Medical relevance of common protein-altering variants in GPCR genes across 337,205 individuals in the UK Biobank study

Christopher DeBoever[1+*], AJ Venkatakrishnan[2,3,4,5+*], Joseph M Paggi[2,3,4,5*], Franziska M. Heydenreich[3,6], Suli-Anne Laurin[6], Matthieu Masureel[3], Yosuke Tanigawa[1], Guhan Venkataraman[1], Michel Bouvier[6], Ron Dror[2,3,4,5+], Manuel A. Rivas[1+]

Affiliations

[1]Department of Biomedical Data Science, School of Medicine, Stanford University, Stanford, CA, 94305, USA.
[2]Department of Computer Science, Stanford University, Stanford, CA, 94305, USA.
[3]Department of Molecular and Cellular Physiology, School of Medicine, Stanford, CA 94305, USA.
[4]Department of Structural Biology, School of Medicine, Stanford University, Stanford, CA 94305, USA.
[5]Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA.
[6]Institute for Research in Immunology and Cancer, Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, QC, H3T 1J4, Canada.

Abstract

G protein-coupled receptors (GPCRs) drive an array of important physiological functions and are the targets of nearly one-third of all FDA approved drugs. Large scale genomic initiatives are mapping the genetic diversity in GPCRs, however, a map of which GPCR genetic variants are associated with phenotypic variation and disease is lacking. Furthermore, the mechanistic basis of how the individual GPCR genetic variants regulate molecular function is also largely unknown. We performed a phenome-wide association analysis for 269 common protein-altering variants in 156 GPCRs and 275 phenotypes using data from 337,205 unrelated white British UK Biobank participants and identified 138 associations at a false discovery rate of 5%. We found a novel association between rs12295710 in *MRGPRE*, a member of the Mas-related receptor family involved in nociception, and migraine risk. We also identified an association between rs3732378, a missense mutation in the binding pocket of CX3CR1, and hypothyroidism. Five orphan GPCRs had eight genetic associations, highlighting novel biology for these receptors of unknown function. We found several associations between GPCR variants and food intake phenotypes, including an association between the variants in *TAS2R38* known to affect the ability to taste phenylthiocarbamide and tea intake as well as a non-additive associations between variants in *TAS2R19* and *TAS2R31* and coffee and tea intake. Finally, we tested whether genetic variants in *ADRB2* associated with immune cell amounts and pulmonary function affect downstream signaling pathways and found that two *ADRB2* haplotypes are associated with differential signaling relative to the most common haplotype. Overall, this study provides a map of genetic associations for GPCR coding variants across a wide variety of phenotypes that can inform future drug discovery efforts targeting GPCRs.

+Correspondence to Christopher DeBoever (cdeboeve@stanford.edu), AJ Venkatakrishnan (ajvenkat@stanford.edu), Ron Dror (rondror@stanford.edu) and Manuel A. Rivas (mrivas@stanford.edu).
* These authors contributed equally.

Introduction

G Protein Coupled Receptors (GPCRs) are allosteric molecular machines that detect extracellular messages, such as small molecules, peptides, light, or ions, and transmit a reciprocal signal into the cell. There are over 800 GPCRs in the human proteome, and they play key roles in diverse physiological functions and are the target of over a third of FDA approved medications (Katritch et al., 2013). Interest from the pharmaceutical industry and basic science has fueled intensive study of GPCRs, yet there still exist many GPCRs for which the primary physiological function is unknown, and even for well studied receptors the full range of physiological impacts may not be known. Understanding the role of GPCRs is particularly relevant because their positioning on the cell surface make them relatively easy to target with drugs, making it realistic to translate an association into a medication.

The UK Biobank dataset contains genetic and phenotypic data for nearly 500,000 individuals and provides a unique opportunity to study the physiological impact of GPCRs by performing a phenome-wide association study (PheWAS) to identify associations between variants in GPCRs and diverse phenotypes such as immune cell measurements, vital signs, and disease risk (Sudlow et al., 2015). A PheWAS focused on protein-altering variants in GPCRs can connect GPCRs to specific phenotypes and generate novel therapeutic hypotheses as they are likely candidates for causal variants (Altshuler et al., 2008). We can also leverage the increasing number of GPCR protein structures and shared structural scaffold across the GPCR family to identify specific genetic variants that are likely to impact protein structure and explain the genetic associations. Additionally, while the vast array of GPCRs recognize a diverse set of extracellular signals, they share a common set of intracellular binding partners: G Proteins and arrestins (Venkatakrishnan et al., 2013). G protein and arrestin coupling are directly measurable, providing insight into the molecular impact of GPCR variants.

In this study, we performed a PheWAS for 269 missense and protein-truncating variants in 156 GPCR genes across 275 phenotypes in the UK Biobank and identified 138 associations at a false discovery rate of 5%. The associations spanned 52 coding variants in 41 GPCRs and 46 quantitative or binary phenotypes and included both novel and previously reported associations. We identified five associations with binary phenotypes including an association between rs12295710 in *MRGPRE*, a gene whose paralogs are involved in nociception, and migraine (Dong et al., 2001). The remaining 133 associations were between coding GPCR variants and quantitative phenotypes including immune cell measurements, body size, pulmonary function, food intake, and others. Five orphan receptors of unknown function had associations with quantitative traits including *GPR35*, a known drug target for inflammatory bowel disease and cardiovascular disease (Divorty et al., 2015). We also found associations between variants in GPCR taste receptors known to affect bitter taste perception and self-reported tea and coffee intake. We tested genetic variants in *ADRB2* associated with several phenotypes for effects on downstream signaling pathways and identified two *ADRB2* haplotypes with differential signaling relative to the most frequent *ADRB2* haplotype. This study links coding variation in GPCRs to a range of diverse phenotypes and identifies novel biology for this important class of drug targets.

Results

GPCR genetic associations across 275 phenotypes

To assess the impact of genetic variation in GPCRs on disease risk and other phenotypes, we performed genetic association analyses between GPCR variants and phenotypes for 337,205 individuals of unrelated white British ancestry in the UK Biobank (see Methods for description). Beginning with a list of 226 non-olfactory human GPCRs, we identified 251 missense and 18 protein-truncating variants in 156 human GPCR genes that were genotyped in the UK Biobank arrays, passed quality control filters (Methods), and had minor allele frequency (MAF) greater than 1% (Table S1). We tested for associations between these 269 variants and 146 quantitative phenotypes with at least 3,000 observations and 129 binary phenotypes with at least 2,000 cases in the studied cohort (Methods, Table S1). The quantitative phenotypes include continuous or ordinal phenotypes such as weight, waist size, and forced expiratory volume. The binary phenotypes include the presence or absence of health conditions such as hyperthyroidism, migraine, or lung cancer. Due to power differences between quantitative and binary phenotype association tests, we corrected for multiple hypothesis testing separately for the quantitative and binary phenotypes using the Benjamini-Yekutieli (BY) procedure at FDR rate of 5%. Overall, we identified 52 coding variants in 41 GPCRs that were associated with at least one of 46 quantitative or binary phenotypes for a total of 138 associations (Table S2).

Taking advantage of GPCRs sharing a common fold, we annotated the tested variants with across-family conservation scores, GPCR functional regions, and PolyPhen-2 pathogenicity scores (Methods) (Adzhubei et al., 2010; Pándy-Szekeres et al., 2018; Venkatakrishnan et al., 2013). We compared the variants with and without significant associations to 86,601 rare protein-altering GPCR variants from gnomAD (Karczewski et al., 2019) (MAF<1%) and 318 GPCR variants reported in ClinVar (Landrum et al., 2018) and found that both the significant and non-significant UK Biobank variants had significantly lower PolyPhen-2 scores than the ClinVar and rare gnomAD GPCR variants (Wilcoxon, $p<1x10^{-5}$), consistent with the fact that the GPCR variants tested here are common (MAF > 1%) variants (Figure S1a-d, Table S3). Interestingly, there was not a significant difference between the UK Biobank variants with or without significant associations (Wilcoxon, p=0.69) suggesting that differences with other reference datasets have more to do with ascertainment of GPCR common variants in UK Biobank. We also compared family conservation scores for variants in class A GPCRs and found that the conservation scores generally agreed between the different variant sets except that ClinVar variants had higher conservation scores than gnomAD rare variants (Wilcoxon, $p=3.8x10^{-6}$) and variants without significant associations had lower conservation scores than ClinVar variants (Wilcoxon, p=0.013, Figure S1e, Table S3). Among the 52 variants with significant associations, we identified eight variants with associations that have PolyPhen-2 scores greater than 0.9, four variants located in the ligand binding pocket, and two variants located in the intracellular coupling interface (Table 1). For instance, we identified an association between rs3732378 (p.Thr280Met, MAF=17.3%) in *CX3CR1* and hypothyroidism ($p=1.3x10^{-7}$, OR=1.07, 95% CI: 1.05-1.10). rs3732378 has a PolyPhen-2 score of 0.774 and is

located in the binding pocket of CX3CR1. This GPCR is the sole receptor for the ligand fractalkine (CX3CL1), a chemokine with anti-apoptotic properties that has been implicated in oncogenesis (Pinato and Mauri, 2014; Rosen and Privalsky, 2011). We also identified associations between rs3732378 and monocyte counts, monocyte percentages, and lymphocyte counts (Table S2), consistent with previous studies (Astle et al., 2016). Expression of CX3CR1 is upregulated in the absence of triiodothyronine (T3) by thyroid receptors TRα1 and TRβ1 with mutations observed in human hepatocellular carcinoma suggesting that thyroid receptors may regulate CX3CR1 in some circumstances (Chan and Privalsky, 2009). Overall, these results indicate that although the common variants tested are generally not predicted to have a high impact on proteins, as is expected for common variants, they associate with phenotypic variation and some indeed have large predicted impact on proteins (Table 1).

Gene	Variant	MAF	HGVSp	PolyPhen-2	Function	Number of Associations
GIPR	rs1800437	19.4%	p.Glu354Gln	1		13
ADRB2	rs1800888	1.5%	p.Thr164Ile	0	Binding	6
ACKR2	rs2228467	6.1%	p.Val41Ala	0.958		5
MC4R	rs2229616	2.0%	p.Val103Ile	0.025	Binding	5
CX3CR1	rs3732378	17.3%	p.Thr280Met	0.774	Binding	4
P2RY2	Affx-5881601	24.0%	p.Arg312Ser	0.015	Coupling	4
CELSR3	rs3821875	11.2%	p.Ser805Thr	0.999		4
LGR4	rs34804482	2.6%	p.Asp844Gly	1		2
P2RY13	rs1466684	18.0%	p.Thr179Met	0.001	Binding	2
GPR35	rs3749171	18.2%	p.Thr139Met	0.956		2
GPR45	rs35946826	14.3%	p.Leu312Phe	1		1
NPFFR1	rs3812694	5.9%	p.Ile145Leu	0.999	Coupling	1
S1PR3	rs34075341	3.6%	p.Arg243Gln	1		1

Table 1. GPCR variants with significant associations and PolyPhen-2 scores greater than 0.9 or known functional annotations. “Number of Associations” indicates the number of significant associations identified for each variant in this study across both quantitative and binary phenotypes.

