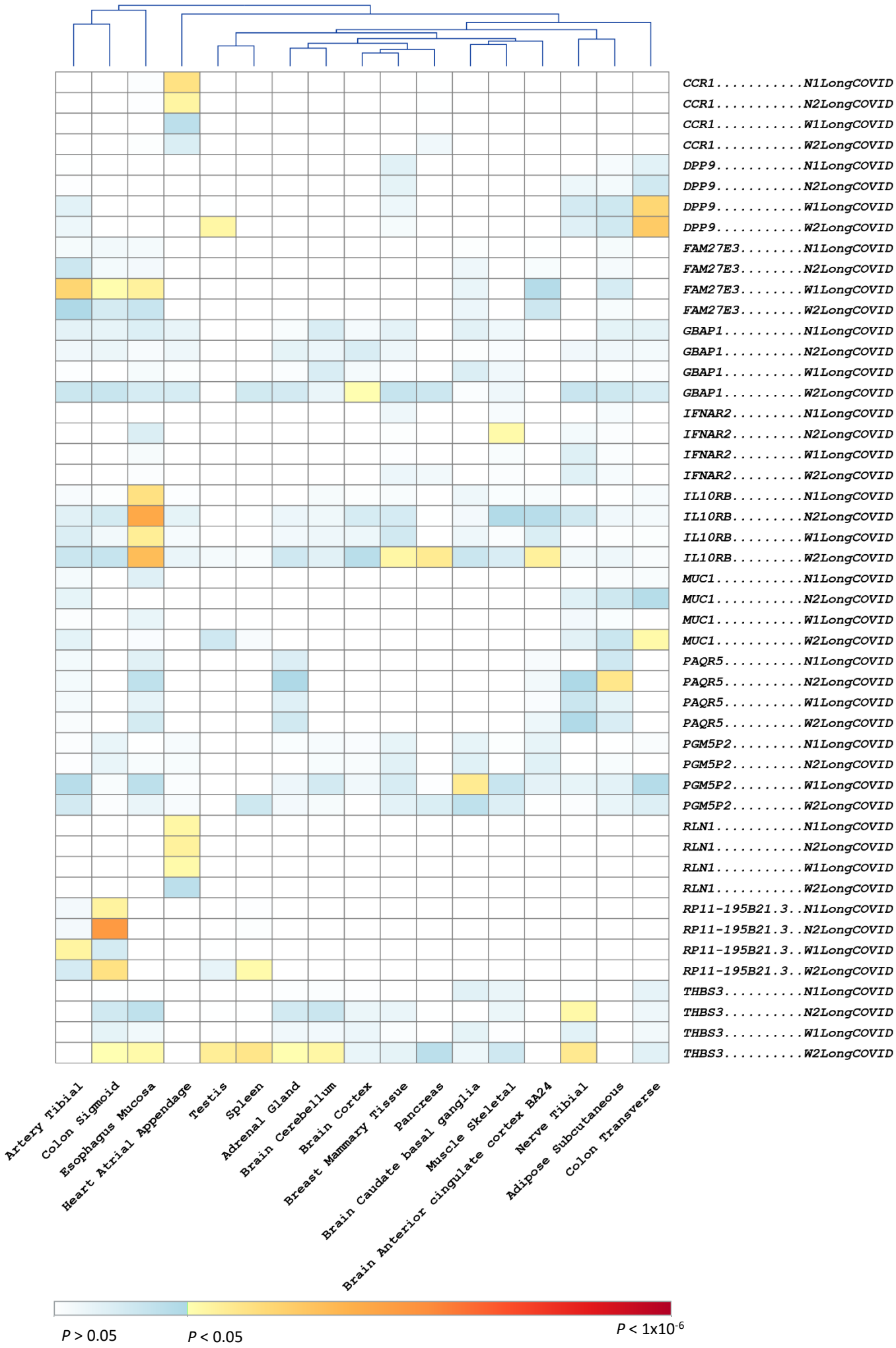

CCR1.....N1LongCOVID
CCR1.....N2LongCOVID
CCR1.....W1LongCOVID
CCR1.....W2LongCOVID
DPP9.....N1LongCOVID
DPP9.....N2LongCOVID
DPP9.....W1LongCOVID
DPP9.....W2LongCOVID
FAM27E3.....N1LongCOVID
FAM27E3.....N2LongCOVID
FAM27E3.....W1LongCOVID
FAM27E3.....W2LongCOVID
GBAP1.....N1LongCOVID
GBAP1.....N2LongCOVID
GBAP1.....W1LongCOVID
GBAP1.....W2LongCOVID
IFNAR2.....N1LongCOVID
IFNAR2.....N2LongCOVID
IFNAR2.....W1LongCOVID
IFNAR2.....W2LongCOVID
IL10RB.....N1LongCOVID
IL10RB.....N2LongCOVID
IL10RB.....W1LongCOVID
IL10RB.....W2LongCOVID
MUC1.....N1LongCOVID
MUC1.....N2LongCOVID
MUC1.....W1LongCOVID
MUC1.....W2LongCOVID
PAQR5.....N1LongCOVID
PAQR5.....N2LongCOVID
PAQR5.....W1LongCOVID
PAQR5.....W2LongCOVID
PGM5P2.....N1LongCOVID
PGM5P2.....N2LongCOVID
PGM5P2.....W1LongCOVID
PGM5P2.....W2LongCOVID
RLN1.....N1LongCOVID
RLN1.....N2LongCOVID
RLN1.....W1LongCOVID
RLN1.....W2LongCOVID
RP11-195B21.3..N1LongCOVID
RP11-195B21.3..N2LongCOVID
RP11-195B21.3..W1LongCOVID
RP11-195B21.3..W2LongCOVID
THBS3.....N1LongCOVID
THBS3.....N2LongCOVID
THBS3.....W1LongCOVID
THBS3.....W2LongCOVID
Artery Tibial
Colon Sigmoid
Esophagus Mucosa
Heart Atrial Appendage
Testis
Spleen
Adrenal Gland
Brain Cerebellum
Brain Cortex
Breast Mammary Tissue
Pancreas
Brain Caudate basal ganglia
Muscle Skeletal
Brain Anterior cingulate cortex BA24
Nerve Tibial
Adipose Subcutaneous
Colon Transverse
P > 0.05
P < 0.05
P < 1x10-6

Figure S15

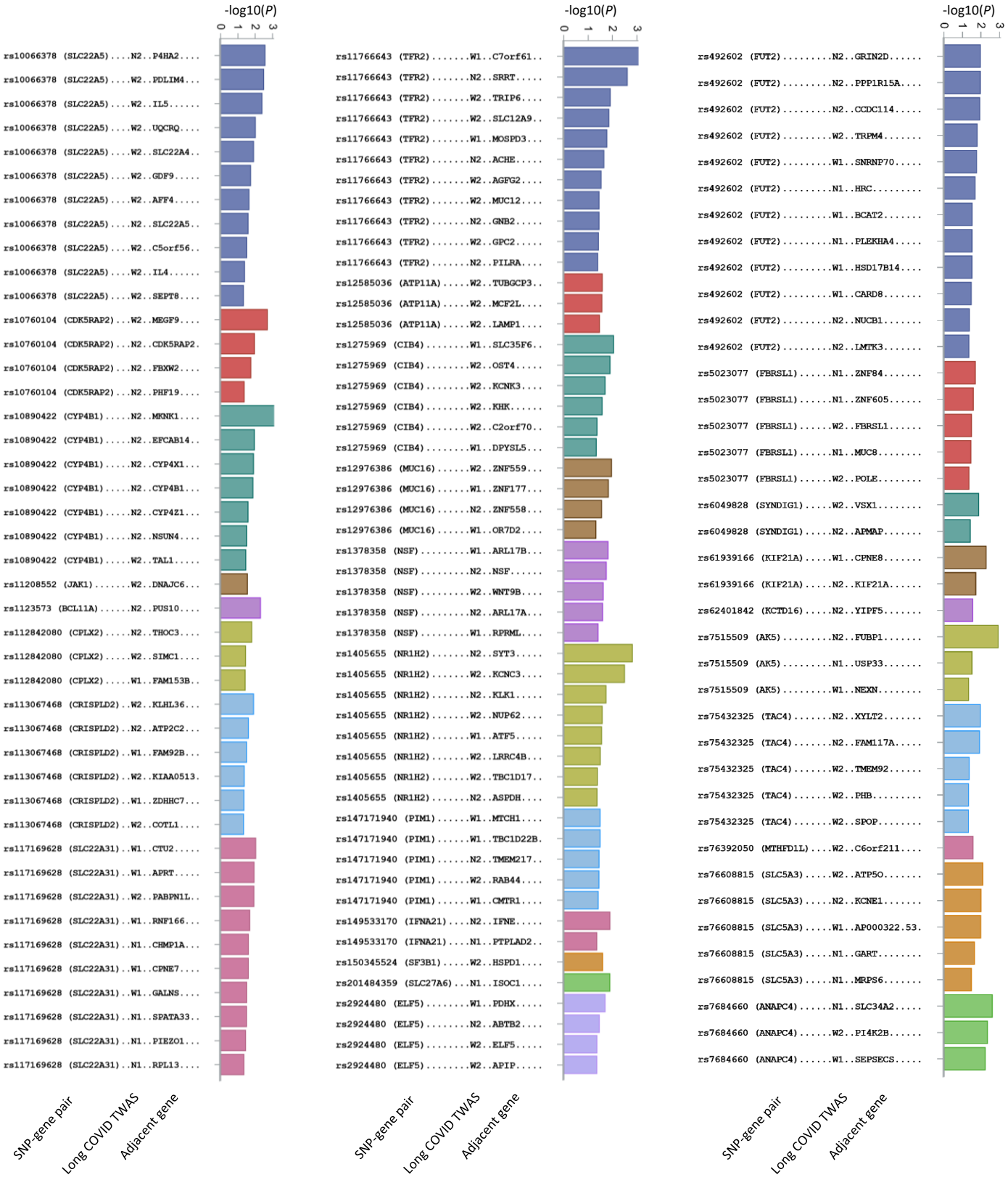
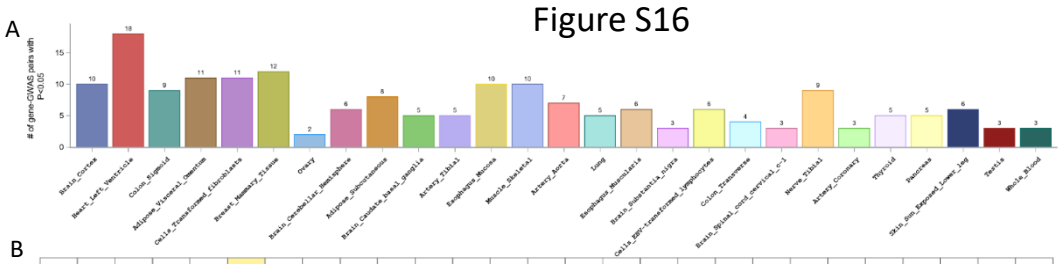
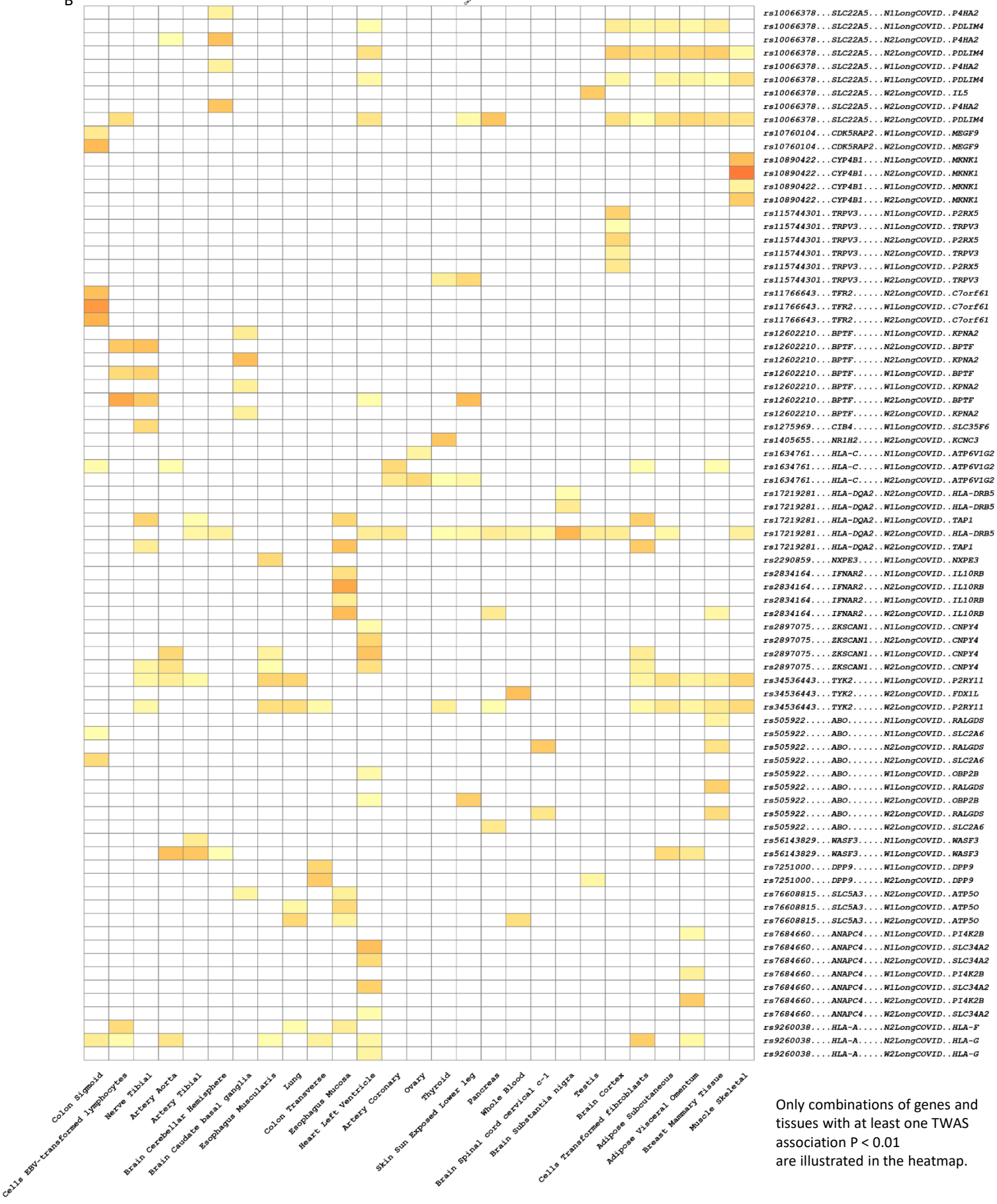

