

Supplementary Appendix
for
Expression quantitative trait locus mapping in pulmonary arterial
hypertension

Anna Ulrich *et al.*

RNA sequencing and transcript abundance estimation	2
eQTL validation procedure.....	2
eQTL studies used for validation.....	3
Supplementary Tables	4

RNA sequencing and transcript abundance estimation

Whole blood (3ml) was collected in Tempus™ Blood RNA Tubes, which were stored at -80 °C until required. RNA was extracted using a Maxwell robotic system (Promega). Samples with a 260/230 ratio >1.5 and a 260/280 ratio in the range 1.9-2.1 were further quality checked by Bioanalyser and those achieving a minimum RNA Integrity Number (RIN) of 7 were submitted for sequencing. Globin-Zero Gold rRNA Removal Kits (Illumina Inc, San Diego, CA) were used to remove ribosomal RNA contamination from whole blood RNA samples. 75bp paired-end sequencing on a Hiseq4000 was performed on pooled libraries of ~80 samples.

Fastq files (raw reads from RNaseq) were analysed using Salmon v0.9.1 (Patro et al., 2017) and GENCODE release 28 to produce transcript abundance estimates which were converted to gene expression data using tximport in R (Soneson et al., 2015). Salmon, the first transcriptome-wide quantifier to correct for fragment GC-content bias was used, which substantially improves the accuracy of abundance estimates and the sensitivity of subsequent differential expression analysis (Patro et al., 2017).

eQTL validation procedure

To assess the extent to which expression quantitative trait locus (eQTL) - transcript pairs in PAH overlap with previously reported eQTL-transcript pairs described in healthy populations, we calculated the validation rate of our findings in the two largest published eQTL studies to date and the Genotype-Tissue Expression (GTEx) Project (Westra et al., 2013, Joehanes et al., 2017, Aguet et al., 2019). Validation rate was defined as the number of significant eQTL-transcript pairs in this study confirmed by the published study divided by the total number of significant eQTL-transcript pairs that were tested by the published study and multiplied by one hundred.

We extracted all eQTL with effects below the study specific significance threshold from the published studies' results for all significant transcripts in this study. Ensembl identifiers were used for the transcripts and genomic coordinates (chromosome and base pair position) on the Genome Reference Consortium Human Build 37 for eQTL when matching my eQTL-transcript pairs to those published by the external studies.

An eQTL-transcript pair was considered confirmed if the lead variant was in linkage disequilibrium ($r^2 \geq 40\%$) with the lead eQTL of the same transcript in the published study. For this purpose, a list of correlated variants was obtained from the European population of the 1000 Genomes Project (Durbin et al., 2010) - using the R package 'proxysnps' - for each lead eQTL reported by the studies used for validation. Each eQTL-transcript pair that reached the study-specific significance threshold in at least one of the studies was considered confirmed.

Since the complete list of variants or transcripts that passed study-specific quality control is not usually made available by published studies, we had to make assumptions when determining the total number of my significant eQTL-transcript pairs tested in the external studies. This did not apply to GTEx results where all tested eQTL-transcript pairs could be retrieved.

All transcripts present on the expression array used in the other two studies were assumed to have been tested. Annotation files were downloaded from the manufacturer's website for the complete list of transcripts present on the expression arrays. Additionally, all eQTL in this study were assumed

to have been available for analysis in the studies used for validation which used genotyping array data imputed to high-density reference panels. We restricted our analyses to common variants with a minimum minor allele frequency of 5% which is at least as high as those of the published eQTL studies used for validation.

eQTL studies used for validation

We selected two of the largest published eQTL studies and the GTEx Project to compare our results to. Westra et al. (Westra et al., 2013) meta-analysed eQTL effects from seven studies totaling 5,311 individuals to identify cis-acting effects genome-wide as well as trans-acting effects of 4,542 variants implicated in diseases and traits from the GWAS Catalog (Buniello et al., 2019) at the time of their study. The other eQTL study used for validation was published by Joehanes et al. (Joehanes et al., 2017) who conducted the largest to-date single cohort transcriptome-wide analysis testing both cis- and trans-acting elements genome-wide in the whole blood samples of 5,257 individuals from the Framingham Heart Study. The GTEx Project aims to create a database of genotype and gene expression correlations in multiple human tissues of consenting donors to aid the scientific community in understanding inherited disease susceptibility. The GTEx website provides a browser (<https://gtexportal.org/home/testyourown>) for retrieving results for any given variant and transcript pair based on Ensembl identifiers in the tissue of interest.

