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. Despite its complexity, genome-wide association studies (GWAS) offer a pathway to uncovering genetic risk factors, though progress has been hindered by the disorder's symptom heterogeneity and the limited power of available datasets. Recent long COVID GWASs have highlighted key genetic associations, such as variants close to *FOXP4*, *ABO*, and *HLA-DQA1*, while underscoring the limitations posed by small sample sizes, restricted diversity, and misclassification biases. Here, an integrative analysis using COVID-19 GWAS data from the Host Genetics Initiative (HGI, release 7) was performed by leveraging proxy phenotypes to overcome the challenges of limited sample size or heterogeneity of long COVID cohorts, resulting in 62 independent single nucleotide polymorphisms (SNPs) prioritized as potentially associated with long COVID. These SNPs were categorized into three groups: (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), representing a catalog of SNPs predispose to both acute COVID-19 and chronic long COVID; and (3) non-hospitalization-specific SNPs, such as variants in *KCTD16* (rs62401842) and *WASF3* (rs56143829), highlighting genetic contributors specific to mild COVID-19 cases that might also contribute to long COVID. Further transcriptome-wide association studies (TWASs) across 48 GTEx tissues, leveraging GWAS data on COVID-19 hospitalization, susceptibility, and long COVID from the HGI consortium, revealed distinct tissue-specific patterns of association. Compared to the acute COVID-19 phenotypes, long COVID exhibited weaker association signals across heart, brain, and muscle-related tissues, as determined by correlations between gene expression of adjacent genes of candidate SNPs (43 out of 62 SNPs) and different COVID phenotypes. Notable TWAS hits included *DPP9* (rs7251000), *CCR1* (rs17078348), and *THBS3* (rs41264915). Phenome-wide TWASs also identified additional significant associations with long COVID related phenotypes, such as *HLA-DQA1* (rs17219281), *HLA-A* (rs9260038), and *HLA-C* (rs1634761) associated with immune-related diseases, *GSDMB* (rs9916158) associated with asthma, *FOXP4* (rs12660421) linked with sleep duration, highlighting their potential roles in long COVID pathophysiology. Therefore, current integrative approach offers a scalable framework for long COVID research by maximizing the statistical power of existing large-scale COVID-19 GWASs and provides novel insights into the genetic underpinnings of long COVID.

Introduction

Long COVID remains a critical and complex public health issue, with a diverse array of over 200 reported symptoms affecting multiple organ systems. 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 conducted by Lama et al.[1], which analyzed approximately 3,500 long COVID cases but not powerful enough to reveal the full genetic landscape due to several key limitations. Primarily, the study drew cases from early phases of the COVID-19 pandemic based on more immediate post-acute symptoms rather than on late-onset long COVID, which can emerge years after initial infection. As a result, potential long COVID cases with delayed symptom onset may be classified as controls, resulting in dilution of statistical power needed to detect true genetic associations. Despite these limitations, Lama et al.[1] 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. However, the stronger association 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. Given that long COVID can also follow mild COVID-19, additional genetic contributors associated specifically with mild cases likely remain undiscovered.

Another long COVID GWAS conducted by Taylor et al.[2] 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. They suggested that it is beneficial to analyze long COVID subtypes separately, as different symptom profiles could be associated with distinct genetic and biological pathways. However, their GWAS faced limitations due to its relatively small sample size ($n < 1{,}000$ long COVID cases) with restricted statistical power. Additionally, the study only reported SNPs derived by network analysis and did not identify any SNPs that reached genome-wide significance for long COVID, emphasizing the need for larger, more statistically robust studies. The absence of genome-

wide significant SNPs underscores a major challenge in long COVID research: even well-defined subtypes require substantial sample sizes to detect genetic variants with confidence.

More recently, a multi-ancestry GWAS was deposited at the preprint server medRxiv by Chaudhary et al.[3], 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 regarding the heterogeneity of long COVID symptoms, leading to the lack of sufficient power to detect associations for less common symptom clusters or subtle genetic variants. Furthermore, the study relied on self-reported cases, which may introduce misclassification biases, particularly given that long COVID lacks a single diagnostic marker. Self-reported symptoms, though valuable for capturing diverse experiences, can introduce noise that complicates statistical analysis and gene identification.

To advance insights into long COVID genetics, additional large-scale GWAS or alternative analyses that utilize existing COVID-19 GWAS summary statistics are essential to uncover genetic factors masked by current sample size limitations. Integrative approaches leveraging large publicly available GWAS of COVID-19 hospitalization and SARS-CoV-2 susceptibility may provide the statistical power required to detect SNPs linked to long COVID.

In this study, an integrative analysis of COVID-19 GWASs from the Host Genetics Initiative (HGI, release 7) was performed with the assumption that these phenotypes of COVID-19 hospitalization and susceptibility considered together to be a proxy phenotype of long COVID. Hypothesizing that long COVID may develop after either mild or severe COVID-19, this study was aimed to identify potential shared or hospitalization-specific genetic elements, as well as non-hospitalization-specific genetic elements. The current analysis prioritized 62 top independent SNPs potentially associated with long COVID into three SNP categories, containing SNPs mapped to two well-known loci involved in SARS-CoV2 entry (*ACE2* [rs190509934] and *TMPRSS2* [rs12329760]): (1) “Severe COVID-19-Specific SNPs”, such as these previously reported long COVID risk SNPs rs12660421 (*FOXP4*), rs10774679 (*OAS3*), and rs17219281 (*HLA-DQA1*), which show significant associations in both GWASs of hospitalized COVID-19 vs. general population and SARS-CoV-2 infection vs. general population with relatively weaker associations in hospitalized vs. non-hospitalized COVID-19; (2) “SNPs Associated with Both Severe and Mild COVID-19”, including the well-known long COVID risk SNPs, rs505922 (*ABO*), covering SNPs that may contribute to long COVID risk across a range of COVID-19 severities; (3) an extra SNP

category specifically emerged from non-hospitalized COVID-19 patients ("Non-hospitalization-specific long COVID SNPs") was further explored to prioritize 20 potential long COVID SNPs, such as the reported long COVID SNP rs62401842 (*KCTD16*), as well as rs56143829 (*WASF3*), with a relative lower association significance threshold. These prioritized SNPs were evaluated in association with long COVID based on published long COVID GWASs, and transcriptome-wide association studies (TWAS) were performed for these HGI COVID-19 hospitalization/susceptibility/long COVID GWASs to determine the expression level association of genes close to these candidate SNPs with these phenotypes, resulting in the conclusion that long COVID is likely to have different tissue etiology, particularly occurring in heart, brain, and muscle related tissues. Further phenome-wide TWASs were conducted to evaluate additional candidate genes potentially involved in long COVID. This integrative approach prioritizes SNPs as well as genes close to them with potential relevance to long COVID, offering a scalable alternative to screen long COVID risk loci by maximizing the statistical power of existing large-sample COVID-19 GWAS data from HGI.

Methods

Genome-wide Screening of Genetic SNPs Potentially Associated with Long COVID

Given the challenge of identifying genetic associations with long COVID due to its diverse symptomatology and the limited cohort sizes in GWAS, proxy phenotypes such as SARS-CoV-2 susceptibility (also refer to COVID-19 susceptibility in this study) and COVID-19 hospitalization were leveraged to prioritize relevant SNPs. This strategy is supported by evidence that long COVID can develop following both mild and severe COVID-19, suggesting shared or distinct genetic risk factors may influence acute-phase outcomes of COVID-19 (e.g., hospitalization, non-hospitalization, or susceptibility to SARS-CoV2 infection) and subsequent long COVID susceptibility.

Publicly available COVID-19 GWAS datasets 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]), and analyzed using an integrative analysis pipeline to prioritize potential long COVID SNPs. Although HGI also shared the long COVID GWASs with 4 slightly different case-control designs that were published by Lammi et al[1], these data were not initially utilized to prioritize long COVID SNPs due to the fact that these long COVID GWASs contain potential bias to severe COVID-19, which may lead to the missing of mild COVID-19 related long COVID risk SNPs. Again, the previously published HGI long COVID GWASs only identified one genome-wide significant locus represented by *FOXP4*. Therefore, in the current study the SNP prioritization workflow primarily focused on two key long COVID risk SNP categories related to severe COVID-19 that are potentially

relevant to long COVID. The first SNP categories is "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 cases. These SNPs are hypothesized to influence long COVID risk based on the epidemiological observation that multiple SARS-CoV2 infections and COVID-19 hospitalizations elevate the likelihood of developing the long COVID condition[4]. The following selection criteria was applied, including (a) $P \leq 0.05$ in HGI-C2, capturing SNPs linked to COVID-19 susceptibility; (b) $P \leq 0.05$ in HGI-B1, identifying SNPs that predispose individuals to COVID-19 hospitalization rather than mild COVID-19; (c) $P \leq 0.05$ in HGI-B2, identifying SNPs distinguishing hospitalized COVID-19 individuals from the general population; (d) $P > 0.05$ in the differential GWAS of HGI-B1 vs. HGI-B2, excluding SNPs with differential effect sizes between these comparisons; (e) all require to display consistent effect directions across HGI-C2, HGI-B1, and HGI-B2. Meanwhile, the second SNP categories is "SNPs Associated with Both Severe and Mild COVID-19", which influence both mild and severe COVID-19 outcomes that are plausible candidates for long COVID risk, given that non-hospitalized cases can also develop long COVID. These SNPs are unlikely to show significant differentiation between hospitalized and non-hospitalized cases but may exhibit 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 (a) $P \leq 0.05$ in HGI-C2, identifying COVID-19 susceptibility-related SNPs; (b) $P \leq 0.05$ in HGI-B2, identifying SNPs associated with COVID-19 hospitalization; (c) $P > 0.05$ in HGI-B1, indicating no significant difference in SNP effects between severe and mild COVID-19; (d) $P \leq 0.05$ in the differential GWAS of HGI-B1 vs. HGI-B2, capturing SNPs with distinct effects in these datasets; (e) all prioritized SNPs need to be fulfill consistent effect size directions across HGI-C2 and HGI-B2. Finally, SNPs met either of the above two sets of conditions were further evaluated based on the criterion of passing the $P < 1E\text{-}7$ in any of the three GWAS datasets (HGI-B1, HGI-B2, or HGI-C2). This prioritization framework provides a systematic approach to identifying genetic variants potentially predisposing individuals to long COVID by leveraging large-scale COVID-19 GWAS data, resulting in 42 top independent SNPs that were subjected to further evaluation on its potential involvement in long COVID by COVID19_GWAS_Analyzer[5, 6] (https://github.com/chengzhongshan/COVID19_GWAS_Analyzer).

Identification of SNPs Potentially Associated with Long COVID in Non-Hospitalized Cases

Current COVID-19 GWAS analyses from the HGI are predominantly biased toward hospitalized cases, potentially overlooking SNPs associated with long COVID that emerge specifically in non-hospitalized individuals. To address this gap, SNPs were filtered using the following criteria: (1) $P > 0.05$ in the GWAS comparing hospitalized cases with the general population (HGI-B2), to exclude SNPs associated with

COVID-19 hospitalization; (2) $P \leq 0.05$ in the GWAS comparing hospitalized cases with non-hospitalized cases (HGI-B1), to capture SNPs significantly differentiating these two groups; (3) $P \leq 0.05$ in the differential GWAS between HGI-B1 and HGI-B2, to retain SNPs with differential association signals between these datasets. To further refine the selection, independent SNPs met a suggestive significance threshold (P<1E-5) were prioritized, ensuring no other genome-wide significant SNPs were located within a 500-kb window in either HGI-B2 or HGI-C2. The use of a relaxed significance threshold 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 long COVID risk in non-hospitalized individuals.

Comparing Two COVID-19 Hospitalization GWASs Using the Differential Z-Score Method

To assess differences in effect sizes of SNPs between two COVID-19 GWAS datasets, 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), as well as comparison SNP effect size between sex-stratified COVID-19 hospitalization GWASs, differential z-score method[6] 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.

Transcriptome-Wide Association Analysis Using S-MetaXcan

Transcriptome-wide association analysis (TWAS) was conducted using S-MetaXcan[7], which infers gene expression association with a phenotype based on its corresponding GWAS summary statistics and GTEx

gene expression datasets across 48 tissues. Four COVID-19 GWAS datasets from the HGI were analyzed, including HGI-B1, HGI-B2, HGI-C2, and HGI-B1-vs-B2, alongside four long COVID GWAS datasets with subtly distinct case-control definitions from HGI[1] and an additional set of 221 GWASs from the UK Biobank (UKB) (all sample sizes >5,000) shared by Benneal's lab (https://www.nealelab.is/uk-biobank). Reference transcriptomic data from 48 tissues in GTEx v7[8] were utilized for the analyses.

S-MetaXcan[7] employs multivariate regression to integrate evidence across different GTEx tissues (n=48), enhancing detection power. Default settings were applied, with the conditional number for eigenvalue set at 30 during singular value decomposition (SVD) truncation and genes on autosomal chromosomes were included in the analysis. First, TWAS results for COVID-19 (n=8, collectively referred to as COVID TWASs) were evaluated for genes proximal to prioritized long COVID-associated SNPs (±500kb to the candidate SNP termed as tag SNP). Genes showing nominal significance ($P<0.05$) in any of the COVID TWASs across the 48 GTEx tissues were prioritized for further investigation.

To expand these findings, phenome-wide association analyses were conducted for the prioritized genes. Associations were examined across a TWAS database aggregating 193 previously published TWAS results[7] and 229 TWASs generated in this study, yielding a total of 418 TWAS datasets.

Statistical analysis software and availability of codes and data

Unless otherwise specified, all analyses were performed using SAS OnDemand for Academics (freely available at https://www.sas.com/en_us/software/on-demand-for-academics.html). The scripts and datasets used in this study are available from the corresponding author upon reasonable request.

Local Manhattan plots for the top candidate long COVID-associated SNPs and all supplementary tables are publicly accessible at the GitHub link (https://github.com/chengzhongshan/COVID19_GWAS_Analyzer/tree/main/LongCOVID_data_and_scripts/LongCOVID_SNP_LocalManhattanPlots).

Results

Prioritizing Genetic Variants Associated with Long COVID Using COVID-19 GWAS Data

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[1, 3]. To uncover additional genetic long COVID risk loci, an integrative workflow was developed (Fig. 1a) by leveraging publicly available COVID-19 GWAS data. Acute COVID-19 phenotypes, such as SARS-CoV-2 susceptibility, COVID-19 hospitalization, and non-hospitalization, were used as proxy traits for the long COVID phenotype to prioritize genetic variants potentially linked to

long COVID. The analysis utilized data from the HGI release 7, including GWAS of hospitalized vs. non-hospitalized cases (HGI-B1), hospitalized cases vs. the general population (HGI-B2), and SARS-CoV-2 infection vs. the general population (HGI-C2). These GWASs comprise large sample sizes (16,512, 44,986, and 159,840 COVID-19 cases for HGI-B1, HGI-B2, and HGI-C2, respectively; Fig. 1a) and exhibit minimal genomic inflation (λ close to 1). As shown in Supplementary Fig. S1 and Fig. 1a, loci near *SLC6A20* and *ABO* displayed notable significance in association signals. These regions, previously validated as COVID-19 risk loci in the literature[9], are well-established contributors to COVID-19 susceptibility and severity. Taken together, the robustness of these three large COVID-19 GWAS datasets, coupled with the strong correlation between long COVID, SARS-CoV-2 susceptibility, and COVID-19 severity, underscores their utility in identifying potential genetic risk variants for long COVID.

Categorization of Prioritized Long COVID Risk SNPs Related to COVID-19 Severity

To explore the genetic basis of long COVID, we hypothesized that large, previously published COVID-19 GWAS datasets could identify SNPs associated with long COVID. Using an integrative workflow (Fig. 1), we identified 42 prioritized SNPs (see Table 1), categorized into two groups: severe COVID-19-specific SNPs (n=27) and SNPs associated with both severe and mild COVID-19 (n=15). The first category includes SNPs linked predominantly to SARS-CoV-2 susceptibility and severe disease outcomes, such as hospitalization, and are more prevalent in hospitalized cases than in non-hospitalized cases. The second category consists of SNPs associated with both severe and mild COVID-19, showing consistent associations with SARS-CoV-2 susceptibility and hospitalization, but no significant differentiation between hospitalized and non-hospitalized cases, i.e., $P>0.05$ in HGI-B1, $P<0.05$ in HGI-B2 and HGI-C2.

Of the 42 SNPs, except two SNPs, rs150345524 (*SF3B1*, P=5.6E-8 in HGI-B2) and rs11766643 (*SLC22A5*, P=7.5E-8 in HGI-C2), all SNPs passed the conventional genome-wide significance threshold (P<5E-8) in at least one HGI COVID-19 GWAS. This underscores the robustness of the identified associations. These 42 independent variants, derived by leveraging acute-phase COVID-19 GWAS data, highlight the potential genetic contributors to long COVID while addressing the challenges of cohort size limitations and phenotype heterogeneity in long COVID studies.

To reveal biological insights from SNP-gene associations, each severe COVID-19-associated SNP was linked with its corresponding nearest gene to create a SNP-gene pair, providing an intuitive framework for assessing their biological relevance in long COVID. However, proximity-based assignments may overlook distal regulatory effects, which will be further addressed in subsequent analyses. Notable SNPs in the severe COVID-19-specific category include rs10774679 (*OAS3*), rs17169628 (*SLC22A31*), rs12329760 (*TMPRSS2*), rs12660421 (*FOXP4*), rs17078348 (*SLC6A20*), rs190509934 (*ACE2*),

rs2834164 (*IFNAR2*), rs3785884 (*MAPT*), rs7251000 (*DPP9*), and rs76608815 (*SLC5A3*), among others. These SNPs show stronger associations with severe COVID-19 than with mild cases (n=27). In the common SNP category, 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. Notably, *ABO* (rs505922) has been reported as a genetic locus associated with long COVID[3]. Several immune-related genes were found close to these SNPs, including *IFNA21* (rs149533170), *TYK2* (rs34536443), *GSDMA/B* (rs9916158), *BCL11A* (rs1123573), *ELF5* (rs2924480), and *JAK1* (rs11208552), highlighting the role of immune dysregulation in long COVID. Additionally, *FUT2* is found close to rs492602, and the protein of FUT2 encodes fucosyltransferase 2, a key enzyme in the ABO antigen synthesis pathway, further underscores the involvement of ABO biology in COVID-19 outcomes[10]. Additionally, four genes—*TAC4*, *NXPE3*, *NSF*, and *FOXG1*—linked to neuronal functions were close to rs75432325, rs2290859, rs1378358, and rs150497803, respectively, while three genes—*ZKSCAN1* (rs2897075), *BCL11A* (rs1123573), and *FOXG1* (rs150497803)—identified as transcription factors.

Evaluation of Odds Ratios for 42 Prioritized Long COVID Risk SNPs

To characterize the 42 candidate long COVID risk SNPs, we evaluated their odds ratios (ORs) across three GWAS datasets: SARS-CoV-2 susceptibility (HGI-C2), COVID-19 hospitalization (HGI-B1 and HGI-B2; the first used mild COVID-19 and the second used general population as controls). The forest plot (Fig. 1B) illustrates the ORs and 95% confidence intervals (CIs) of these SNPs, stratified by their classification as severe COVID-19-specific SNPs or SNPs associated with both severe and mild COVID-19. This stratification enabled detailed comparisons of effect sizes across the SNP categories and datasets.

Distinct differences in OR distributions were observed between the two SNP categories (Fig. 1B). SNPs in the severe COVID-19-specific category exhibited greater deviation in ORs from 1 compared to those associated with both severe and mild COVID-19. To quantify these differences, absolute effect sizes (β) corresponding to each OR were used for statistical comparisons, as β values account for both protective (OR<1) and risk (OR>1) effects in a unified framework. As shown in Supplementary Fig. S2, SNPs classified as severe COVID-19-specific displayed a higher frequency of statistical significance ($P<0.05$) in absolute effect sizes across GWAS datasets. These SNPs also demonstrated significantly larger absolute effect sizes in comparisons between HGI-B1/HGI-B2 and HGI-C2, underscoring their stronger association with COVID-19 hospitalization. Notably, only in HGI-B1 did the absolute effect sizes differ significantly between the two SNP categories, a result reflecting the selection criteria: severe COVID-19-specific SNPs required $P<0.05$ in both HGI-B1 and HGI-B2, whereas SNPs associated with both severe and mild COVID-19 needed $P<0.05$ in HGI-B2 but not HGI-B1. These findings emphasize the expected

differences stemming from category-specific filtering criteria while highlighting the relevance of severe COVID-19-specific SNPs to hospitalization outcomes and their potential role in long COVID.

To further dissect the associations, SNPs were classified into detailed sub-categories based on their association p-values, as reflected by the color-coded ORs in Fig. 1B for 4 categories, including category 1 (Hosp-specific, *P*<5E-8): SNPs achieving genome-wide significance specifically for hospitalization, suggesting these loci are risk factors for severe COVID-19 and long COVID, category 2 (Common, *P*<5E-8): SNPs achieving genome-wide significance for general susceptibility or hospitalization without significant allele frequency differences between hospitalized and non-hospitalized cases, indicating broader applicability, category 3 (Hosp-specific, suggestive/non-significant): SNPs significant only at a suggestive and non-significant associations in hospitalized cases, and category 4 (Common, suggestive/non-significant): SNPs with suggestive or non-significant associations across both case types. Among the 42 SNPs, three—including rs17078348 (*SLC6A20*), rs12660421 (*FOXP4*), and rs2834164 (*IFNAR2*)—achieved genome-wide significance (*P*<5E-8) across all three GWAS datasets. Thirty-six SNPs met this threshold in HGI-B2 alone. SNPs such as rs17078348 (*SLC6A20*), rs12660421 (*FOXP4*), rs1405655 (*NR1H2*), rs61860402 (*SFTPD*), rs10774679 (*OAS3*), and rs7251000 (*DPP9*), with significant ORs confined to hospitalization (Category 1), are particularly relevant to severe COVID-19 outcomes potentially predisposing individuals to long COVID. Conversely, SNPs such as rs505922 (*ABO*), rs492602 (*FUT2*), rs2290859 (*NXPE3*), and rs1123573 (*BCL11A*), which exhibit non-specific associations across patient groups (Category 2), may indicate general infection susceptibility factors exacerbating long COVID pathology. Notably, *ABO* and *FUT2* loci have established roles in immune function and host susceptibility[10]. In conclusion, these results provide a comprehensive evaluation of SNPs associated with long COVID, leveraging COVID-19 GWAS data to elucidate their contributions to disease susceptibility and severity while identifying biologically relevant loci for further long COVID investigation.