-log10(P)
0 1 2 3
rs10066378 (SLC22A5)....N2..P4HA2....
rs10066378 (SLC22A5)....W2..PDLIM4...
rs10066378 (SLC22A5)....W2..IL5.....
rs10066378 (SLC22A5)....W2..UQCRQ....
rs10066378 (SLC22A5)....W2..SLC22A4..
rs10066378 (SLC22A5)....W2..GDF9.....
rs10066378 (SLC22A5)....W2..AFF4.....
rs10066378 (SLC22A5)....N2..SLC22A5..
rs10066378 (SLC22A5)....W2..C5orf56..
rs10066378 (SLC22A5)....W2..IL4.....
rs10066378 (SLC22A5)....W2..SEPT8....
rs10760104 (CDK5RAP2)...W2..MEGF9....
rs10760104 (CDK5RAP2)...N2..CDK5RAP2.
rs10760104 (CDK5RAP2)...N2..FBXW2....
rs10760104 (CDK5RAP2)...N2..PHF19....
rs10890422 (CYP4B1).....N2..MKNK1....
rs10890422 (CYP4B1).....N2..EFCAB14..
rs10890422 (CYP4B1).....N2..CYP4X1...
rs10890422 (CYP4B1).....N2..CYP4B1...
rs10890422 (CYP4B1).....N2..CYP4Z1...
rs10890422 (CYP4B1).....N2..NSUN4....
rs10890422 (CYP4B1).....W2..TAL1.....
rs11208552 (JAK1).....W2..DNAJC6...
rs1123573 (BCL11A).....N2..PUS10....
rs112842080 (CPLX2).....N2..THOC3....
rs112842080 (CPLX2).....W2..SIMC1....
rs112842080 (CPLX2).....W1..FAM153B..
rs113067468 (CRISPLD2)..W2..KLHL36...
rs113067468 (CRISPLD2)..N2..ATP2C2...
rs113067468 (CRISPLD2)..W1..FAM92B...
rs113067468 (CRISPLD2)..W2..KIAA0513.
rs113067468 (CRISPLD2)..W1..ZDHHC7...
rs113067468 (CRISPLD2)..W2..COTL1....
rs117169628 (SLC22A31)..W1..CTU2.....
rs117169628 (SLC22A31)..W1..APRT.....
rs117169628 (SLC22A31)..W2..PABPN1L..
rs117169628 (SLC22A31)..W1..RNF166...
rs117169628 (SLC22A31)..N1..CHMP1A...
rs117169628 (SLC22A31)..W1..CPNE7....
rs117169628 (SLC22A31)..W1..GALNS....
rs117169628 (SLC22A31)..N1..SPATA33..
rs117169628 (SLC22A31)..N1..PIEZO1...
rs117169628 (SLC22A31)..N1..RPL13....
SNP-gene pair
Long COVID TWAS
Adjacent gene
-log10(P)
0 1 2 3
rs11766643 (TFR2).....W1..C7orf61..
rs11766643 (TFR2).....N2..SRRT.....
rs11766643 (TFR2).....W2..TRIP6....
rs11766643 (TFR2).....W2..SLC12A9..
rs11766643 (TFR2).....W1..MOSPD3...
rs11766643 (TFR2).....N2..ACHE.....
rs11766643 (TFR2).....W2..AGFG2....
rs11766643 (TFR2).....W2..MUC12....
rs11766643 (TFR2).....N2..GNB2.....
rs11766643 (TFR2).....W2..GPC2.....
rs11766643 (TFR2).....N2..PILRA....
rs12585036 (ATP11A).....W2..TUBGCP3..
rs12585036 (ATP11A).....W2..MCF2L....
rs12585036 (ATP11A).....W2..LAMP1....
rs1275969 (CIB4).....W1..SLC35F6..
rs1275969 (CIB4).....W2..OST4.....
rs1275969 (CIB4).....W2..KCNK3....
rs1275969 (CIB4).....W2..KHK.....
rs1275969 (CIB4).....W2..C2orf70..
rs1275969 (CIB4).....W1..DPYSL5...
rs12976386 (MUC16).....W2..ZNF559...
rs12976386 (MUC16).....W1..ZNF177...
rs12976386 (MUC16).....N2..ZNF558...
rs12976386 (MUC16).....W1..OR7D2....
rs1378358 (NSF).....W1..ARL17B...
rs1378358 (NSF).....N2..NSF.....
rs1378358 (NSF).....W2..WNT9B....
rs1378358 (NSF).....N2..ARL17A...
rs1378358 (NSF).....W1..RPRML....
rs1405655 (NR1H2).....N2..SYT3.....
rs1405655 (NR1H2).....W2..KCNC3....
rs1405655 (NR1H2).....N2..KLK1.....
rs1405655 (NR1H2).....W2..NUP62....
rs1405655 (NR1H2).....W1..ATF5.....
rs1405655 (NR1H2).....W2..LRRC4B...
rs1405655 (NR1H2).....W2..TBC1D17..
rs1405655 (NR1H2).....N2..ASPDH....
rs147171940 (PIM1).....W1..MTCH1....
rs147171940 (PIM1).....W1..TBC1D22B.
rs147171940 (PIM1).....N2..TMEM217..
rs147171940 (PIM1).....W2..RAB44....
rs147171940 (PIM1).....W1..CMTR1....
rs149533170 (IFNA21)....N2..IFNE.....
rs149533170 (IFNA21)....N1..PTPLAD2..
rs150345524 (SF3B1).....W2..HSPD1....
rs201484359 (SLC27A6)...N1..ISOC1....
rs2924480 (ELF5).....W1..PDHX.....
rs2924480 (ELF5).....N2..ABTB2....
rs2924480 (ELF5).....W2..ELF5.....
rs2924480 (ELF5).....W2..APIP.....
SNP-gene pair
Long COVID TWAS
Adjacent gene
-log10(P)
0 1 2 3
rs492602 (FUT2).....N2..GRIN2D.....
rs492602 (FUT2).....N2..PPP1R15A....
rs492602 (FUT2).....N2..CCDC114.....
rs492602 (FUT2).....W2..TRPM4.....
rs492602 (FUT2).....W1..SNRNP70.....
rs492602 (FUT2).....N1..HRC.....
rs492602 (FUT2).....W1..BCAT2.....
rs492602 (FUT2).....N1..PLEKHA4.....
rs492602 (FUT2).....W1..HSD17B14....
rs492602 (FUT2).....W1..CARD8.....
rs492602 (FUT2).....N2..NUCB1.....
rs492602 (FUT2).....N2..LMTK3.....
rs5023077 (FBRSL1).....N1..ZNF84.....
rs5023077 (FBRSL1).....N1..ZNF605.....
rs5023077 (FBRSL1).....W2..FBRSL1.....
rs5023077 (FBRSL1).....N1..MUC8.....
rs5023077 (FBRSL1).....W2..POLE.....
rs6049828 (SYNDIG1).....W2..VSX1.....
rs6049828 (SYNDIG1).....N2..APMAP.....
rs61939166 (KIF21A).....W1..CPNE8.....
rs61939166 (KIF21A).....N2..KIF21A.....
rs62401842 (KCTD16).....N2..YIPF5.....
rs7515509 (AK5).....N2..FUBP1.....
rs7515509 (AK5).....N1..USP33.....
rs7515509 (AK5).....W1..NEXN.....
rs75432325 (TAC4).....N2..XYLT2.....
rs75432325 (TAC4).....N2..FAM117A.....
rs75432325 (TAC4).....W2..TMEM92.....
rs75432325 (TAC4).....W2..PHB.....
rs75432325 (TAC4).....W2..SPOP.....
rs76392050 (MTHFD1L)....W2..C6orf211....
rs76608815 (SLC5A3).....W2..ATP5O.....
rs76608815 (SLC5A3).....N2..KCNE1.....
rs76608815 (SLC5A3).....W1..AP000322.53.
rs76608815 (SLC5A3).....N1..GART.....
rs76608815 (SLC5A3).....N1..MRPS6.....
rs7684660 (ANAPC4).....N1..SLC34A2.....
rs7684660 (ANAPC4).....W2..PI4K2B.....
rs7684660 (ANAPC4).....W1..SEPSECS.....
SNP-gene pair
Long COVID TWAS
Adjacent gene

Figure S16



B



Only combinations of genes and tissues with at least one TWAS association P < 0.01 are illustrated in the heatmap.