The current release (V8) of the GTEx Project has 838 post-mortem donor samples with genetic data of which 670 contributed to the eQTL mapping in whole blood (GTExPortal, Aguet et al., 2019). Only tissue samples that passed histological examination were accepted for the project, however tissue exclusions were not made based on cause of death. Demographics and cause of death statistics can be found on the GTEx Portal (<https://gtexportal.org/>).

Westra et al. meta-analysed seven studies that measured gene expression in peripheral blood on one of Illumina's whole-genome Expression BeadChips (HT12v3, HT12v4 or H8v2 arrays). Since most of the cohorts used the HT12v3 array, the analyses were restricted to transcripts present on this array (Westra et al., 2013). Joehanes et al. used the Affymetrix Human Exon ST 1.0 array for the whole cohort (Joehanes et al., 2017). The complete list of transcripts was obtained from the manifest files, namely HumanHt-12_V3_0_R3_11283641_A.bgx for Illumina and HuEx-1_0-st-v2.na33.1.hg19.probeset.csv for Affymetrix, available on the manufacturers' websites.

The genotype data from Westra et al. and Joehanes et al. were imputed to the largest haplotype reference panels available at the time of conducting their analyses. Westra et al. mapped cis-eQTL differently from our and the other two studies' approach, using a 250 kilobases (kb) maximum distance from the probe midpoint to demarcate cis effects, while eQTLs with a distance greater than 5 Mb were defined as trans-eQTLs. The validation rate of our trans eQTL was not assessed using Westra's trans eQTL results as they have only tested a small set of selected variants. Also, cis-eQTL in this study were restricted to the ones not farther than 250 kB from the transcript's TSS when validating cis-eQTL with Westra's results.

Supplementary Tables

Supplementary Table 1 Patient characteristics and white blood cell fractions in the PAH Cohort study.....	5
Supplementary Table 2 Lead variants for novel eQTL in the PAH Cohort ordered by variance explained in gene expression.....	6
Supplementary Table 3 Phenotypes from the NHGRI-EBI GWAS Catalog of published genome-wide association studies reported for novel and confirmed eQTL.	21
Supplementary Table 4 Significant KEGG pathways from the enrichment analysis of genes with novel eQTL in the PAH Cohort.	27

Characteristic	Males	Females
N (%)	82 (30)	194 (70)
Age at diagnosis, yr	45.6 (24.5)	47.7 (22.1)
Age at sampling, yr	54.4 (22.8)	52.9 (21.8)
B cells	0.028 (0.014)	0.028 (0.016)
Alternatively activated macrophages	0.022 (0.008)	0.020 (0.011)
Neutrophils	0.265 (0.098)	0.285 (0.100)
Natural killer cells	0.020 (0.017)	0.020 (0.020)
CD8+ T cells	0.050 (0.057)	0.050 (0.038)
Regulatory CD4+ T cells	0.150 (0.128)	0.160 (0.128)
Other uncharacterized cells	0.397 (0.117)	0.392 (0.141)

Supplementary Table 1 Patient characteristics and white blood cell fractions in the PAH Cohort study.
Age and white blood cell fractions are presented as median (interquartile range).

Supplementary Table 2 Lead variants for novel eQTL in the PAH Cohort ordered by variance explained in gene expression.

Column names: eQTL rs ID = dbSNP (150) variant identifier; eQTL chr:pos = chromosome and genomic position (Ensembl GRCh37 coordinates); ref = reference allele in the human genome reference GRCh37; alt = alternative allele in the human genome reference GRCh37; Ensembl gene ID = Ensembl stable gene identifiers; variant type = dbSNP version 150 variant annotation; R² = percent (%) variance explained in gene expression by the genetic variant; Gene type: PC = protein coding, TC- = transcribed, PPG = processed pseudogene, unPPG = unprocessed pseudogene, lncRNA = long non-coding RNA, TR V gene = T-cell receptor variable gene, IG V gene = immunoglobulin variable gene, snRNA = small nuclear RNA, snoRNA = small nucleolar RNA, miRNA = micro RNA.