Associations with binary medical phenotypes

We identified 5 associations between GPCR variants and binary phenotypes (BY-adjusted $p < 0.05$, Table 2, Table S2). We identified a novel association between rs12295710 (p.Gly15Ser, MAF=47.1%) in *MRGPRE*, a member of the Mas-related receptor family, and migraine ($p=3.8\times10^{-8}$, OR=1.07, 95% CI: 1.05-1.10). Neither this variant nor other variants in linkage disequilibrium (LD) with it ($R^2>0.8$, British in England and Scotland (GBR)) have been previously reported as associated with migraine in the GWAS Catalog, in a meta-analysis of migraine GWAS including 375,000 individuals, or in a study of broadly-defined headaches in the UK Biobank (Buniello et al., 2019; International Headache Genetics Consortium et al., 2016; Meng

et al., 2018). Though the endogenous ligand of MRGPRE is unknown, Mas-related receptors are expressed in nociceptive sensory neurons and play a role in pain response. MRGPRE is expressed throughout the brains of macaque, mouse, and human including sensory neurons in mouse (Dong et al., 2001; Manteniotis et al., 2013; Zhang et al., 2005). Suggestive associations between variants in *MRGPRE* and white matter mean diffusivity and brain lesion distribution in multiple sclerosis have been reported indicating a possible role for *MRGPRE* in neurological disorders (Gourraud et al., 2013; Raffield et al., 2015).

We found an association between rs2234919 (p.Pro52Thr, MAF=5.5%) in the Thyroid Stimulating Hormone Receptor (*TSHR*) and hyperthyroidism ($p=3.6x10^{-7}$, OR=0.74, 95% CI: 0.66-0.83). Previous studies have found conflicting evidence regarding the association between rs2234919 and hyperthyroidism (Pujol-Borrell et al., 2015), and the variant is flagged as likely benign in ClinVar consistent with our finding that the minor allele lowers hyperthyroidism risk. We also identified an association between rs72703203 (p.Arg2015Lys, MAF=4.2%) in *CELSR2* and high cholesterol ($p=1.5x10^{-17}$, OR=0.85, 95% CI: 0.81-0.88). Variants in or near *CELSR2* have been previously associated with cholesterol, response to statins, lipoprotein-associated phospholipase A2 activity, and coronary artery disease (Arvind et al., 2014; Grallert et al., 2012; Ma et al., 2010; Postmus et al., 2014; van der Harst and Verweij, 2018). While rs72703203 is not in high LD with previously reported variants (Figure S2), it is not significant (p=0.21) in a model that includes rs12740374, a 3' UTR variant in *SORT1* proposed as the causal variant in this region (Musunuru and Kathiresan, 2019). We identified an association between rs2274911 (p.Pro91Ser, MAF=26.7%) in *GPRC6A* and prostate cancer ($p=8.2x10^{-8}$, OR=1.12, 95% CI: 1.07-1.17). rs2274911 is in LD with rs339331 ($R^2=0.846$, GBR) which has previously been associated with prostate cancer, and rs2274911 is not significant in a model that includes rs339331 as a covariate (p=0.47) (Hoffmann et al., 2015; Takata et al., 2010; Wang et al., 2015).

Gene	Variant	MAF	HGVSp	Phenotype	p-value	OR	OR 95% CI
CELSR2	rs72703203	4.2%	p.Arg2015Lys	High cholesterol	$1.5x10^{-17}$	0.85	0.81 - 0.88
MRGPRE	rs12295710	47.2%	p.Gly15Ser	Migraine	3.8E-08	1.07	1.05 - 1.1
GPRC6A	rs2274911	26.7%	p.Pro91Ser	Prostate cancer	$8.2x10^{-8}$	1.12	1.07 - 1.17
CX3CR1	rs3732378	17.3%	p.Thr280Met	Hypothyroidism/myxoedema	$1.3x10^{-7}$	1.07	1.05 - 1.1
TSHR	rs2234919	5.5%	p.Pro52Thr	Hyperthyroidism/thyrotoxicosis	$3.6x10^{-7}$	0.74	0.66 - 0.83

Table 2. Significant associations between coding GPCR variants and binary phenotypes.

Associations with quantitative phenotypes

We identified 133 quantitative phenotype associations for 49 protein-altering variants in 38 genes and 41 distinct phenotypes (BY-adjusted $p < 0.05$, Figure 1A-B, Table S2). 48 of these associations are present in this GWAS Catalog or in LD (R^2>0.8, GBR) with GWAS Catalog associations (Buniello et al., 2019). We identified novel associations for variants in three GPCRs with no entries in the GWAS catalog: rs7570797 in *PROKR1* (p.Ser40Gly, MAF=4.6%) was associated with platelet crit ($p=1.59 \times 10^{-7}$, β=-0.023, 95% CI: -0.032 - -0.015); rs1466684 in *P2RY13* (p.Thr179Met, MAF=18.0%) was associated with neutrophil count ($p=1.68 \times 10^{-6}$, β =0.016, 95% CI: 0.0094-0.023); and rs4274188 in *MRGPRX3* (p.Asn169Asp, MAF=23.4%) was associated with standing height ($p=7.69 \times 10^{-6}$, β=0.0090, 95% CI: 0.0051-0.013). We stratified the significant quantitative phenotype associations into the following phenotype categories (where n indicates the number of phenotypes with at least one association for each category): immune cell measurements (n=17), body size (n=6), lung function (n=6), food intake (n=4), physical ability (n=3), urine biomarkers (n=2), vital signs (n=2), and intelligence (n=1) (Table S1). We describe these associations in the sections below.

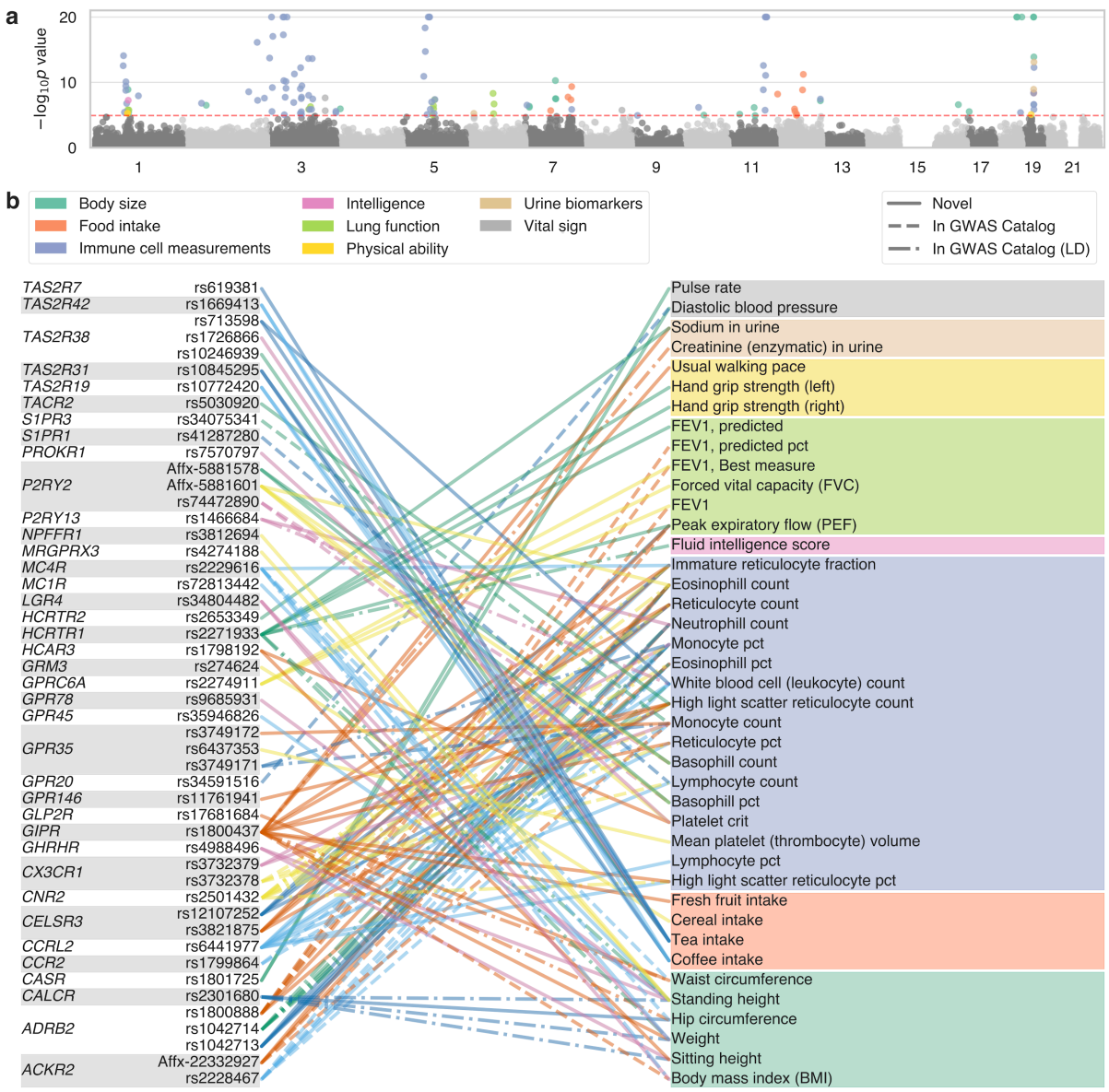

Figure 1. (a) Manhattan plot for quantitative phenotype associations across 269 variants and 146 phenotypes. Dashed red line indicates significance threshold after applying Benjamini-Yekutieli procedure. Scatter point colors indicate phenotype category (see legend in (b)) for significant associations. P-values less than $1x10^{-20}$ are plotted at $1x10^{-20}$. (b) Line plot showing associations between coding variants in indicated GPCRs (left) and quantitative phenotypes (right). Colors indicate phenotype category. Dashed lines indicate associations that are present in the GWAS Catalog and dot-dashed lines indicate that the variant is in LD with a variant associated with the phenotype in the GWAS Catalog.

Genetic variation in orphan GPCRs associated with quantitative phenotypes

The 269 variants tested for associations included 44 variants in 26 orphan GPCRs whose ligands and function are generally unknown (Tang et al., 2012). We identified quantitative phenotype associations for variants in five orphan receptors (Table 3). We found associations between rs3749171 in *GPR35* (p.Thr139Met, MAF=18.2%) and monocyte count ($p=7.5x10^{-17}$, β =-0.024, 95% CI: -0.031 - -0.019) and monocyte percentage ($p=5.7x10^{-8}$, β=-0.25, 95% CI: -0.016 - -0.034). rs3749171 is in LD with the 5' UTR variant rs34236350 (R^2=1, GBR) that has previously been reported as associated with monocyte count (Figure S3) (Astle et al., 2016). rs3749171 is in high LD with the intron variant rs4676410 (R^2=0.82, GBR), and these two variants have previously been associated with ankylosing spondylitis, ulcerative colitis/inflammatory bowel disease, and pediatric autoimmune diseases (Figure S3) (Ellinghaus et al., 2013; (igas) and International Genetics of Ankylosing Spondylitis Consortium (IGAS), 2013; International Multiple Sclerosis Genetics Consortium et al., 2015; Li et al., 2015; The International IBD Genetics Consortium (IIBDGC) et al., 2016, 2012; Venkateswaran et al., 2018). We also found associations between rs6437353 (p.Arg13His, MAF=45.6%) and mean platelet (thrombocyte) volume ($p=2.8x10^{-9}$, β=-0.021, 95% CI: -0.028 - -0.014) and rs3749172 (p.Ser294Arg, MAF=42.6%) and monocyte count ($p=2.6x10^{-8}$, β=0.013, 95% CI: 0.0084-0.017). These variants are not in LD with rs3749171 or rs34236350 (Figure S3), and neither these variants nor variants in LD with them have been previously reported as associated with these phenotypes in the GWAS Catalog (Buniello et al., 2019).