Figure S17

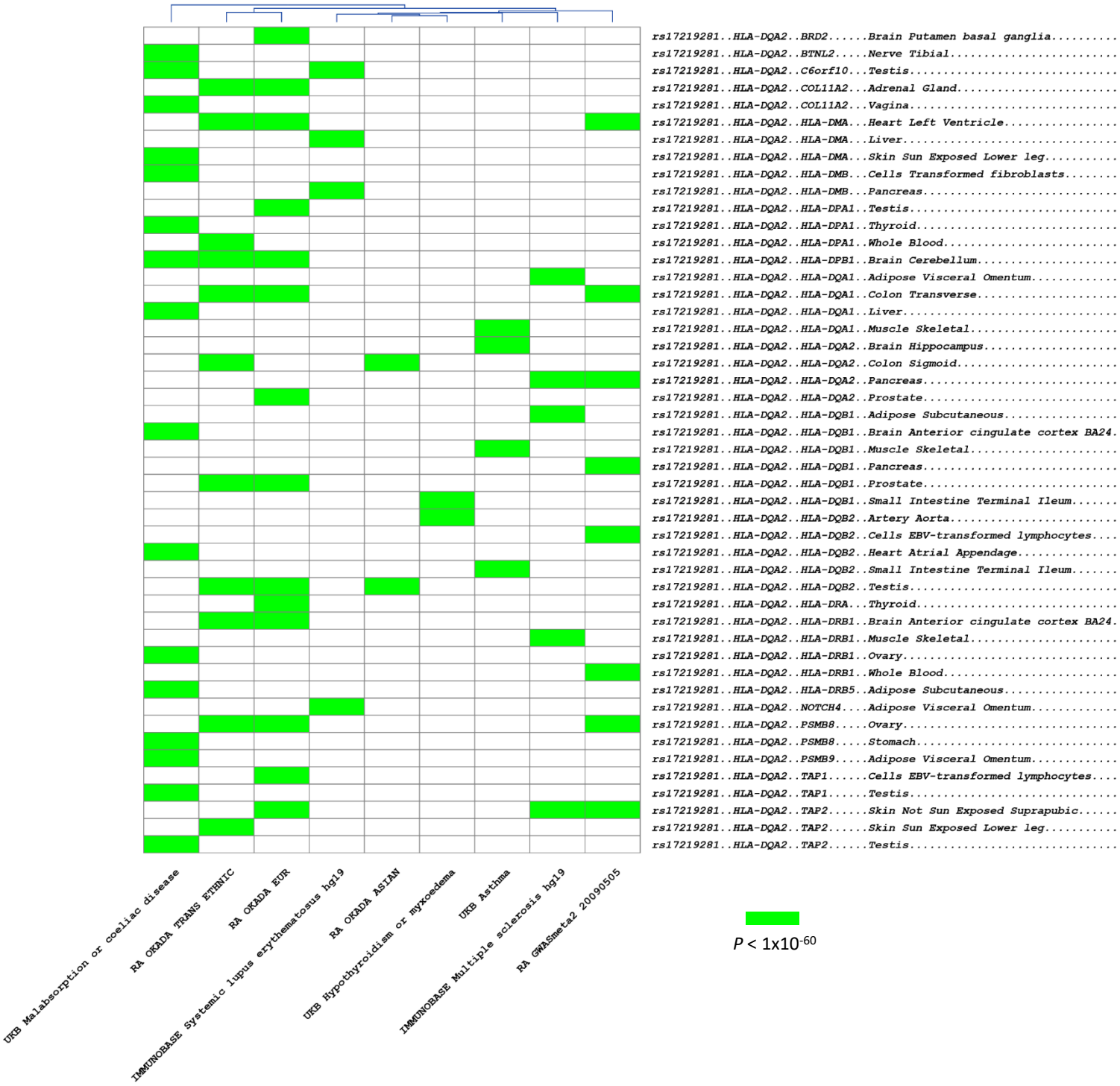

rs17219281..HLA-DQA2..BRD2.....Brain Putamen basal ganglia.....
rs17219281..HLA-DQA2..BTNL2.....Nerve Tibial.....
rs17219281..HLA-DQA2..C6orf10...Testis.....
rs17219281..HLA-DQA2..COL11A2...Adrenal Gland.....
rs17219281..HLA-DQA2..COL11A2...Vagina.....
rs17219281..HLA-DQA2..HLA-DMA...Heart Left Ventricle.....
rs17219281..HLA-DQA2..HLA-DMA...Liver.....
rs17219281..HLA-DQA2..HLA-DMA...Skin Sun Exposed Lower leg.....
rs17219281..HLA-DQA2..HLA-DMB...Cells Transformed fibroblasts.....
rs17219281..HLA-DQA2..HLA-DMB...Pancreas.....
rs17219281..HLA-DQA2..HLA-DPA1..Testis.....
rs17219281..HLA-DQA2..HLA-DPA1..Thyroid.....
rs17219281..HLA-DQA2..HLA-DPA1..Whole Blood.....
rs17219281..HLA-DQA2..HLA-DPB1..Brain Cerebellum.....
rs17219281..HLA-DQA2..HLA-DQA1..Adipose Visceral Omentum.....
rs17219281..HLA-DQA2..HLA-DQA1..Colon Transverse.....
rs17219281..HLA-DQA2..HLA-DQA1..Liver.....
rs17219281..HLA-DQA2..HLA-DQA1..Muscle Skeletal.....
rs17219281..HLA-DQA2..HLA-DQA2..Brain Hippocampus.....
rs17219281..HLA-DQA2..HLA-DQA2..Colon Sigmoid.....
rs17219281..HLA-DQA2..HLA-DQA2..Pancreas.....
rs17219281..HLA-DQA2..HLA-DQA2..Prostate.....
rs17219281..HLA-DQA2..HLA-DQB1..Adipose Subcutaneous.....
rs17219281..HLA-DQA2..HLA-DQB1..Brain Anterior cingulate cortex BA24.
rs17219281..HLA-DQA2..HLA-DQB1..Muscle Skeletal.....
rs17219281..HLA-DQA2..HLA-DQB1..Pancreas.....
rs17219281..HLA-DQA2..HLA-DQB1..Prostate.....
rs17219281..HLA-DQA2..HLA-DQB1..Small Intestine Terminal Ileum.....
rs17219281..HLA-DQA2..HLA-DQB2..Artery Aorta.....
rs17219281..HLA-DQA2..HLA-DQB2..Cells EBV-transformed lymphocytes....
rs17219281..HLA-DQA2..HLA-DQB2..Heart Atrial Appendage.....
rs17219281..HLA-DQA2..HLA-DQB2..Small Intestine Terminal Ileum.....
rs17219281..HLA-DQA2..HLA-DQB2..Testis.....
rs17219281..HLA-DQA2..HLA-DRA...Thyroid.....
rs17219281..HLA-DQA2..HLA-DRB1..Brain Anterior cingulate cortex BA24.
rs17219281..HLA-DQA2..HLA-DRB1..Muscle Skeletal.....
rs17219281..HLA-DQA2..HLA-DRB1..Ovary.....
rs17219281..HLA-DQA2..HLA-DRB1..Whole Blood.....
rs17219281..HLA-DQA2..HLA-DRB5..Adipose Subcutaneous.....
rs17219281..HLA-DQA2..NOTCH4....Adipose Visceral Omentum.....
rs17219281..HLA-DQA2..PSMB8.....Ovary.....
rs17219281..HLA-DQA2..PSMB8.....Stomach.....
rs17219281..HLA-DQA2..PSMB9.....Adipose Visceral Omentum.....
rs17219281..HLA-DQA2..TAP1......Cells EBV-transformed lymphocytes....
rs17219281..HLA-DQA2..TAP1......Testis.....
rs17219281..HLA-DQA2..TAP2......Skin Not Sun Exposed Suprapubic.....
rs17219281..HLA-DQA2..TAP2......Skin Sun Exposed Lower leg.....
rs17219281..HLA-DQA2..TAP2......Testis.....
UKB Malabsorption or coeliac disease
RA OKADA TRANS ETHNIC
RA OKADA EUR
IMMUNOBASE Systemic lupus erythematosus hg19
RA OKADA ASIAN
UKB Hypothyroidism or myxoedema
UKB Asthma
IMMUNOBASE Multiple sclerosis hg19
RA GWASmeta2 20090505
P < 1x10-60

Figure S18

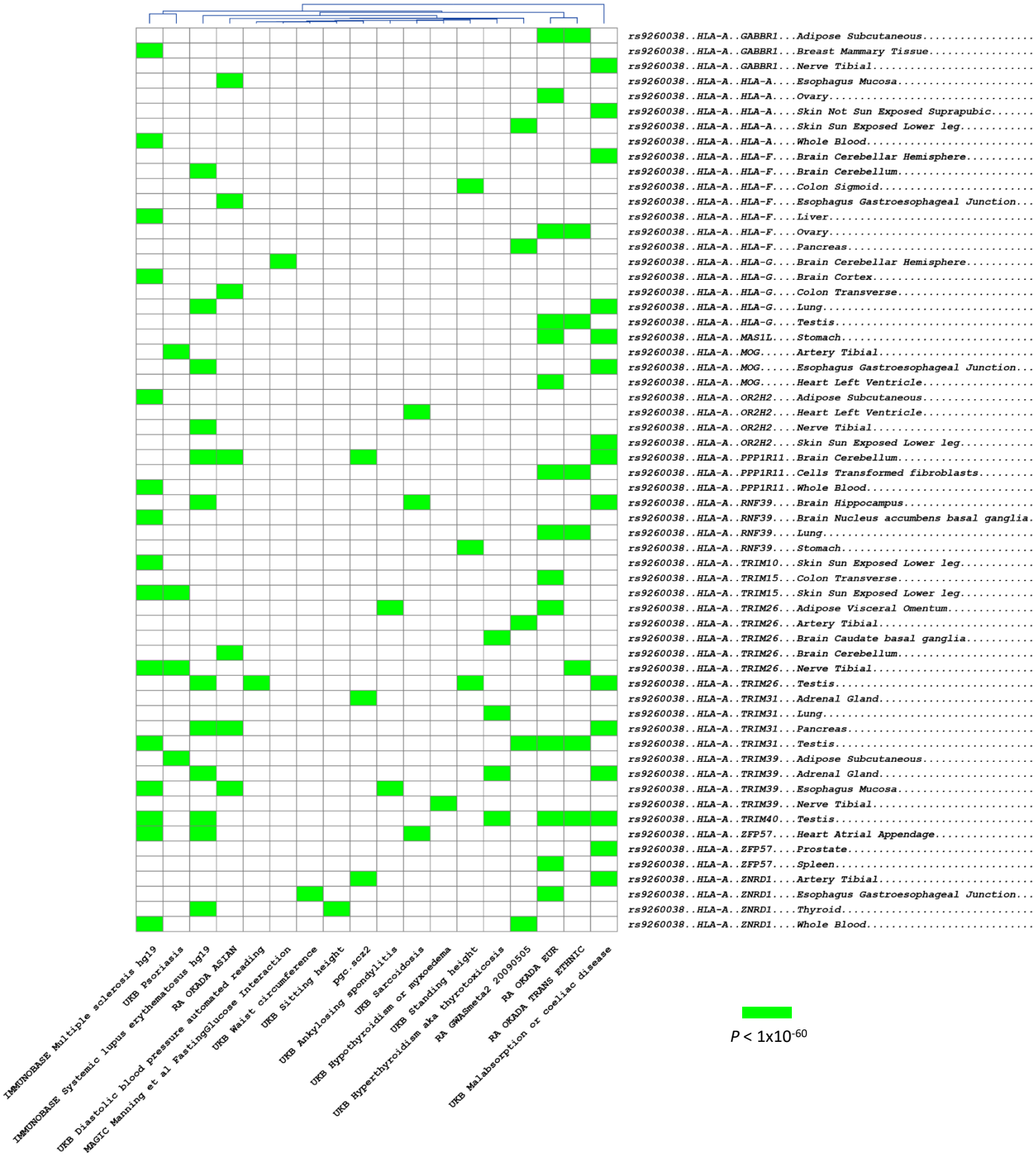

rs9260038..HLA-A..GABBR1...Adipose Subcutaneous
rs9260038..HLA-A..GABBR1...Breast Mammary Tissue
rs9260038..HLA-A..GABBR1...Nerve Tibial
rs9260038..HLA-A..HLA-A....Esophagus Mucosa
rs9260038..HLA-A..HLA-A....Ovary
rs9260038..HLA-A..HLA-A....Skin Not Sun Exposed Suprapubic
rs9260038..HLA-A..HLA-A....Skin Sun Exposed Lower leg
rs9260038..HLA-A..HLA-A....Whole Blood
rs9260038..HLA-A..HLA-F....Brain Cerebellar Hemisphere
rs9260038..HLA-A..HLA-F....Brain Cerebellum
rs9260038..HLA-A..HLA-F....Colon Sigmoid
rs9260038..HLA-A..HLA-F....Esophagus Gastroesophageal Junction
rs9260038..HLA-A..HLA-F....Liver
rs9260038..HLA-A..HLA-F....Ovary
rs9260038..HLA-A..HLA-F....Pancreas
rs9260038..HLA-A..HLA-G....Brain Cerebellar Hemisphere
rs9260038..HLA-A..HLA-G....Brain Cortex
rs9260038..HLA-A..HLA-G....Colon Transverse
rs9260038..HLA-A..HLA-G....Lung
rs9260038..HLA-A..HLA-G....Testis
rs9260038..HLA-A..MAS1L....Stomach
rs9260038..HLA-A..MOG.....Artery Tibial
rs9260038..HLA-A..MOG.....Esophagus Gastroesophageal Junction
rs9260038..HLA-A..MOG.....Heart Left Ventricle
rs9260038..HLA-A..OR2H2....Adipose Subcutaneous
rs9260038..HLA-A..OR2H2....Heart Left Ventricle
rs9260038..HLA-A..OR2H2....Nerve Tibial
rs9260038..HLA-A..OR2H2....Skin Sun Exposed Lower leg
rs9260038..HLA-A..PPP1R11..Brain Cerebellum
rs9260038..HLA-A..PPP1R11..Cells Transformed fibroblasts
rs9260038..HLA-A..PPP1R11..Whole Blood
rs9260038..HLA-A..RNF39....Brain Hippocampus
rs9260038..HLA-A..RNF39....Brain Nucleus accumbens basal ganglia
rs9260038..HLA-A..RNF39....Lung
rs9260038..HLA-A..RNF39....Stomach
rs9260038..HLA-A..TRIM10...Skin Sun Exposed Lower leg
rs9260038..HLA-A..TRIM15...Colon Transverse
rs9260038..HLA-A..TRIM15...Skin Sun Exposed Lower leg
rs9260038..HLA-A..TRIM26...Adipose Visceral Omentum
rs9260038..HLA-A..TRIM26...Artery Tibial
rs9260038..HLA-A..TRIM26...Brain Caudate basal ganglia
rs9260038..HLA-A..TRIM26...Brain Cerebellum
rs9260038..HLA-A..TRIM26...Nerve Tibial
rs9260038..HLA-A..TRIM26...Testis
rs9260038..HLA-A..TRIM31...Adrenal Gland
rs9260038..HLA-A..TRIM31...Lung
rs9260038..HLA-A..TRIM31...Pancreas
rs9260038..HLA-A..TRIM31...Testis
rs9260038..HLA-A..TRIM39...Adipose Subcutaneous
rs9260038..HLA-A..TRIM39...Adrenal Gland
rs9260038..HLA-A..TRIM39...Esophagus Mucosa
rs9260038..HLA-A..TRIM39...Nerve Tibial
rs9260038..HLA-A..TRIM40...Testis
rs9260038..HLA-A..ZFP57....Heart Atrial Appendage
rs9260038..HLA-A..ZFP57....Prostate
rs9260038..HLA-A..ZFP57....Spleen
rs9260038..HLA-A..ZNRD1....Artery Tibial
rs9260038..HLA-A..ZNRD1....Esophagus Gastroesophageal Junction
rs9260038..HLA-A..ZNRD1....Thyroid
rs9260038..HLA-A..ZNRD1....Whole Blood
IMMUNOBASE Multiple sclerosis hg19
UKB Psoriasis
IMMUNOBASE Systemic lupus erythematosus hg19
RA OKADA ASIAN
UKB Diastolic blood pressure automated reading
MAGIC Manning et al FastingGlucose interaction
UKB Waist circumference
UKB Sitting height
pgc.scz2
UKB Ankylosing spondylitis
UKB Sarcoidosis
UKB Hypothyroidism or myxoedema
UKB Standing height
UKB Hyperthyroidism aka thyrotoxicosis
RA GWASmeta2 20090505
RA OKADA EUR
RA OKADA TRANS ETHNIC
UKB Malabsorption or coeliac disease
P < 1x10-60