eQTL rs ID	eQTL chr:pos	ref	alt	Gene symbol	Ensembl gene ID	cis/trans	eQTL P-value	R ²	Gene type
rs2295768	6:2954602	C	T	HTATSF1P2	ENSG00000271361	cis	7.1x10-83	57.4	PPG
rs2240101	15:31115709	G	C	RP5.1086D14.5	ENSG00000259845	cis	4.5x10-79	56.2	unPPG
rs2302559	7:105903904	T	C	NAMPTL	ENSG00000229644	trans	6.35x10-78	55.9	PPG
rs745666	17:72744798	G	C	RP11.452I5.2	ENSG00000266036	cis	1.9x10-70	53.3	PC
rs7635	11:721570	C	T	RPLP2	ENSG00000177600	cis	2.05x10-70	53.3	lncRNA
rs471521	12:10115647	A	T	CLEC12B	ENSG00000256660	cis	1.94x10-68	52.5	PC
rs67933574	12:11226237	A	G	TAS2R43	ENSG00000255374	cis	8.58x10-65	51.1	PC
rs11079015	17:39975164	C	T	NT5C3B	ENSG00000141698	cis	2.28x10-64	51	PC
rs2280370	16:89627460	T	G	RPL13P12	ENSG00000215030	trans	1.72x10-62	50.2	TCunPPG
rs9892882	17:18501596	A	G	USP32P1	ENSG00000188933	trans	2.03x10-62	50.2	PPG
rs9892882	17:18501596	A	G	CCDC144A	ENSG00000170160	trans	3.71x10-62	50.1	PC
rs72785089	16:70005967	A	G	CLEC18A	ENSG00000157322	cis	1.14x10-61	49.9	PC
rs10947675	6:37479972	G	A	RP1.153P14.8	ENSG00000204110	cis	4.66x10-60	49.2	lncRNA
rs1455979	21:44524006	G	A	U2AF1	ENSG00000160201	cis	2.5x10-59	48.9	PC
rs12975591	19:10627814	A	G	CTC.539A10.4	ENSG00000267100	cis	1.6x10-58	48.5	lncRNA
rs10790962	11:128385169	G	A	Sep-03	ENSG00000100167	trans	4.37x10-57	47.9	PC
rs12225372	11:74604594	A	G	XRRNA1	ENSG00000166435	cis	4.79x10-56	47.4	PC
rs2732603	17:44367288	G	C	RP11.197O15.1	ENSG00000214401	cis	5.07x10-55	46.9	lncRNA
rs79240608	1:149585945	T	C	NBPF26	ENSG00000273136	trans	6.4x10-55	46.9	PC
rs2336136	14:74030518	T	A	ACOT1	ENSG00000184227	cis	1.12x10-54	46.8	PC
rs2286466	16:2014283	A	G	RPS2P5	ENSG00000240342	trans	4.46x10-54	46.5	PPG
rs2656417	17:18475630	T	C	USP32P2	ENSG00000233327	cis	7.58x10-54	46.4	TCunPPG
rs55969088	18:9024224	G	T	NDUFV2P1	ENSG00000267809	trans	7.55x10-54	46.4	PPG
rs9264607	6:31237605	G	A	HLA.L	ENSG00000243753	trans	2.13x10-53	46.2	TCunPPG
rs61744075	8:23288346	C	A	NKX3.1	ENSG00000167034	cis	1.92x10-52	45.7	PC
rs17847967	12:8082367	T	C	SLC2A14	ENSG00000173262	cis	2.4x10-52	45.6	PC
rs11677238	2:114424974	T	G	AC002055.4	ENSG00000184319	trans	4.55x10-52	45.5	TCunPPG
rs73287153	7:12139272	A	G	VWDE	ENSG00000146530	cis	7.37x10-52	45.4	PC
rs216257	6:24934664	A	T	LSAMP	ENSG00000185565	trans	1x10-50	44.8	PC
rs10489762	1:183570545	G	T	CTD.2058B24.2	ENSG00000259048	trans	1.84x10-50	44.7	lncRNA
rs2732603	17:44367288	G	C	RP11.583F2.1	ENSG00000264057	trans	2.02x10-50	44.7	TCunPPG
rs111839387	7:65285169	C	T	PSPH	ENSG00000146733	trans	5.95x10-50	44.4	PC
rs12447763	16:23753594	G	T	RP11.151E14.1	ENSG00000261035	trans	8.42x10-50	44.4	lncRNA
rs9934565	16:87764267	A	T	RP11.278A23.2	ENSG00000260671	cis	1.73x10-49	44.2	lncRNA
rs2732617	17:44363632	A	C	RP11.359G22.2	ENSG00000230400	trans	4.78x10-49	44	lncRNA