Expanded Screening of Potential Causal Genes Linked to 42 Prioritized Long COVID Risk SNPs

To identify potential causal genes associated with 42 prioritized long COVID risk SNPs, we employed local Manhattan plots, centering each candidate SNP within a 500-kb genomic window (Fig. 2 & 3). Given that 41 of these SNPs, except rs12329760 (*TMPRSS2*), are located in non-coding regions, manual evaluation was required to prioritize genes proximal to each SNP with robust association signals.

The first category of long COVID SNPs (n=27) includes variants associated with severe COVID-19 hospitalization but not non-hospitalized cases. Based on the evaluation of association patterns, nine SNPs were prioritized (Fig. 2): rs17078348 (*SLC6A20*), rs12660421 (*FOXP4*), rs12329760 (*TMPRSS2*),

rs190509934 (*ACE2*), rs61860402 (*SFTPD*), rs10774679 (*OAS3*), rs7251000 (*DPP9*), rs3785884 (*MAPT*), and rs2834164 (*IFNAR2*). Strong association signals (P <1E-30) were observed for rs17078348 (*SLCA20*), with the locus encompassing several genes, including *LZTFL1*, *CCR9*, *FYCO1*, *CCR1* and *XCR1*, previously implicated in COVID-19 severity and susceptibility. Similarly, rs12660421 (*FOXP4*) showed significant association signals near *FOXP4*, a known long COVID risk gene, with the SNP residing within its promoter region and relatively distant from other nearby genes. For rs12329760 (*TMPRSS2*), a missense variant, nominally significant associations with COVID-19 hospitalization (HGI-B2) and SARS-CoV-2 susceptibility (HGI-C2) were detected. Notably, nearby SNPs exhibited weaker signals (P<0.05) in HGI-B1 and differential effect sizes between HGI-B1 and HGI-B2, suggesting a complex genetic architecture. Several antiviral genes, including *MX1* and *MX2*, were identified near this locus. Other prioritized loci included rs61860402 (*SFTPD*) and rs7251000 (*DPP9*), two genes relevant to lung function, and rs10774679 (*OAS3*) and rs2834164 (*IFNAR2*), which are located within interferon-related gene clusters. While rs3785884 (*MAPT*) was initially implicated, its association may be driven by *CRHR1*, a gene associated with age-related neurodegenerative processes[11], aligning with the increased risk of severe COVID-19 and long COVID in older populations.

The second category of SNPs (n=15), associated with both severe and mild COVID-19 cases, yielded nine prioritized loci (Fig. 3): rs149533170 (*IFNA21*), rs34536443 (*TYK2*), rs75432325 (*TAC4*), rs505922 (*ABO*), rs1120591 (*RAB2A*), rs3897075 (*ZKSCAN1*), rs9916158 (*GSDMA/B*), rs1378358 (*NSF*), and rs2924480 (*ELF5*). Among these, rs505922 (*ABO*) and rs34536443 (*TYK2*) displayed genome-wide significant associations with both hospitalized and non-hospitalized COVID-19, as evidenced by consistent signals in HGI-B2 and HGI-C2 datasets. The remaining SNPs were located in gene-dense regions, complicating the identification of specific causal genes. For instance, rs149533170 (*IFNA21*) is located near a cluster of interferon genes, including *IFNB1* and multiple *IFNAs*, which are central to interferon-mediated antiviral responses[12]. Similarly, rs9916158 (*GSDMA/B*) is proximal to genes implicated in asthma, while rs75432325 (*TAC4*) resides near *DLX3* and *DLX4*, which are involved in developmental pathways[13]. Less complex loci, such as rs1120591 (*RAB2A*) and rs2924480 (*ELF5*), contained fewer neighboring genes, facilitating clearer prioritization. *RAB2A* is an insulin regulator[14], while *ELF5* encodes a transcription factor relevant to epithelial differentiation and innate immune response[15].

In total, 18 SNPs were prioritized based on their local Manhattan plots, representing two categories of long COVID risk loci. These SNPs are associated with biologically plausible candidate genes involved in pathways related to COVID-19 severity, immune response, and tissue-specific functions. Further experimental validation is required to determine their roles in long COVID.

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[16, 17], potentially due to stronger immune responses to SARS-CoV-2 infection in women. To investigate whether the top 42 candidate SNPs associated with long COVID exhibit sex-biased associations with COVID-19 hospitalization, we analyzed publicly available sex-stratified GWAS summary statistics from the GRASP COVID-19 database[18] generated based on UK Biobank samples.

The largest sex-stratified COVID-19 hospitalization GWASs (cases=1,343 and controls=262,886 for females; cases=1,917 and controls=221,174 for males) were selected, which includes mixed ancestry populations comparable to the four HGI COVID-19 GWAS datasets used in our study. Using the delta-z-score method to compare male and female GWAS results, SNPs with differential effect sizes for COVID-19 hospitalization were identified (Fig. 4A). Genomic inflation factors for both sex-stratified GWASs and the sex-biased differential GWAS were close to 1, indicating no evidence of inflation (Fig. 4B).

For these sex-stratified COVID-19 hospitalization GWASs, compact local Manhattan plots (Fig. 4C) and forest plots (Fig. 4D) were generated for these 42 candidate SNPs. Due to the smaller sample sizes in the sex-stratified GWASs, only 15 SNPs reached nominal significance (P<0.05) in at least one of the three analyses. Of these, rs17078348 (*SLC6A20*), a well-established severe COVID-19 risk SNP, showed nominal significance in both male and female cohorts. Five SNPs displayed nominal significance exclusively in females, including rs75432325 (*TAC4*), rs505922 (*ABO*), rs12660421 (*FOXP4*), rs7251000 (*DPP9*), and rs11766643 (*TFR2*). Eight SNPs were nominally significant only in males, including rs897075 (*ZKSCAN1*), rs117169628 (*SLC22A31*), rs7515509 (*AK5*), rs76608815 (*SLC5A3*), rs12585036 (*ATP11A*), rs5023077 (*FBRSL1*), rs10774679 (*IFNAR21*), and rs12329760 (*HLA-C*). None of the SNPs passed the stringent multiple testing correction threshold (P<6E-4), yet local Manhattan plots for 12 of the 42 SNPs revealed relatively weaker association signals in the sex-stratified GWAS compared to the stronger signals observed in HGI-B2 (Fig. 5). This reduction in signal strength likely reflects the smaller sample sizes in the sex-stratified datasets, which reduced statistical power to detect sex-biased associations. Finally, aggregating the absolute effect sizes of SNPs across the four groups (female-SNP-category 1, female-SNP-category 2, male-SNP-category 1, and male-SNP-category 2) based on the nominal association P<0.05 revealed no significant differences (ANOVA P=0.15). In summary, while the sex-biased GWAS analysis did not identify SNPs with statistically significant sex-biased associations with COVID-19 hospitalization, it highlighted 15 SNPs with nominal significance, including rs505922 (*ABO*) and rs12660421 (*FOXP4*), which have been previously implicated as risk variants for both severe COVID-19 and long COVID. These findings warrant further investigation in larger sex-stratified COVID-

19 and long COVID GWAS.

Evaluation of Previously Reported Long COVID-Associated SNPs

To assess whether the 42 prioritized long COVID SNPs identified in our study overlap with or are proximal to those reported in two independent long COVID GWASs by Taylor et al.[2] and Chaudhary et al.[3], a set of 83 SNPs derived from SNP network analyses and 3 SNPs identified via traditional GWAS methods from these studies was created. These SNPs were analyzed for their association signals in three HGI GWAS datasets (supplementary Table S1). Notably, rs12660421 (*FOXP4*), identified by the HGI long COVID GWAS, was excluded from this evaluation as it displayed stronger associations with general COVID-19 phenotypes in the HGI datasets analyzed here. This SNP was already included among the 42 prioritized long COVID risk SNPs.

Our evaluation was guided by the hypothesis that true long COVID associations reported by Taylor et al.[2] and Chaudhary et al.[3]. would exhibit at least nominal significance in one of the HGI GWAS datasets. The larger sample sizes and higher statistical power of the HGI datasets enhance their ability to detect associations with traits closely related to long COVID, such as severe COVID-19 or susceptibility to SARS-CoV-2 infection. Moreover, overlap between our prioritized SNPs and those reported in previous studies would strengthen the hypothesis that the loci in question are implicated in long COVID pathogenesis.

Among the three SNPs reported by Chaudhary et al.[3] (Fig. 6), rs643434 (*ABO*) and rs2080090 (*BPTF*) demonstrated nominally significant associations ($P<0.05$) with COVID-19 hospitalization (HGI-B2) and susceptibility to COVID-19 (HGI-C2). rs643434 (*ABO*) surpassed genome-wide significance thresholds in both HGI-B2 (P= 6.65E-24) and HGI-C2 (P=8.69E-85) and showed non-significant association in HGI-B2 (P=0.6) but demonstrated significantly differential effect sizes (P=6E-5) between HGI-B1 (hospitalized vs. non-hospitalized COVID-19) and HGI-B2. This supported its classification as a "category 2" long COVID SNP (associated with both severe and mild COVID-19 phenotypes). Similarly, rs2080090 (*BPTF*) showed stronger associations in HGI-B2 (P= 1.9E-4) than in HGI-C2 (P=2.5E-2), with no significant association in HGI-B1. This SNP was also categorized as a suggestive "category 2" long COVID SNP. In contrast, rs9273363 (*HLA-DQB1*) did not exhibit significant associations across any of the HGI GWAS datasets analyzed ($P>0.2$). Nevertheless, among these 42 candidate SNPs, rs17219281 (*HLA-DQA1*), a “category 1” long COVID SNP that tends to be more 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 long COVID. Taken together, the three recently published long COVID loci, including *ABO*, *BPTF*, and *HLA-DQB1*, were completely covered

by these 42 prioritized long COVID SNPs.

Meanwhile, of the 80 SNPs reported by Taylor et al.[2], 12 displayed nominally significant associations in at least one HGI GWAS dataset (Supplementary Table S2 and Fig. 6). These included four SNPs associated with HGI-B2, five with HGI-B1, and seven with HGI-C2. Notably, rs4303401 (*FRMD6*) surpassed the multiple-testing correction threshold (P<1.5E-4), representing a strong candidate SNP. Expanding the analysis to include a ±500kb genomic window around the reported SNPs identified additional suggestive genome-wide associations in the HGI datasets. For example, in Fig. 6, rs62401842 (proximal to *KCTD16*, which shows brain-specific expression) and rs7671107 and rs1017716 (adjacent to *ANAPC4*, a gene implicated in aging and strongly supported as eQTLs accordingly to GTEx eQTL database) emerged as promising candidates. Other noteworthy SNPs in Fig. 6 included rs12602210 (*BPTF*) and rs17412601 and rs2290859 (both mapped to *NXPE3*), which demonstrated suggestive or genome-wide significance in at least one HGI GWAS dataset. In summary, four loci, including *KCTD16*, *ANAPC4*, *BPTF*, and *NXPE3*, revealed by Taylor et al[2]. derived from SNP network analysis are independently confirmed to be nominally associated with COVID-19 related phenotypes in the current study; among them *BPTF* is a loci confirmed to be associated with long COVID by Chaudhary et al.[3], Taylor et al., and current study. Therefore, these findings underscore the potential of integrating data from diverse GWAS to elucidate the genetic architecture of long COVID and highlight candidate loci for further investigation.

Potential Long COVID SNPs Specifically Identified in Non-Hospitalized COVID-19 Cases

To identify potential long COVID SNPs specific to non-hospitalized COVID-19 cases, we screened SNPs that showed suggestive genome-wide association signals in HGI-B1 (hospitalized vs. non-hospitalized, P< 1E-5) and HGI-B1-vs-B2 (differential GWAS between HGI-B1 and HGI-B2, P<0.05) but were not significant in HGI-B2 (hospitalized vs. general population, P>0.05) and were not in proximity to any genome-wide significant SNPs in HGI-B2. This selection identified 20 independent SNPs potentially associated with long COVID that specifically emerged from mild (non-hospitalized) COVID-19 cases (Table 2; local Manhattan plots accessible at GitHub link provided in the Method section; Fig. 7). All 20 SNPs showed suggestive genome-wide significant differences in effect sizes between HGI-B1 and HGI-B2 (P<6E-5). Although none met the genome-wide significance threshold (P<5E-8), notable findings were observed. For example, in Fig. 7, rs62401842 (*KCTD16*) exhibited the strongest association in HGI-B1 (P=8.53E-7), while showing no association in HGI-B2 (P=0.38) and a significant difference between HGI-B1 and HGI-B2 (P=3.18E-7). Gene *KCTD16*, encoding potassium channel tetramerization domain-containing 16[19], is highly expressed in brain tissues according to the GTEx database. In addition, eight SNPs illustrated in Fig. 7 were mapped to genes with brain-related functions, including rs9799354

(*NLGN1*), rs112842080 (*CPLX2*), rs61858037 (*NRG3*), rs61939166 (*KIF21A*), rs367777 (*NAV3*), rs56143829 (*WASF3*), rs11454577 (*AKAP6*), and rs6049828 (*SYNDIG1*). Among these, *WASF3* has been implicated in mitochondrial dysfunction and may mediate exercise intolerance in myalgic encephalomyelitis/chronic fatigue syndrome[20]. Additionally, among the 20 candidate SNPs identified, only three (rs62401842 [*KCTD16*], rs9799354 [*NLGN1*], and rs112842080 [*CPLX2*]) exhibit male-biased associations with COVID-19 hospitalization at nominal significance levels (Fig. 8A, B). 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. 8C-H). 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 provide suggestive evidence of association with long COVID in mild COVID-19.

Transcriptome-Wide Association Study Identifies Candidate Genes near Prioritized SNPs for COVID-19 and Long COVID Phenotypes

To elucidate genetic mechanisms underlying COVID-19 and long COVID phenotypes, we performed a transcriptome-wide association study (TWAS) using S-MetaXcan across 48 GTEx tissues. This approach utilized COVID-19 GWAS summary statistics from HGI: HGI-B1, HGI-B2, HGI-C2, and HGI-B1-vs-B2. The analysis targeted 62 candidate SNPs previously prioritized for their potential association with long COVID and aimed to identify their adjacent genes whose expression levels may be associated with COVID-19 hospitalization or susceptibility phenotypes, as well as long COVID.

Of the 62 candidate SNPs, two SNPs (rs190509934 near *ACE2* and rs9570861 near *OR7E156P*) not included in the TWAS analysis. The former resides on chromosome X, which S-MetaXcan excludes due to its focus on autosomal genes, while the latter SNP had no protein-coding genes within a 500 kb window. The remaining 60 SNPs corresponded to 753 adjacent genes tested by these COVID TWASs (Fig. S5) with detectable expression across the 48 GTEx tissues.

To simplify interpretation, we retained only the strongest association (smallest p-value) for each gene across 48 GTEx tissues. This approach was justified by the hypothesis that specific tissues, rather than all tissues, are likely to drive the pathogenicity of COVID-19. Applying varying significance thresholds (P<0.05, 0.01, 0.001, 0.0001, 0.00001), the analysis revealed 422, 152, 43, 19, and 11 adjacent genes meeting these criteria in at least one of the four HGI datasets (Fig. S3, Table S2). Using a genome-wide

significance threshold (*P*<3.5E-7), three loci displayed significant associations: (1) *DPP9*, showing the strongest association signal (*P*=4.6E-18) in adipose visceral omentum tissue; (2) *CXCR6* gene cluster, which included multiple genes (*CXCR6*, *CCR3*, *CCR1*, and *LZTFL1*), displaying the most significant association (*P*=1.3E-10) in nerve tibial tissue; (3) *MUC1* gene cluster, comprising *MUC1* and *FAM189B*, highlighted in multiple GTEx tissues, including lung and colon. Additional loci with suggestive TWAS significance (*P*<3.5E-6) included *FYCO1*, *SLC6A20*, *XCR1*, and *THBS3*, with specific associations to HGI datasets B1, B2, or C2 in distinct GTEx tissues.

To investigate overlap with long COVID, TWAS was extended to four long COVID GWASs using the same S-MetaXcan framework. Genome-wide evaluation of TWAS signals across the eight HGI GWAS datasets (Fig. S4) confirmed that the top hits with association *P*<1E-5 originated exclusively from COVID-19 hospitalization and susceptibility GWAS, with no contributions from long COVID GWAS (Fig. S4). Notably, the candidate genes proximal to these prioritized SNPs accounted for most of the top signals across the eight HGI GWAS (Fig. 9), with the exception of *CCR5*, *EXOSC7*, *FLT1P1*, *RLN1*, *GBAP1*, *ARHGEF2*, *ZSCAN31*, *ZKSCAN3*, *PGMSP2*, *PAQR5*, and *TMIE*. Further in-depth evaluation of these outlier genes in long COVID GWAS revealed most of them lack of significance (Fig. S5), reinforcing the utility of prioritizing top candidate SNPs and their adjacent genes for downstream analyses. Interestingly, although the strongest associations identified in COVID-19 hospitalization or susceptibility GWASs (*P*<1E-5) were not genome-wide significant in long COVID datasets, four genes (*CCR1*, *DPP9*, *MUC1*, and *THBS3*; Fig. S5) showed nominally significant associations (*P*<0.05) in long COVID GWASs across specific tissues, such as colon transverse, testis, and heart atrial appendage. These associations were generally weaker than those observed for acute COVID-19 phenotypes, suggesting tissue-specific contributions to long COVID pathophysiology. In details, *DPP9* demonstrated strong associations in adipose-related tissues for COVID-19 hospitalization but weaker signals in colon transverse for long COVID. *THBS3* showed widespread associations across 19 GTEx tissues for COVID-19 hospitalization but exhibited weaker signals in long COVID, particularly in colon sigmoid and spleen. *MUC1* and *CCR1* exhibited similar patterns, with weaker long COVID associations in colon transverse and heart atrial appendage, respectively. Furthermore, using a relaxed threshold (*P*<0.05) identified additional adjacent genes associated with long COVID (Fig. 10 & Fig. S6-7). These included *OAS3* (rs10774679), *TMPRSS2* (rs12329760), *HLA-A* (rs4s9260038), *HLA-C* (rs1634761), *HLA-DQA1* (rs17219281), *WASF3* (rs65143829), *IFNAR2* (rs2834164), *SFTPD* (rs61860402), and *GSDMA/B* (rs9916158), with associations enriched in tissues such as heart, brain, colon, adipose, and muscle. Notably, the well-established *ABO* locus, a COVID-19 susceptibility locus, did not show significant TWAS signals. This suggests that its effects may stem from non-expression-related mechanisms, such as genetic mutations influencing blood types.

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

Phenome-Wide Analysis of Gene Expression Associations with Candidate Genes

To investigate gene expression associations with long COVID-related phenotypes, we performed a phenome-wide analysis of gene expression associations (Pheno-TWAS) analysis targeting 918 tested TWAS genes located within a 1Mb window of 62 candidate SNPs previously linked to COVID-19. Recognizing the limitations of published long COVID GWASs, which often rely on imprecisely defined phenotypes, this analysis extended beyond prior long COVID studies by utilizing a manually curated TWAS database encompassing 418 phenotypes across 48 GTEx tissues. Of the 918 tested TWAS genes, 400 (43.6%), including 87 associated with *HLA* loci (e.g., *HLA-DQA2*, *HLA-A*, and *HLA-C*; Fig. S8-10), passed a stringent bonferroni-adjusted significance threshold (P<2.7E-9). Genes located near *HLA* loci emerged as top hits, demonstrating significant associations with immune-related disorders such as asthma, rheumatoid arthritis, psoriasis, coeliac disease, and systemic lupus erythematosus—conditions that may contribute to long COVID susceptibility. In addition to *HLA*-associated loci, 313 genes across 44 non-*HLA* loci surpassed the genome-wide significance threshold in TWASs. Key findings include: (1) *IFNAR2*: strongly associated with COVID-19 hospitalization (P=2.6E-10; Fig. S11); (2) *GSDMB*: linked to asthma (P=2.7E-57; Fig. S12); (3) *BPTF*: associated with trunk fat percentage (P=1E-18; Fig. S13); (4) *CCDC171*: associated with body fat percentage (P=2E-35; Fig. S14). Interestingly, *FOXP4* expression, which did not exhibit significant associations with long COVID phenotypes in prior TWASs, demonstrated a suggestive association with sleep duration (P=1.1E-5; Fig. S15). This finding highlights the potential for phenome-wide approaches to uncover previously uncharacterized associations relevant to long COVID. In total, pheno-TWAS identified 45 candidate SNPs associated with genome-wide significant expression-linked phenotypes. These include three major *HLA* loci (*HLA-DQA2*, *HLA-A*, and *HLA-C*) as well as non-*HLA* loci such as *IFNAR2*, *BPTF*, *CCDC171*, and *GSDMB*. The implicated phenotypes—ranging from COVID-19 hospitalization and immune disorders to body composition traits and sleep duration—provide a broad spectrum of potential biological pathways contributing to long COVID pathophysiology. Full results, including nominally significant associations, are provided in Supplementary Table S3.

Discussion

Our integrative analysis identifies 62 candidate loci potentially associated with long COVID, leveraging large-scale COVID-19 GWAS datasets from the HGI. This approach addresses the inherent limitations of long COVID-specific GWAS, such as limited sample sizes and phenotype heterogeneity, by utilizing proxy phenotypes like SARS-CoV-2 susceptibility and COVID-19 hospitalization. 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 provide a scalable framework to investigate the genetic underpinnings of long COVID.