We found associations for missense variants in orphan GPCRs *GPR146*, *GPR20*, *GPR45*, and *GPR78* as well (Table 3). The association between rs34591516 (p.Gly313Ser, MAF=4.6%) in *GPR20* and diastolic blood pressure ($p=1.8x10^{-6}$, β=0.028, 95% CI: 0.017-0.04) agrees with a previous study that found the same variant associated with diastolic and systolic blood pressure (CHARGE-Heart Failure Consortium et al., 2016). We find that rs9685931 (p.Arg342His, MAF=11.3%) in *GPR78* is associated with standing height ($p=1.1x10^{-6}$, β=-0.012, 95% CI: -0.0077 - -0.017). A different variant in the 3' UTR of *GPR78*, rs3775887, is in weak LD with rs9685931 (LD=0.18, GBR) and has previously been associated with FVC. Overall these associations indicate potential functions or relevant pathways for these orphan GPCRs.

Gene	Variant	MAF	HGVSp	Phenotype	p-value	BETA	BETA 95% CI
GPR35	rs3749171	18.2%	p.Thr139Met	Monocyte count	$7.5x10^{-17}$	-0.025	-0.031 - -0.019
GPR35	rs6437353	45.6%	p.Arg13His	Mean platelet (thrombocyte) volume	$2.8x10^{-9}$	-0.021	-0.028 - -0.014
GPR35	rs3749172	42.6%	p.Ser294Arg	Monocyte count	$2.6x10^{-8}$	0.013	0.0084 - 0.017
GPR35	rs3749171	18.2%	p.Thr139Met	Monocyte percentage	$5.7x10^{-8}$	-0.025	-0.034 - -0.016
GPR146	rs11761941	14.8%	p.Gly11Glu	High light scatter reticulocyte count	$3.1x10^{-7}$	-0.017	-0.023 - -0.01

GPR45	rs35946826	14.3%	p.Leu312Phe	Standing height	$3.2x10^{-7}$	-0.012	-0.017 - -0.0077
GPR78	rs9685931	11.3%	p.Arg342His	Standing height	$1.1x10^{-6}$	0.013	0.0078 - 0.018
GPR20	rs34591516	4.6%	p.Gly313Ser	Diastolic blood pressure, automated reading	$1.8x10^{-6}$	0.028	0.017 - 0.04

Table 3. Significant quantitative phenotype associations for orphan GPCRs.

Genetic variation in GPCRs associated with food intake

We identified 12 associations between coding variants in GPCRs and food intake phenotypes. We found that variants in several taste receptors were associated with either coffee or tea intake (Table 4). Variants in the bitter taste receptor genes *TAS2R19*, *TAS2R31*, and *TAS2R42* were associated with coffee intake, though the variants in *TAS2R19* and *TAS2R31* are in high LD (R^2=0.9, GBR). rs10772420 (p.Arg299Cys, MAF=46.8%) in *TAS2R19* has been shown to affect perception of bitterness, though it is unclear whether rs10772420, rs10845295 (p.Arg35Trp, MAF=48.6%) in *TAS2R31*, or a different variant is causal (Hayes et al., 2015, 2011). We also found associations between variants in the taste receptors *TAS2R19*, *TAS2R31*, *TAS2R38*, *TAS2R42*, and *TAS2R7* and tea intake. The same variants in *TAS2R19* and *TAS2R31* that are in high LD and were associated with coffee intake are also associated with tea intake. The variants in *TAS2R7* and *TAS2R42* associated with tea intake are in moderate LD (R^2=0.59, GBR). The variants in *TAS2R38* that are associated with tea intake are in high LD and are known to affect the ability to taste phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP) (Figure S4) (Kim, 2003; Risso et al., 2016). The ability to taste PTC follows a dominant model of inheritance where the alternate “taster” allele C is dominant to the “non-taster” reference T allele for rs10246939 (Kim, 2003). We therefore fit additive and genotypic models for tea intake and the variants in *TAS2R38* to test whether the observed association with tea intake is dominant. Interestingly, we found that the association between variants in *TAS2R38* and tea intake is consistent with an additive effect where each additional copy of the taster C allele is associated with increased tea intake (Figure 2A). We did find evidence, however, that the association between rs10772420 in *TAS2R19* and coffee intake is non-additive, where only one copy of the alternate allele accounts for the majority of the decrease in coffee intake (Figure 2B). We found weaker evidence that the association between variants in *TAS2R19* and tea intake is non-additive, where one copy of the alternate alleles accounts for most of the increase in tea intake (Figure 2C). As noted above, rs10772420 is in LD with the rs10845295 in *TAS2R31* (0.895, GBR). Overall, we identified three independent associations between variants in taste receptors and tea intake and two independent associations for coffee intake.

We also observed two associations between non-taste related GPCRs and self-reported food intake. We found that rs1800437 (p.Glu354Gln, MAF=19.4%) in *GIPR* is associated with fresh fruit intake ($p=4.2x10^{-9}$, β=0.017, 95% CI: 0.011-0.023). GIP-R is a receptor of gastric inhibitory polypeptide (GIP), which stimulates insulin release in the presence of elevated glucose (Dupre et al., 1973). Mice lacking GIP-R exhibit higher blood glucose levels and impaired initial insulin response following oral glucose load (Miyawaki et al., 1999). A previous pharmacological study

of GIP-R found that there is a marked decrease in basal GIP-R signaling for the alternate rs1800437 allele compared to WT (Fortin et al., 2010). We also identified several other associations for rs1800437 in *GIPR* (Figure 1b) that agree with reported associations in the GWAS Catalog including associations with body mass index, weight, waist circumference, and hip circumference, though the reported association with fresh fruit intake has not been previously reported (Buniello et al., 2019). We also found that the splice donor variant rs274624 (MAF=37.0%) in *GRM3* is associated with cereal intake ($p=2.1x10^{-6}$, β=-0.011, 95% CI: -0.015 - -0.0064). *GRM3* encodes a metabotropic glutamate receptor that is involved in brain function. Missense variants in *GRM3* have been previously associated with cognitive performance and self-reported math ability and non-coding variants near *GRM3* have been associated with several cognitive and psychiatric phenotypes including schizophrenia and neuroticism (23andMe Research Team et al., 2018; Buniello et al., 2019). A non-coding variant near *GRM3* has also been associated with body mass index, but no associations between variants in or near *GRM3* and food intake have been reported previously (Kichaev et al., n.d.). These results demonstrate that genetic variation in GPCRs can be linked to food intake using self-reported intake measurements in a population biobank.

Gene	Variant	MAF	HGVSp	Phenotype	p-value	BETA	BETA 95% CI
TAS2R19	rs10772420	46.8%	p.Arg299Cys	Coffee intake	$2.6x10^{-6}$	-0.011	-0.016 - -0.0065
TAS2R31	rs10845295	48.6%	p.Arg35Trp	Coffee intake	$1.1x10^{-5}$	-0.01	-0.015 - -0.0057
TAS2R42	rs1669413	16.5%	p.Gly255Trp	Coffee intake	$1.5x10^{-9}$	-0.019	-0.025 - -0.013
TAS2R19	rs10772420	46.8%	p.Arg299Cys	Tea intake	$1.2x10^{-6}$	0.012	0.0069 - 0.016
TAS2R31	rs10845295	48.6%	p.Arg35Trp	Tea intake	$7.2x10^{-6}$	0.011	0.006 - 0.015
TAS2R38	rs10246939	44.6%	p.Ile296Val	Tea intake	$1.8x10^{-8}$	0.013	0.0087 - 0.018
TAS2R38	rs1726866	44.6%	p.Ala262Val	Tea intake	$4.6x10^{-8}$	-0.013	-0.018 - -0.0084
TAS2R38	rs713598	40.0%	p.Ala49Pro	Tea intake	$4.5x10^{-10}$	0.015	0.01 - 0.02
TAS2R42	rs1669413	16.5%	p.Gly255Trp	Tea intake	$6.0x10^{-12}$	0.022	0.016 - 0.028
TAS2R7	rs619381	12.7%	p.Met304Ile	Tea intake	$6.4x10^{-9}$	-0.021	-0.028 - -0.014

Table 4. Associations between variants in taste receptors and coffee or tea intake.

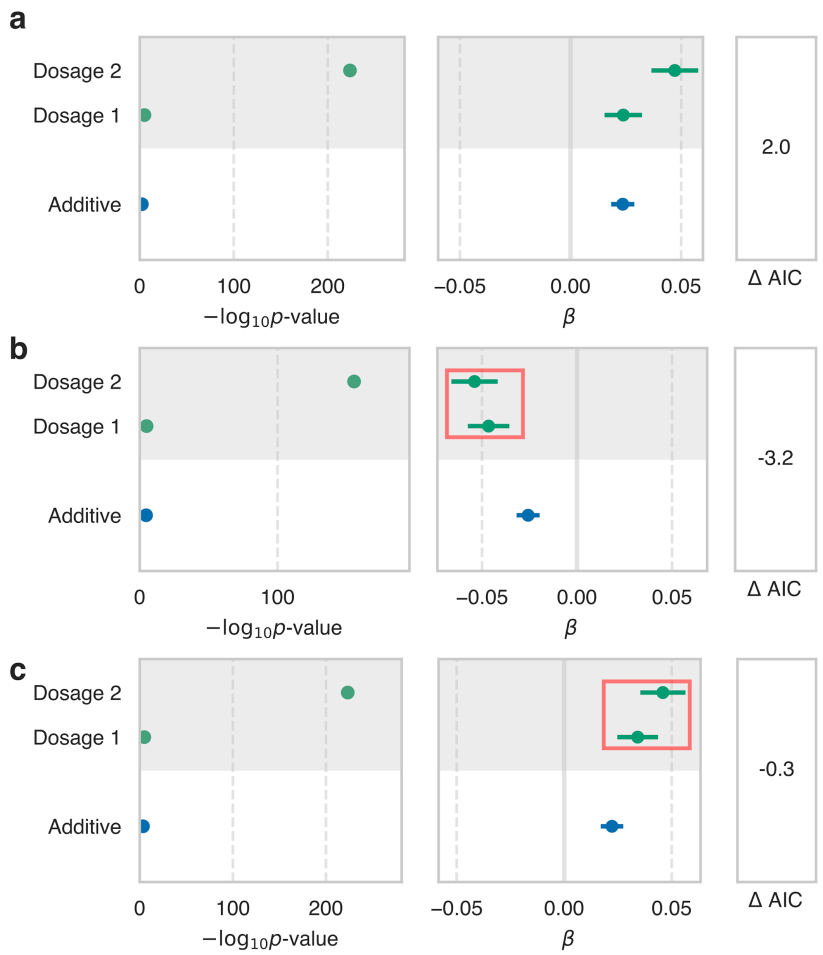

Figure 2. -$\log_{10}$p-values and effect sizes for additive (white background) and genotypic (grey background) models of association between (a) rs10246939 in *TAS2R38* and tea intake, (b) rs10772420 in *TAS2R19* and coffee intake, and (c) rs10772420 in *TAS2R19* and tea intake. The genotypic model includes separate terms for heterozygous and homozygous alternate genotypes. Red boxes indicate non-additive effects. Δ AIC is the difference in AIC between the two models (genotypic minus additive).