eQTL rs ID	eQTL chr:pos	ref	alt	Gene symbol	Ensembl gene ID	cis/trans	eQTL P-value	R ²	Gene type
rs61809245	1:148203960	G	A	<i>NBPF20</i>	ENSG00000162825	trans	8.88x10-13	15.6	PC
rs62364993	5:80514325	G	C	<i>CKMT2</i>	ENSG00000131730	cis	8.58x10-13	15.6	PC
rs6488334	12:11140444	T	C	<i>TAS2R30</i>	ENSG00000256188	cis	9.02x10-13	15.6	lncRNA
rs72905231	2:191303525	A	G	<i>NEMP2</i>	ENSG00000189362	cis	9.05x10-13	15.6	PC
rs77391378	21:29947914	T	A	<i>AF131217.1</i>	ENSG00000232855	cis	9.87x10-13	15.6	lncRNA
rs776296	7:63804337	T	A	<i>ZNF736</i>	ENSG00000234444	cis	8.84x10-13	15.6	PC
rs1233578	6:28712247	A	G	<i>BTN3A2</i>	ENSG00000186470	trans	1.14x10-12	15.5	PC
rs2697291	2:198111312	A	G	<i>RP11.794P6.1</i>	ENSG00000255428	trans	1.12x10-12	15.5	PC
rs61979250	14:76668905	C	G	<i>GPATCH2L</i>	ENSG00000089916	cis	1.09x10-12	15.5	PC
rs34019730	16:89812930	G	A	<i>FANCA</i>	ENSG00000187741	cis	1.24x10-12	15.4	lncRNA
rs35084227	2:87277202	T	C	<i>PLG</i>	ENSG00000122194	trans	1.25x10-12	15.4	PC
rs35785032	4:175697250	G	A	<i>ADAM29</i>	ENSG00000168594	cis	1.49x10-12	15.4	PC
rs372161775	5:175398456	A	G	<i>THOC3</i>	ENSG00000051596	cis	1.29x10-12	15.4	PC
rs41265179	1:153601828	G	T	<i>NUP210L</i>	ENSG00000143552	cis	1.4x10-12	15.4	PC
rs10584	15:90328551	G	C	<i>ANPEP</i>	ENSG00000166825	cis	1.72x10-12	15.3	PC
rs11246329	11:843713	C	T	<i>TSPAN4</i>	ENSG00000214063	cis	1.72x10-12	15.3	PC
rs12812329	12:9167788	T	C	<i>RP11.436I9.3</i>	ENSG00000255776	cis	1.54x10-12	15.3	PC
rs2684556	17:44580533	A	G	<i>RP11.156P1.1</i>	ENSG00000263142	cis	1.56x10-12	15.3	PC
rs3807020	6:41747048	G	C	<i>OR56A1</i>	ENSG00000180934	trans	1.62x10-12	15.3	PC
rs56175172	11:77245728	A	G	<i>AP000580.1</i>	ENSG00000219529	cis	1.51x10-12	15.3	PPG
rs61945858	12:120879461	A	G	<i>DYNLL1</i>	ENSG00000088986	cis	1.62x10-12	15.3	PPG
rs62414240	6:76274825	C	T	<i>RNU6.1016P</i>	ENSG00000252498	cis	1.72x10-12	15.3	snRNA
rs7448600	5:21365623	C	A	<i>RP11.823P9.3</i>	ENSG00000233974	cis	1.82x10-12	15.3	PPG
rs10435105	7:3330994	C	T	<i>SDK1</i>	ENSG00000146555	cis	1.97x10-12	15.2	TCunPPG
rs112161250	15:58440089	G	GCA	<i>RP1</i>	ENSG00000104237	trans	1.94x10-12	15.2	PC
rs9567377	13:44720217	G	T	<i>SMIM2.AS1</i>	ENSG00000227258	cis	2x10-12	15.2	PC

Supplementary Table 3 Phenotypes from the NHGRI-EBI GWAS Catalog of published genome-wide association studies reported for novel and confirmed eQTL.