The analysis highlights 42 genome-wide or suggestive loci, 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, underscoring the complexity of long COVID's etiology.

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 long COVID's heterogeneity.

Our findings underscore the utility of integrative approaches to prioritize genetic risk loci for complex traits like long COVID. The discovery of loci such as *FOXP*4, *ABO*, and *HLA-DQA1*, as well as *BPTF*, suggests shared biological pathways between acute COVID-19 severity and its chronic manifestations. 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 data. No individual-identifiable information, such as personal identifiers or images, was 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. Additionally, supplementary data are available at the GitHub repository: https://github.com/chengzhongshan/COVID19_GWAS_Analyzer.

Competing interest

The author has no competing interests or potential conflicts of interest.

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 without their generosities in freely sharing these COVID-19 hospitalization GWASs and other COVID-19/long COVID related GWASs as these were essential for the current finding.

TABLES

Table 1. Prioritized long COVID SNPs based on integrative analysis of published COVID-19 GWASs

Variant information						HGI-B1			HGI-B2			HGI-C2			HGI-B1-vs-B2	Long COVID
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

Note: HGI: the COVID-19 host genetics initiative; chr: chromosome; pos: hg38 genomic position; allele: alternative allele; AF: allele frequency for alternative allele; diffPval: statistical significance *P* for differential effect size of each common SNP by comparing the two HGI GWASs, including HGI-B1 and HGI-B2; Hosp: long COVID SNPs tend to be specifically emerged in hospitalized COVID-19 cases; Both: long COVID SNPs tend to be emerged in both hospitalized and non-hospitalized COVID-19 cases. Association Ps passed the suggestive genome-wide association significance threshold of P < 1E-7 are highlighted in light red color, and differential effect size significance P < 0.05 is colored with light green. Note: HGI-C2 (SASR-CoV2 susceptibility), HGI-B1 (hospitalized vs. non-hospitalized COVID-19), and HGI-B2 (hospitalized vs. general population).

Table 2. Potential long COVID risk SNPs emerged specifically from 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

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

SUPPLEMENTARY TABLES

Table S1. Association signals of published long COVID SNPs across 4 HGI COVID-19 GWASs

Table S2. Adjacent genes of these 62 candidate SNPs that pass the nominal significance threshold in the transcriptome-wide association studies (TWAS) of COVID-19 hospitalization and susceptibility

Table S3. Adjacent genes of these 62 candidate SNPs that pass the nominal significance threshold in the phenome-wide transcriptome association studies (pheno-TWASs)

FIGURES

Figure 1. Workflow to prioritize genetic single nucleotide polymorphisms (SNPs) potentially associated with long COVID

(A) 42 top prioritized long COVID risk SNPs, with 27 SNPs associated specifically with hospitalization risk and 15 showing common risk patterns across both hospitalized and non-hospitalized cases. These SNPs were selected based on a significance threshold of $P \leq 1E\text{-}7$ in at least one of the GWAS studies (HGI-C2, HGI-B2, or HGI-B1). Each SNP is centered within a 1Mb window in local Manhattan plots to highlight the genomic context, including known COVID-19 and long COVID risk loci. Key genes near these SNPs, including *FOXP4*—a locus previously implicated in long COVID risk—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 and long COVID risk. (B) Forest plot of 42 SNPs identified from COVID-19 GWAS on susceptibility (HGI-C2) and hospitalization (HGI-B1 and HGI-B2), showing suggestive ($P<1E\text{-}7$ or genome-wide significant associations ($P<5E\text{-}8$). 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, $P<5E\text{-}8$); (2) common across cases, genome-wide significant (Common, $P<5E\text{-}8$); (3) hospitalization-specific, suggestive or non-significant (hosp-specific, P = (5E-8, 1]); (4) common across cases, suggestive or non-significant (Common, P = (5E-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) are non-significant.

Figure 2. Local Manhattan plots for potential long COVID risk SNPs more likely associated with hospitalized than non-hospitalized COVID-19

Figure 3. Local Manhattan plots for potential long COVID risk SNPs likely associated with both hospitalized and non-hospitalized COVID-19

Figure 4. Evaluating top potential long COVID risk SNPs 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 long COVID SNPs that tend to be specific in hospitalized cases and that more frequent in both hospitalized and non-hospitalized cases. All association signals shown in the figure are within a 100Mb genomic window for each potential long COVID risk SNP residing at the

center. Genes close to these 42 SNPs are added into the upper track. (C) Forest plots for these 42 potential long COVID risk SNPs in sex-stratified COVID-19 hospitalization GWASs, as well as the differential effect size GWAS between females and males. Odd Ratio (OR) 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 tagged 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 GWASs.

Figure 5. Local Manhattan plots for potential long COVID SNPs displaying suggestively sex-biased association with COVID-19 hospitalization

Figure 6. Evaluating published long COVID SNPs in the current HGI COVID-19 GWASs

(A) Forest plots for these 15 SNPs in currently studied 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 S1) that show nominal significant association in at least one of the 3 current HGI COVID-19 GWASs. (B-G) Local Manhattan plots illustrate published long COVID SNPs approximate to these top candidate long COVID SNPs identified in current study with each candidate SNP is located in the center of each genomic window; GWASs labels from bottom to top in each panel are HGI-B1, HGI-B2, HGI-B1-vs-B2, and HGI-C2, respectively.

Figure 7. Local Manhattan plots for potential long COVID risk SNPs specifically associated with non-hospitalized COVID-19

Figure 8. Evaluation of potential long COVID risk SNPs associated with non-hospitalized COVID-19 in sex-biased COVID-19 hospitalization GWAS from UK Biobank

(A) Local Manhattan plots for candidate long COVID-associated SNPs identified exclusively in the non-hospitalized COVID-19 group from sex-stratified COVID-19 hospitalization GWAS (the differential effect size GWAS between females and males at the top, female GWAS in the middle, and male GWAS shown at the bottom). (B) Forest plots displaying the odds ratios (OR) and corresponding 95% confidence intervals (CI) for these SNPs in non-hospitalized COVID-19 patients from the sex-stratified GWAS (male results at 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 identified specifically in the non-hospitalized COVID-19 group within the sex-stratified GWAS.

Figure 9. Evaluation of top genes adjacent to prioritized long COVID risk 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 long COVID risk SNPs based on transcriptome-wide association studies (TWAS) for COVID-19 susceptibility and hospitalization. Genes are categorized by the GTEx tissue in which they display the most significant association. Besides rs1078348, other two SNPs have multiple genes with the smallest TWAS *P* are highlighted. (B) Heatmap depicting the significance (*P*) of gene-tissue associations for prioritized long COVID 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 ($P<1E-6$). Genes are sorted by TWAS, and GTEx tissues are clustered to highlight patterns of tissue specificity.

Figure 10. Top genes adjacent to prioritized long COVID risk SNPs (n=23) in long COVID-19 transcriptome-wide association studies (TWASs)

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

SUPPLEMENTARY FIGURES

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

Figure S2. Differential effect size analysis for candidate long COVID SNPs by SNP categories and GWAS

Figure S3. Total number of genes adjacent to prioritized long COVID risk SNPs (n=62) in COVID-19 transcriptome-wide association studies (TWASs)

Bar plots illustrating the number of genes adjacent to prioritized long COVID risk SNPs (n=62) as well as number of TWASs out of 4 COVID-19 related GWAS across 48 GTEx tissues passed different

association *P* thresholds. Note that the 62 SNPs are colored according to its corresponding long COVID SNP category as depicted in Figure 1.

Figure S4. 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 derived from four distinct case-control designs. Each dot represents a gene, with its genomic position along the x-axis and -log10(*P*) on the y-axis.

Figure S5. Evaluation of top genes identified from transcriptome-wide association studies (TWAS) of COVID-19 susceptibility and hospitalization in 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 ($P<1E-6$). Rows are sorted by genes, and columns represent tissues, allowing visualization of tissue-specific patterns of gene expression relevance.

Figure S6. Other top genes adjacent to prioritized long COVID risk SNPs (n=29) in long COVID-19 transcriptome-wide association studies (TWASs)

Bar plots illustrating the best *P* for genes adjacent to prioritized long COVID risk SNPs (n=29) in 4 long COVID TWASs across 48 GTEx tissues.

Figure S7. Evaluation of top genes adjacent to prioritized long COVID risk SNPs (n=62) in transcriptome-wide association studies (TWAS) 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 long COVID risk 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 S8. Evaluation of genes adjacent to prioritized long COVID risk 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*<1E-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*<1E-60.

Figure S9. Evaluation of genes adjacent to prioritized long COVID risk SNP rs9260038 (*HLA-A*) in phenome-transcriptome-wide association studies (pheno-TWAS) across 48 GTEx tissues

To the sake of clarity, similar heatmap as that in Figure S8 only depicting the significance of SNP-gene pairs with *P*<1E-60 across 48 GTEx tissues in TWASs.

Figure S10. Evaluation of genes adjacent to prioritized long COVID risk SNP rs1634761 (*HLA-C*) in phenome-transcriptome-wide association studies (pheno-TWAS) across 48 GTEx tissues

To the sake of clarity, similar heatmap as that in Figure S8 only depicting the significance of SNP-gene pairs with *P*<1E-60 across 48 GTEx tissues in TWASs.

Figure S11. Evaluation of genes adjacent to prioritized long COVID risk 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*<1E-6). Rows are sorted by genes, and columns represent TWAS, allowing visualization of tissue-specific patterns of gene expression relevance.across 48 GTEx tissues in TWASs.

Figure S12. Evaluation of genes adjacent to prioritized long COVID risk SNP rs9916158 (*GSDMA/B*) in phenome-transcriptome-wide association studies (pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S11.

Figure S13. Evaluation of genes adjacent to prioritized long COVID risk SNP rs12602210 (*BPTF*) in phenome-transcriptome-wide association studies (pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S11.

Figure S14. Evaluation of genes adjacent to prioritized long COVID risk SNP rs79611697 (*CCDC171*) in phenome-transcriptome-wide association studies (pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S11.

Figure S15. Evaluation of genes adjacent to prioritized long COVID risk SNP rs12660421 (*FOXP4*) in phenome-transcriptome-wide association studies (pheno-TWAS) across 48 GTEx tissues

Heatmap legend as that in Figure S11.

References

1. Lammi V, Nakanishi T, Jones SE, Andrews SJ, Karjalainen J, Cortés B, O'Brien HE, Fulton-Howard BE, Haapaniemi HH, Schmidt A *et al*: **Genome-wide Association Study of Long COVID**. *medRxiv* 2023:2023.2006.2029.23292056.
2. Taylor K, Pearson M, Das S, Sardell J, Chocian K, Gardner S: **Genetic risk factors for severe and fatigue dominant long COVID and commonalities with ME/CFS identified by combinatorial analysis**. *Journal of translational medicine* 2023, **21**(1):775.
3. Chaudhary NS, Weldon CH, Nandakumar P, Team aR, Holmes MV, Aslibekyan S: **Multi-ancestry GWAS of Long COVID identifies immune-related loci and etiological links to chronic fatigue syndrome, fibromyalgia and depression**. *medRxiv* 2024:2024.2010.2007.24315052.
4. Cai M, Xie Y, Topol EJ, Al-Aly Z: **Three-year outcomes of post-acute sequelae of COVID-19**. *Nat Med* 2024, **30**(6):1564-1573.
5. Zhang K, Lin S, Luo YS, Cheng Z: **Protocol to search for genetic factors related to severe COVID-19 by analyzing publicly available genome-wide association studies**. *STAR protocols* 2024, **5**(3):103028.
6. Cheng Z, Cai Y, Zhang K, Zhang J, Gui H, Luo YS, Zhou J, DeVeale B: **MAP3K19 regulatory variation in populations with African ancestry may increase COVID-19 severity**. *iScience* 2023, **26**(9):107555.
7. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, Torstenson ES, Shah KP, Garcia T, Edwards TL *et al*: **Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics**. *Nat Commun* 2018, **9**(1):1825.
8. GTEx-Consortium: **The GTEx Consortium atlas of genetic regulatory effects across human tissues**. *Science* 2020, **369**(6509):1318-1330.
9. Severe-COVID-GWAS-Group, Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, Fernandez J, Prati D, Baselli G *et al*: **Genomewide Association Study of Severe Covid-19 with Respiratory Failure**. *N Engl J Med* 2020, **383**(16):1522-1534.
10. Galeev A, Suwandi A, Cepic A, Basu M, Baines JF, Grassl GA: **The role of the blood group-related glycosyltransferases FUT2 and B4GALNT2 in susceptibility to infectious disease**. *International journal of medical microbiology : IJMM* 2021, **311**(3):151487.
11. Polanczyk G, Caspi A, Williams B, Price TS, Danese A, Sugden K, Uher R, Poulton R, Moffitt TE: **Protective effect of CRHR1 gene variants on the development of adult depression following childhood maltreatment: replication and extension**. *Archives of general psychiatry* 2009, **66**(9):978-985.

12. Samuel CE: **Antiviral actions of interferons**. *Clinical microbiology reviews* 2001, **14**(4):778-809, table of contents.
13. Tamaoki N, Takahashi K, Aoki H, Iida K, Kawaguchi T, Hatakeyama D, Inden M, Chosa N, Ishisaki A, Kunisada T *et al*: **The homeobox gene DLX4 promotes generation of human induced pluripotent stem cells**. *Sci Rep* 2014, **4**:7283.
14. Sugawara T, Kano F, Murata M: **Rab2A is a pivotal switch protein that promotes either secretion or ER-associated degradation of (pro)insulin in insulin-secreting cells**. *Scientific Reports* 2014, **4**(1):6952.
15. Pietzner M, Chua RL, Wheeler E: **ELF5 is a potential respiratory epithelial cell-specific risk gene for severe COVID-19**. 2022, **13**(1):4484.
16. Bai F, Tomasoni D, Falcinella C, Barbanotti D, Castoldi R, Mulè G, Augello M, Mondatore D, Allegrini M, Cona A *et al*: **Female gender is associated with long COVID syndrome: a prospective cohort study**. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2022, **28**(4):611.e619-611.e616.
17. Neuhouser ML, Butt HI, Hu C, Shadyab AH, Garcia L, Follis S, Mouton C, Harris HR, Wactawski-Wende J, Gower EW *et al*: **Risk factors for long COVID syndrome in postmenopausal women with previously reported diagnosis of COVID-19**. *Annals of epidemiology* 2024, **98**:36-43.
18. Thibord F, Chan MV, Chen MH, Johnson AD: **A year of Covid-19 GWAS results from the GRASP portal reveals potential genetic risk factors**. *HGG Adv* 2022:100095.
19. Zuo H, Glaaser I, Zhao Y, Kurinov I, Mosyak L, Wang H, Liu J, Park J, Frangaj A, Sturchler E *et al*: **Structural basis for auxiliary subunit KCTD16 regulation of the GABA(B) receptor**. *Proc Natl Acad Sci U S A* 2019, **116**(17):8370-8379.
20. Wang PY, Ma J, Kim YC, Son AY, Syed AM, Liu C, Mori MP, Huffstutler RD, Stolinski JL, Talagala SL *et al*: **WASF3 disrupts mitochondrial respiration and may mediate exercise intolerance in myalgic encephalomyelitis/chronic fatigue syndrome**. *Proc Natl Acad Sci U S A* 2023, **120**(34):e2302738120.

Figure 1

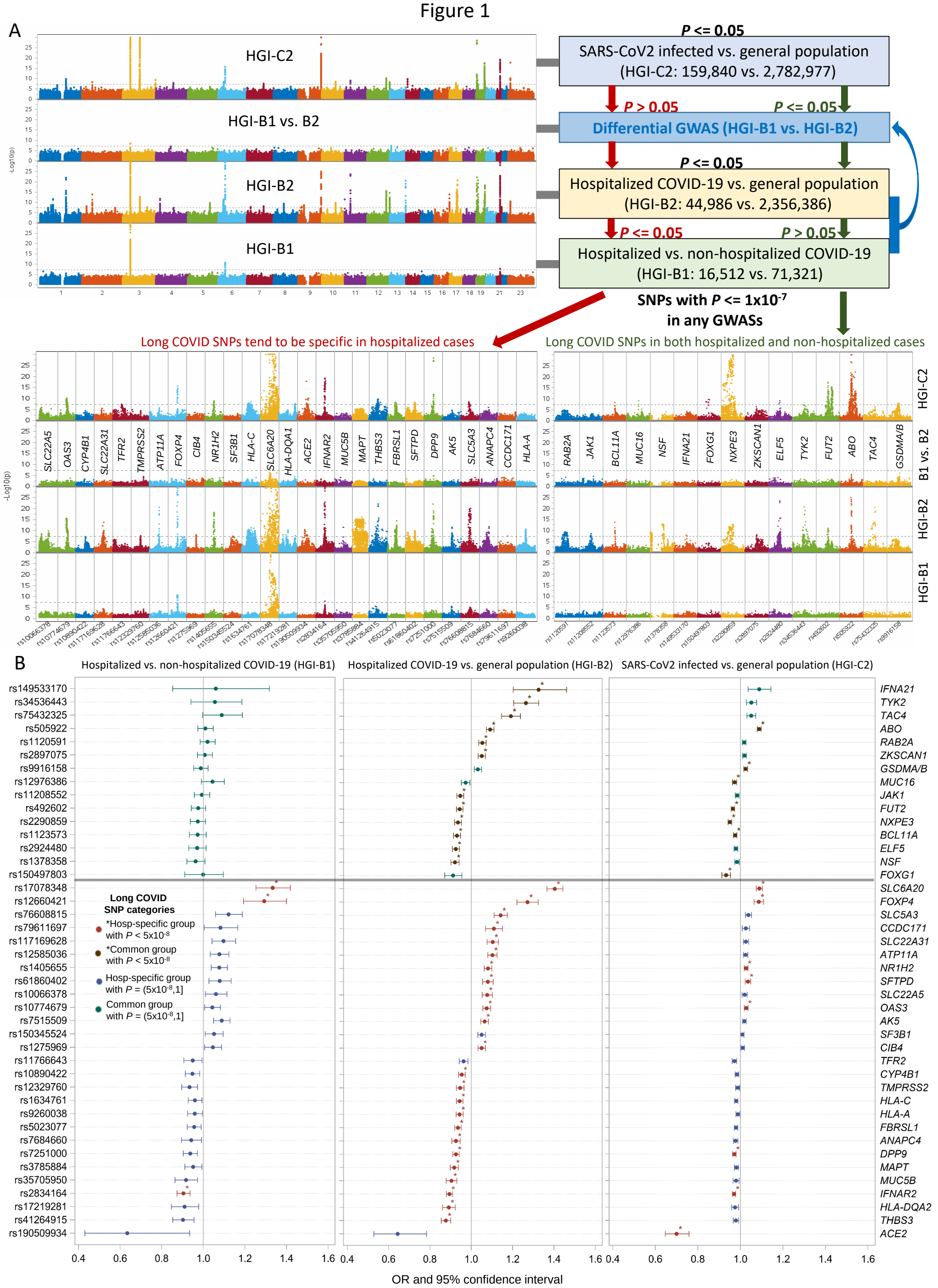

A
HGI-C2
HGI-B1 vs. B2
HGI-B2
HGI-B1
-Log10(p)
P <= 0.05
SARS-CoV2 infected vs. general population
(HGI-C2: 159,840 vs. 2,782,977)
P > 0.05
P <= 0.05
Differential GWAS (HGI-B1 vs. HGI-B2)
P <= 0.05
Hospitalized COVID-19 vs. general population
(HGI-B2: 44,986 vs. 2,356,386)
P <= 0.05
P > 0.05
Hospitalized vs. non-hospitalized COVID-19
(HGI-B1: 16,512 vs. 71,321)
SNPs with P <= 1x10-7
in any GWASs
Long COVID SNPs tend to be specific in hospitalized cases
Long COVID SNPs in both hospitalized and non-hospitalized cases
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
HGI-C2
B1 vs. B2
HGI-B2
HGI-B1
B
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 2

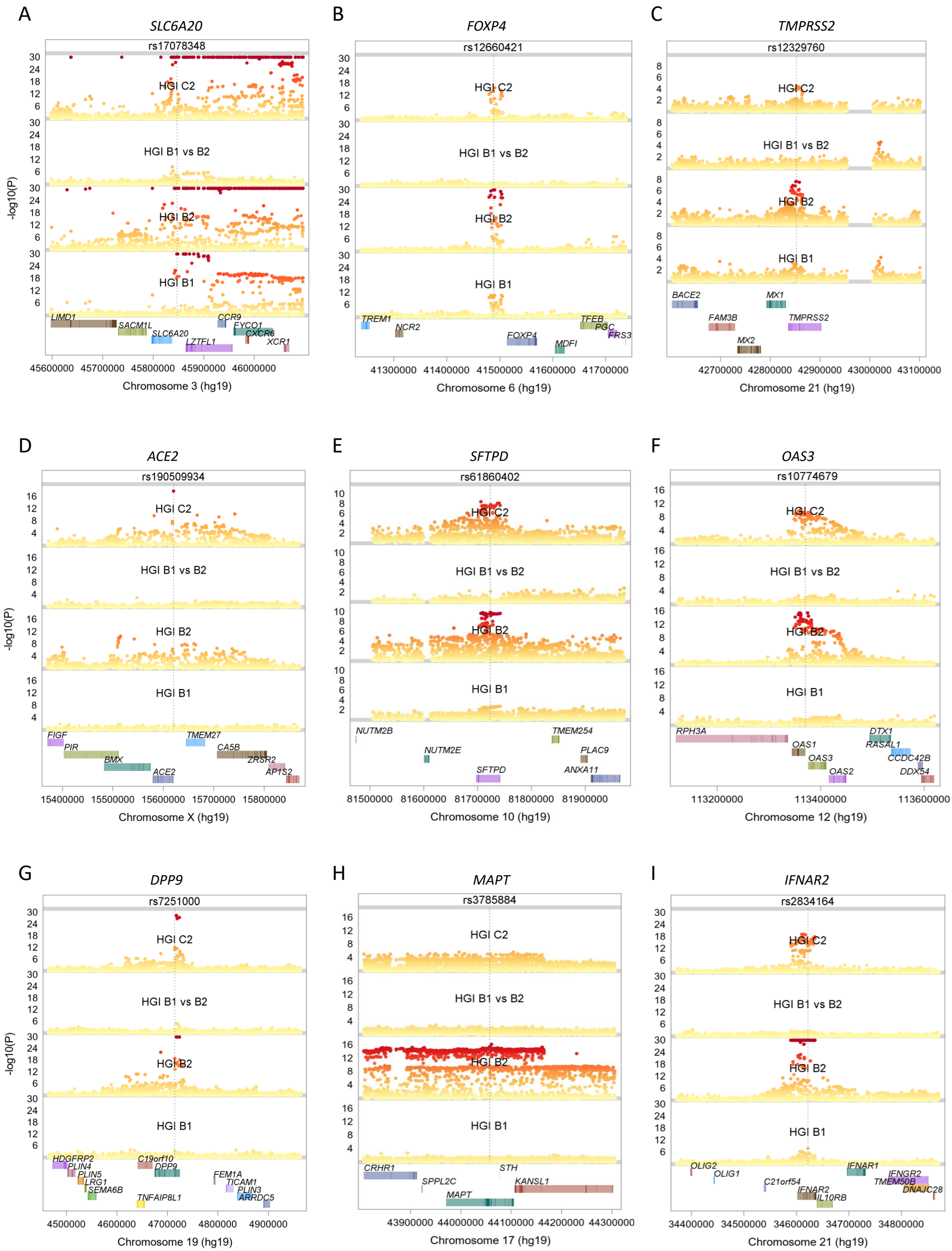

A
SLC6A20
rs17078348
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
LIMD1
SACM1L
SLC6A20
LZTFL1
CCR9
FYCO1
CXCR6
XCR1
45600000 45700000 45800000 45900000 46000000
Chromosome 3 (hg19)
B
FOXP4
rs12660421
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
TREM1
NCR2
FOXP4
MDFI
TFEB
PGC
FRS3
41300000 41400000 41500000 41600000 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 3