Genetic variation in *ADRB2* alters arrestin signaling

We identified associations between three missense variants in *ADRB2*, the β2 adrenergic receptor, and immune cell counts and percentages as well as forced expiratory volume in 1-second (FEV1) and peak expiratory flow (PEF) (Table 5). Several of these associations have been reported previously in the GWAS Catalog or are in LD with reported associations, though associations between rs1800888 (p.Thr164Ile, MAF=1.5%) and eosinophil counts ($p=5.6\times10^{-6}$, β =0.045, 95% CI: 0.026-0.065) and eosinophil percentage ($p=4.3\times10^{-8}$, β=0.061, 95% CI: 0.039-0.082) were not in the GWAS Catalog. Similarly, the associations between rs1042713 (p.Gly16Arg, MAF=36.1%) and immune cell counts and percentages were not in the GWAS Catalog or in LD with associations in the GWAS Catalog (Table 5) (Buniello et al., 2019). Since

rs1042713 and rs1042714 are in moderate LD (R^2=0.469, GBR), we fit models for eosinophil counts, eosinophil percentage, monocyte percentage, and neutrophil count using all three variants in *ADRB2*. For all four phenotypes, only one of rs1042713 and rs1042714 were significant ($p<0.05$) indicating that these two variants likely do not have independent effects on these immune cell phenotypes. rs1800888 remained significant ($p<0.05$) conditioned on both rs1042713 and rs1042714 for eosinophil counts and percentages.

To further explore how genetic variation in *ADRB2* might impact the function of the receptor, we assayed cAMP signaling, which is downstream of Gs, and β-arrestin-1 and η-arrestin-2 recruitment for the RQT, GET, GQI, and GQT *ADRB2* haplotypes upon stimulation with the endogenous ligands epinephrine (epi) and norepinephrine (norepi). We refer to the *ADRB2* haplotypes by the residues at the 3 variable positions; for instance, the most common haplotype is RQT where rs1042713 codes for arginine (Arg16Gly), rs1042714 codes for glutamine (Gln27Glu), and rs1800888 codes for threonine (Thr164Ile). We found that two of the haplotypes displayed significantly different signaling profiles than the most frequent haplotype RQT. The GET haplotype retained full arrestin recruitment but significantly affected Gs signaling; for epi, GET has a decreased pEC50 in the cAMP assay ($p=0.02$, Tukey's test), indicating that a higher concentration of ligand is required to induce the same level of signaling through cAMP (Figure S5, Table S4). Meanwhile, the GQI haplotype displayed decreased arrestin pEC50 and full Gs signaling relative to RQT ($p < 0.004$, Tukey's test).

Variant	MAF	HGVSp	Phenotype	GWAS Catalog	p-value	BETA	BETA 95% CI
rs1042714	45.3%	p.Glu27Gln	Eosinophil count	TRUE	2.5×10^{-33}	0.029	0.024 - 0.034
rs1042713	36.1%	p.Gly16Arg	Eosinophil count	FALSE	1.9×10^{-15}	0.02	0.015 - 0.025
rs1800888	1.5%	p.Thr164Ile	Eosinophil count	FALSE	5.6×10^{-6}	0.045	0.026 - 0.065
rs1042714	45.3%	p.Glu27Gln	Eosinophil percentage	TRUE	9.9×10^{-43}	0.037	0.032 - 0.042
rs1042713	36.1%	p.Gly16Arg	Eosinophil percentage	FALSE	4.6×10^{-19}	0.025	0.019 - 0.03
rs1800888	1.5%	p.Thr164Ile	Eosinophil percentage	FALSE	4.3×10^{-8}	0.061	0.039 - 0.082
rs1800888	1.5%	p.Thr164Ile	FEV1	TRUE	4.6×10^{-8}	-0.042	-0.057 - -0.027
rs1800888	1.5%	p.Thr164Ile	FEV1, Best measure	TRUE	1.1×10^{-6}	-0.04	-0.056 - -0.024
rs1800888	1.5%	p.Thr164Ile	FEV1, predicted percentage	TRUE	9.2×10^{-6}	-0.075	-0.11 - -0.042
rs1042714	45.3%	p.Glu27Gln	Monocyte	TRUE	1.0×10^{-7}	0.019	0.012 - 0.026

			percentage				
rs1042713	36.1%	p.Gly16Arg	Monocyte percentage	FALSE	$8.6x10^{-6}$	0.016	0.0092 - 0.024
rs1042714	45.3%	p.Glu27Gln	Neutrophil count	TRUE	$9.2x10^{-23}$	-0.027	-0.033 - -0.022
rs1042713	36.1%	p.Gly16Arg	Neutrophil count	FALSE	$1.2x10^{-11}$	-0.019	-0.025 - -0.014
rs1800888	1.5%	p.Thr164Ile	Peak expiratory flow (PEF)	TRUE	$3.8x10^{-7}$	-0.043	-0.06 - -0.026
rs1042714	45.3%	p.Glu27Gln	White blood cell (leukocyte) count	TRUE	$1.3x10^{-6}$	-0.016	-0.023 - -0.0098

Table 5. Associations between missense variants in *ADRB2* and quantitative phenotypes. FEV1 is forced expiratory volume in one second.

Discussion

This study provides a systematic catalog of associations between coding variants in GPCRs and 275 diverse phenotypes in the UK Biobank. We replicated known associations between variants associated with disease risk and quantitative phenotypes and identified novel associations that in some cases indicate novel functions for GPCRs. The association between rs12295710 in *MRGPRE* and migraine is particularly interesting given that the genes in this family play a role in sensing pain (Dong et al., 2001). We found that rs3749171 in *GPR35* is associated with monocyte counts and percentage and is in LD with*GPR35* variants associated with ankylosing spondylitis, inflammatory bowel disease, and pediatric autoimmune diseases (refs above). Interestingly, we also found a novel association between rs3749172 in *GPR35* and monocyte counts, though this variant does not appear to be in LD with loci previously associated with diseases. The novel associations between GPCR taste receptors and self-reported tea and coffee intake demonstrate the utility of questionnaire data in population biobanks, even for relatively difficult-to-estimate phenotypes like food intake.

We found that the GET *ADRB2* haplotype had decreased Gs signaling (as measured by cAMP levels) relative to the RQT haplotype and that the GQI haplotype had decreased arrestin recruitment relative to the RQT haplotype. The observed genetic association between decreased forced expiratory volume and the rs1800888 allele that codes for isoleucine and our results showing that GQI has decreased arrestin signaling might inform the role of B2AR in pulmonary function. Since arrestin recruitment induces B2AR desensitization, and, somewhat paradoxically, B2AR agonist activity is required for the development of asthma in mouse models, it is possible that the GQI haplotype may decrease lung function by reducing agonist-induced B2AR desensitization (Moore et al., 2007; Thanawala et al., 2013). It is also possible,

however, that the acute differences observed in these experiments may differ from long-term effects of differential signaling associated with genetic variants.

The genetic associations reported here link GPCRs to phenotypes and will be useful for assessing the function of GPCRs and potential effects of modulating these genes with therapeutics. Future studies utilizing exome or genome sequencing that can better ascertain rare variants in these genes will likely identify new associations that are not observed for common variants that have faced stronger selective pressure. In particular, studies that focus on populations with unique genetic histories such as founder populations or groups with high consanguinity offer an important opportunity to identify high-impact, rare variants in GPCRs that may have large impacts on phenotypic variation and disease risk.

Methods

Quality Control of Genotype Data

We used genotype data from UK Biobank dataset release version 2 for all aspects of the study. To minimize the impact of cofounders and unreliable observations, we used a subset of individuals that satisfied all of the following criteria: (1) self-reported white British ancestry, (2) used to compute principal components, (3) not marked as outliers for heterozygosity and missing rates, (4) do not show putative sex chromosome aneuploidy, and (5) have at most 10 putative third-degree relatives. These criteria are reported by the UK Biobank in the file "ukb_sqc_v2.txt" in the following columns respectively: (1) "in_white_British_ancestry_subset," (2) "used_in_pca_calculation," (3) "het_missing_outliers," (4) "putative_sex_chromosome_aneuploidy", and (5) "excess_relatives." We removed 151,169 individuals that did not meet these criteria. For the remaining 337,205 individuals, we used PLINK v1.90b4.4 76 to compute the following statistics for each variant: (a) genotyping missingness rate, (b) p-values of Hardy-Weinberg test, and (c) allele frequencies (calculated separately for the two genotyping arrays used by the UK Biobank).

GPCR Variants

UK Biobank Variant Annotation

Variants on the UK Biobank arrays were annotated using VEP version 87 and the GRCh37 reference (McLaren et al., 2016). Protein truncating variants were predicted using the VEP LOFTEE plugin (https://github.com/konradjk/loftee). Linkage disequilibrium (LD) was calculated for the GBR British population from 1000 Genomes Phase 3 (Version 5) using LDmatrix from LDlink (https://ldlink.nci.nih.gov/) unless otherwise stated(Machiela and Chanock, 2015).

GPCR Gene Definition

A list of GPCRs was obtained from the International Union of Basic and Clinical Pharmacology (IUPHAR) website in December 2017 (Table S1). This list excludes olfactory GPCRs.

GPCR Variant Filtering

We identified 1,263 coding variants in GPCRs that (1) were genotyped on the UK Biobank genotyping arrays, (2) not filtered out during quality control of the genotyping data, (3) were not in the major histocompatibility complex (chr6:28477797-33448354) (Church et al., 2011), and (4) were annotated as one of the following variant types by VEP: missense_variant, frameshift_variant, stop_gained, splice_region_variant, splice_donor_variant, splice_acceptor_variant, stop_lost, inframe_deletion, start_lost, inframe_insertion, incomplete_terminal_codon_variant. We filtered the variants to include only those with minor allele frequency greater than 1% in the 337,205 white British subjects used in this study resulting in 269 variants in 156 genes (Table S1). 251 of the 269 variants are missense variants and the remaining 18 variants are predicted protein truncating variants.

GPCR Variant Annotation

86,601 rare protein-altering GPCR variants were obtained from from gnomAD (MAF<1%) and 318 GPCR variants were obtained from ClinVar (Karczewski et al., 2019; Landrum et al., 2018). Family conservation scores were created using multiple sequence alignments from GPCRdb (Pándy-Szekeres et al., 2018). The scores were defined as the frequency of the most common residue at each structurally equivalent position. PolyPhen-2 (Adzhubei et al., 2010) scores were obtained through the web portal at http://genetics.bwh.harvard.edu/pph2/bgi.shtml. Position of variants on the GPCR fold were determined using the multiple alignment from GPCRdb (Pándy-Szekeres et al., 2018).

Phenotype Definitions

Hospital Record and Verbal Questionnaire

As previously described, we used the following procedure to define cases and controls for non-cancer phenotypes (DeBoever et al., 2018). For a given phenotype, ICD-10 codes (Data-Field 41202) were grouped with self-reported non-cancer illness codes from verbal questionnaires (Data-Field 20002) that were closely related. This was done by first creating a computationally generated candidate list of closely related ICD-10 codes and self-reported non-cancer illness codes, then manually curating the matches. The computational mapping was performed by calculating the token set ratio between the ICD-10 code description and the self-reported illness code description using the FuzzyWuzzy python package. The high scoring ICD-10 matches for each self-reported illness were then manually curated to ensure high confidence mappings. Manual curation was required to validate the matches because fuzzy string matching may return words that are similar in spelling but not in meaning. For example, to create a hypertension cohort the code description from Data-Field 20002 ("Hypertension") was mapped to all ICD-10

code descriptions and all closely related codes were returned ("I10: Essential (primary) hypertension" and "I95: Hypotension"). After manual curation code I10 would be kept and code I95 would be discarded. After matching ICD-10 codes and with self-reported illness codes, cases were identified for each phenotype using only the associated ICD-10 codes, only the associated self-reported illness codes, or both the associated ICD-10 codes and self-reported illness codes.