Figure S19

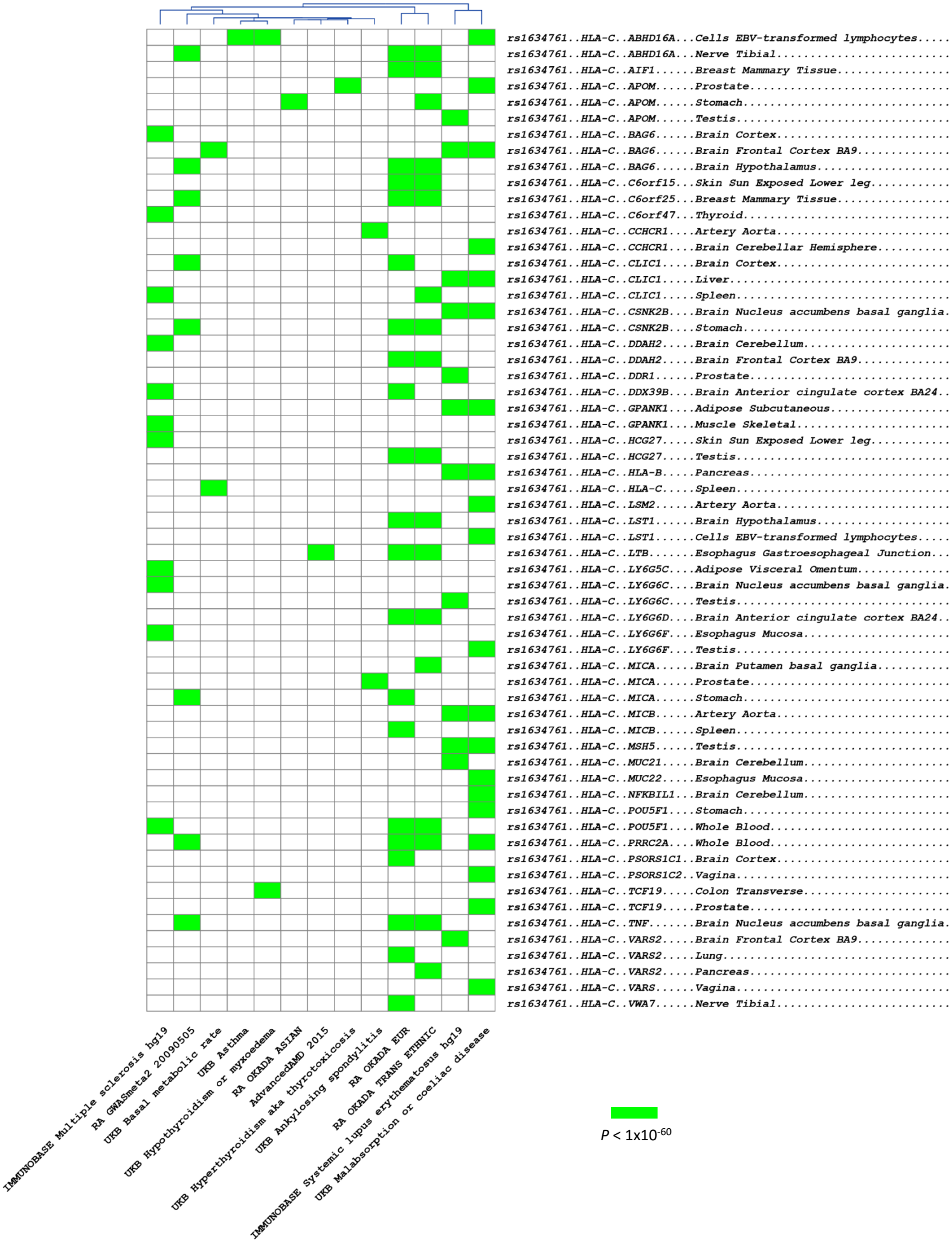

rs1634761..HLA-C..ABHD16A...Cells EBV-transformed lymphocytes.....
rs1634761..HLA-C..ABHD16A...Nerve Tibial.....
rs1634761..HLA-C..AIF1.....Breast Mammary Tissue.....
rs1634761..HLA-C..APOM.....Prostate.....
rs1634761..HLA-C..APOM.....Stomach.....
rs1634761..HLA-C..APOM.....Testis.....
rs1634761..HLA-C..BAG6.....Brain Cortex.....
rs1634761..HLA-C..BAG6.....Brain Frontal Cortex BA9.....
rs1634761..HLA-C..BAG6.....Brain Hypothalamus.....
rs1634761..HLA-C..C6orf15...Skin Sun Exposed Lower leg.....
rs1634761..HLA-C..C6orf25...Breast Mammary Tissue.....
rs1634761..HLA-C..C6orf47...Thyroid.....
rs1634761..HLA-C..CCHCR1....Artery Aorta.....
rs1634761..HLA-C..CCHCR1....Brain Cerebellar Hemisphere.....
rs1634761..HLA-C..CLIC1.....Brain Cortex.....
rs1634761..HLA-C..CLIC1.....Liver.....
rs1634761..HLA-C..CLIC1.....Spleen.....
rs1634761..HLA-C..CSNK2B....Brain Nucleus accumbens basal ganglia.
rs1634761..HLA-C..CSNK2B....Stomach.....
rs1634761..HLA-C..DDAH2.....Brain Cerebellum.....
rs1634761..HLA-C..DDAH2.....Brain Frontal Cortex BA9.....
rs1634761..HLA-C..DDR1......Prostate.....
rs1634761..HLA-C..DDX39B....Brain Anterior cingulate cortex BA24..
rs1634761..HLA-C..GPANK1....Adipose Subcutaneous.....
rs1634761..HLA-C..GPANK1....Muscle Skeletal.....
rs1634761..HLA-C..HCG27.....Skin Sun Exposed Lower leg.....
rs1634761..HLA-C..HCG27.....Testis.....
rs1634761..HLA-C..HLA-B.....Pancreas.....
rs1634761..HLA-C..HLA-C.....Spleen.....
rs1634761..HLA-C..LSM2......Artery Aorta.....
rs1634761..HLA-C..LST1......Brain Hypothalamus.....
rs1634761..HLA-C..LST1......Cells EBV-transformed lymphocytes.....
rs1634761..HLA-C..LTB.......Esophagus Gastroesophageal Junction...
rs1634761..HLA-C..LY6G5C....Adipose Visceral Omentum.....
rs1634761..HLA-C..LY6G6C....Brain Nucleus accumbens basal ganglia.
rs1634761..HLA-C..LY6G6C....Testis.....
rs1634761..HLA-C..LY6G6D....Brain Anterior cingulate cortex BA24..
rs1634761..HLA-C..LY6G6F....Esophagus Mucosa.....
rs1634761..HLA-C..LY6G6F....Testis.....
rs1634761..HLA-C..MICA......Brain Putamen basal ganglia.....
rs1634761..HLA-C..MICA......Prostate.....
rs1634761..HLA-C..MICA......Stomach.....
rs1634761..HLA-C..MICB......Artery Aorta.....
rs1634761..HLA-C..MICB......Spleen.....
rs1634761..HLA-C..MSH5......Testis.....
rs1634761..HLA-C..MUC21.....Brain Cerebellum.....
rs1634761..HLA-C..MUC22.....Esophagus Mucosa.....
rs1634761..HLA-C..NFKBIL1...Brain Cerebellum.....
rs1634761..HLA-C..POU5F1....Stomach.....
rs1634761..HLA-C..POU5F1....Whole Blood.....
rs1634761..HLA-C..PRRC2A....Whole Blood.....
rs1634761..HLA-C..PSORS1C1..Brain Cortex.....
rs1634761..HLA-C..PSORS1C2..Vagina.....
rs1634761..HLA-C..TCF19.....Colon Transverse.....
rs1634761..HLA-C..TCF19.....Prostate.....
rs1634761..HLA-C..TNF.......Brain Nucleus accumbens basal ganglia.
rs1634761..HLA-C..VARS2.....Brain Frontal Cortex BA9.....
rs1634761..HLA-C..VARS2.....Lung.....
rs1634761..HLA-C..VARS2.....Pancreas.....
rs1634761..HLA-C..VARS......Vagina.....
rs1634761..HLA-C..VWA7......Nerve Tibial.....
IMMUNOBASE Multiple sclerosis hg19
RA GWASmeta2 20090505
UKB Basal metabolic rate
UKB Asthma
UKB Hypothyroidism or myxoedema
RA OKADA ASIAN
AdvancedAMD 2015
UKB Hyperthyroidism aka thyrotoxicosis
UKB Ankylosing spondylitis
RA OKADA EUR
RA OKADA TRANS ETHNIC
IMMUNOBASE Systemic lupus erythematosus hg19
UKB Malabsorption or coeliac disease
P < 1x10-60