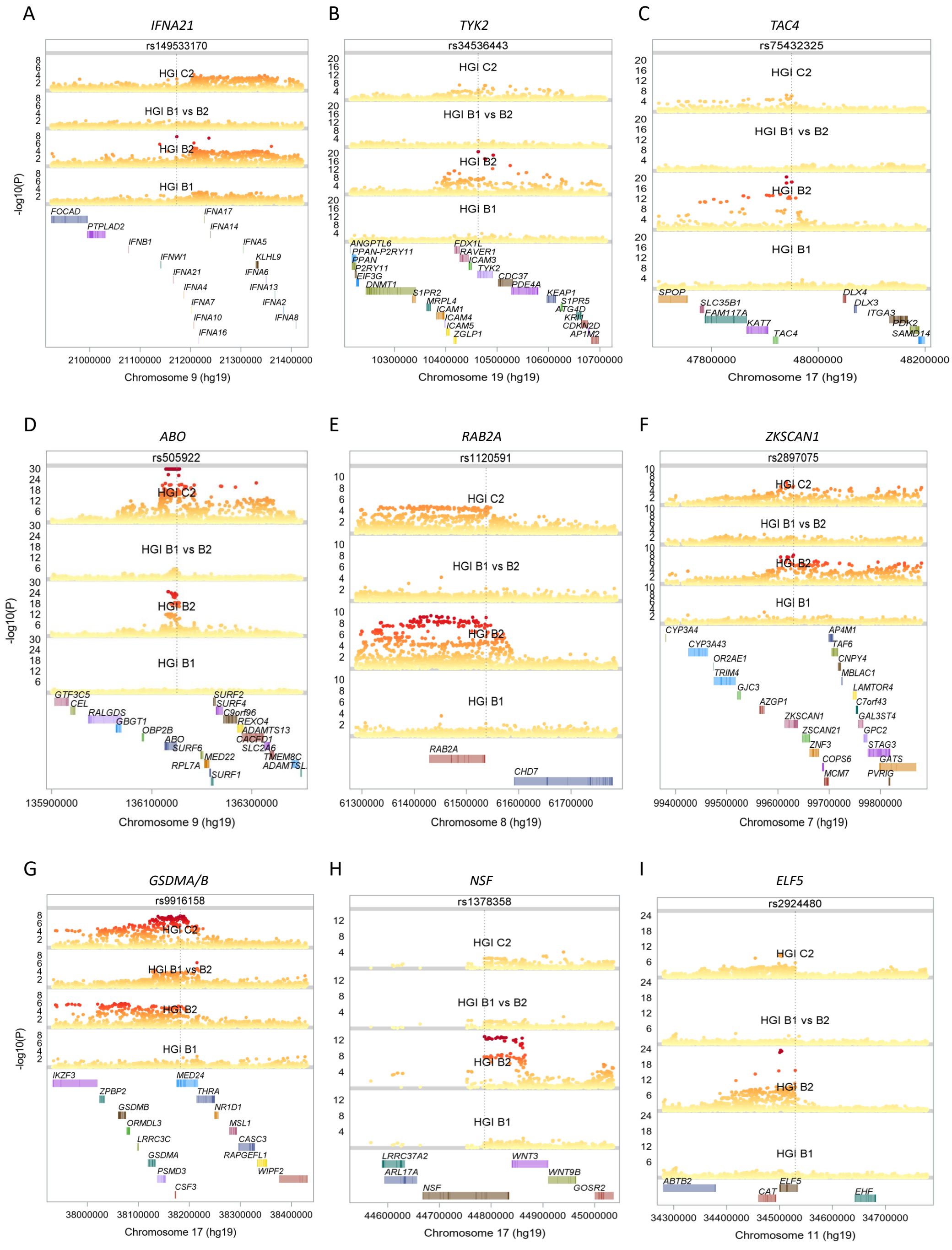

A
IFNA21
rs149533170
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
FOCAD
PTPLAD2
IFNB1
IFNW1
IFNA21
IFNA4
IFNA7
IFNA10
IFNA16
IFNA17
IFNA14
IFNA5
KLHL9
IFNA6
IFNA13
IFNA2
IFNA8
21000000 21100000 21200000 21300000 21400000
Chromosome 9 (hg19)
B
TYK2
rs34536443
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
ANGPTL6
PPAN-P2RY11
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 4

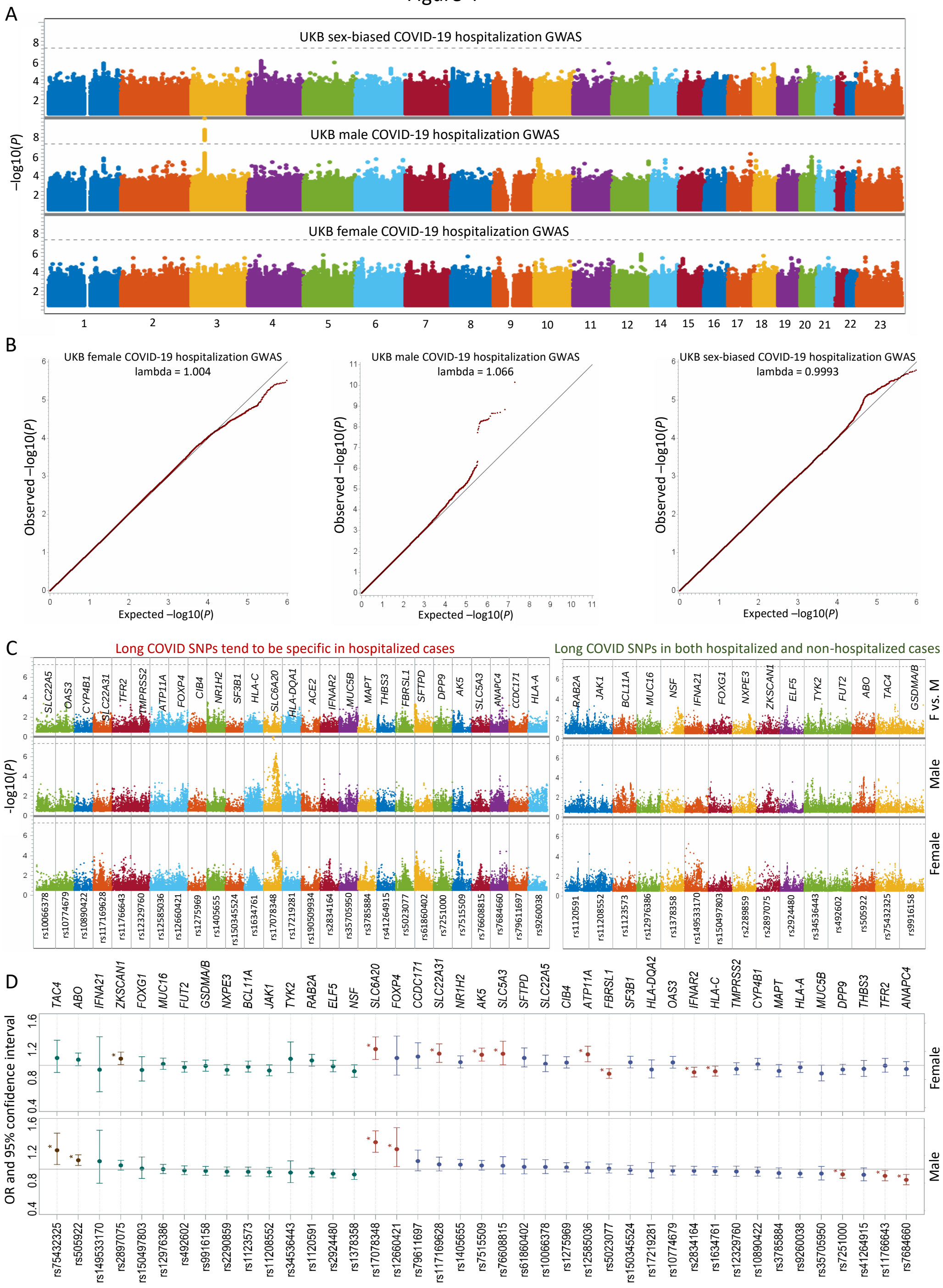

A
UKB sex-biased COVID-19 hospitalization GWAS
UKB male COVID-19 hospitalization GWAS
UKB female COVID-19 hospitalization GWAS
−log10(P)
B
UKB female COVID-19 hospitalization GWAS
lambda = 1.004
UKB male COVID-19 hospitalization GWAS
lambda = 1.066
UKB sex-biased COVID-19 hospitalization GWAS
lambda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

Figure 5

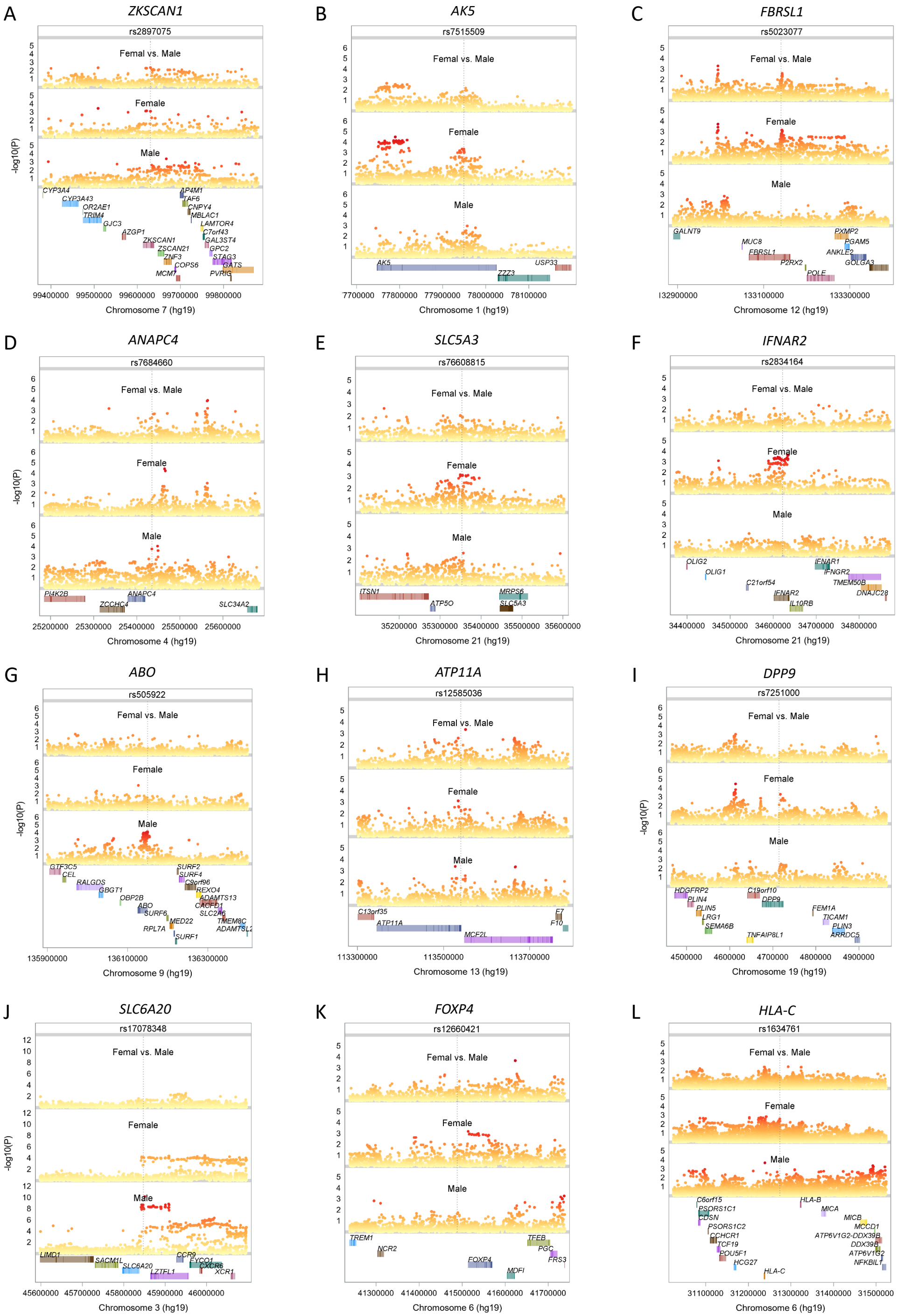
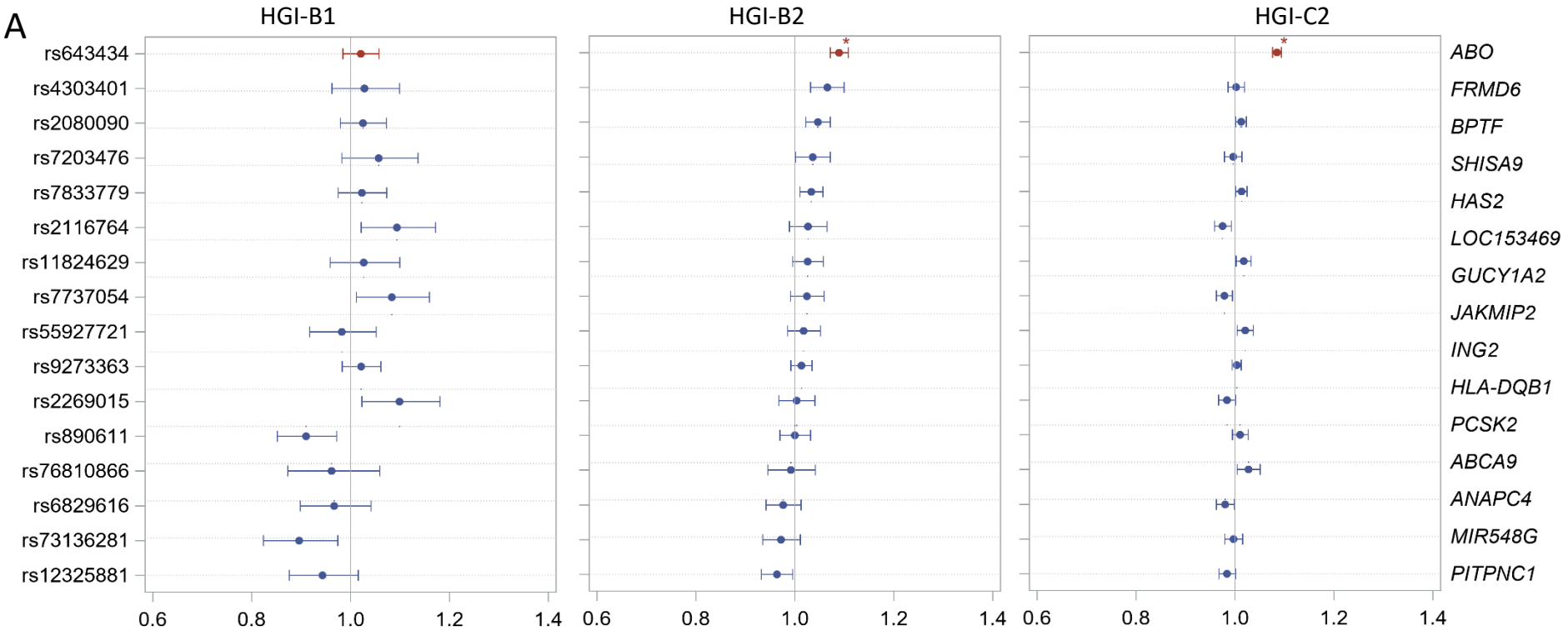
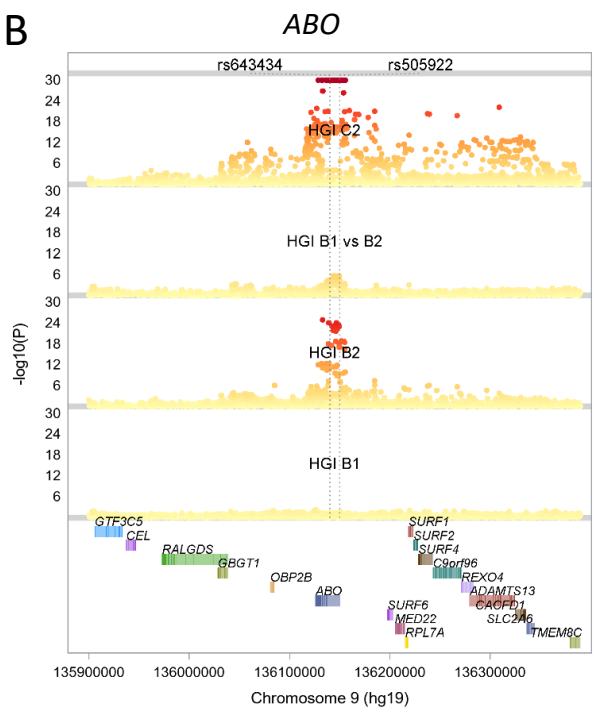
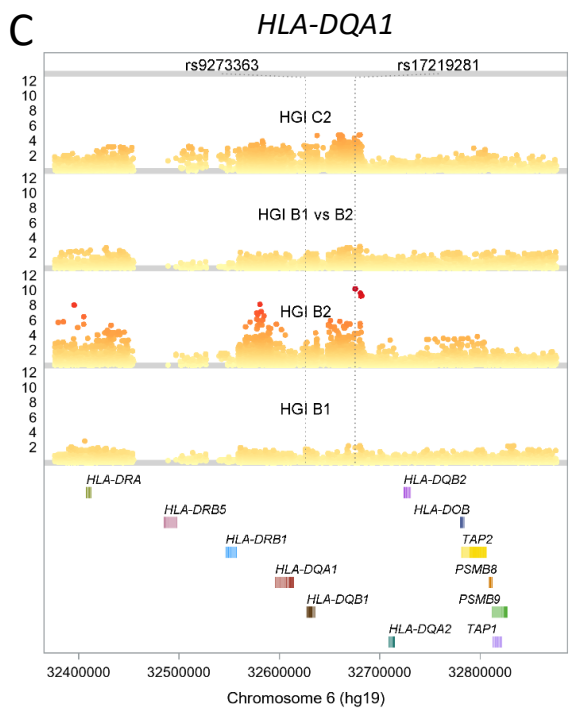
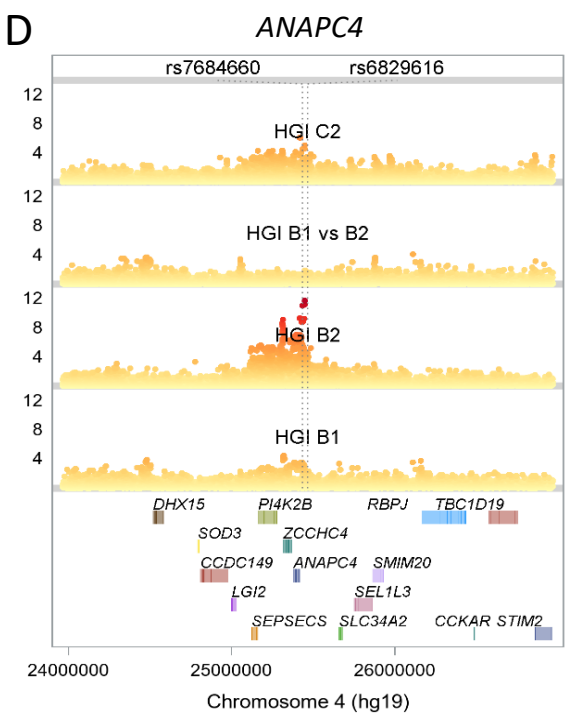

A
ZKSCAN1
rs2897075
Femal vs. Male
Female
Male
-log10(P)
CYP3A4
CYP3A43
OR2AE1
TRIM4
GJC3
AZGP1
ZKSCAN1
ZSCAN21
ZNF3
COPS6
MCM7
AP4M1
TAF6
CNPY4
MBLAC1
LAMTOR4
C7orf43
GAL3ST4
GPC2
STAG3
GATS
PVRIG
99400000 99500000 99600000 99700000 99800000
Chromosome 7 (hg19)
B
AK5
rs7515509
Femal vs. Male
Female
Male
AK5
ZZZ3
USP33
77700000 77800000 77900000 78000000 78100000
Chromosome 1 (hg19)
C
FBRSL1
rs5023077
Femal vs. Male
Female
Male
GALNT9
MUC8
FBRSL1
P2RX2
POLE
PXMP2
PGAM5
ANKLE2
GOLGA3
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
OBP2B
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
FYCO1
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-B
MICA
MICB
MCCD1
ATP6V1G2-DDX39B
DDX39B
ATP6V1G2
NFKBIL1
HLA-C
31100000 31200000 31300000 31400000 31500000
Chromosome 6 (hg19)