The full list of phenotypes reported for eQTL in the PAH Cohort study including sixty-three eQTL reported for at least one lung-related phenotype in the GWAS Catalog are listed in the column ‘GWAS Catalog phenotypes’. The column ‘confirmation’ indicates if the eQTL was found in the population-based studies used for confirmation.

eQTL rs ID	Gene symbol	confirmation	GWAS Catalog phenotypes
rs10790962	<i>SEPT3</i>	novel	Glaucoma; Glaucoma (multi-trait analysis); Intraocular pressure; Lung function (FVC); Psoriasis; Multiple sclerosis; Psoriasis vulgaris; Celiac disease; Lymphocyte percentage of white cells
rs7947900	<i>IFITM2</i>	novel	Eosinophil counts; Red blood cell count; Lung function (FEV1/FVC); White blood cell count; Granulocyte percentage of myeloid white cells; Monocyte percentage of white cells; Monocyte count; Neutrophil percentage of white cells; Lymphocyte percentage of white cells; Myeloid white cell count; White blood cell count (basophil); Sum neutrophil eosinophil counts; Granulocyte count; Neutrophil count; Sum basophil neutrophil counts; Medication use (thyroid preparations)
rs1451659	<i>OR51R1P</i>	novel	Lung cancer
rs2008345	<i>RP11.298I3.1</i>	novel	Lung function (FEV1/FVC); Household income (MTAG); Cognitive ability; years of educational attainment or schizophrenia (pleiotropy); Cognitive ability (MTAG); Educational attainment (years of education); Educational attainment (MTAG); Highest math class taken; Cognitive performance (MTAG); Self-reported math ability; Highest math class taken (MTAG)
rs10136164	<i>PAX9</i>	novel	Lung function (FVC)
rs72785089	<i>CLEC18A</i>	novel	Lung function (FVC)
rs8058597	<i>NPIP15</i>	novel	Lung function (FVC)
rs111164082	<i>RP11.798G7.3</i>	novel	Interstitial lung disease; Progressive supranuclear palsy; Risk-taking tendency (4-domain principal component model); Depressed affect; Eosinophil counts; Worry; Neuroticism; Number of sexual partners; Red blood cell count; General cognitive ability; Lung function (FVC); Male-pattern baldness; Alcohol consumption; Ease of getting up in the morning; Daytime nap; White matter microstructure (fractional anisotropy); Parkinson's disease; Alcohol use disorder (total score); Alcohol use disorder (consumption score); Osteoarthritis (hip); Sleep duration; Automobile speeding propensity; Brain region volumes; General factor of neuroticism; Primary biliary cirrhosis; Neuroticism; Irritable mood; Feeling guilty; Feeling miserable; Experiencing mood swings; Worry too long after an embarrassing experience; Feeling hurt; Feeling worry; Type 1 diabetes; Waist-hip ratio
rs3785884	<i>LRRC37A2</i>	novel	Interstitial lung disease; Progressive supranuclear palsy; Risk-taking tendency (4-domain principal component model); Depressed affect; Eosinophil counts; Worry; Neuroticism; Number of sexual partners; Red blood cell count; General cognitive ability; Lung function (FVC); Household income; Male-pattern baldness; Alcohol consumption; Ease of getting up in the morning; Daytime nap; White matter microstructure (fractional anisotropy); Parkinson's disease; Alcohol use disorder (total score); Alcohol use disorder (consumption score); Osteoarthritis (hip); Sleep duration; Automobile speeding propensity; Brain region volumes; General factor of neuroticism; Neuroticism; Irritable mood; Feeling guilty; Feeling miserable; Experiencing mood swings; Worry