Figure S20

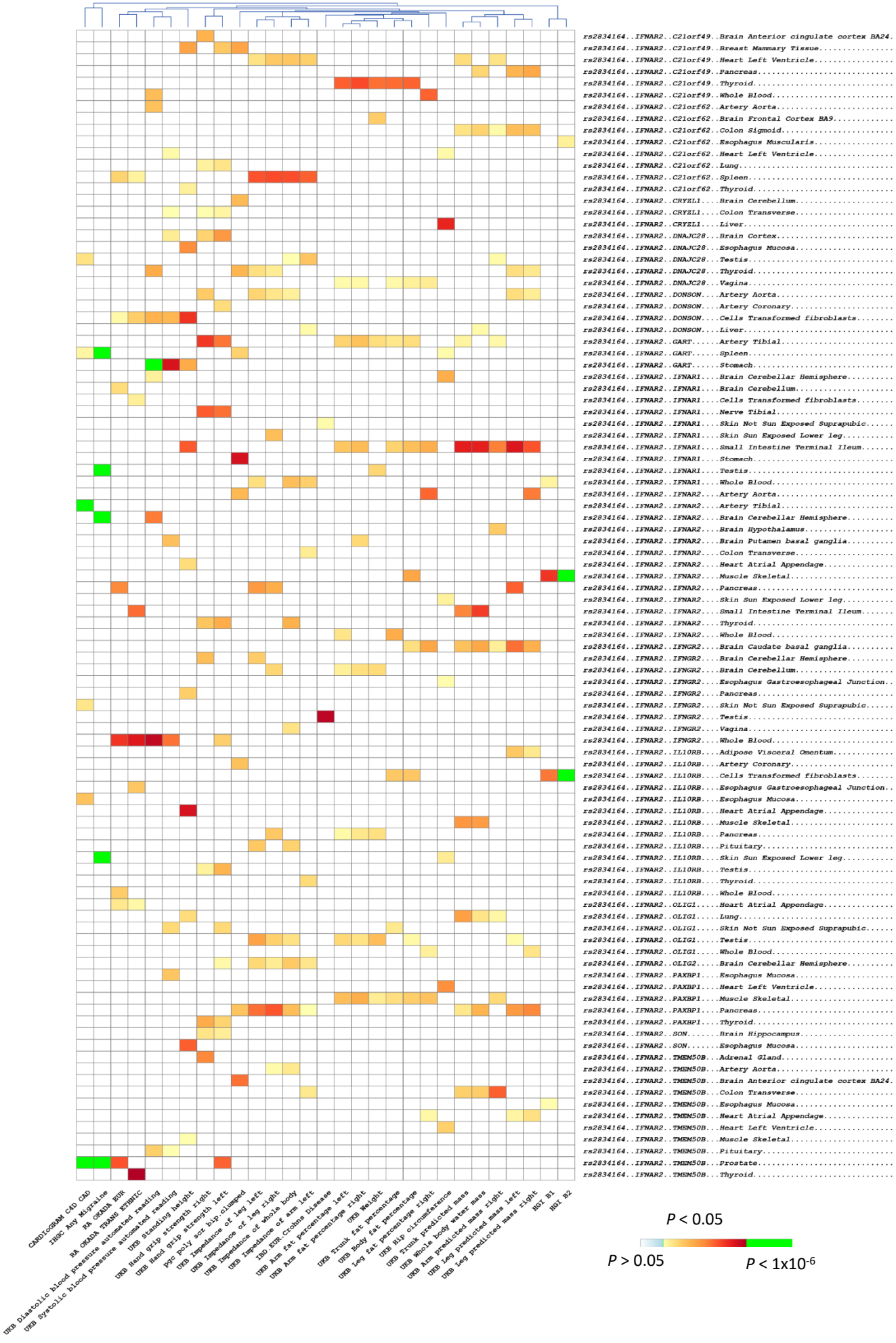

rs2834164..IFNAR2..C21orf49..Brain Anterior cingulate cortex BA24.
rs2834164..IFNAR2..C21orf49..Breast Mammary Tissue
rs2834164..IFNAR2..C21orf49..Heart Left Ventricle
rs2834164..IFNAR2..C21orf49..Pancreas
rs2834164..IFNAR2..C21orf49..Thyroid
rs2834164..IFNAR2..C21orf49..Whole Blood
rs2834164..IFNAR2..C21orf62..Artery Aorta
rs2834164..IFNAR2..C21orf62..Brain Frontal Cortex BA9
rs2834164..IFNAR2..C21orf62..Colon Sigmoid
rs2834164..IFNAR2..C21orf62..Esophagus Muscularis
rs2834164..IFNAR2..C21orf62..Heart Left Ventricle
rs2834164..IFNAR2..C21orf62..Lung
rs2834164..IFNAR2..C21orf62..Spleen
rs2834164..IFNAR2..C21orf62..Thyroid
rs2834164..IFNAR2..CRYZL1....Brain Cerebellum
rs2834164..IFNAR2..CRYZL1....Colon Transverse
rs2834164..IFNAR2..CRYZL1....Liver
rs2834164..IFNAR2..DNAJC28...Brain Cortex
rs2834164..IFNAR2..DNAJC28...Esophagus Mucosa
rs2834164..IFNAR2..DNAJC28...Testis
rs2834164..IFNAR2..DNAJC28...Thyroid
rs2834164..IFNAR2..DNAJC28...Vagina
rs2834164..IFNAR2..DONSON....Artery Aorta
rs2834164..IFNAR2..DONSON....Artery Coronary
rs2834164..IFNAR2..DONSON....Cells Transformed fibroblasts
rs2834164..IFNAR2..DONSON....Liver
rs2834164..IFNAR2..GART......Artery Tibial
rs2834164..IFNAR2..GART......Spleen
rs2834164..IFNAR2..GART......Stomach
rs2834164..IFNAR2..IFNAR1....Brain Cerebellar Hemisphere
rs2834164..IFNAR2..IFNAR1....Brain Cerebellum
rs2834164..IFNAR2..IFNAR1....Cells Transformed fibroblasts
rs2834164..IFNAR2..IFNAR1....Nerve Tibial
rs2834164..IFNAR2..IFNAR1....Skin Not Sun Exposed Suprapubic
rs2834164..IFNAR2..IFNAR1....Skin Sun Exposed Lower leg
rs2834164..IFNAR2..IFNAR1....Small Intestine Terminal Ileum
rs2834164..IFNAR2..IFNAR1....Stomach
rs2834164..IFNAR2..IFNAR1....Testis
rs2834164..IFNAR2..IFNAR1....Whole Blood
rs2834164..IFNAR2..IFNAR2....Artery Aorta
rs2834164..IFNAR2..IFNAR2....Artery Tibial
rs2834164..IFNAR2..IFNAR2....Brain Cerebellar Hemisphere
rs2834164..IFNAR2..IFNAR2....Brain Hypothalamus
rs2834164..IFNAR2..IFNAR2....Brain Putamen basal ganglia
rs2834164..IFNAR2..IFNAR2....Colon Transverse
rs2834164..IFNAR2..IFNAR2....Heart Atrial Appendage
rs2834164..IFNAR2..IFNAR2....Muscle Skeletal
rs2834164..IFNAR2..IFNAR2....Pancreas
rs2834164..IFNAR2..IFNAR2....Skin Sun Exposed Lower leg
rs2834164..IFNAR2..IFNAR2....Small Intestine Terminal Ileum
rs2834164..IFNAR2..IFNAR2....Thyroid
rs2834164..IFNAR2..IFNAR2....Whole Blood
rs2834164..IFNAR2..IFNGR2....Brain Caudate basal ganglia
rs2834164..IFNAR2..IFNGR2....Brain Cerebellar Hemisphere
rs2834164..IFNAR2..IFNGR2....Brain Cerebellum
rs2834164..IFNAR2..IFNGR2....Esophagus Gastroesophageal Junction
rs2834164..IFNAR2..IFNGR2....Pancreas
rs2834164..IFNAR2..IFNGR2....Skin Not Sun Exposed Suprapubic
rs2834164..IFNAR2..IFNGR2....Testis
rs2834164..IFNAR2..IFNGR2....Vagina
rs2834164..IFNAR2..IFNGR2....Whole Blood
rs2834164..IFNAR2..IL10RB....Adipose Visceral Omentum
rs2834164..IFNAR2..IL10RB....Artery Coronary
rs2834164..IFNAR2..IL10RB....Cells Transformed fibroblasts
rs2834164..IFNAR2..IL10RB....Esophagus Gastroesophageal Junction
rs2834164..IFNAR2..IL10RB....Esophagus Mucosa
rs2834164..IFNAR2..IL10RB....Heart Atrial Appendage
rs2834164..IFNAR2..IL10RB....Muscle Skeletal
rs2834164..IFNAR2..IL10RB....Pancreas
rs2834164..IFNAR2..IL10RB....Pituitary
rs2834164..IFNAR2..IL10RB....Skin Sun Exposed Lower leg
rs2834164..IFNAR2..IL10RB....Testis
rs2834164..IFNAR2..IL10RB....Thyroid
rs2834164..IFNAR2..IL10RB....Whole Blood
rs2834164..IFNAR2..OLIG1.....Heart Atrial Appendage
rs2834164..IFNAR2..OLIG1.....Lung
rs2834164..IFNAR2..OLIG1.....Skin Not Sun Exposed Suprapubic
rs2834164..IFNAR2..OLIG1.....Testis
rs2834164..IFNAR2..OLIG1.....Whole Blood
rs2834164..IFNAR2..OLIG2.....Brain Cerebellar Hemisphere
rs2834164..IFNAR2..PAXBP1....Esophagus Mucosa
rs2834164..IFNAR2..PAXBP1....Heart Left Ventricle
rs2834164..IFNAR2..PAXBP1....Muscle Skeletal
rs2834164..IFNAR2..PAXBP1....Pancreas
rs2834164..IFNAR2..PAXBP1....Thyroid
rs2834164..IFNAR2..SON.......Brain Hippocampus
rs2834164..IFNAR2..SON.......Esophagus Mucosa
rs2834164..IFNAR2..TMEM50B...Adrenal Gland
rs2834164..IFNAR2..TMEM50B...Artery Aorta
rs2834164..IFNAR2..TMEM50B...Brain Anterior cingulate cortex BA24.
rs2834164..IFNAR2..TMEM50B...Colon Transverse
rs2834164..IFNAR2..TMEM50B...Esophagus Mucosa
rs2834164..IFNAR2..TMEM50B...Heart Atrial Appendage
rs2834164..IFNAR2..TMEM50B...Heart Left Ventricle
rs2834164..IFNAR2..TMEM50B...Muscle Skeletal
rs2834164..IFNAR2..TMEM50B...Pituitary
rs2834164..IFNAR2..TMEM50B...Prostate
rs2834164..IFNAR2..TMEM50B...Thyroid
CARDIOGRAM C4D CAD
IBGC Any Migraine
RA OKADA TRANS ETHNIC
RA OKADA EUR
UKB Diastolic blood pressure automated reading
UKB Systolic blood pressure automated reading
UKB Standing height
UKB Hand grip strength right
UKB Hand grip strength left
PGC poly scz hip clumped
UKB Impedance of leg left
UKB Impedance of leg right
UKB Impedance of whole body
UKB Impedance of arm left
IBD.EUR.Crohns Disease
UKB Arm fat percentage left
UKB Arm fat percentage right
UKB Weight
UKB Trunk fat percentage
UKB Body fat percentage
UKB Leg fat percentage right
UKB Hip circumference
UKB Trunk predicted mass
UKB Whole body water mass
UKB Arm predicted mass right
UKB Leg predicted mass left
UKB Leg predicted mass right
RGI B1
RGI B2
P < 0.05
P > 0.05
P < 1x10-6