Family History

We used data from Category 100034 (Family history - Touchscreen - UK Biobank Assessment Centre) to define "cases" and controls for family history phenotypes. This category contains data from the touchscreen questionnaire on questions related to family size, sibling order, family medical history (of parents and siblings), and age of parents (age of death if died). We focused on Data Coding 20107: Illness of father and 20110: Illness of mother.

Genetic Association Analyses

Quantitative phenotypes

We identified 146 quantitative phenotypes with at least 3,000 observations among the 337,205 white British subjects in the UK Biobank. As previously described, we took non-NA median of multiple measurements across up to three time points (Tanigawa et al., 2019). We focused on food intake, immune cell measurements, gross body measurements, behavioral phenotypes, and several other phenotypes (Table S1). We performed linear regression association analysis with v2.00a (20 Sep, 2017) for these 146 quantitative phenotypes. We performed quantile normalization for each phenotype (--pheno-quantile-normalize option), where we fit the linear model with covariates and transform the residuals to Normal distribution N(0, 1) while preserving the original rank in the residuals. We used the following covariates in our analysis: age, sex, array type, and the first four principal components, where array type is a binary variable that represents whether an individual was genotyped with UK Biobank Axiom Array or UK BiLEVE Axiom Array. For variants that were specific to one array, we did not use array as a covariate. We corrected p-values using the Benjamin-Yekutieli approach implemented in R's p.adjust. We considered associations with BY-corrected p-values less than 0.05 as significant which controls the false discovery rate at 5% (Yekutieli and Benjamini, 2001).

Binary phenotypes

Wd identified 129 binary phenotypes with at least 2,000 cases among the 337,205 white British subjects in the UK Biobank and performed logistic regression association analysis with Firth-fallback using PLINK v2.00a (17 July 2017) (Chang et al., 2015). Firth-fallback is a hybrid algorithm which normally uses the logistic regression code described in (Hill et al., 2017), but switches to a port of logistf() (https://cran.r-project.org/web/packages/logistf/index.html) in two cases: (1) one of the cells in the 2x2 allele count by case/control status contingency table is empty (2) logistic regression was attempted since all the contingency table cells were nonzero, but it failed to converge within the usual number of steps. We used the following covariates in

our analysis: age, sex, array type, and the first four principal components, where array type is a binary variable that represents whether an individual was genotyped with UK Biobank Axiom Array or UK BiLEVE Axiom Array. For variants that were specific to one array, we did not use array as a covariate. We corrected p-values using the Benjamin-Yekutieli approach implemented in R's p.adjust. We considered associations with BY-corrected p-values less than 0.05 as significant which controls the false discovery rate at 5%.

Comparison to GWAS Catalog

We compared significant associations to the GWAS Catalog (Buniello et al., 2019) to determine whether the associations we identified had been previously reported. We downloaded the GWAS Catalog v1.0.2 on August 23, 2019. We matched phenotypes in the GWAS Catalog to our phenotypes using the FuzzyWuzzy python package and manually reviewed matches. We erred on the side of including similar phenotypes from the GWAS Catalog for matches that were not perfect. We then matched the associated variants to variants in the GWAS Catalog using rs identifiers and chromosome:position identifiers. In order to match variants using chromosome:position, we used dbSNP or UCSC LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver) to convert UK Biobank genotyping array coordinates in hg19 to GRCh38 coordinates to match with the GWAS Catalog.

We used LDlink (https://ldlink.nci.nih.gov/) (Machiela and Chanock, 2015) to identify variants in LD with the variants that we found associations for using the GBR British population from 1000 Genomes Phase 3 (Version 5). We matched these variants in LD with the associated variants to variants in the GWAS Catalog using rs identifiers and chromosome:position identifiers when rs identifiers failed.

B2AR Conditional Analyses

We fit a model for each of eosinophil counts, eosinophil percentage, monocyte percentage, and neutrophil count with the genotypes of rs1042713, rs1042714, and rs1800888 as independent variables. We included a constant and age, sex, array type, and the first four principal components as covariates for these models.

Coffee and Tea Intake Additivity Analysis

We fit two models for each of coffee and tea intake to test whether the associations were consistent with non-additive effects. We performed quantile normalization for each phenotype to transform the values to standard normal distribution. For the additive model, we encoded genotypes as 0,1,2 and included a single term for the effect of genotype on tea/coffee intake. For the genotypic model we included a term to indicate whether a subject was heterozygous and a separate term to indicate whether a subject was homozygous alternate. We included a constant and age, sex, array type, and the first four principal components as covariates for these regressions.

Supplementary Information

Supplementary Figures

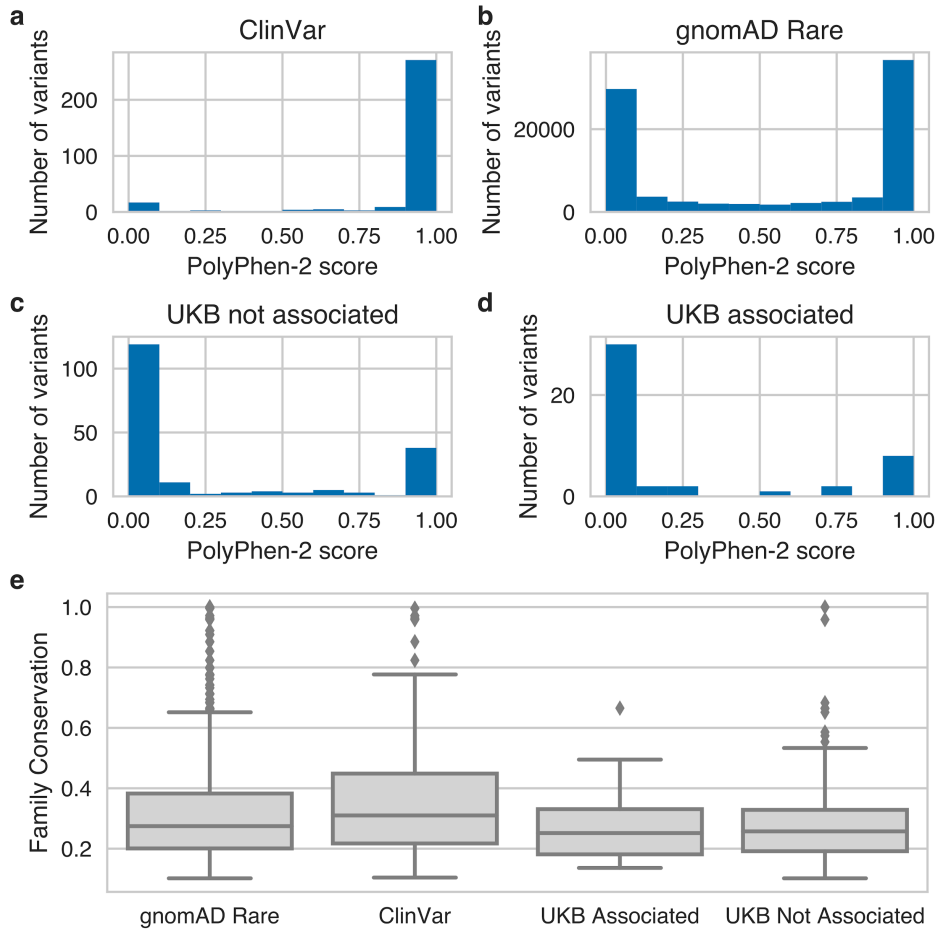

Figure S1. Distribution of PolyPhen-2 scores for (a) GPCR variants in ClinVar; (b) rare (MAF<1%) GPCR variants in gnomAD, (c) variants analyzed in this study that were not significantly associated with any phenotype, and (d) variants analyzed in this study that were significantly associated with at least one phenotype. (e) Distribution of family conservation scores for variants in class A GPCRs across the four groups of variants in a-d. The distribution of conservation scores is significantly different (Wilcoxon $p<0.05$) between the gnomAD Rare and ClinVar variants and between the ClinVar and UKB Not Associated variants.

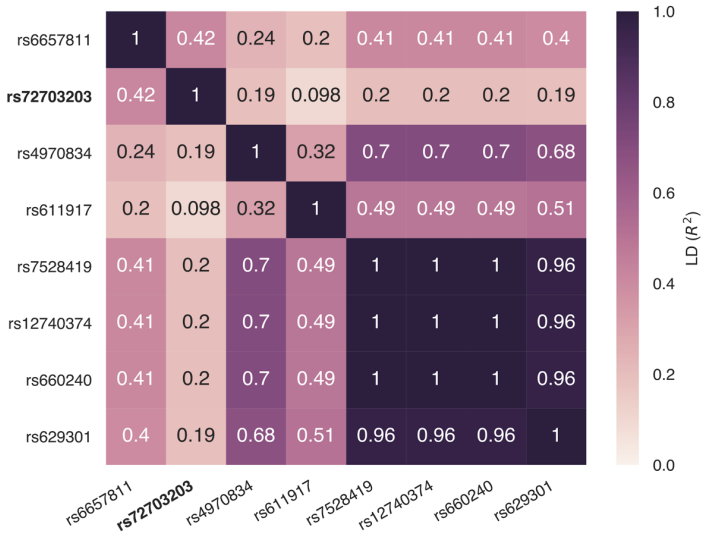

Figure S2. Linkage disequilibrium in the GBR British population between rs72703203 (bold) and other *CELSR2* variants reported in the GWAS Catalog as significantly associated with various lipid and cardiac phenotypes.

Figure S3. Linkage disequilibrium in the GBR British population between variants in *GPR35* with significant associations in this study (bold) and variants reported in the GWAS Catalog.

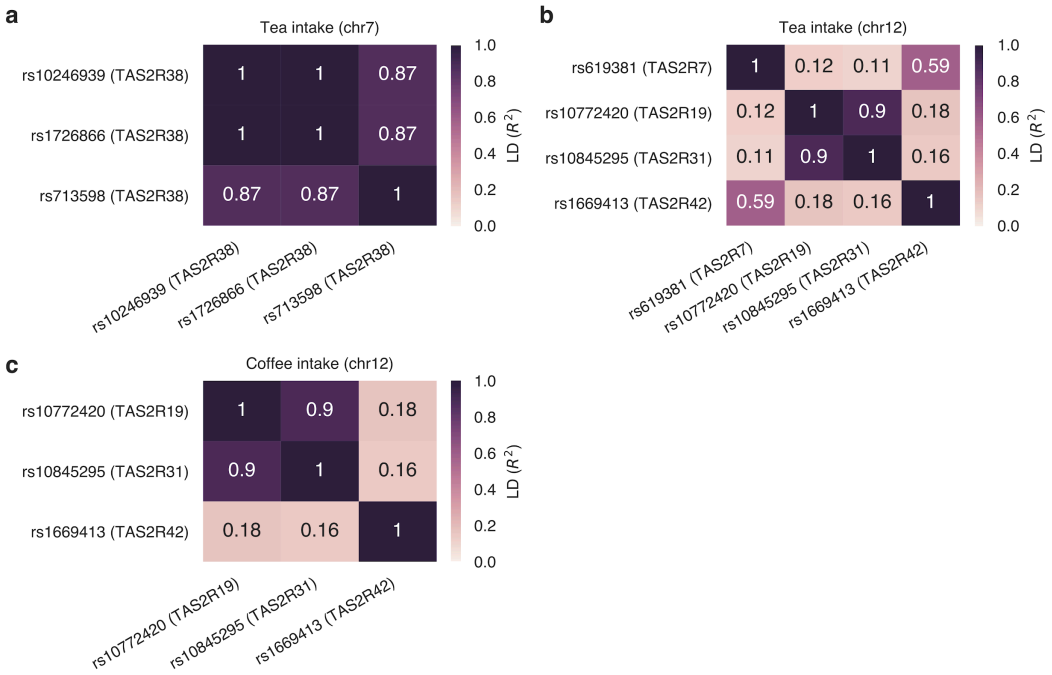

Figure S4. Linkage disequilibrium in the GBR British population between variants in taste receptors associated with (a) tea intake on chromosome 7, (b) tea intake on chromosome 12, or (c) coffee intake on chromosome 12.