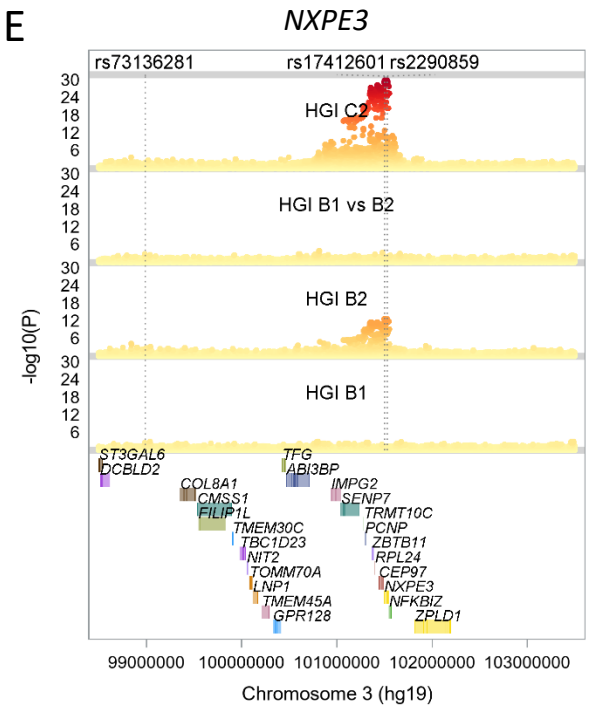
Figure 6

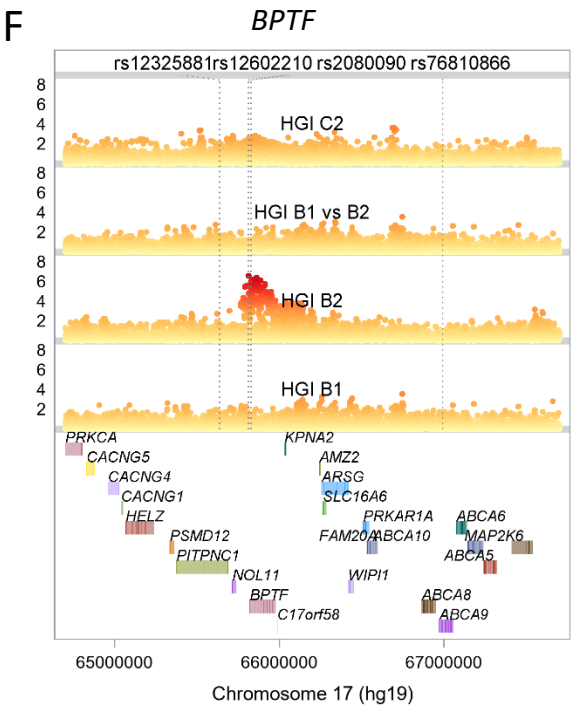

A
HGI-B1
HGI-B2
HGI-C2
rs643434
rs4303401
rs2080090
rs7203476
rs7833779
rs2116764
rs11824629
rs7737054
rs55927721
rs9273363
rs2269015
rs890611
rs76810866
rs6829616
rs73136281
rs12325881
ABO
FRMD6
BPTF
SHISA9
HAS2
LOC153469
GUCY1A2
JAKMIP2
ING2
HLA-DQB1
PCSK2
ABCA9
ANAPC4
MIR548G
PITPNC1
0.6
0.8
1.0
1.2
1.4
OR and 95% confidence interval


B
ABO
rs643434
rs505922
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
GTF3C5
CEL
RALGDS
GBGT1
OBP2B
ABO
SURF1
SURF2
SURF4
C9orf96
REXO4
ADAMTS13
CACFD1
SLC2A6
SURF6
MED22
RPL7A
TMEM8C
135900000
136000000
136100000
136200000
136300000
Chromosome 9 (hg19)


C
HLA-DQA1
rs9273363
rs17219281
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
HLA-DRA
HLA-DRB5
HLA-DRB1
HLA-DQA1
HLA-DQB1
HLA-DQB2
HLA-DOB
TAP2
PSMB8
PSMB9
HLA-DQA2
TAP1
32400000
32500000
32600000
32700000
32800000
Chromosome 6 (hg19)


D
ANAPC4
rs7684660
rs6829616
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
DHX15
PI4K2B
RBPJ
TBC1D19
SOD3
ZCCHC4
CCDC149
ANAPC4
SMIM20
LGI2
SEL1L3
SEPSECS
SLC34A2
CCKAR
STIM2
24000000
25000000
26000000
Chromosome 4 (hg19)


E
NXPE3
rs73136281
rs17412601
rs2290859
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
ST3GAL6
DCBLD2
TFG
ABI3BP
COL8A1
CMSS1
FILIP1L
IMPG2
SENP7
TRMT10C
PCNP
ZBTB11
RPL24
CEP97
NXPE3
NFKBIZ
ZPLD1
TMEM30C
TBC1D23
NIT2
TOMM70A
LNP1
TMEM45A
GPR128
99000000
100000000
101000000
102000000
103000000
Chromosome 3 (hg19)


F
BPTF
rs12325881
rs12602210
rs2080090
rs76810866
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
PRKCA
CACNG5
CACNG4
CACNG1
HELZ
PSMD12
PITPNC1
NOL11
BPTF
C17orf58
KPNA2
AMZ2
ARSG
SLC16A6
PRKAR1A
FAM20A
ABCA10
WIPI1
ABCA6
MAP2K6
ABCA5
ABCA8
ABCA9
65000000
66000000
67000000
Chromosome 17 (hg19)

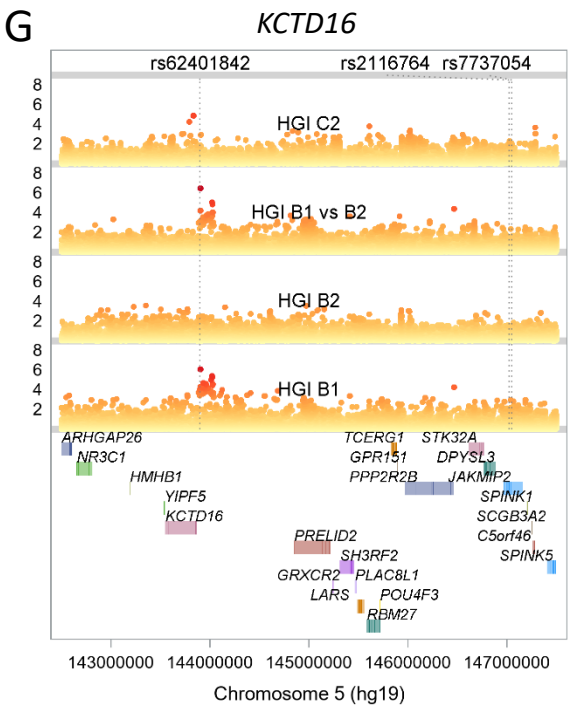

G
KCTD16
rs62401842
rs2116764
rs7737054
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
ARHGAP26
NR3C1
HMHB1
YIPF5
KCTD16
TCERG1
GPR151
PPP2R2B
STK32A
DPYSL3
JAKMIP2
SPINK1
SCGB3A2
C5orf46
SPINK5
PRELID2
SH3RF2
GRXCR2
PLAC8L1
LARS
POU4F3
RBM27
143000000
144000000
145000000
146000000
147000000
Chromosome 5 (hg19)

Figure 7

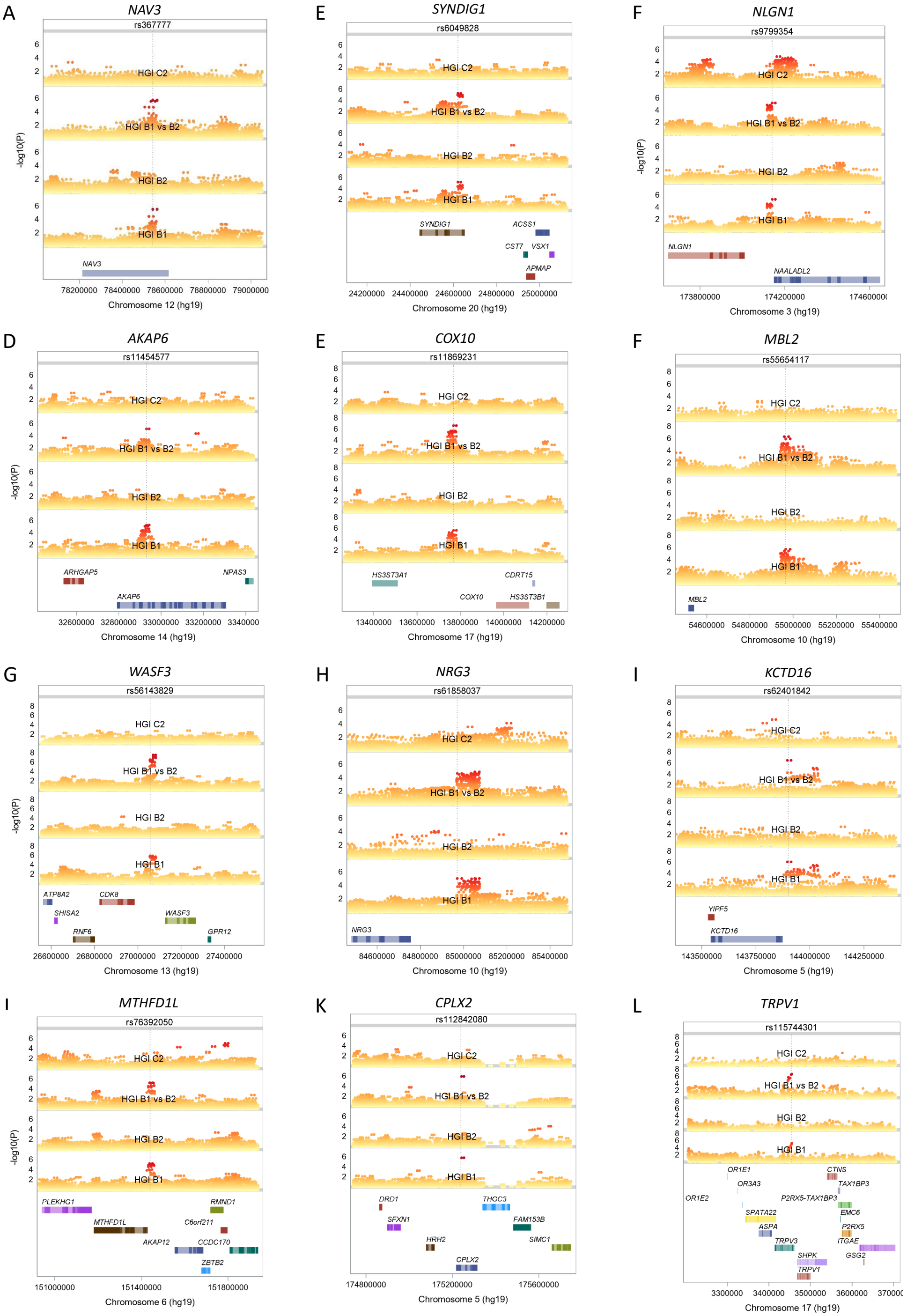

A
NAV3
rs367777
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
-log10(P)
NAV3
78200000 78400000 78600000 78800000 79000000
Chromosome 12 (hg19)
E
SYNDIG1
rs6049828
HGI C2
HGI B1 vs B2
HGI B2
HGI B1
SYNDIG1
ACSS1
CST7
VSX1
APMAP
24200000 24400000 24600000 24800000 25000000
Chromosome 20 (hg19)
F
NLGN1
rs9799354
HGI 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)
I
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
174800000 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 8

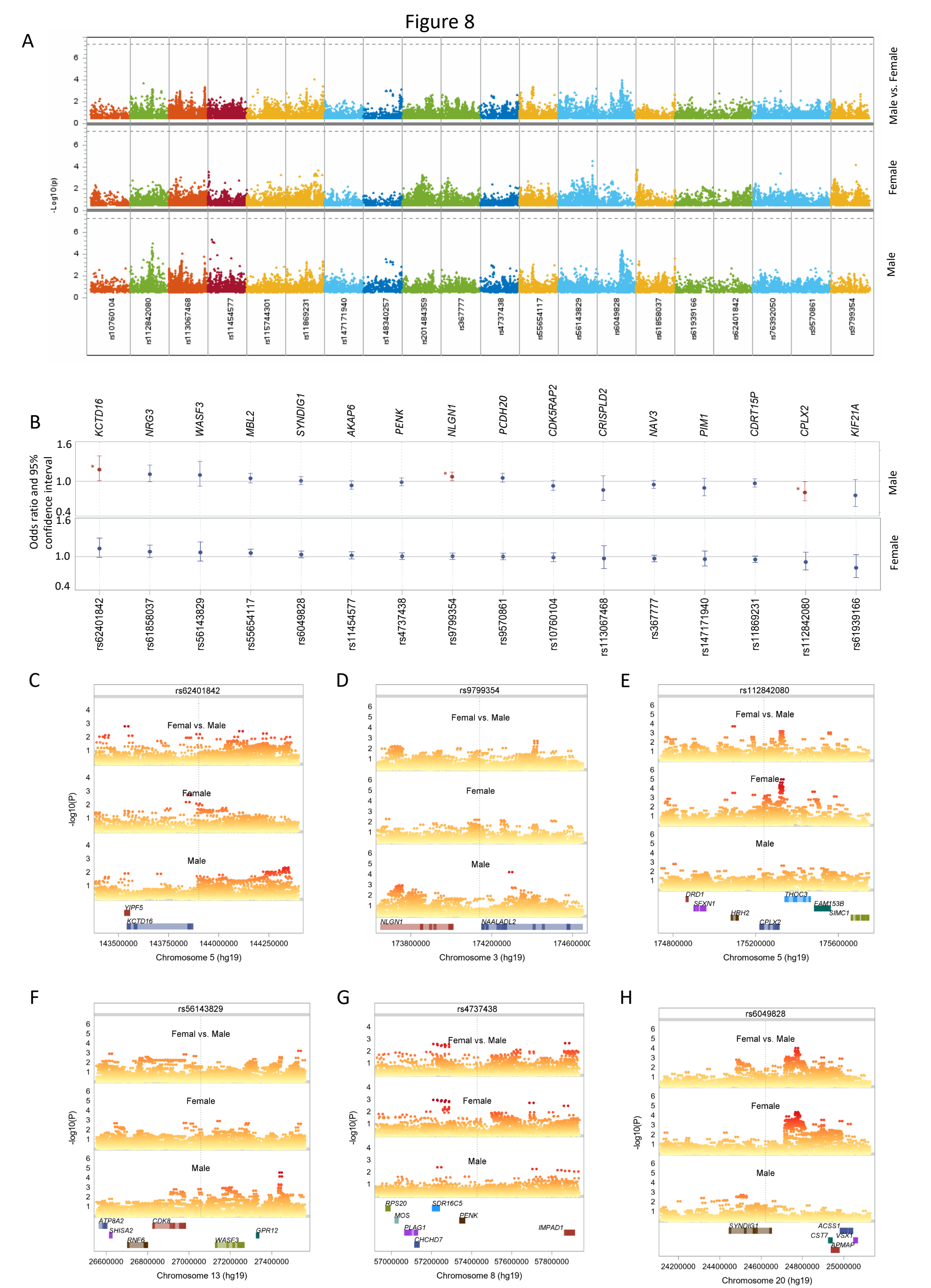

A
-Log10(p)
Male vs. Female
Female
Male
rs10760104
rs112842080
rs113067468
rs11454577
rs115744301
rs11869231
rs147171940
rs148340257
rs201484359
rs367777
rs4737438
rs55654117
rs56143829
rs6049828
rs61858037
rs61939166
rs62401842
rs76392050
rs9570861
rs9799354
B
KCTD16
NRG3
WASF3
MBL2
SYNDIG1
AKAP6
PENK
NLGN1
PCDH20
CDK5RAP2
CRISPLD2
NAV3
PIM1
CDRT15P
CPLX2
KIF21A
Odds ratio and 95% confidence interval
Male
Female
C
rs62401842
Femal vs. Male
Female
Male
-log10(P)
YIPF5
KCTD16
143500000
143750000
144000000
144250000
Chromosome 5 (hg19)
D
rs9799354
Femal vs. Male
Female
Male
NLGN1
NAALADL2
173800000
174200000
Chromosome 3 (hg19)
E
rs112842080
Femal vs. Male
Female
Male
DRD1
SFXN1
HRH2
THOC3
FAM153B
SIMC1
CPLX2
174800000
175200000
175600000
Chromosome 5 (hg19)
F
rs56143829
Femal vs. Male
Female
Male
ATP8A2
SHISA2
RNF6
CDK8
WASF3
GPR12
26600000
26800000
27000000
27200000
27400000
Chromosome 13 (hg19)
G
rs4737438
Femal vs. Male
Female
Male
RPS20
MOS
SDR16C5
PENK
PLAG1
CHCHD7
IMPAD1
57000000
57200000
57400000
57600000
57800000
Chromosome 8 (hg19)
H
rs6049828
Femal vs. Male
Female
Male
SYNDIG1
ACSS1
CST7
VSX1
APMAP
24200000
24400000
24600000
24800000
25000000
Chromosome 20 (hg19)

Figure 9

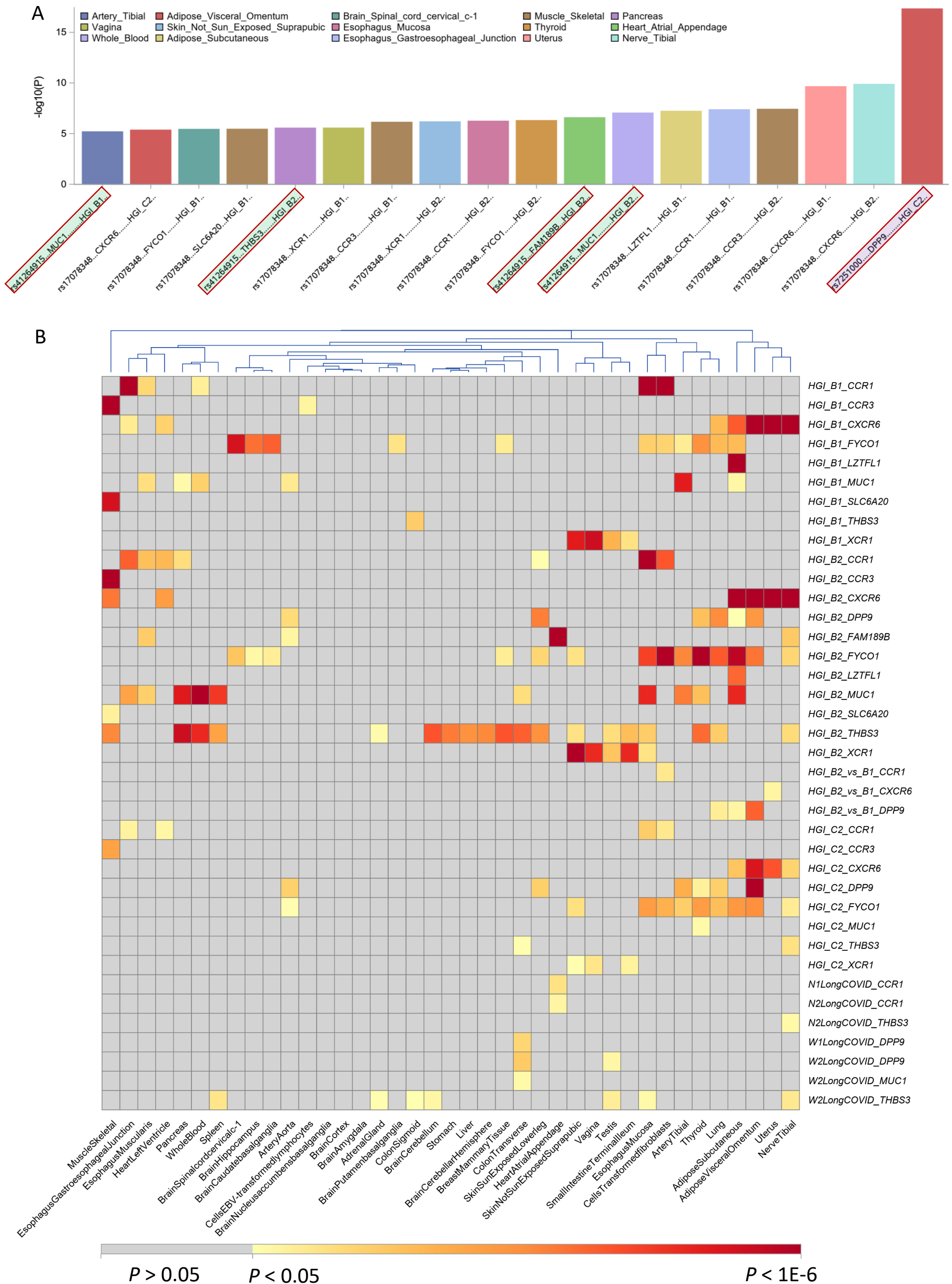

A
Artery_Tibial
Vagina
Whole_Blood
Adipose_Visceral_Omentum
Skin_Not_Sun_Exposed_Suprapubic
Adipose_Subcutaneous
Brain_Spinal_cord_cervical_c-1
Esophagus_Mucosa
Esophagus_Gastroesophageal_Junction
Muscle_Skeletal
Thyroid
Uterus
Pancreas
Heart_Atrial_Appendage
Nerve_Tibial
-log10(P)
0
5
10
15
rs41264915...MUC1.....HGI_B1..
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 < 1E-6