Figure S21

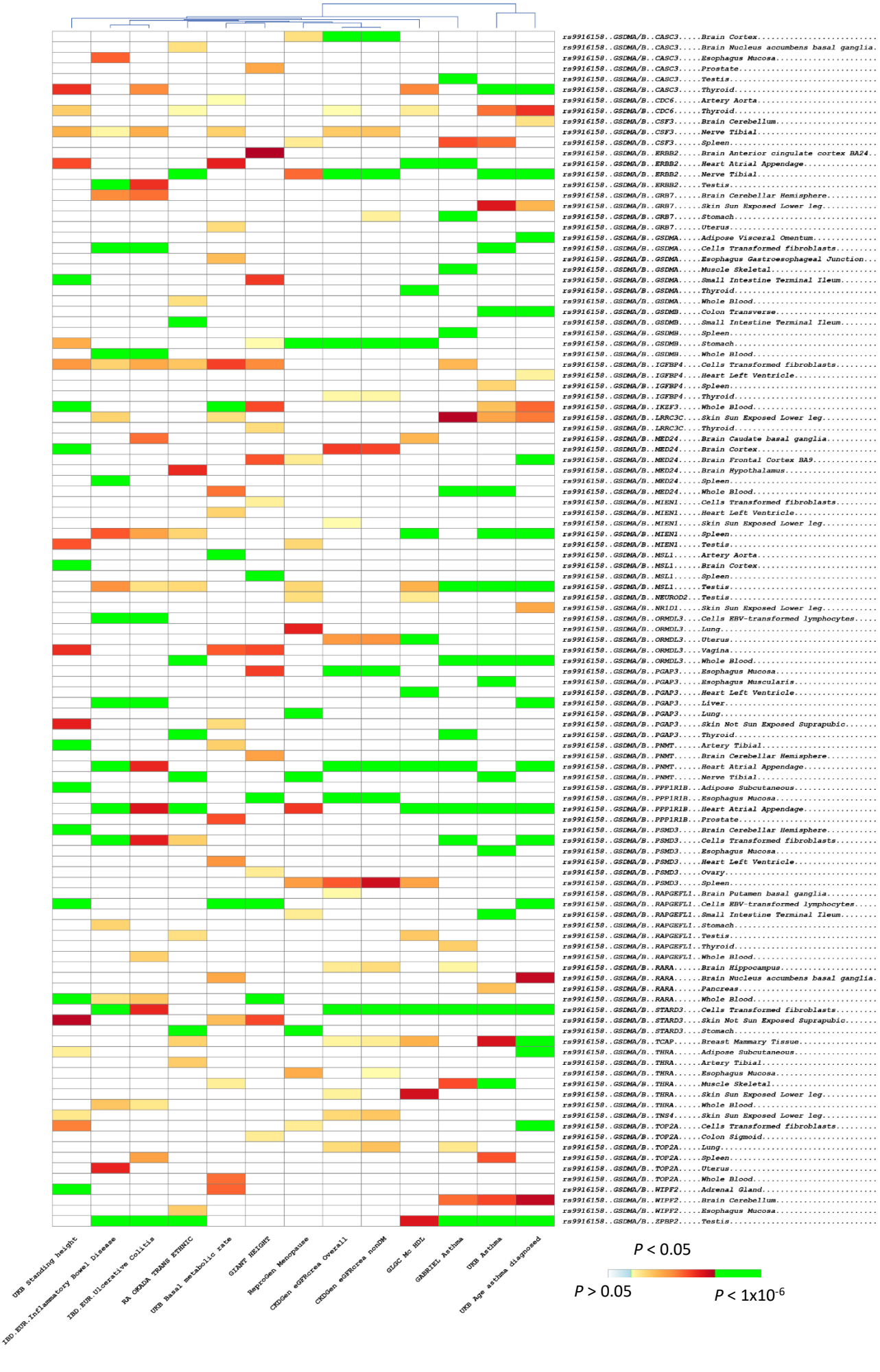

rs9916158..GSDMA/B..CASC3.....Brain Cortex
rs9916158..GSDMA/B..CASC3.....Brain Nucleus accumbens basal ganglia
rs9916158..GSDMA/B..CASC3.....Esophagus Mucosa
rs9916158..GSDMA/B..CASC3.....Prostate
rs9916158..GSDMA/B..CASC3.....Testis
rs9916158..GSDMA/B..CASC3.....Thyroid
rs9916158..GSDMA/B..CDC6.....Artery Aorta
rs9916158..GSDMA/B..CDC6.....Thyroid
rs9916158..GSDMA/B..CSF3.....Brain Cerebellum
rs9916158..GSDMA/B..CSF3.....Nerve Tibial
rs9916158..GSDMA/B..CSF3.....Spleen
rs9916158..GSDMA/B..ERBB2.....Brain Anterior cingulate cortex BA24
rs9916158..GSDMA/B..ERBB2.....Heart Atrial Appendage
rs9916158..GSDMA/B..ERBB2.....Nerve Tibial
rs9916158..GSDMA/B..ERBB2.....Testis
rs9916158..GSDMA/B..GRB7.....Brain Cerebellar Hemisphere
rs9916158..GSDMA/B..GRB7.....Skin Sun Exposed Lower leg
rs9916158..GSDMA/B..GRB7.....Stomach
rs9916158..GSDMA/B..GRB7.....Uterus
rs9916158..GSDMA/B..GSDMA.....Adipose Visceral Omentum
rs9916158..GSDMA/B..GSDMA.....Cells Transformed fibroblasts
rs9916158..GSDMA/B..GSDMA.....Esophagus Gastroesophageal Junction
rs9916158..GSDMA/B..GSDMA.....Muscle Skeletal
rs9916158..GSDMA/B..GSDMA.....Small Intestine Terminal Ileum
rs9916158..GSDMA/B..GSDMA.....Thyroid
rs9916158..GSDMA/B..GSDMA.....Whole Blood
rs9916158..GSDMA/B..GSDMB.....Colon Transverse
rs9916158..GSDMA/B..GSDMB.....Small Intestine Terminal Ileum
rs9916158..GSDMA/B..GSDMB.....Spleen
rs9916158..GSDMA/B..GSDMB.....Stomach
rs9916158..GSDMA/B..GSDMB.....Whole Blood
rs9916158..GSDMA/B..IGFBP4....Cells Transformed fibroblasts
rs9916158..GSDMA/B..IGFBP4....Heart Left Ventricle
rs9916158..GSDMA/B..IGFBP4....Spleen
rs9916158..GSDMA/B..IGFBP4....Thyroid
rs9916158..GSDMA/B..IKZF3.....Whole Blood
rs9916158..GSDMA/B..LRRC3C....Skin Sun Exposed Lower leg
rs9916158..GSDMA/B..LRRC3C....Thyroid
rs9916158..GSDMA/B..MED24.....Brain Caudate basal ganglia
rs9916158..GSDMA/B..MED24.....Brain Cortex
rs9916158..GSDMA/B..MED24.....Brain Frontal Cortex BA9
rs9916158..GSDMA/B..MED24.....Brain Hypothalamus
rs9916158..GSDMA/B..MED24.....Spleen
rs9916158..GSDMA/B..MED24.....Whole Blood
rs9916158..GSDMA/B..MIEN1.....Cells Transformed fibroblasts
rs9916158..GSDMA/B..MIEN1.....Heart Left Ventricle
rs9916158..GSDMA/B..MIEN1.....Skin Sun Exposed Lower leg
rs9916158..GSDMA/B..MIEN1.....Spleen
rs9916158..GSDMA/B..MIEN1.....Testis
rs9916158..GSDMA/B..MSL1.....Artery Aorta
rs9916158..GSDMA/B..MSL1.....Brain Cortex
rs9916158..GSDMA/B..MSL1.....Spleen
rs9916158..GSDMA/B..MSL1.....Testis
rs9916158..GSDMA/B..NEUROD2...Testis
rs9916158..GSDMA/B..NR1D1.....Skin Sun Exposed Lower leg
rs9916158..GSDMA/B..ORMDL3....Cells EBV-transformed lymphocytes
rs9916158..GSDMA/B..ORMDL3....Lung
rs9916158..GSDMA/B..ORMDL3....Uterus
rs9916158..GSDMA/B..ORMDL3....Vagina
rs9916158..GSDMA/B..ORMDL3....Whole Blood
rs9916158..GSDMA/B..PGAP3.....Esophagus Mucosa
rs9916158..GSDMA/B..PGAP3.....Esophagus Muscularis
rs9916158..GSDMA/B..PGAP3.....Heart Left Ventricle
rs9916158..GSDMA/B..PGAP3.....Liver
rs9916158..GSDMA/B..PGAP3.....Lung
rs9916158..GSDMA/B..PGAP3.....Skin Not Sun Exposed Suprapubic
rs9916158..GSDMA/B..PGAP3.....Thyroid
rs9916158..GSDMA/B..PNMT.....Artery Tibial
rs9916158..GSDMA/B..PNMT.....Brain Cerebellar Hemisphere
rs9916158..GSDMA/B..PNMT.....Heart Atrial Appendage
rs9916158..GSDMA/B..PNMT.....Nerve Tibial
rs9916158..GSDMA/B..PPP1R1B...Adipose Subcutaneous
rs9916158..GSDMA/B..PPP1R1B...Esophagus Mucosa
rs9916158..GSDMA/B..PPP1R1B...Heart Atrial Appendage
rs9916158..GSDMA/B..PPP1R1B...Prostate
rs9916158..GSDMA/B..PSMD3.....Brain Cerebellar Hemisphere
rs9916158..GSDMA/B..PSMD3.....Cells Transformed fibroblasts
rs9916158..GSDMA/B..PSMD3.....Esophagus Mucosa
rs9916158..GSDMA/B..PSMD3.....Heart Left Ventricle
rs9916158..GSDMA/B..PSMD3.....Ovary
rs9916158..GSDMA/B..PSMD3.....Spleen
rs9916158..GSDMA/B..RAPGEFL1..Brain Putamen basal ganglia
rs9916158..GSDMA/B..RAPGEFL1..Cells EBV-transformed lymphocytes
rs9916158..GSDMA/B..RAPGEFL1..Small Intestine Terminal Ileum
rs9916158..GSDMA/B..RAPGEFL1..Stomach
rs9916158..GSDMA/B..RAPGEFL1..Testis
rs9916158..GSDMA/B..RAPGEFL1..Thyroid
rs9916158..GSDMA/B..RAPGEFL1..Whole Blood
rs9916158..GSDMA/B..RARA.....Brain Hippocampus
rs9916158..GSDMA/B..RARA.....Brain Nucleus accumbens basal ganglia
rs9916158..GSDMA/B..RARA.....Pancreas
rs9916158..GSDMA/B..RARA.....Whole Blood
rs9916158..GSDMA/B..STARD3....Cells Transformed fibroblasts
rs9916158..GSDMA/B..STARD3....Skin Not Sun Exposed Suprapubic
rs9916158..GSDMA/B..STARD3....Stomach
rs9916158..GSDMA/B..TCAP.....Breast Mammary Tissue
rs9916158..GSDMA/B..THRA.....Adipose Subcutaneous
rs9916158..GSDMA/B..THRA.....Artery Tibial
rs9916158..GSDMA/B..THRA.....Esophagus Mucosa
rs9916158..GSDMA/B..THRA.....Muscle Skeletal
rs9916158..GSDMA/B..THRA.....Skin Sun Exposed Lower leg
rs9916158..GSDMA/B..THRA.....Whole Blood
rs9916158..GSDMA/B..TNS4.....Skin Sun Exposed Lower leg
rs9916158..GSDMA/B..TOP2A.....Cells Transformed fibroblasts
rs9916158..GSDMA/B..TOP2A.....Colon Sigmoid
rs9916158..GSDMA/B..TOP2A.....Lung
rs9916158..GSDMA/B..TOP2A.....Spleen
rs9916158..GSDMA/B..TOP2A.....Uterus
rs9916158..GSDMA/B..TOP2A.....Whole Blood
rs9916158..GSDMA/B..WIPF2.....Adrenal Gland
rs9916158..GSDMA/B..WIPF2.....Brain Cerebellum
rs9916158..GSDMA/B..WIPF2.....Esophagus Mucosa
rs9916158..GSDMA/B..ZPBP2.....Testis
UKB Standing height
IBD.EUR.Inflammatory Bowel Disease
IBD.EUR.Ulcerative Colitis
RA OKADA TRANS ETHNIC
UKB Basal metabolic rate
GIANT HEIGHT
ReprodGen Menopause
CKDGen eGFRcrea Overall
CKDGen eGFRcrea nonDM
GLGC Mc HDL
GABRIEL Asthma
UKB Asthma
UKB Age asthma diagnosed
P < 0.05
P > 0.05
P < 1x10-6