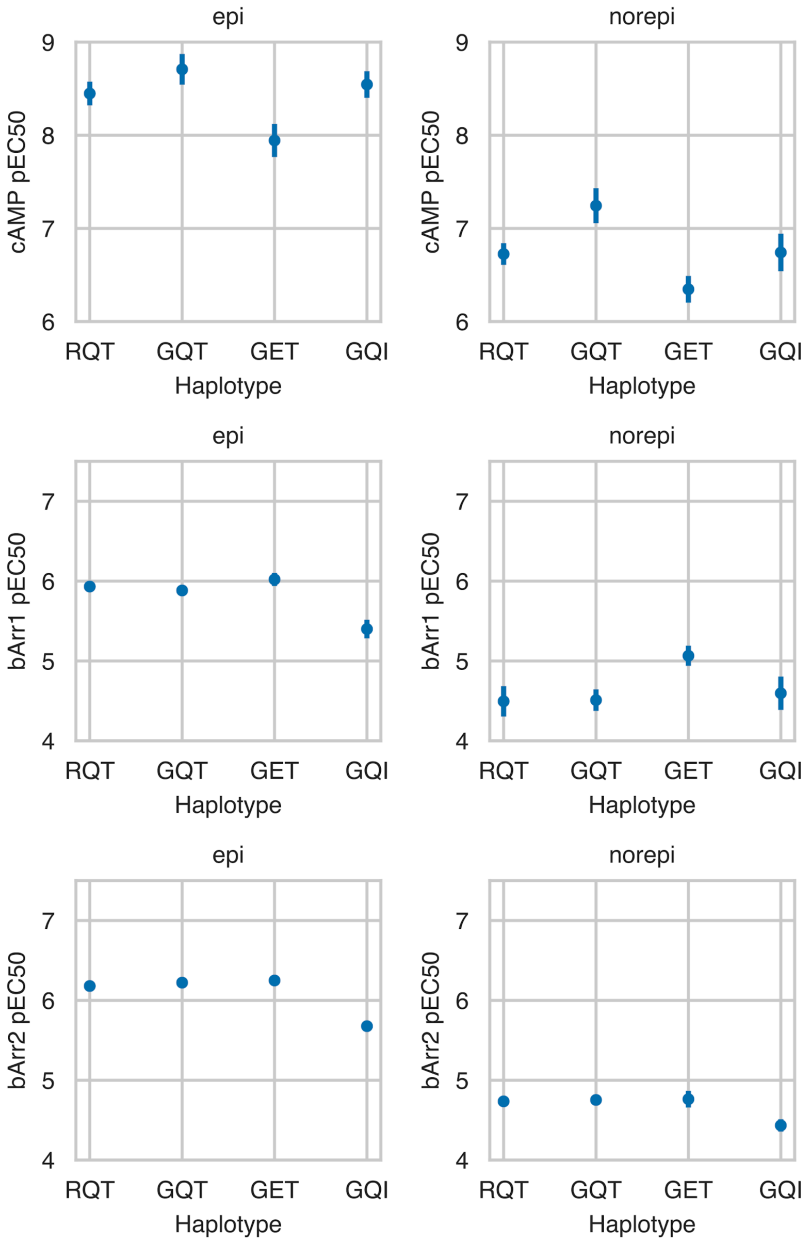

Figure S5. pEC50 for cAMP signaling, beta-arrestin-1 recruitment, and beta-arrestin-2 recruitment for four *ADRB2* haplotypes upon stimulation with the endogenous ligands epinephrine (epi) and norepinephrine (norepi).

Supplementary Tables

Table S1. Phenotypes and GPCR genes and variants used for PheWAS. The “GPCR genes” tab contains 398 GPCR genes that we used to identify GPCR genetic variants. The “GPCR tested variants” tab contains the 268 variants that we tested in the PheWAS analysis. The “GPCR variant summary” tab shows the number of each variant type included in the pheWAS

analysis. The “Quantitative phenotypes” and “Binary phenotypes” tabs contain the phenotypes analyzed and the number of observations and cases, respectively.

Table S2. GPCR PheWAS results. Summary statistics for quantitative and binary associations with BY-adjusted p-value less than 0.05.

Table S3. GPCR variant annotations. Across-family conservation scores, GPCR functional region annotations, and PolyPhen-2 pathogenicity scores for GPCR variants assessed in this study, 86,601 rare GPCR variants from gnomAD (MAF<1%) and 318 GPCR variants reported in ClinVar.

Table S4. p-values from Tukey’s test for differences in pEC50 for cAMP signaling, beta-arrestin-1 recruitment, and beta-arrestin-2 recruitment for four *ADRB2* haplotypes upon stimulation with the endogenous ligands epinephrine (epi) and norepinephrine (norepi).

Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number 24983, “Generating effective therapeutic hypotheses from genomic and hospital linkage data” (http://www.ukbiobank.ac.uk/wp-content/uploads/2017/06/24983-Dr-Manuel-Rivas.pdf). Based on the information provided in Protocol 44532 the Stanford IRB has determined that the research does not involve human subjects as defined in 45 CFR 46.102(f) or 21 CFR 50.3(g). All participants of UK Biobank provided written informed consent (more information is available at https://www.ukbiobank.ac.uk/2018/02/gdpr/). This work was supported by National Human Genome Research Institute (NHGRI) of the National Institutes of Health (NIH) under awards R01HG010140 (M.A.R.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. M.A.R. is supported by Stanford University and a National Institute of Health center for Multi- and Trans-ethnic Mapping of Mendelian and Complex Diseases grant (5U01 HG009080). C.D. and A.J.V are supported by the Stanford ChEM-H Postdocs at the Interface Seed Grant. The authors would like to thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at https://gnomad.broadinstitute.org/about.

References

23andMe Research Team, COGENT (Cognitive Genomics Consortium), Social Science Genetic Association Consortium, Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linnér R, Fontana MA, Kundu T, Lee C, Li H, Li R, Royer R, Timshel PN, Walters RK, Willoughby EA, Yengo L, Alver M, Bao Y, Clark DW, Day FR, Furlotte NA, Joshi PK, Kemper KE, Kleinman A, Langenberg C, Mägi R, Trampush JW, Verma SS, Wu Y, Lam M, Zhao JH, Zheng Z, Boardman JD, Campbell H, Freese J, Harris KM, Hayward C, Herd P, Kumari M, Lencz T, Luan J, Malhotra AK, Metspalu A, Milani L, Ong KK, Perry JRB, Porteous DJ, Ritchie MD, Smart MC, Smith BH,

Tung JY, Wareham NJ, Wilson JF, Beauchamp JP, Conley DC, Esko T, Lehrer SF, Magnusson PKE, Oskarsson S, Pers TH, Robinson MR, Thom K, Watson C, Chabris CF, Meyer MN, Laibson DI, Yang J, Johannesson M, Koellinger PD, Turley P, Visscher PM, Benjamin DJ, Cesarini D. 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* **50**:1112–1121.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods* **7**:248–249.

Altshuler D, Daly MJ, Lander ES. 2008. Genetic mapping in human disease. *Science* **322**:881–888.

Arvind P, Nair J, Jambunathan S, Kakkar VV, Shanker J. 2014. CELSR2–PSRC1–SORT1 gene expression and association with coronary artery disease and plasma lipid levels in an Asian Indian cohort. *J Cardiol* **64**:339–346.

Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, Marchini J, Martens JHA, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N, Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang S-Y, Wheeler E, Wilder SP, Iotchkova V, Moore C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D, Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martín D, Watt S, Yang Y, Guigo R, Beck S, Paul DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS, Soranzo N. 2016. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* **167**:1415–1429.e19.

Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, Suveges D, Vrousgou O, Whetzel PL, Amode R, Guillen JA, Riat HS, Trevanion SJ, Hall P, Junkins H, Flicek P, Burdett T, Hindorff LA, Cunningham F, Parkinson H. 2019. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*. doi:10.1093/nar/gky1120

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. doi:10.1186/s13742-015-0047-8

Chan IH, Privalsky ML. 2009. Thyroid hormone receptor mutants implicated in human hepatocellular carcinoma display an altered target gene repertoire. *Oncogene* **28**:4162–4174.

CHARGE-Heart Failure Consortium, EchoGen Consortium, METASTROKE Consortium, GIANT Consortium, EPIC-InterAct Consortium, Lifelines Cohort Study, Wellcome Trust Case Control Consortium, Understanding Society Scientific Group, EPIC-CVD Consortium, CHARGE+ Exome Chip Blood Pressure Consortium, T2D-GENES Consortium, GoT2DGenes Consortium, ExomeBP Consortium, CHD Exome+ Consortium, Surendran P, Drenos F, Young R, Warren H, Cook JP, Manning AK, Grarup N, Sim X, Barnes DR, Witkowska K, Staley JR, Tragante V, Tukiainen T, Yaghootkar H, Masca N, Freitag DF, Ferreira T, Giannakopoulou O, Tinker A, Harakalova M, Mihailov E, Liu C, Kraja AT, Nielsen SF, Rasheed A, Samuel M, Zhao W, Bonnycastle LL, Jackson AU, Narisu N, Swift AJ, Southam L, Marten J, Huyghe JR, Stančáková A, Fava C, Ohlsson T, Matchan A, Stirrups KE, Bork-Jensen J, Gjesing AP, Kontto J, Perola M, Shaw-Hawkins S, Havulinna AS, Zhang H, Donnelly LA, Groves CJ, Rayner NW, Neville MJ, Robertson NR, Yiorkas AM, Herzig K-H, Kajantie E, Zhang W, Willems SM, Lannfelt L, Malerba G, Soranzo N,

Trabetti E, Verweij N, Evangelou E, Moayyeri A, Vergnaud A-C, Nelson CP, Poveda A, Varga TV, Caslake M, de Craen AJM, Trompet S, Luan J, Scott RA, Harris SE, Liewald DCM, Marioni R, Menni C, Farmaki A-E, Hallmans G, Renström F, Huffman JE, Hassinen M, Burgess S, Vasan RS, Felix JF, Uria-Nickelsen M, Malarstig A, Reilly DF, Hoek M, Vogt TF, Lin H, Lieb W, Traylor M, Markus HS, Highland HM, Justice AE, Marouli E, Lindström J, Uusitupa M, Komulainen P, Lakka TA, Rauramaa R, Polasek O, Rudan I, Rolandsson O, Franks PW, Dedoussis G, Spector TD, Jousilahti P, Männistö S, Deary IJ, Starr JM, Langenberg C, Wareham NJ, Brown MJ, Dominiczak AF, Connell JM, Jukema JW, Sattar N, Ford I, Packard CJ, Esko T, Mägi R, Metspalu A, de Boer RA, van der Meer P, van der Harst P, Gambaro G, Ingelsson E, Lind L, de Bakker PIW, Numans ME, Brandslund I, Christensen C, Petersen ERB, Korpi-Hyövälti E, Oksa H, Chambers JC, Kooner JS, Blakemore AIF, Franks S, Jarvelin M-R, Husemoen LL, Linneberg A, Skaaby T, Thuesen B, Karpe F, Tuomilehto J, Doney ASF, Morris AD, Palmer CNA, Holmen OL, Hveem K, Willer CJ, Tuomi T, Groop L, Käräjämäki A, Palotie A, Ripatti S, Salomaa V, Alam DS, Majumder A al S, Di Angelantonio E, Chowdhury R, McCarthy MI, Poulter N, Stanton AV, Sever P, Amouyel P, Arveiler D, Blankenberg S, Ferrières J, Kee F, Kuulasmaa K, Müller-Nurasyid M, Veronesi G, Virtamo J, Deloukas P, Elliott P, Zeggini E, Kathiresan S, Melander O, Kuusisto J, Laakso M, Padmanabhan S, Porteous DJ, Hayward C, Scotland G, Collins FS, Mohlke KL, Hansen T, Pedersen O, Boehnke M, Stringham HM, Frossard P, Newton-Cheh C, Tobin MD, Nordestgaard BG, Caulfield MJ, Mahajan A, Morris AP, Tomaszewski M, Samani NJ, Saleheen D, Asselbergs FW, Lindgren CM, Danesh J, Wain LV, Butterworth AS, Howson JMM, Munroe PB. 2016. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nat Genet* **48**:1151–1161.