Figure 10

-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

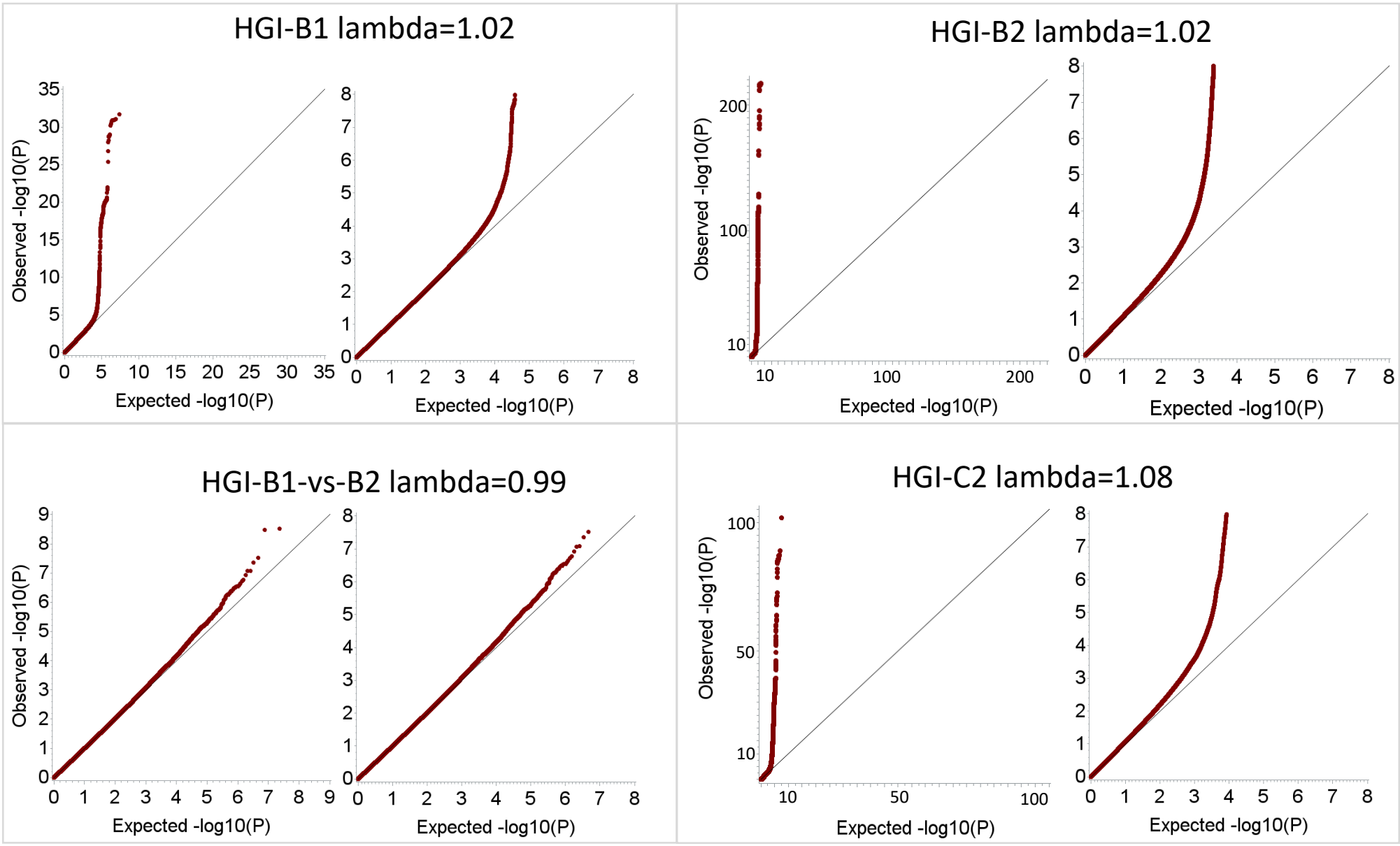


Figure S2

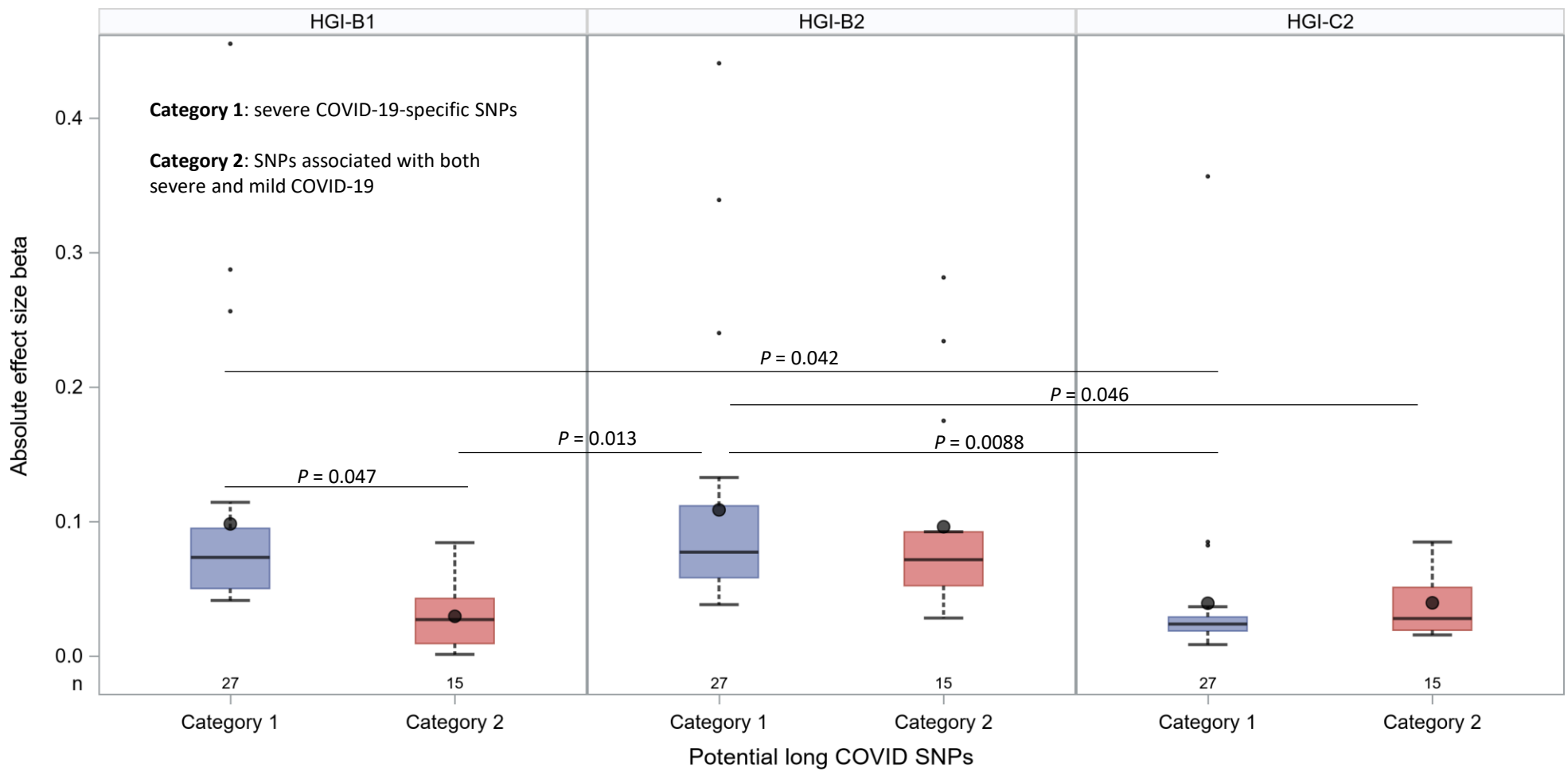


Figure S3

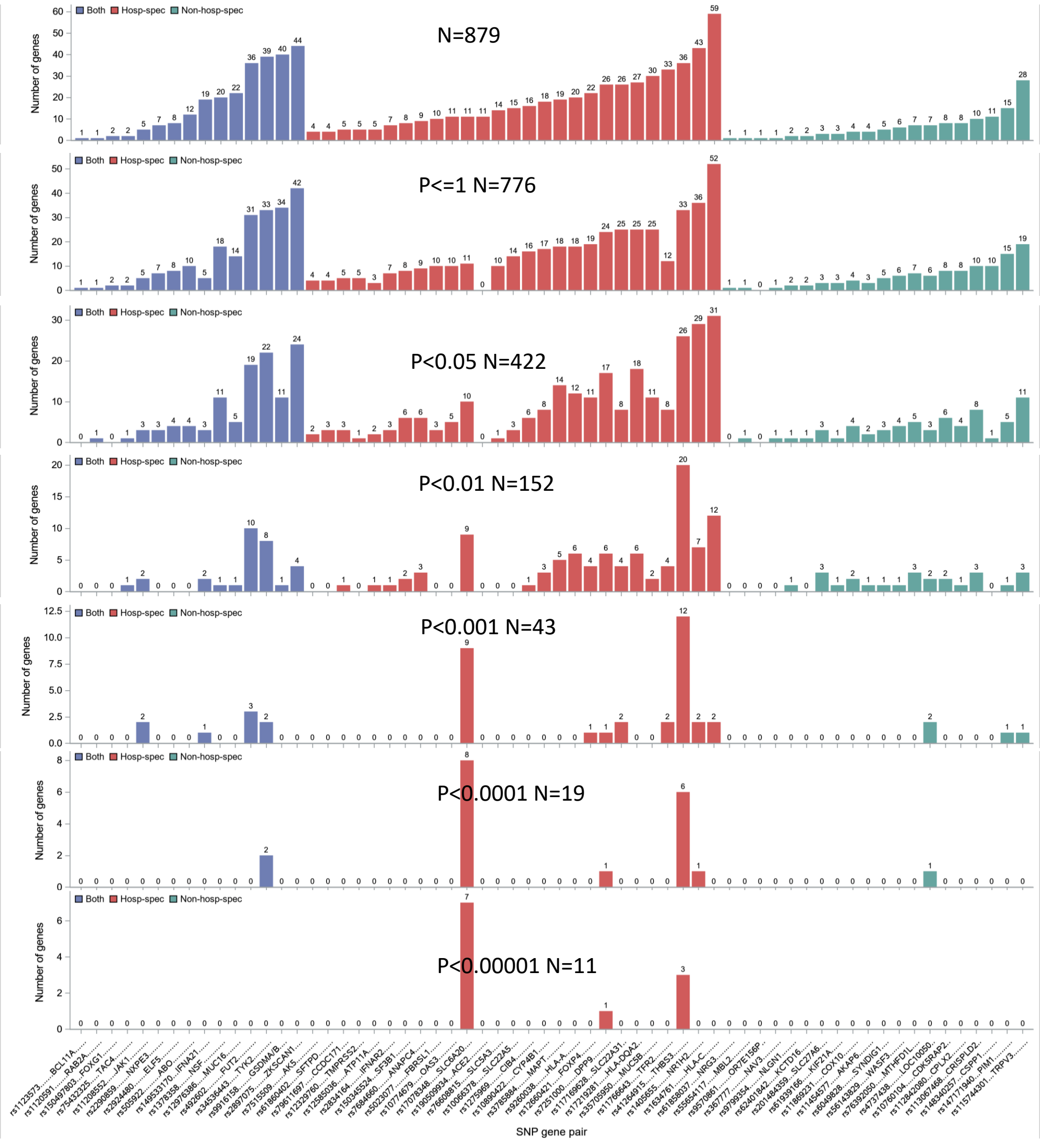

Both
Hosp-spec
Non-hosp-spec
Number of genes
N=879
P<=1 N=776
P<0.05 N=422
P<0.01 N=152
P<0.001 N=43
P<0.0001 N=19
P<0.00001 N=11
SNP gene pair

Figure S4

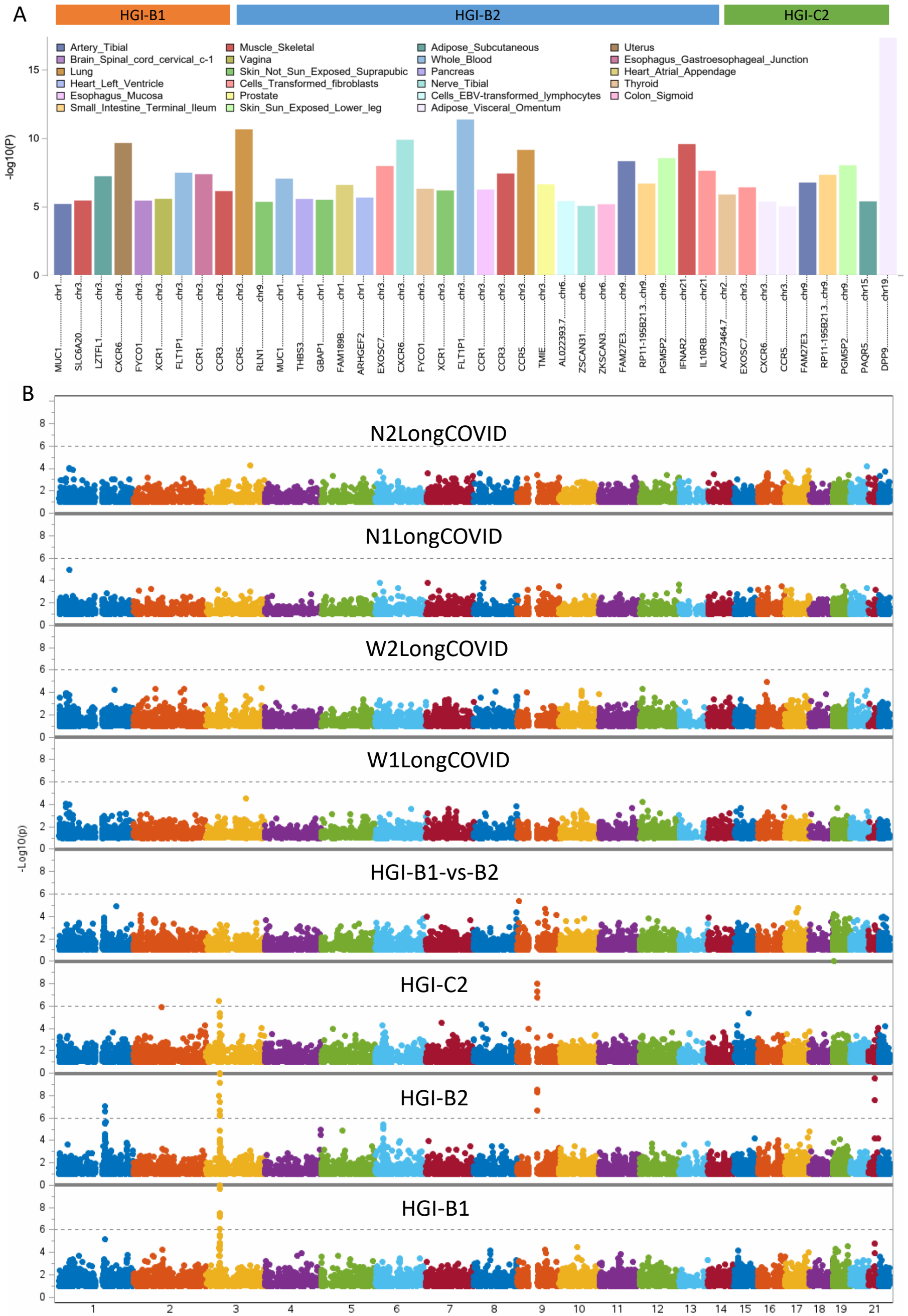

A
HGI-B1
HGI-B2
HGI-C2
Artery_Tibial
Brain_Spinal_cord_cervical_c-1
Lung
Heart_Left_Ventricle
Esophagus_Mucosa
Small_Intestine_Terminal_Ileum
Muscle_Skeletal
Vagina
Skin_Not_Sun_Exposed_Suprapubic
Cells_Transformed_fibroblasts
Prostate
Skin_Sun_Exposed_Lower_leg
Adipose_Subcutaneous
Whole_Blood
Pancreas
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 S5

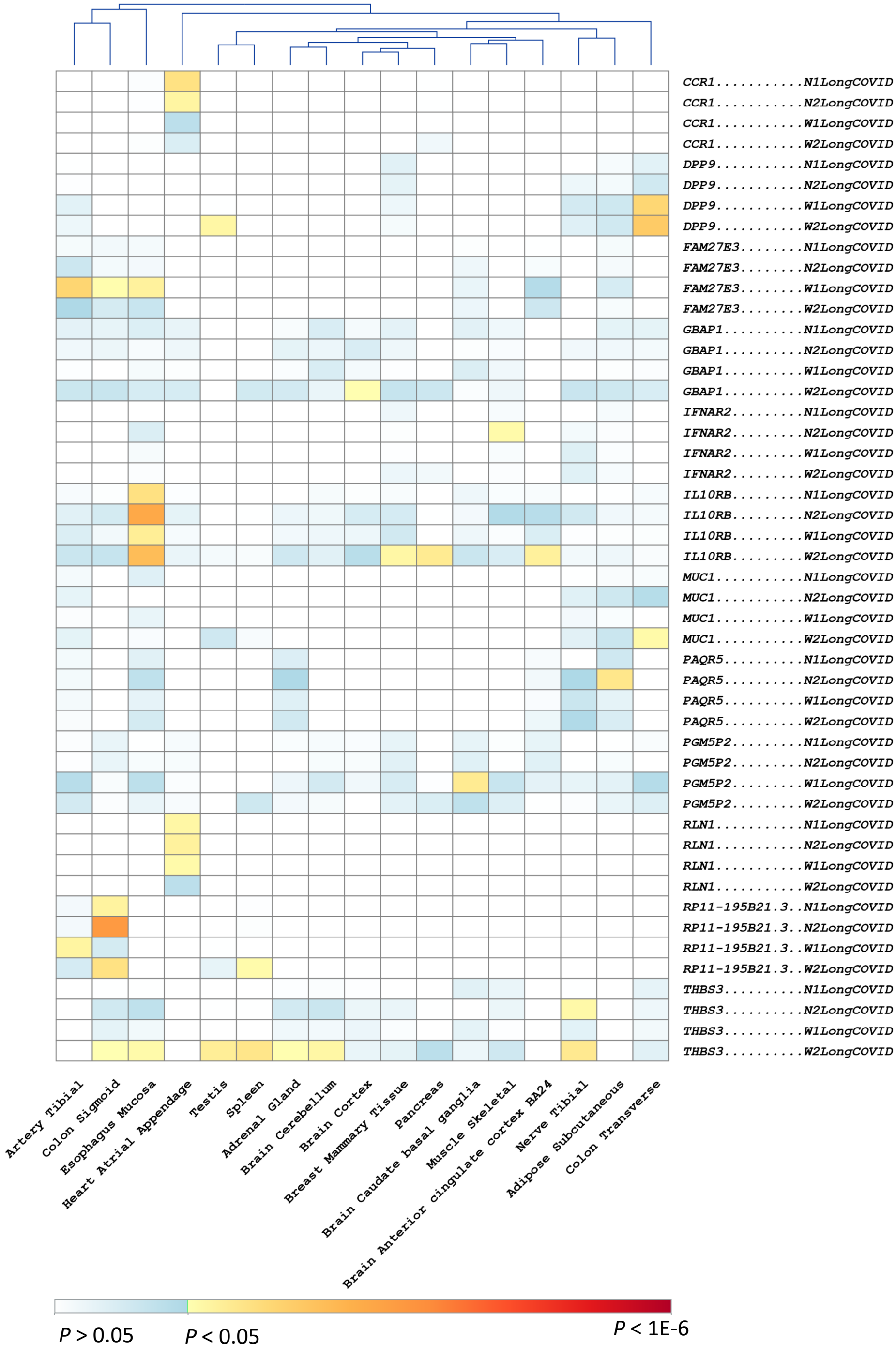
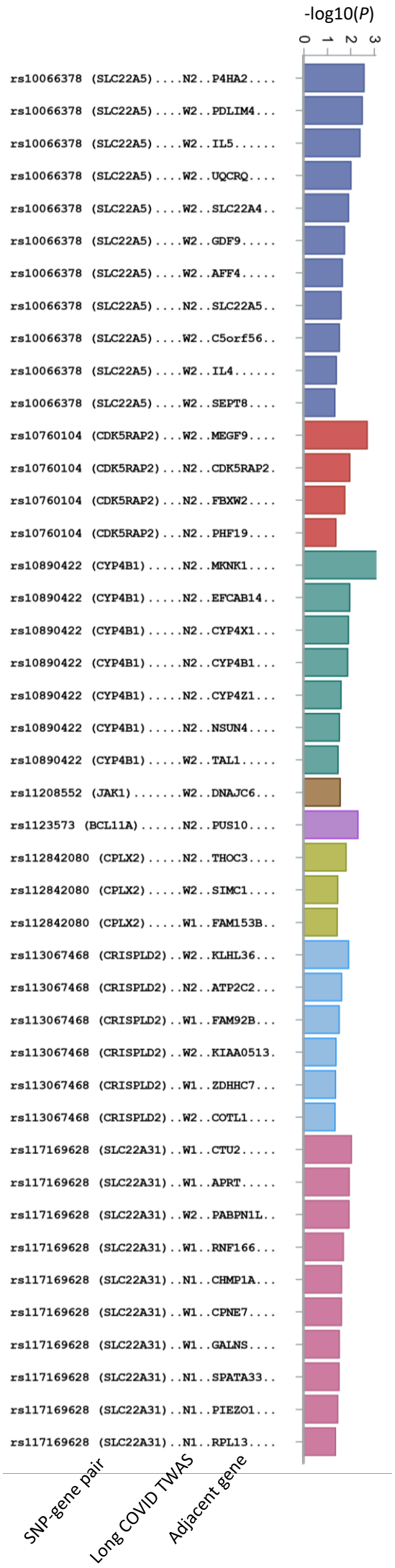
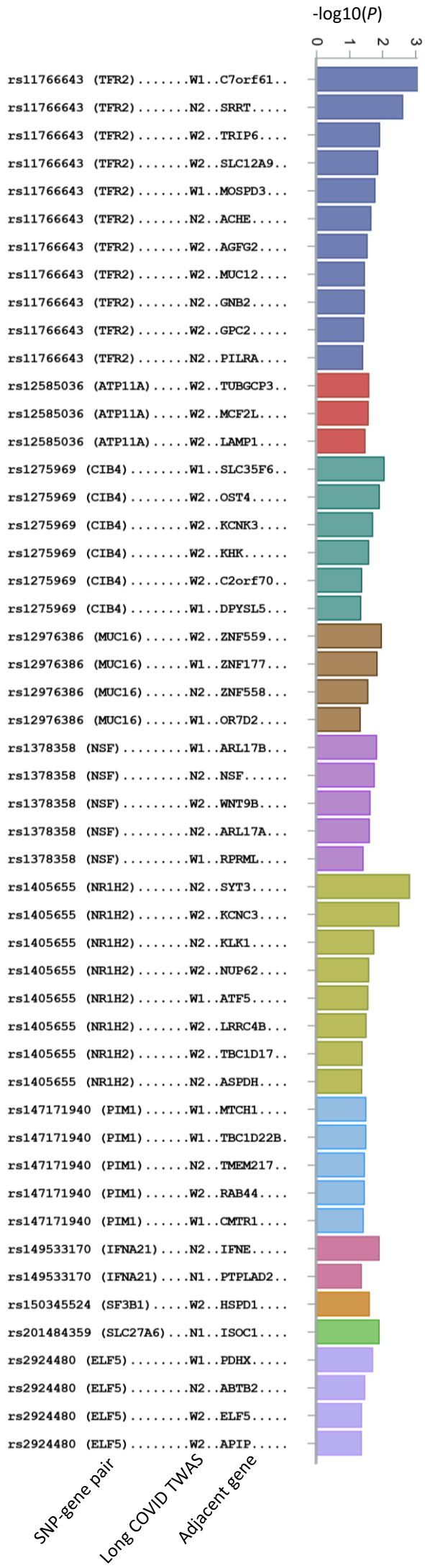