Figure S22

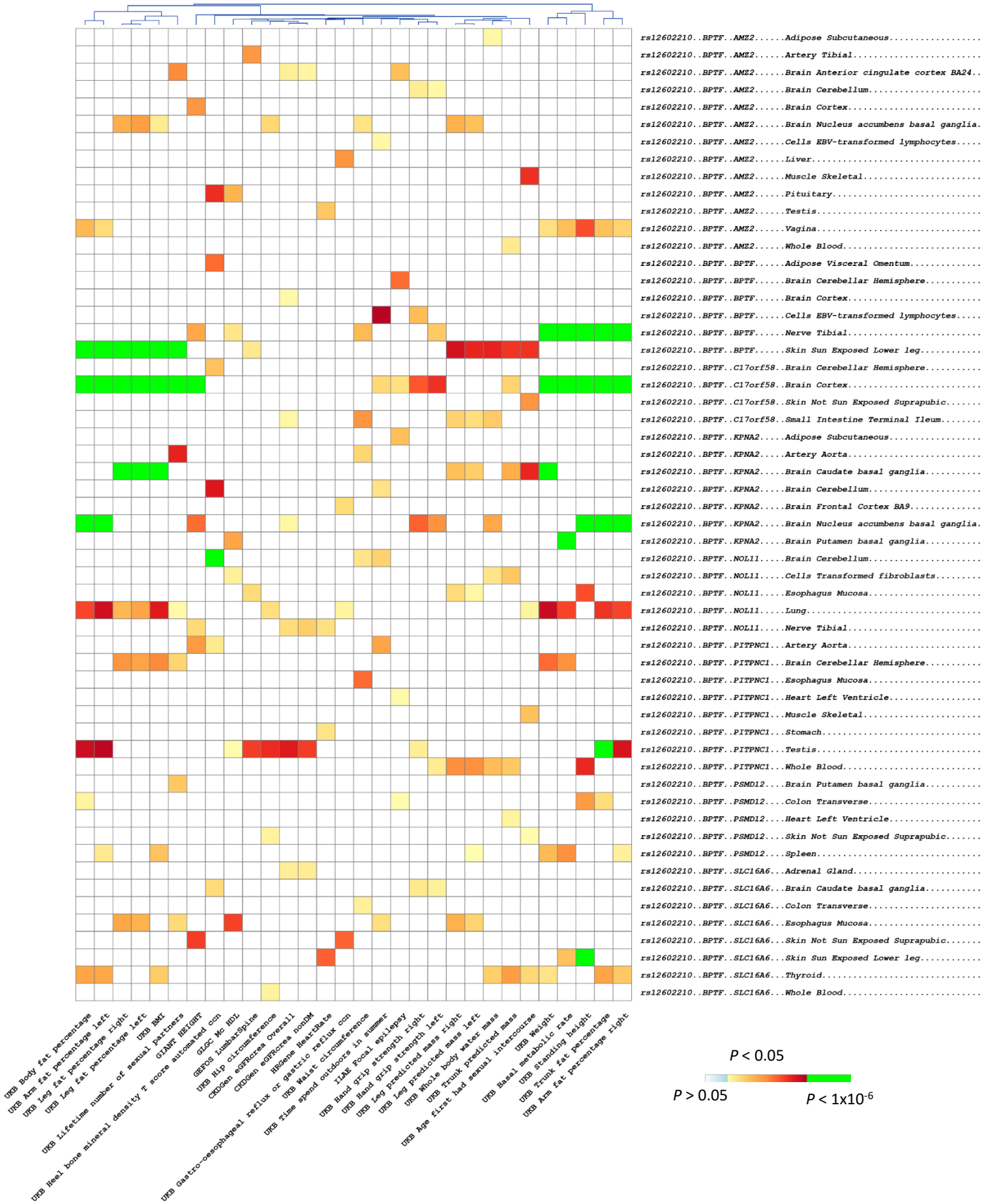

rs12602210..BPTF..AMZ2.....Adipose Subcutaneous
rs12602210..BPTF..AMZ2.....Artery Tibial
rs12602210..BPTF..AMZ2.....Brain Anterior cingulate cortex BA24
rs12602210..BPTF..AMZ2.....Brain Cerebellum
rs12602210..BPTF..AMZ2.....Brain Cortex
rs12602210..BPTF..AMZ2.....Brain Nucleus accumbens basal ganglia
rs12602210..BPTF..AMZ2.....Cells EBV-transformed lymphocytes
rs12602210..BPTF..AMZ2.....Liver
rs12602210..BPTF..AMZ2.....Muscle Skeletal
rs12602210..BPTF..AMZ2.....Pituitary
rs12602210..BPTF..AMZ2.....Testis
rs12602210..BPTF..AMZ2.....Vagina
rs12602210..BPTF..AMZ2.....Whole Blood
rs12602210..BPTF..BPTF.....Adipose Visceral Omentum
rs12602210..BPTF..BPTF.....Brain Cerebellar Hemisphere
rs12602210..BPTF..BPTF.....Brain Cortex
rs12602210..BPTF..BPTF.....Cells EBV-transformed lymphocytes
rs12602210..BPTF..BPTF.....Nerve Tibial
rs12602210..BPTF..BPTF.....Skin Sun Exposed Lower leg
rs12602210..BPTF..C17orf58..Brain Cerebellar Hemisphere
rs12602210..BPTF..C17orf58..Brain Cortex
rs12602210..BPTF..C17orf58..Skin Not Sun Exposed Suprapubic
rs12602210..BPTF..C17orf58..Small Intestine Terminal Ileum
rs12602210..BPTF..KPNA2.....Adipose Subcutaneous
rs12602210..BPTF..KPNA2.....Artery Aorta
rs12602210..BPTF..KPNA2.....Brain Caudate basal ganglia
rs12602210..BPTF..KPNA2.....Brain Cerebellum
rs12602210..BPTF..KPNA2.....Brain Frontal Cortex BA9
rs12602210..BPTF..KPNA2.....Brain Nucleus accumbens basal ganglia
rs12602210..BPTF..KPNA2.....Brain Putamen basal ganglia
rs12602210..BPTF..NOL11.....Brain Cerebellum
rs12602210..BPTF..NOL11.....Cells Transformed fibroblasts
rs12602210..BPTF..NOL11.....Esophagus Mucosa
rs12602210..BPTF..NOL11.....Lung
rs12602210..BPTF..NOL11.....Nerve Tibial
rs12602210..BPTF..PITPNC1...Artery Aorta
rs12602210..BPTF..PITPNC1...Brain Cerebellar Hemisphere
rs12602210..BPTF..PITPNC1...Esophagus Mucosa
rs12602210..BPTF..PITPNC1...Heart Left Ventricle
rs12602210..BPTF..PITPNC1...Muscle Skeletal
rs12602210..BPTF..PITPNC1...Stomach
rs12602210..BPTF..PITPNC1...Testis
rs12602210..BPTF..PITPNC1...Whole Blood
rs12602210..BPTF..PSMD12....Brain Putamen basal ganglia
rs12602210..BPTF..PSMD12....Colon Transverse
rs12602210..BPTF..PSMD12....Heart Left Ventricle
rs12602210..BPTF..PSMD12....Skin Not Sun Exposed Suprapubic
rs12602210..BPTF..PSMD12....Spleen
rs12602210..BPTF..SLC16A6...Adrenal Gland
rs12602210..BPTF..SLC16A6...Brain Caudate basal ganglia
rs12602210..BPTF..SLC16A6...Colon Transverse
rs12602210..BPTF..SLC16A6...Esophagus Mucosa
rs12602210..BPTF..SLC16A6...Skin Not Sun Exposed Suprapubic
rs12602210..BPTF..SLC16A6...Skin Sun Exposed Lower leg
rs12602210..BPTF..SLC16A6...Thyroid
rs12602210..BPTF..SLC16A6...Whole Blood
UKB Body fat percentage
UKB Arm fat percentage left
UKB Leg fat percentage right
UKB Leg fat percentage left
UKB BMI
GIANT HEIGHT
UKB Lifetime number of sexual partners
UKB Heel bone mineral density T score automated ccn
GLGC Mc HDL
GEFOS LumbarSpine
UKB Hip circumference
CKDGen eGFRcrea Overall
CKDGen eGFRcrea nonDM
HRGene HeartRate
UKB Gastro-oesophageal reflux or gastric reflux ccn
UKB Waist circumference
UKB Time spend outdoors in summer
ILAE Focal epilepsy
UKB Hand grip strength right
UKB Hand grip strength left
UKB Leg predicted mass right
UKB Leg predicted mass left
UKB Whole body water mass
UKB Trunk predicted mass
UKB Age first had sexual intercourse
UKB Weight
UKB Basal metabolic rate
UKB Standing height
UKB Trunk fat percentage
UKB Arm fat percentage right
P < 0.05
P > 0.05
P < 1x10-6

Figure S23

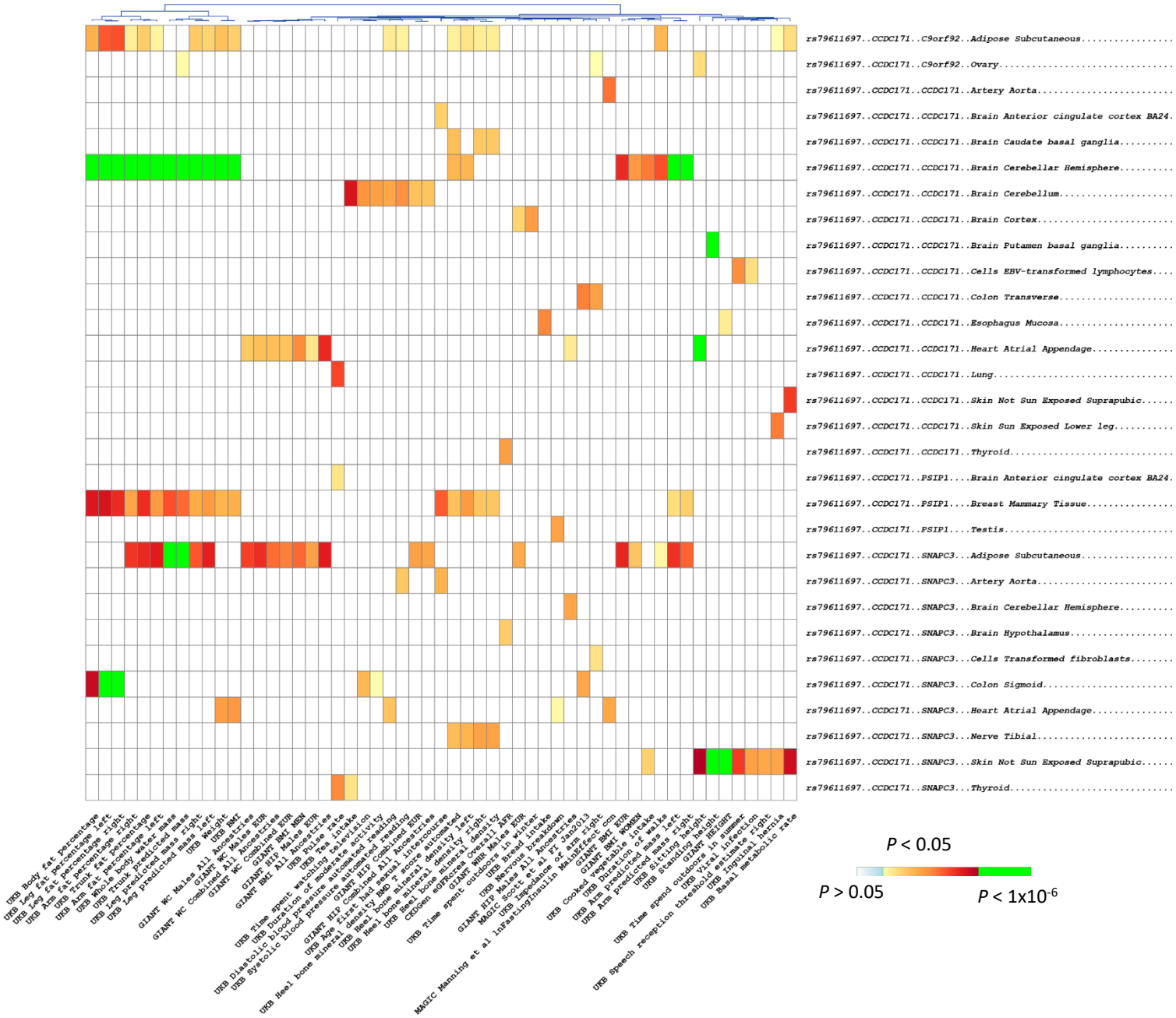
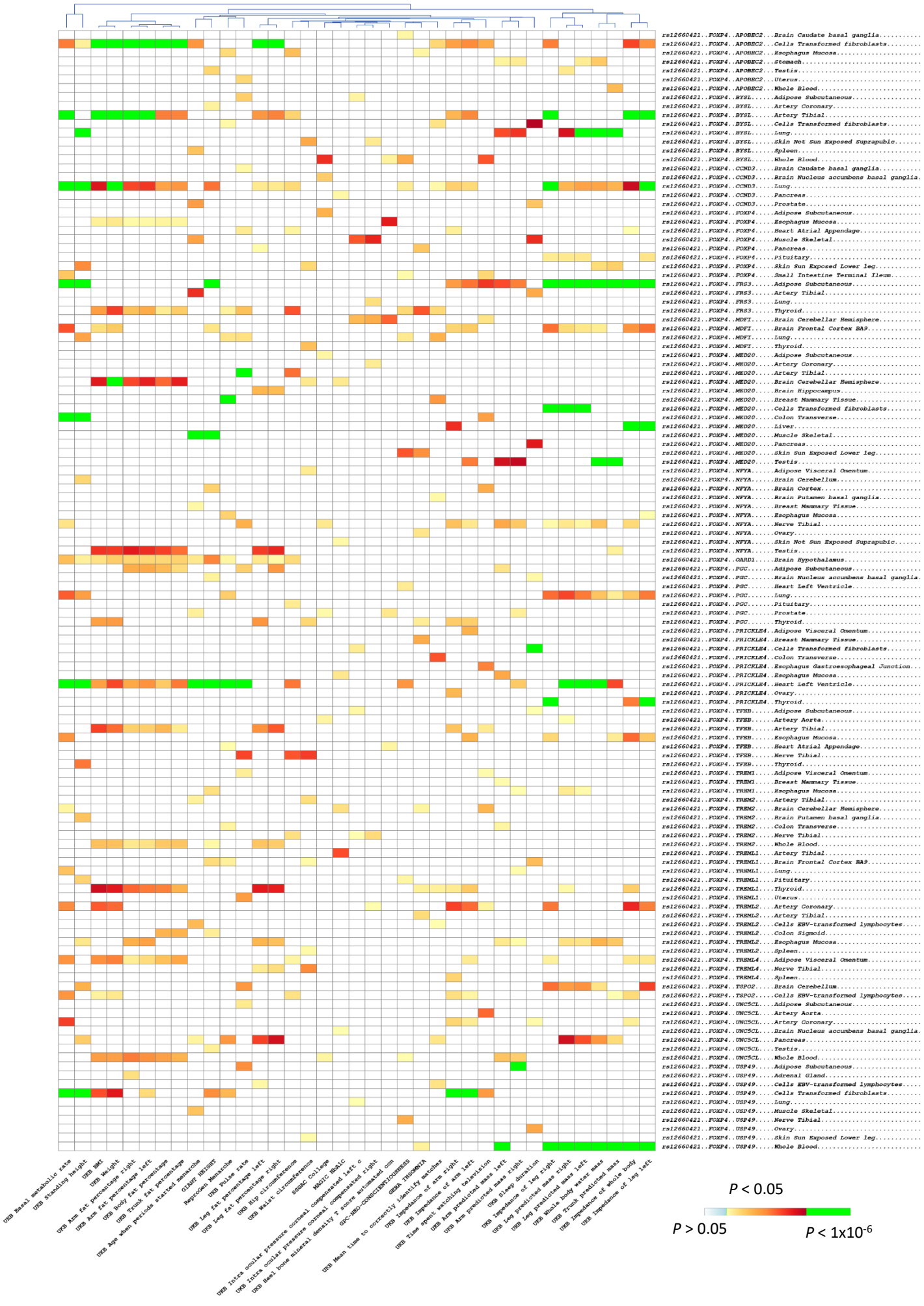