Church DM, Schneider VA, Graves T, Auger K, Cunningham F, Bouk N, Chen H-C, Agarwala R, McLaren WM, Ritchie GRS, Albracht D, Kremitzki M, Rock S, Kotkiewicz H, Kremitzki C, Wollam A, Trani L, Fulton L, Fulton R, Matthews L, Whitehead S, Chow W, Torrance J, Dunn M, Harden G, Threadgold G, Wood J, Collins J, Heath P, Griffiths G, Pelan S, Grafham D, E. Eichler E, Weinstock G, Mardis ER, Wilson RK, Howe K, Flicek P, Hubbard T. 2011. Modernizing Reference Genome Assemblies. doi:10.1371/journal.pbio.1001091

DeBoever C, Tanigawa Y, Lindholm ME, McInnes G, Lavertu A, Ingelsson E, Chang C, Ashley EA, Bustamante CD, Daly MJ, Rivas MA. 2018. Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. *Nat Commun* **9**:1066.

Divorty N, Mackenzie AE, Nicklin SA, Milligan G. 2015. G protein-coupled receptor 35: an emerging target in inflammatory and cardiovascular disease. *Front Pharmacol* **6**:19189.

Dong X, Han S, Zylka MJ, Simon MI, Anderson DJ. 2001. A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons. *Cell* **106**:619–632.

Dupre J, Ross SA, Watson D, Brown JC. 1973. STIMULATION OF INSULIN SECRETION BY GASTRIC INHIBITORY POLYPEPTIDE IN MAN. 1. *J Clin Endocrinol Metab* **37**:826–828.

Ellinghaus D, Folseraas T, Holm K, Ellinghaus E, Melum E, Balschun T, Laerdahl JK, Shiryaev A, Gotthardt DN, Weismüller TJ, Schramm C, Wittig M, Bergquist A, Björnsson E, Marschall H-U, Vatn M, Teufel A, Rust C, Gieger C, Wichmann H-E, Runz H, Sterneck M, Rupp C, Braun F, Weersma RK, Wijmenga C, Ponsioen CY, Mathew CG, Rutgeerts P, Vermeire S, Schrumpf E, Hov JR, Manns MP, Boberg KM, Schreiber S, Franke A, Karlsen TH. 2013. Genome-wide association analysis in Primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology* **58**:1074–1083.

Fortin J-P, Schroeder JC, Zhu Y, Beinborn M, Kopin AS. 2010. Pharmacological Characterization of Human Incretin Receptor Missense Variants. *J Pharmacol Exp Ther* **332**:274–280.

Gourraud P-A, Sdika M, Khankhanian P, Henry RG, Beheshtian A, Matthews PM, Hauser SL, Oksenberg JR, Pelletier D, Baranzini SE. 2013. A genome-wide association study of brain

lesion distribution in multiple sclerosis. *Brain* **136**:1012–1024.
Grallert H, Dupuis J, Bis JC, Dehghan A, Barbalic M, Baumert J, Lu C, Smith NL, Uitterlinden AG, Roberts R, Khuseyinova N, Schnabel RB, Rice KM, Rivadeneira F, Hoogeveen RC, Fontes JD, Meisinger C, Keaney JF Jr, Lemaitre R, Aulchenko YS, Vasan RS, Ellis S, Hazen SL, van Duijn CM, Nelson JJ, März W, Schunkert H, McPherson RM, Stirnadel-Farrant HA, Psaty BM, Gieger C, Siscovick D, Hofman A, Illig T, Cushman M, Yamamoto JF, Rotter JI, Larson MG, Stewart AFR, Boerwinkle E, Witteman JCM, Tracy RP, Koenig W, Benjamin EJ, Ballantyne CM. 2012. Eight genetic loci associated with variation in lipoprotein-associated phospholipase A2 mass and activity and coronary heart disease: meta-analysis of genome-wide association studies from five community-based studies. *Eur Heart J* **33**:238–251.
Hayes JE, Feeney EL, Nolden AA, McGeary JE. 2015. Quinine Bitterness and Grapefruit Liking Associate with Allelic Variants in TAS2R31. *Chem Senses* **40**:437–443.
Hayes JE, Wallace MR, Knopik VS, Herbstman DM, Bartoshuk LM, Duffy VB. 2011. Allelic Variation in TAS2R Bitter Receptor Genes Associates with Variation in Sensations from and Ingestive Behaviors toward Common Bitter Beverages in Adults. *Chem Senses* **36**:311–319.
Hill A, Loh P-R, Bharadwaj RB, Pons P, Shang J, Guinan E, Lakhani K, Kilty I, Jelinsky SA. 2017. Stepwise Distributed Open Innovation Contests for Software Development: Acceleration of Genome-Wide Association Analysis. *Gigascience* **6**:1–10.
Hoffmann TJ, Van Den Eeden SK, Sakoda LC, Jorgenson E, Habel LA, Graff RE, Passarelli MN, Cario CL, Emami NC, Chao CR, Ghai NR, Shan J, Ranatunga DK, Quesenberry CP, Aaronson D, Presti J, Wang Z, Berndt SI, Chanock SJ, McDonnell SK, French AJ, Schaid DJ, Thibodeau SN, Li Q, Freedman ML, Penney KL, Mucci LA, Haiman CA, Henderson BE, Seminara D, Kvale MN, -Y. Kwok P, Schaefer C, Risch N, Witte JS. 2015. A Large Multiethnic Genome-Wide Association Study of Prostate Cancer Identifies Novel Risk Variants and Substantial Ethnic Differences. *Cancer Discovery*. doi:10.1158/2159-8290.cd-15-0315
(igas) IG of ASC, International Genetics of Ankylosing Spondylitis Consortium (IGAS). 2013. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nature Genetics*. doi:10.1038/ng.2667
International Headache Genetics Consortium, Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, Farh K-H, Cuenca-Leon E, Muona M, Furlotte NA, Kurth T, Ingason A, McMahon G, Ligthart L, Terwindt GM, Kallela M, Freilinger TM, Ran C, Gordon SG, Stam AH, Steinberg S, Borck G, Koiranen M, Quaye L, Adams HHH, Lehtimäki T, Sarin A-P, Wedenoja J, Hinds DA, Buring JE, Schürks M, Ridker PM, Hrafnsdottir MG, Stefansson H, Ring SM, Hottenga J-J, Penninx BWJH, Färkkilä M, Artto V, Kaunisto M, Vepsäläinen S, Malik R, Heath AC, Madden PAF, Martin NG, Montgomery GW, Kurki MI, Kals M, Mägi R, Pärn K, Hämäläinen E, Huang H, Byrnes AE, Franke L, Huang J, Stergiakouli E, Lee PH, Sandor C, Webber C, Cader Z, Muller-Myhsok B, Schreiber S, Meitinger T, Eriksson JG, Salomaa V, Heikkilä K, Loehrer E, Uitterlinden AG, Hofman A, van Duijn CM, Cherkas L, Pedersen LM, Stubhaug A, Nielsen CS, Männikkö M, Mihailov E, Milani L, Göbel H, Esserlind A-L, Christensen AF, Hansen TF, Werge T, Kaprio J, Aromaa AJ, Raitakari O, Ikram MA, Spector T, Järvelin M-R, Metspalu A, Kubisch C, Strachan DP, Ferrari MD, Belin AC, Dichgans M, Wessman M, van den Maagdenberg AMJM, Zwart J-A, Boomsma DI, Smith GD, Stefansson K, Eriksson N, Daly MJ, Neale BM, Olesen J, Chasman DI, Nyholt DR, Palotie A. 2016. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet* **48**:856–866.
International Multiple Sclerosis Genetics Consortium, International IBD Genetics Consortium, Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L, Shah T, Abedian S, Cheon JH, Cho J, Daryani NE, Franke L, Fuyuno Y, Hart A,

Juyal RC, Juyal G, Kim WH, Morris AP, Poustchi H, Newman WG, Midha V, Orchard TR, Vahedi H, Sood A, Sung JJY, Malekzadeh R, Westra H-J, Yamazaki K, Yang S-K, Barrett JC, Franke A, Alizadeh BZ, Parkes M, B K T, Daly MJ, Kubo M, Anderson CA, Weersma RK. 2015. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* **47**:979–986.

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, The Genome Aggregation Database Consortium, Neale BM, Daly MJ, MacArthur DG. 2019. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *Genomics*.

Katritch V, Cherezov V, Stevens RC. 2013. Structure-function of the G protein-coupled receptor superfamily. *Annu Rev Pharmacol Toxicol* **53**:531–556.

Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund M, Schoech A, Pasaniuc B, Price AL. n.d. Leveraging polygenic functional enrichment to improve GWAS power. doi:10.1101/222265

Kim U-K. 2003. Positional Cloning of the Human Quantitative Trait Locus Underlying Taste Sensitivity to Phenylthiocarbamide. *Science* **299**:1221–1225.

Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, Karapetyan K, Katz K, Liu C, Maddipatla Z, Malheiro A, McDaniel K, Ovetsky M, Riley G, Zhou G, Holmes JB, Kattman BL, Maglott DR. 2018. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* **46**:D1062–D1067.

Li YR, Li J, Zhao SD, Bradfield JP, Mentch FD, Maggadottir SM, Hou C, Abrams DJ, Chang D, Gao F, Guo Y, Wei Z, Connolly JJ, Cardinale CJ, Bakay M, Glessner JT, Li D, Kao C, Thomas KA, Qiu H, Chiavacci RM, Kim CE, Wang F, Snyder J, Richie MD, Flatø B, Førre Ø, Denson LA, Thompson SD, Becker ML, Guthery SL, Latiano A, Perez E, Resnick E, Russell RK, Wilson DC, Silverberg MS, Annese V, Lie BA, Punaro M, Dubinsky MC, Monos DS, Strisciuglio C, Staiano A, Miele E, Kugathasan S, Ellis JA, Munro JE, Sullivan KE, Wise CA, Chapel H, Cunningham-Rundles C, Grant SFA, Orange JS, Sleiman PMA, Behrens EM, Griffiths AM, Satsangi J, Finkel TH, Keinan A, Prak ETL, Polychronakos C, Baldassano RN, Li H, Keating BJ, Hakonarson H. 2015. Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nat Med* **21**:1018–1027.