CCR1.....N1LongCOVID
CCR1.....N2LongCOVID
CCR1.....W1LongCOVID
CCR1.....W2LongCOVID
DPP9.....N1LongCOVID
DPP9.....N2LongCOVID
DPP9.....W1LongCOVID
DPP9.....W2LongCOVID
FAM27E3.....N1LongCOVID
FAM27E3.....N2LongCOVID
FAM27E3.....W1LongCOVID
FAM27E3.....W2LongCOVID
GBAP1.....N1LongCOVID
GBAP1.....N2LongCOVID
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 < 1E-6

Figure S6


-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

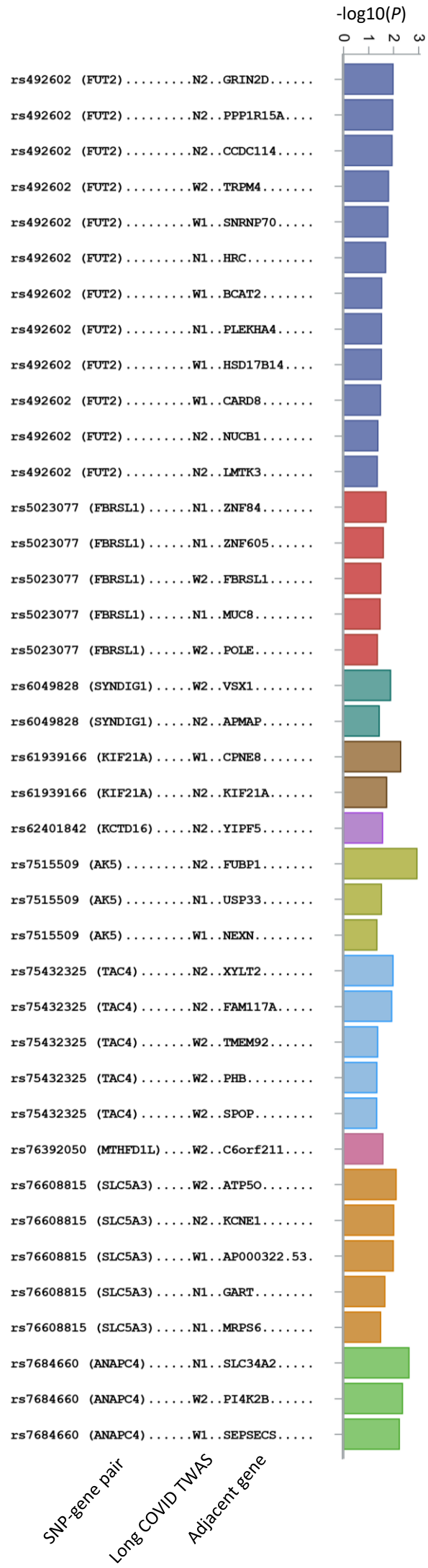
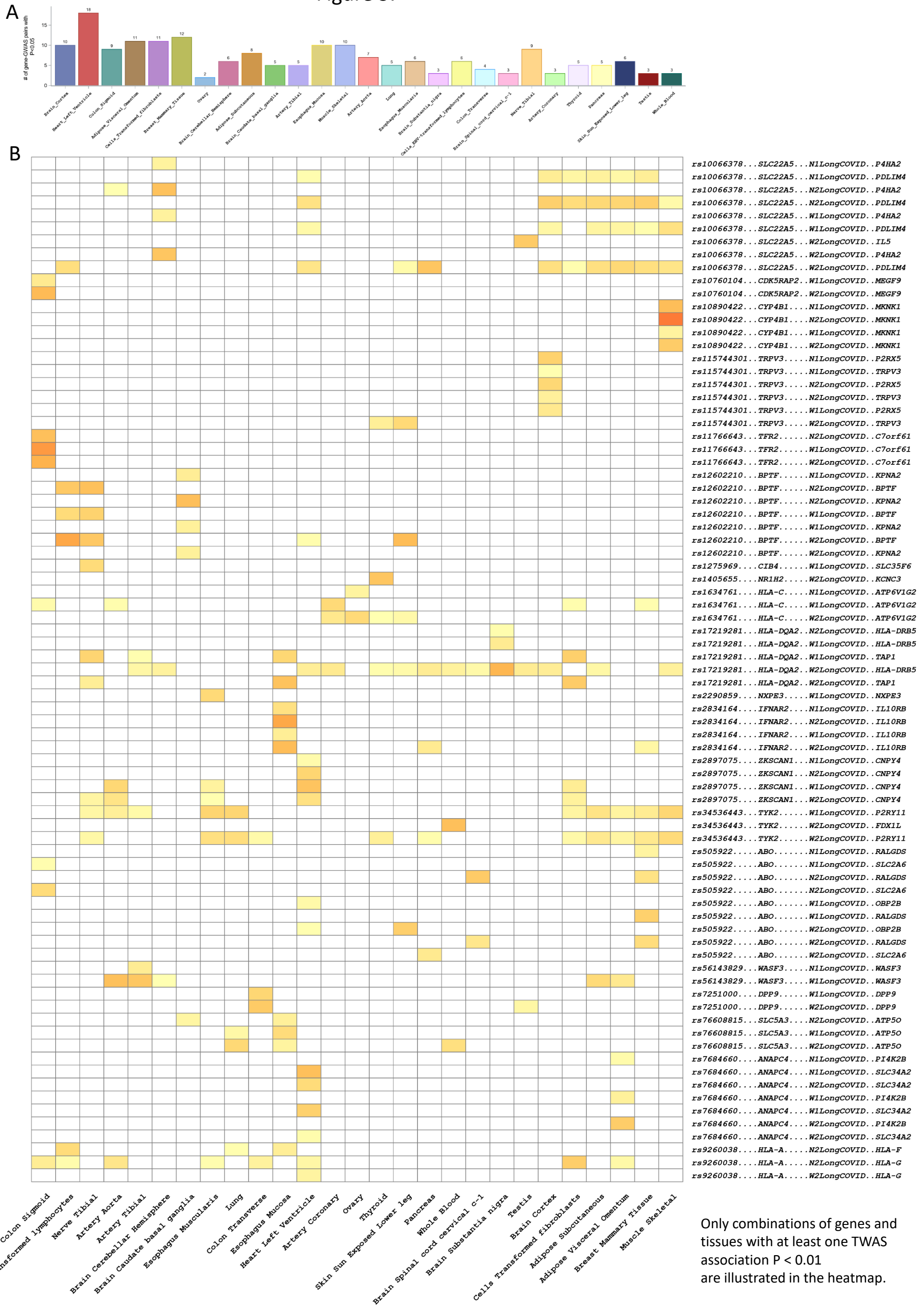

-log10(P)
0
1
2
3
rs492602 (FUT2).....N2..GRIN2D.....
rs492602 (FUT2).....N2..PPP1R15A....
rs492602 (FUT2).....N2..CCDC114.....
rs492602 (FUT2).....W2..TRPM4.....
rs492602 (FUT2).....W1..SNRNP70.....
rs492602 (FUT2).....N1..HRC.....
rs492602 (FUT2).....W1..BCAT2.....
rs492602 (FUT2).....N1..PLEKHA4.....
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 S7



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

Figure S8

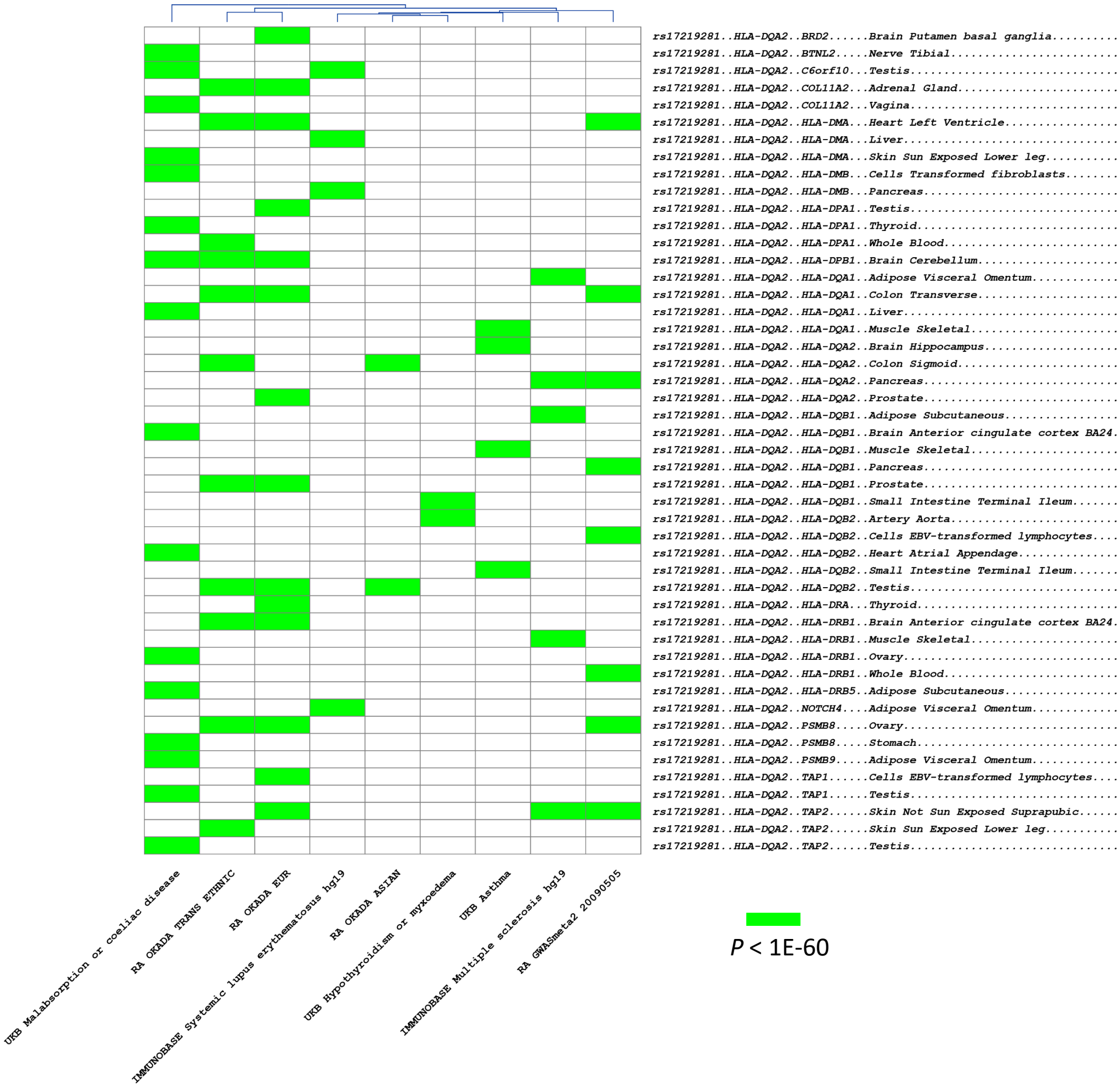

rs17219281..HLA-DQA2..BRD2.....Brain Putamen basal ganglia.....
rs17219281..HLA-DQA2..BTNL2.....Nerve Tibial.....
rs17219281..HLA-DQA2..C6orf10...Testis.....
rs17219281..HLA-DQA2..COL11A2...Adrenal Gland.....
rs17219281..HLA-DQA2..COL11A2...Vagina.....
rs17219281..HLA-DQA2..HLA-DMA...Heart Left Ventricle.....
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 < 1E-60

Figure S9

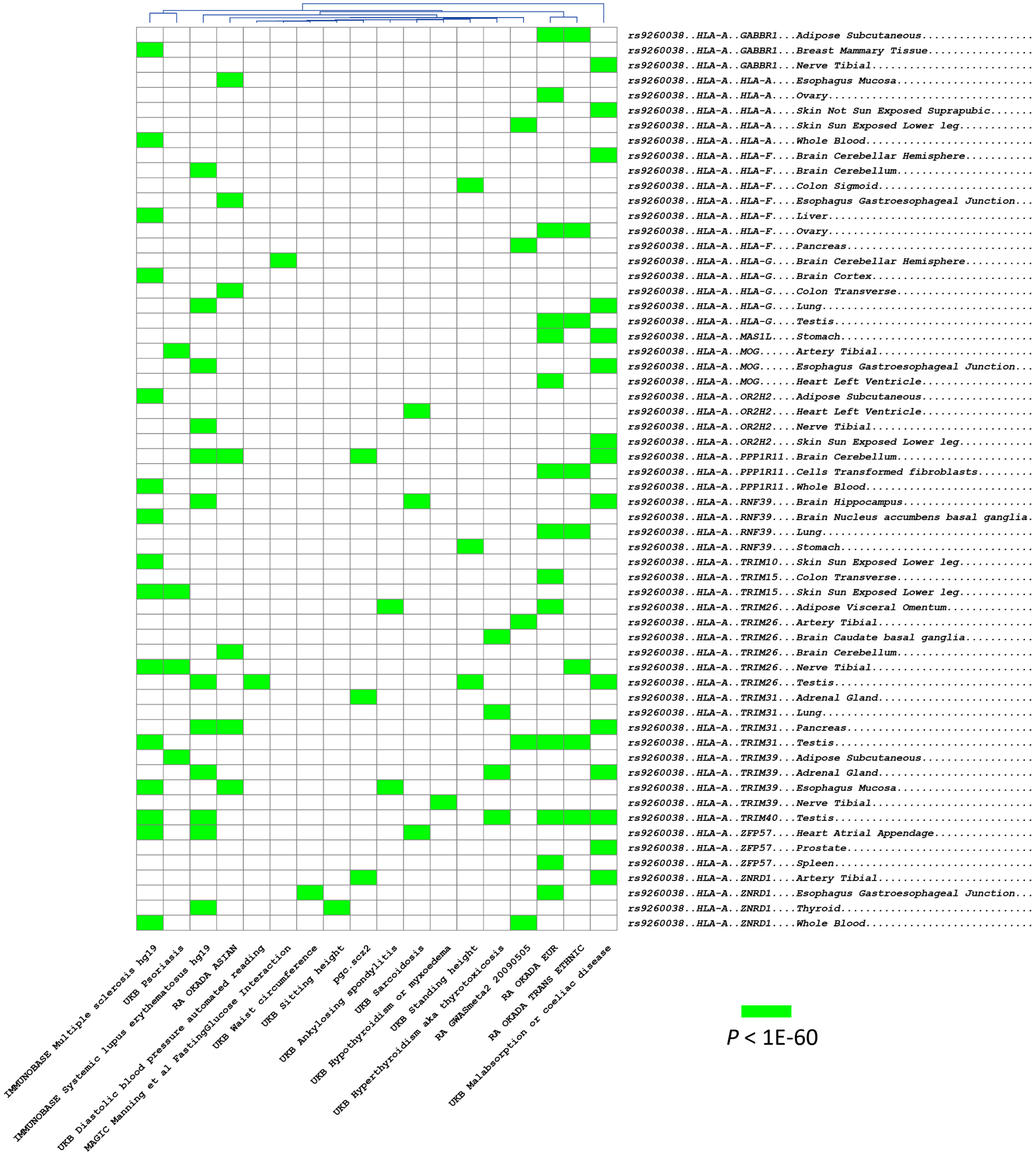

rs9260038..HLA-A..GABBR1...Adipose Subcutaneous.....
rs9260038..HLA-A..GABBR1...Breast Mammary Tissue.....
rs9260038..HLA-A..GABBR1...Nerve Tibial.....
rs9260038..HLA-A..HLA-A....Esophagus Mucosa.....
rs9260038..HLA-A..HLA-A....Ovary.....
rs9260038..HLA-A..HLA-A....Skin Not Sun Exposed Suprapubic.....
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 < 1E-60

Figure S10

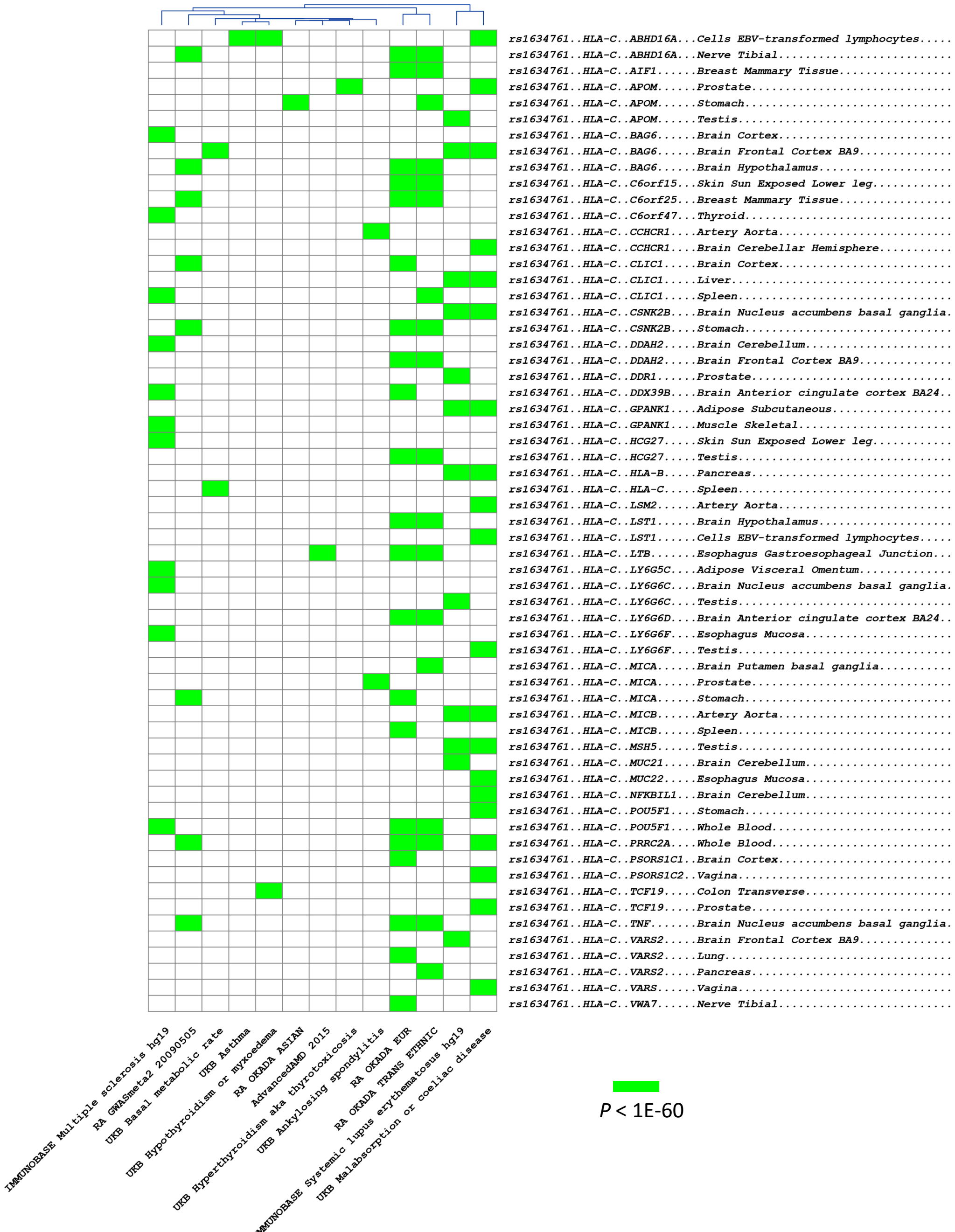

rs1634761..HLA-C..ABHD16A...Cells EBV-transformed lymphocytes.....
rs1634761..HLA-C..ABHD16A...Nerve Tibial.....
rs1634761..HLA-C..AIF1.....Breast Mammary Tissue.....
rs1634761..HLA-C..APOM.....Prostate.....
rs1634761..HLA-C..APOM.....Stomach.....
rs1634761..HLA-C..APOM.....Testis.....
rs1634761..HLA-C..BAG6.....Brain 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 < 1E-60