rs79611697..CCDC171..C9orf92..Adipose Subcutaneous
rs79611697..CCDC171..C9orf92..Ovary
rs79611697..CCDC171..CCDC171..Artery Aorta
rs79611697..CCDC171..CCDC171..Brain Anterior cingulate cortex BA24
rs79611697..CCDC171..CCDC171..Brain Caudate basal ganglia
rs79611697..CCDC171..CCDC171..Brain Cerebellar Hemisphere
rs79611697..CCDC171..CCDC171..Brain Cerebellum
rs79611697..CCDC171..CCDC171..Brain Cortex
rs79611697..CCDC171..CCDC171..Brain Putamen basal ganglia
rs79611697..CCDC171..CCDC171..Cells EBV-transformed lymphocytes
rs79611697..CCDC171..CCDC171..Colon Transverse
rs79611697..CCDC171..CCDC171..Esophagus Mucosa
rs79611697..CCDC171..CCDC171..Heart Atrial Appendage
rs79611697..CCDC171..CCDC171..Lung
rs79611697..CCDC171..CCDC171..Skin Not Sun Exposed Suprapubic
rs79611697..CCDC171..CCDC171..Skin Sun Exposed Lower leg
rs79611697..CCDC171..CCDC171..Thyroid
rs79611697..CCDC171..PSIP1....Brain Anterior cingulate cortex BA24
rs79611697..CCDC171..PSIP1....Breast Mammary Tissue
rs79611697..CCDC171..PSIP1....Testis
rs79611697..CCDC171..SNAPC3...Adipose Subcutaneous
rs79611697..CCDC171..SNAPC3...Artery Aorta
rs79611697..CCDC171..SNAPC3...Brain Cerebellar Hemisphere
rs79611697..CCDC171..SNAPC3...Brain Hypothalamus
rs79611697..CCDC171..SNAPC3...Cells Transformed fibroblasts
rs79611697..CCDC171..SNAPC3...Colon Sigmoid
rs79611697..CCDC171..SNAPC3...Heart Atrial Appendage
rs79611697..CCDC171..SNAPC3...Nerve Tibial
rs79611697..CCDC171..SNAPC3...Skin Not Sun Exposed Suprapubic
rs79611697..CCDC171..SNAPC3...Thyroid
UKB Body fat percentage
UKB Leg fat percentage left
UKB Leg fat percentage right
UKB Arm fat percentage right
UKB Trunk fat percentage
UKB Arm fat percentage left
UKB Whole body water mass
UKB Trunk predicted mass
UKB Leg predicted mass right
UKB Leg predicted mass left
UKB Weight
UKB BMI
GIANT WC Males All Ancestries
GIANT WC Males EUR
GIANT WC Combined All Ancestries
GIANT WC Combined EUR
GIANT BMI MEN
GIANT BMI All Ancestries
GIANT BMI Males EUR
UKB Pulse rate
UKB Tea intake
UKB Time spent watching television
UKB Duration of moderate activity
UKB Diastolic blood pressure automated reading
UKB Systolic blood pressure automated reading
GIANT HIP Combined All Ancestries
GIANT HIP Combined EUR
UKB Age first had sexual intercourse
UKB Heel bone mineral density left
UKB Heel bone mineral density right
UKB Heel bone mineral density
UKB Heel bone mineral density
CKDGen eGFRcrea Overall AFR
GIANT WHR Males EUR
UKB Time spent outdoors in winter
UKB Bread intake
UKB Nervous breakdown
GIANT HIP Males All Ancestries
MAGIC Scott et al FI Jan2013
UKB Impedance of arm right
MAGIC Manning et al InFastingInsulin MainEffect ccn
GIANT BMI EUR
GIANT BMI WOMEN
UKB Cooked vegetable intake
UKB Duration of walks
UKB Arm predicted mass left
UKB Arm predicted mass right
UKB Sitting height
UKB Standing height
GIANT HEIGHT
UKB Time spend outdoors in summer
UKB Viral infection
UKB Speech reception threshold estimate right
UKB Inguinal hernia
UKB Basal metabolic rate
P < 0.05
P > 0.05
P < 1x10-6

Figure S24


rs12660421..FOXP4..APOBEC2...Brain Caudate basal ganglia
rs12660421..FOXP4..APOBEC2...Cells Transformed fibroblasts
rs12660421..FOXP4..APOBEC2...Esophagus Mucosa
rs12660421..FOXP4..APOBEC2...Stomach
rs12660421..FOXP4..APOBEC2...Testis
rs12660421..FOXP4..APOBEC2...Uterus
rs12660421..FOXP4..APOBEC2...Whole Blood
rs12660421..FOXP4..BYSL.....Adipose Subcutaneous
rs12660421..FOXP4..BYSL.....Artery Coronary
rs12660421..FOXP4..BYSL.....Artery Tibial
rs12660421..FOXP4..BYSL.....Cells Transformed fibroblasts
rs12660421..FOXP4..BYSL.....Lung
rs12660421..FOXP4..BYSL.....Skin Not Sun Exposed Suprapubic
rs12660421..FOXP4..BYSL.....Spleen
rs12660421..FOXP4..BYSL.....Whole Blood
rs12660421..FOXP4..CCND3.....Brain Caudate basal ganglia
rs12660421..FOXP4..CCND3.....Brain Nucleus accumbens basal ganglia
rs12660421..FOXP4..CCND3.....Lung
rs12660421..FOXP4..CCND3.....Pancreas
rs12660421..FOXP4..CCND3.....Prostate
rs12660421..FOXP4..FOXP4.....Adipose Subcutaneous
rs12660421..FOXP4..FOXP4.....Esophagus Mucosa
rs12660421..FOXP4..FOXP4.....Heart Atrial Appendage
rs12660421..FOXP4..FOXP4.....Muscle Skeletal
rs12660421..FOXP4..FOXP4.....Pancreas
rs12660421..FOXP4..FOXP4.....Pituitary
rs12660421..FOXP4..FOXP4.....Skin Sun Exposed Lower leg
rs12660421..FOXP4..FOXP4.....Small Intestine Terminal Ileum
rs12660421..FOXP4..FRS3.....Adipose Subcutaneous
rs12660421..FOXP4..FRS3.....Artery Tibial
rs12660421..FOXP4..FRS3.....Lung
rs12660421..FOXP4..FRS3.....Thyroid
rs12660421..FOXP4..MDFI.....Brain Cerebellar Hemisphere
rs12660421..FOXP4..MDFI.....Brain Frontal Cortex BA9
rs12660421..FOXP4..MDFI.....Lung
rs12660421..FOXP4..MDFI.....Thyroid
rs12660421..FOXP4..MED20.....Adipose Subcutaneous
rs12660421..FOXP4..MED20.....Artery Coronary
rs12660421..FOXP4..MED20.....Artery Tibial
rs12660421..FOXP4..MED20.....Brain Cerebellar Hemisphere
rs12660421..FOXP4..MED20.....Brain Hippocampus
rs12660421..FOXP4..MED20.....Breast Mammary Tissue
rs12660421..FOXP4..MED20.....Cells Transformed fibroblasts
rs12660421..FOXP4..MED20.....Colon Transverse
rs12660421..FOXP4..MED20.....Liver
rs12660421..FOXP4..MED20.....Muscle Skeletal
rs12660421..FOXP4..MED20.....Pancreas
rs12660421..FOXP4..MED20.....Skin Sun Exposed Lower leg
rs12660421..FOXP4..MED20.....Testis
rs12660421..FOXP4..NFYA.....Adipose Visceral Omentum
rs12660421..FOXP4..NFYA.....Brain Cerebellum
rs12660421..FOXP4..NFYA.....Brain Cortex
rs12660421..FOXP4..NFYA.....Brain Putamen basal ganglia
rs12660421..FOXP4..NFYA.....Breast Mammary Tissue
rs12660421..FOXP4..NFYA.....Esophagus Mucosa
rs12660421..FOXP4..NFYA.....Nerve Tibial
rs12660421..FOXP4..NFYA.....Ovary
rs12660421..FOXP4..NFYA.....Skin Not Sun Exposed Suprapubic
rs12660421..FOXP4..NFYA.....Testis
rs12660421..FOXP4..OARD1.....Brain Hypothalamus
rs12660421..FOXP4..PGC.....Adipose Subcutaneous
rs12660421..FOXP4..PGC.....Brain Nucleus accumbens basal ganglia
rs12660421..FOXP4..PGC.....Heart Left Ventricle
rs12660421..FOXP4..PGC.....Lung
rs12660421..FOXP4..PGC.....Pituitary
rs12660421..FOXP4..PGC.....Prostate
rs12660421..FOXP4..PGC.....Thyroid
rs12660421..FOXP4..PRICKLE4..Adipose Visceral Omentum
rs12660421..FOXP4..PRICKLE4..Breast Mammary Tissue
rs12660421..FOXP4..PRICKLE4..Cells Transformed fibroblasts
rs12660421..FOXP4..PRICKLE4..Colon Transverse
rs12660421..FOXP4..PRICKLE4..Esophagus Gastroesophageal Junction
rs12660421..FOXP4..PRICKLE4..Esophagus Mucosa
rs12660421..FOXP4..PRICKLE4..Heart Left Ventricle
rs12660421..FOXP4..PRICKLE4..Ovary
rs12660421..FOXP4..PRICKLE4..Thyroid
rs12660421..FOXP4..TFEB.....Adipose Subcutaneous
rs12660421..FOXP4..TFEB.....Artery Aorta
rs12660421..FOXP4..TFEB.....Artery Tibial
rs12660421..FOXP4..TFEB.....Esophagus Mucosa
rs12660421..FOXP4..TFEB.....Heart Atrial Appendage
rs12660421..FOXP4..TFEB.....Nerve Tibial
rs12660421..FOXP4..TFEB.....Thyroid
rs12660421..FOXP4..TREM1.....Adipose Visceral Omentum
rs12660421..FOXP4..TREM1.....Breast Mammary Tissue
rs12660421..FOXP4..TREM1.....Esophagus Mucosa
rs12660421..FOXP4..TREM2.....Artery Tibial
rs12660421..FOXP4..TREM2.....Brain Cerebellar Hemisphere
rs12660421..FOXP4..TREM2.....Brain Putamen basal ganglia
rs12660421..FOXP4..TREM2.....Colon Transverse
rs12660421..FOXP4..TREM2.....Nerve Tibial
rs12660421..FOXP4..TREM2.....Whole Blood
rs12660421..FOXP4..TREML1....Artery Tibial
rs12660421..FOXP4..TREML1....Brain Frontal Cortex BA9
rs12660421..FOXP4..TREML1....Lung
rs12660421..FOXP4..TREML1....Pituitary
rs12660421..FOXP4..TREML1....Thyroid
rs12660421..FOXP4..TREML1....Uterus
rs12660421..FOXP4..TREML2....Artery Coronary
rs12660421..FOXP4..TREML2....Artery Tibial
rs12660421..FOXP4..TREML2....Cells EBV-transformed lymphocytes
rs12660421..FOXP4..TREML2....Colon Sigmoid
rs12660421..FOXP4..TREML2....Esophagus Mucosa
rs12660421..FOXP4..TREML2....Spleen
rs12660421..FOXP4..TREML4....Adipose Visceral Omentum
rs12660421..FOXP4..TREML4....Nerve Tibial
rs12660421..FOXP4..TREML4....Spleen
rs12660421..FOXP4..TSPO2.....Brain Cerebellum
rs12660421..FOXP4..TSPO2.....Cells EBV-transformed lymphocytes
rs12660421..FOXP4..UNC5CL....Adipose Subcutaneous
rs12660421..FOXP4..UNC5CL....Artery Aorta
rs12660421..FOXP4..UNC5CL....Artery Coronary
rs12660421..FOXP4..UNC5CL....Brain Nucleus accumbens basal ganglia
rs12660421..FOXP4..UNC5CL....Pancreas
rs12660421..FOXP4..UNC5CL....Testis
rs12660421..FOXP4..UNC5CL....Whole Blood
rs12660421..FOXP4..USP49.....Adipose Subcutaneous
rs12660421..FOXP4..USP49.....Adrenal Gland
rs12660421..FOXP4..USP49.....Cells EBV-transformed lymphocytes
rs12660421..FOXP4..USP49.....Cells Transformed fibroblasts
rs12660421..FOXP4..USP49.....Lung
rs12660421..FOXP4..USP49.....Muscle Skeletal
rs12660421..FOXP4..USP49.....Nerve Tibial
rs12660421..FOXP4..USP49.....Ovary
rs12660421..FOXP4..USP49.....Skin Sun Exposed Lower leg
rs12660421..FOXP4..USP49.....Whole Blood
UKB Basal metabolic rate
UKB Standing height
UKB BMI
UKB Weight
UKB Arm fat percentage right
UKB Arm fat percentage left
UKB Body fat percentage
UKB Trunk fat percentage
UKB Age when periods started menarche
GIANT HEIGHT
Neuroticism Menarche
UKB Pulse rate
UKB Leg fat percentage right
UKB Leg fat percentage left
UKB Hip circumference
UKB Waist circumference
SSGAC College
MAGIC HbA1C
UKB Intra ocular pressure corneal compensated left c
UKB Intra ocular pressure corneal compensated right
UKB Reel bone mineral density T score automated con
GPC-UEO-CONSCIENTIOUSNESS
GERA INSOMNIA
UKB Mean time to correctly identify matches
UKB Impedance of arm right
UKB Impedance of arm left
UKB Time spent watching television
UKB Arm predicted mass left
UKB Arm predicted mass right
UKB Sleep duration
UKB Impedance of leg right
UKB Leg predicted mass right
UKB Leg predicted mass left
UKB Whole body water mass
UKB Trunk predicted mass
UKB Impedance of whole body
UKB Impedance of leg left
P < 0.05
P > 0.05
P < 1x10-6