Machiela MJ, Chanock SJ. 2015. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants: Fig. 1. *Bioinformatics*. doi:10.1093/bioinformatics/btv402

Ma L, Yang J, Runesha HB, Tanaka T, Ferrucci L, Bandinelli S, Da Y. 2010. Genome-wide association analysis of total cholesterol and high-density lipoprotein cholesterol levels using the Framingham heart study data. *BMC Med Genet* **11**:55.

Manteniotis S, Lehmann R, Flegel C, Vogel F, Hofreuter A, Schreiner BSP, Altmüller J, Becker C, Schöbel N, Hatt H, Gisselmann G. 2013. Comprehensive RNA-Seq expression analysis of sensory ganglia with a focus on ion channels and GPCRs in Trigeminal ganglia. *PLoS One* **8**:e79523.

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl Variant Effect Predictor. *Genome Biol* **17**:S8.

Meng W, Adams MJ, Hebert HL, Deary IJ, McIntosh AM, Smith BH. 2018. A Genome-Wide

Association Study Finds Genetic Associations with Broadly-Defined Headache in UK Biobank (N = 223,773). *EBioMedicine* **28**:180–186.
Miyawaki K, Yamada Y, Yano H, Niwa H, Ban N, Ihara Y, Kubota A, Fujimoto S, Kajikawa M, Kuroe A, Tsuda K, Hashimoto H, Yamashita T, Jomori T, Tashiro F, -i. Miyazaki J, Seino Y. 1999. Glucose intolerance caused by a defect in the entero-insular axis: A study in gastric inhibitory polypeptide receptor knockout mice. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.96.26.14843
Moore CAC, Milano SK, Benovic JL. 2007. Regulation of Receptor Trafficking by GRKs and Arrestins. *Annu Rev Physiol* **69**:451–482.
Musunuru K, Kathiresan S. 2019. Genetics of Common, Complex Coronary Artery Disease. *Cell* **177**:132–145.
Pándy-Szekeres G, Munk C, Tsonkov TM, Mordalski S, Harpsøe K, Hauser AS, Bojarski AJ, Gloriam DE. 2018. GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Res* **46**:D440–D446.
Pinato DJ, Mauri FA. 2014. Metastatic Spread of Lung Cancer to Brain and LiverBrain Metastases from Primary Tumors. pp. 123–129.
Postmus I, Trompet S, Deshmukh HA, Barnes MR, Li X, Warren HR, Chasman DI, Zhou K, Arsenault BJ, Donnelly LA, Wiggins KL, Avery CL, Griffin P, Feng Q, Taylor KD, Li G, Evans DS, Smith AV, de Keyser CE, Johnson AD, de Craen AJM, Stott DJ, Buckley BM, Ford I, Westendorp RGJ, Slagboom PE, Sattar N, Munroe PB, Sever P, Poulter N, Stanton A, Shields DC, O'Brien E, Shaw-Hawkins S, Chen Y-DI, Nickerson DA, Smith JD, Dubé MP, Boekholdt SM, Hovingh GK, Kastelein JJP, McKeigue PM, Betteridge J, Neil A, Durrington PN, Doney A, Carr F, Morris A, McCarthy MI, Groop L, Ahlqvist E, Welcome Trust Case Control Consortium, Bis JC, Rice K, Smith NL, Lumley T, Whitsel EA, Stürmer T, Boerwinkle E, Ngwa JS, O'Donnell CJ, Vasan RS, Wei W-Q, Wilke RA, Liu C-T, Sun F, Guo X, Heckbert SR, Post W, Sotoodehnia N, Arnold AM, Stafford JM, Ding J, Herrington DM, Kritchevsky SB, Eiriksdottir G, Launer LJ, Harris TB, Chu AY, Giulianini F, MacFadyen JG, Barratt BJ, Nyberg F, Stricker BH, Uitterlinden AG, Hofman A, Rivadeneira F, Emilsson V, Franco OH, Ridker PM, Gudnason V, Liu Y, Denny JC, Ballantyne CM, Rotter JI, Adrienne Cupples L, Psaty BM, Palmer CNA, Tardif J-C, Colhoun HM, Hitman G, Krauss RM, Wouter Jukema J, Caulfield MJ. 2014. Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins. *Nat Commun* **5**:5068.
Pujol-Borrell R, Giménez-Barcons M, Marín-Sánchez A, Colobran R. 2015. Genetics of Graves' Disease: Special Focus on the Role of TSHR Gene. *Horm Metab Res* **47**:753–766.
Raffield LM, Cox AJ, Hugenschmidt CE, Freedman BI, Langefeld CD, Williamson JD, Hsu F-C, Maldjian JA, Bowden DW. 2015. Heritability and genetic association analysis of neuroimaging measures in the Diabetes Heart Study. *Neurobiol Aging* **36**:1602.e7–15.
Risso DS, Mezzavilla M, Pagani L, Robino A, Morini G, Tofanelli S, Carrai M, Campa D, Barale R, Caradonna F, Gasparini P, Luiselli D, Wooding S, Drayna D. 2016. Global diversity in the TAS2R38 bitter taste receptor: revisiting a classic evolutionary PROPosal. *Sci Rep* **6**:115.
Rosen MD, Privalsky ML. 2011. Thyroid hormone receptor mutations in cancer and resistance to thyroid hormone: perspective and prognosis. *J Thyroid Res* **2011**:361304.
Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. 2015. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**:e1001779.
Takata R, Akamatsu S, Kubo M, Takahashi A, Hosono N, Kawaguchi T, Tsunoda T, Inazawa J, Kamatani N, Ogawa O, Fujioka T, Nakamura Y, Nakagawa H. 2010. Genome-wide association study identifies five new susceptibility loci for prostate cancer in the Japanese population. *Nat Genet* **42**:751–754.

Tang X-L, Wang Y, Li D-L, Luo J, Liu M-Y. 2012. Orphan G protein-coupled receptors (GPCRs): biological functions and potential drug targets. *Acta Pharmacologica Sinica*. doi:10.1038/aps.2011.210

Tanigawa Y, Li J, Justesen JM, Horn H, Aguirre M, DeBoever C, Chang C, Narasimhan B, Lage K, Hastie T, Park CY, Bejerano G, Ingelsson E, Rivas MA. 2019. Components of genetic associations across 2,138 phenotypes in the UK Biobank highlight adipocyte biology. *Nat Commun* **10**:1001.

Thanawala VJ, Forkuo GS, Al-Sawalha N, Azzegagh Z, Nguyen LP, Eriksen JL, Tuvim MJ, Lowder TW, Dickey BF, Knoll BJ, Walker JKL, Bond RA. 2013. β 2 -Adrenoceptor Agonists Are Required for Development of the Asthma Phenotype in a Murine Model. *Am J Respir Cell Mol Biol* **48**:220–229.

The International IBD Genetics Consortium (IIBDGC), International Genetics of Ankylosing Spondylitis Consortium (IGAS), International PSC Study Group (IPSCSG), Genetic Analysis of Psoriasis Consortium (GAPC), Psoriasis Association Genetics Extension (PAGE), Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, Park YR, Raychaudhuri S, Pouget JG, Hübenthal M, Folseraas T, Wang Y, Esko T, Metspalu A, Westra H-J, Franke L, Pers TH, Weersma RK, Collij V, D'Amato M, Halfvarson J, Jensen AB, Lieb W, Degenhardt F, Forstner AJ, Hofmann A, Schreiber S, Mrowietz U, Juran BD, Lazaridis KN, Brunak S, Dale AM, Trembath RC, Weidinger S, Weichenthal M, Ellinghaus E, Elder JT, Barker JNWN, Andreassen OA, McGovern DP, Karlsen TH, Barrett JC, Parkes M, Brown MA, Franke A. 2016. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat Genet* **48**:510–518.

The International IBD Genetics Consortium (IIBDGC), Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Philip Schumm L, Sharma Y, Anderson CA, Essers J, Mitrovic M, Ning K, Cleynen I, Theatre E, Spain SL, Raychaudhuri S, Goyette P, Wei Z, Abraham C, Achkar J-P, Ahmad T, Amininejad L, Ananthakrishnan AN, Andersen V, Andrews JM, Baidoo L, Balschun T, Bampton PA, Bitton A, Boucher G, Brand S, Büning C, Cohain A, Cichon S, D'Amato M, De Jong D, Devaney KL, Dubinsky M, Edwards C, Ellinghaus D, Ferguson LR, Franchimont D, Fransen K, Gearry R, Georges M, Gieger C, Glas J, Haritunians T, Hart A, Hawkey C, Hedl M, Hu X, Karlsen TH, Kupcinskas L, Kugathasan S, Latiano A, Laukens D, Lawrance IC, Lees CW, Louis E, Mahy G, Mansfield J, Morgan AR, Mowat C, Newman W, Palmieri O, Ponsioen CY, Potocnik U, Prescott NJ, Regueiro M, Rotter JI, Russell RK, Sanderson JD, Sans M, Satsangi J, Schreiber S, Simms LA, Sventoraityte J, Targan SR, Taylor KD, Tremelling M, Verspaget HW, De Vos M, Wijmenga C, Wilson DC, Winkelmann J, Xavier RJ, Zeissig S, Zhang B, Zhang CK, Zhao H, Silverberg MS, Annese V, Hakonarson H, Brant SR, Radford-Smith G, Mathew CG, Rioux JD, Schadt EE, Daly MJ, Franke A, Parkes M, Vermeire S, Barrett JC, Cho JH. 2012. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**:119–124.

van der Harst P, Verweij N. 2018. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res* **122**:433–443.

Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Madan Babu M. 2013. Molecular signatures of G-protein-coupled receptors. *Nature*. doi:10.1038/nature11896

Venkateswaran S, Prince J, Cutler DJ, Marigorta UM, Okou DT, Prahalad S, Mack D, Boyle B, Walters T, Griffiths A, Sauer CG, LeLeiko N, Keljo D, Markowitz J, Baker SS, Rosh J, Pfefferkorn M, Heyman MB, Patel A, Otley A, Baldassano R, Noe J, Rufo P, Oliva-Hemker M, Davis S, Zwick ME, Gibson G, Denson LA, Hyams J, Kugathasan S. 2018. Enhanced Contribution of HLA in Pediatric Onset Ulcerative Colitis. *Inflamm Bowel Dis* **24**:829–838.

Wang M, Takahashi A, Liu F, Ye D, Ding Q, Qin C, Yin C, Zhang Z, Matsuda K, Kubo M, Na R, Lin X, Jiang H, Ren S, Sun J, Zheng SL, Le Marchand L, Isaacs WB, Mo Z, Haiman CA,

Sun Y, Nakagawa H, Xu J. 2015. Large-scale association analysis in Asians identifies new susceptibility loci for prostate cancer. *Nat Commun* **6**:8469.
Yekutieli D, Benjamini Y. 2001. under dependency. *Ann Stat* **29**:1165–1188.
Zhang L, Taylor N, Xie Y, Ford R, Johnson J, Paulsen JE, Bates B. 2005. Cloning and expression of MRG receptors in macaque, mouse, and human. *Brain Res Mol Brain Res* **133**:187–197.