Figure S11

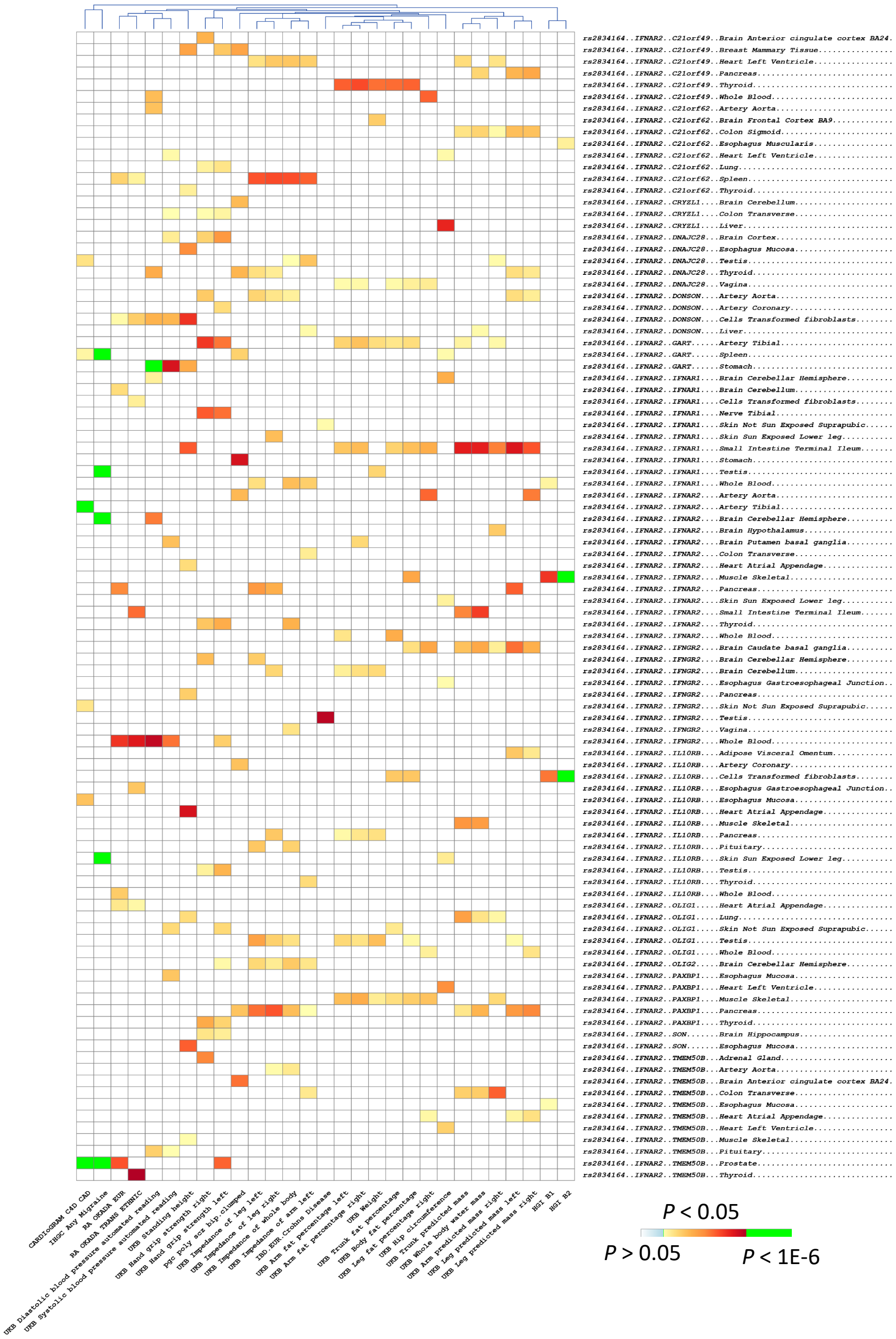

rs2834164..IFNAR2..C21orf49..Brain Anterior cingulate cortex BA24.
rs2834164..IFNAR2..C21orf49..Breast Mammary Tissue.....
rs2834164..IFNAR2..C21orf49..Heart Left Ventricle.....
rs2834164..IFNAR2..C21orf49..Pancreas.....
rs2834164..IFNAR2..C21orf49..Thyroid.....
rs2834164..IFNAR2..C21orf49..Whole Blood.....
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
IHGC Any Migraine
RA OKADA EUR
RA OKADA TRANS ETHNIC
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 bip 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
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
HGI B1
HGI B2
P < 0.05
P > 0.05
P < 1E-6

Figure S12

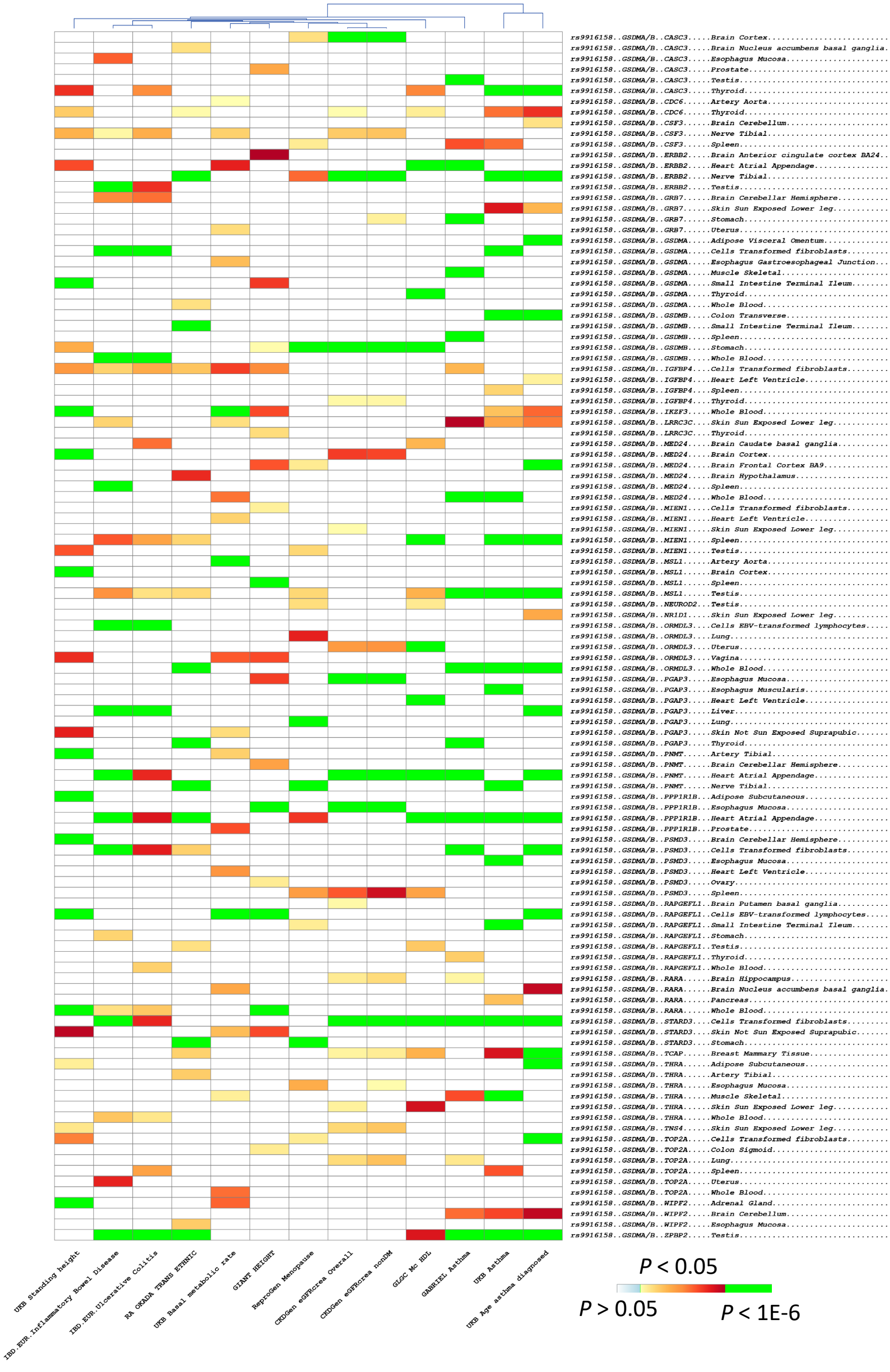

rs9916158..GSDMA/B..CASC3.....Brain Cortex
rs9916158..GSDMA/B..CASC3.....Brain Nucleus accumbens basal ganglia.
rs9916158..GSDMA/B..CASC3.....Esophagus Mucosa
rs9916158..GSDMA/B..CASC3.....Prostate
rs9916158..GSDMA/B..CASC3.....Testis
rs9916158..GSDMA/B..CASC3.....Thyroid
rs9916158..GSDMA/B..CDC6.....Artery 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
ReproGen 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 < 1E-6

Figure S13

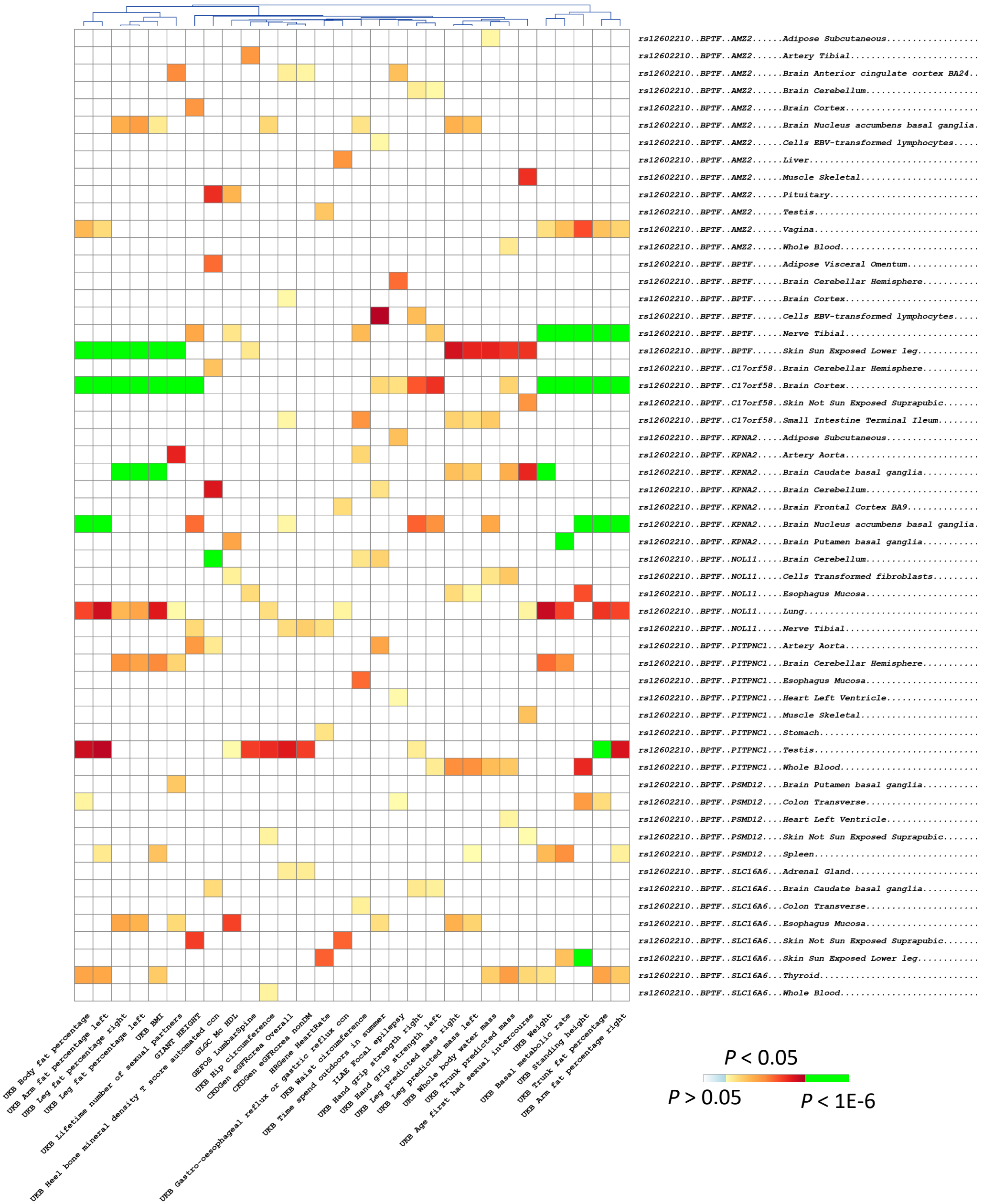

rs12602210..BPTF..AMZ2.....Adipose Subcutaneous.....
rs12602210..BPTF..AMZ2.....Artery Tibial.....
rs12602210..BPTF..AMZ2.....Brain Anterior cingulate cortex BA24..
rs12602210..BPTF..AMZ2.....Brain Cerebellum.....
rs12602210..BPTF..AMZ2.....Brain Cortex.....
rs12602210..BPTF..AMZ2.....Brain Nucleus accumbens 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
UKB Lifetime number of sexual partners
GIANT HEIGHT
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
IIAE 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 < 1E-6

Figure S14

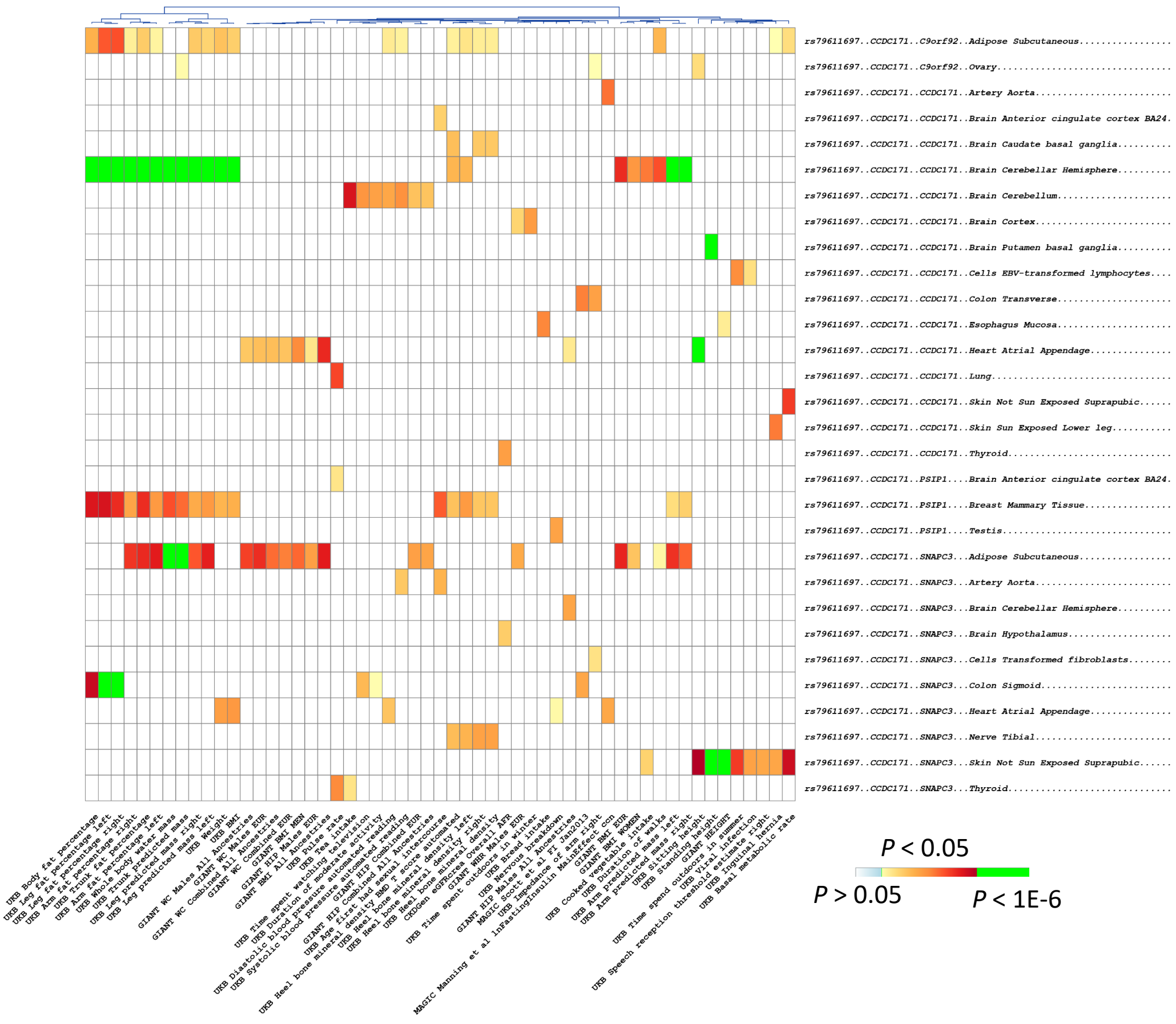
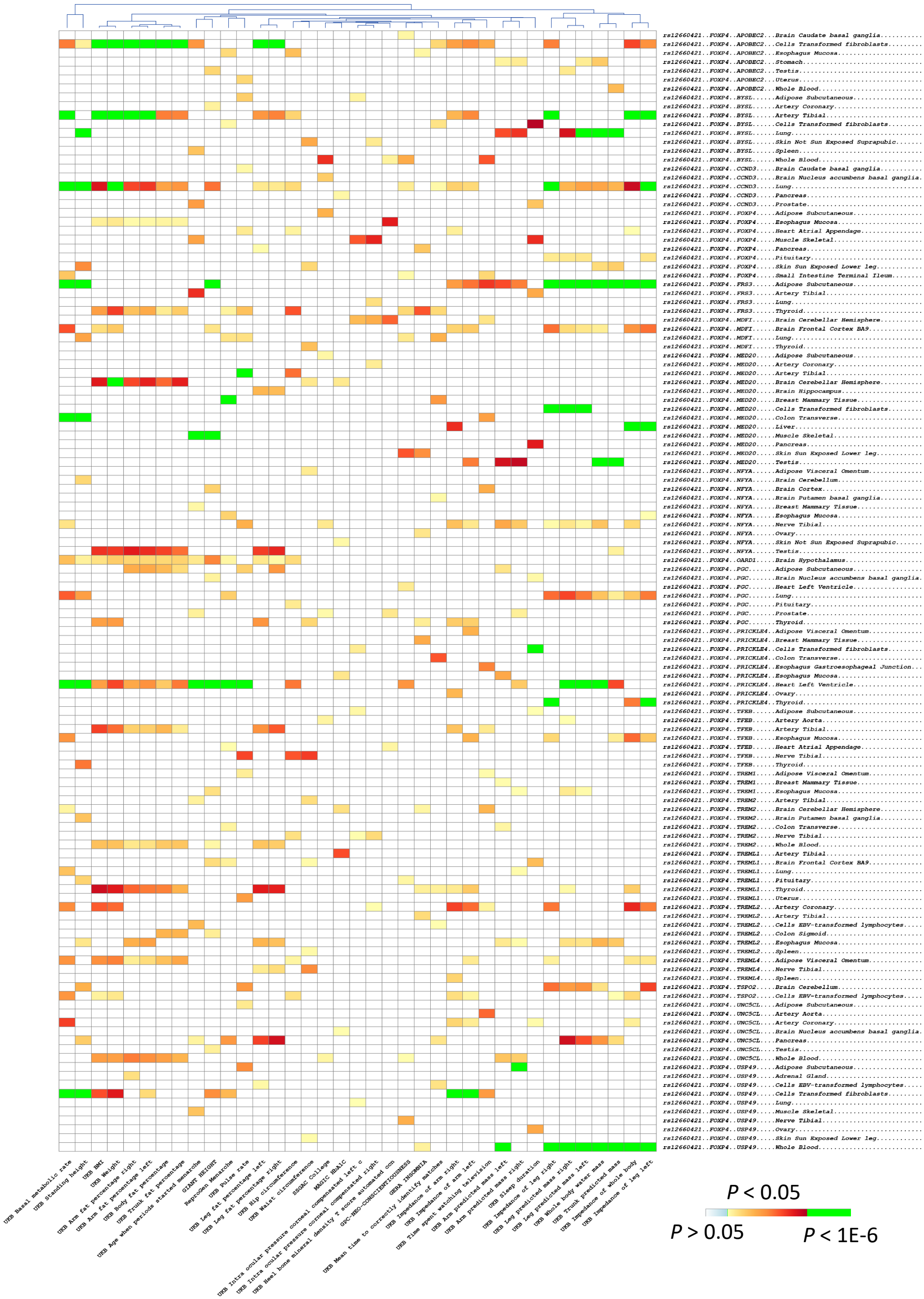

rs79611697..CCDC171..C9orf92..Adipose Subcutaneous
rs79611697..CCDC171..C9orf92..Ovary
rs79611697..CCDC171..CCDC171..Artery Aorta
rs79611697..CCDC171..CCDC171..Brain Anterior cingulate cortex BA24.
rs79611697..CCDC171..CCDC171..Brain Caudate basal ganglia
rs79611697..CCDC171..CCDC171..Brain Cerebellar 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 HIP Males EUR
GIANT BMI All Ancestries
UKB Pulse rate
UKB Tea intake
UKB Time spent watching television
GIANT HIP Combined EUR
UKB Duration of moderate activity
UKB Diastolic blood pressure automated reading
UKB Systolic blood pressure automated reading
GIANT HIP Combined All Ancestries
UKB Age first had sexual intercourse
UKB Heel bone mineral density BMD T score automated
UKB Heel bone mineral density left
UKB Heel bone mineral density right
UKB Heel bone mineral density
CKDGen eGFRcrea overall AFR
UKB Time spent outdoors in winter
GIANT WHR Males EUR
UKB Bread intake
UKB Nervous breakdown
MAGIC Scott et al FI Jan2013
UKB Impedance of arm right
MAGIC Manning et al lnFastingInsulin 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
UKB Inguinal hernia right
UKB Basal metabolic rate
P < 0.05
P > 0.05
P < 1E-6

Figure S15


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
ReproGen Menarche
UKB Pulse rate
UKB Leg fat percentage left
UKB Leg fat percentage right
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 Heel bone mineral density T score automated ccn
GPC-NEO-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 < 1E-6