eQTL rs ID	Gene symbol	confirmation	GWAS Catalog phenotypes
			too long after an embarrassing experience; Feeling hurt; Feeling worry; Type 1 diabetes; Waist-hip ratio; Medication use (anilides)
rs112534607	AC126544.4	novel	Lung function in never smokers (low FEV1 vs high FEV1); Loneliness; White matter microstructure (radial diusivities); White matter microstructure (axial diusivities); White matter microstructure (mean diusivities); White matter microstructure (fractional anisotropy); General cognitive ability; Neuroticism; Feeling guilty; Feeling miserable; Experiencing mood swings; Red cell distribution width; Mood instability; Breast cancer
rs2732617	RP11.359G22.2	novel	Lung function in never smokers (low FEV1 vs high FEV1); Loneliness; White matter microstructure (radial diusivities); White matter microstructure (axial diusivities); White matter microstructure (mean diusivities); White matter microstructure (fractional anisotropy); General cognitive ability; Neuroticism; Feeling guilty; Feeling miserable; Experiencing mood swings; Red cell distribution width; Mood instability; Breast cancer
rs2732603	RP11.583F2.1	novel	Lung function in never smokers (low FEV1 vs high FEV1); Loneliness; White matter microstructure (radial diusivities); White matter microstructure (axial diusivities); White matter microstructure (mean diusivities); White matter microstructure (fractional anisotropy); General cognitive ability; Neuroticism; Feeling guilty; Feeling miserable; Experiencing mood swings; Red cell distribution width; Mood instability; Breast cancer
rs8073963	HOXB.AS1	novel	Lung function (FEV1/FVC); Primary tooth development (number of teeth)
rs35858034	CYP2T2P	novel	Smoking behavior; Chronic obstructive pulmonary disease; Smoking behaviour (cigarettes smoked per day); Blood protein levels; Post bronchodilator FEV1/FVC ratio; Emphysema imaging phenotypes; Post bronchodilator FEV1; Hemoglobin concentration
rs2353408	U2SURP	novel	Red cell distribution width; Lung function (FEV1/FVC); Urate levels
rs9842198	PA2G4P4	novel	Sunburns; Interleukin-2 levels; Nose size; Breast cancer; Chronic obstructive pulmonary disease or resting heart rate (pleiotropy)
rs10079553	RUFY1	novel	Lung cancer
rs6925979	ADGRG6	novel	Height; Scoliosis; Body fat distribution (trunk fat ratio); Body fat distribution (leg fat ratio); Body fat distribution (arm fat ratio); Heel bone mineral density; Chronic obstructive pulmonary disease; Adolescent idiopathic scoliosis; Waist circumference adjusted for body mass index; Waist circumference adjusted for BMI in non-smokers; Waist circumference adjusted for BMI in smokers; Waist circumference adjusted for BMI (adjusted for smoking behaviour); Waist circumference adjusted for BMI (joint analysis main effects and smoking interaction); Chronic obstructive pulmonary disease or high blood pressure (pleiotropy)
rs2769345	RP11.514O12.4	novel	Crohn's disease; Vitiligo; Autoimmune traits; Hypothyroidism; Graves' disease; Inflammatory bowel disease; Lung cancer; Lung adenocarcinoma; Basal cell carcinoma; Medication use (thyroid preparations)
rs9272779	HLA.DQA1	novel	Schizophrenia; Lung function in never smokers (low FEV1 vs high FEV1); Self-reported allergy; Composite immunoglobulin trait (IgA x IgG x IgM); Type 2 diabetes; Childhood steroid-sensitive nephrotic syndrome; Asthma and hay fever; Allergic sensitization; Epstein Barr virus nuclear antigen 1 IgG levels; Epstein Barr virus nuclear antigen 1 IgG seropositivity; Epstein-Barr virus immune response (multivariate analysis); Autism spectrum disorder or schizophrenia; Asthma or allergic disease (pleiotropy); Asthma; Hypothyroidism; Ulcerative colitis
rs9273495	HLA.DQB2	novel	Asthma; Ulcerative colitis; Chronic lymphocytic leukemia; Type 1 diabetes; Type 2 diabetes; Eosinophilic granulomatosis with polyangiitis; Asthma and hay fever; Epstein Barr virus nuclear antigen 1 IgG levels; Asthma or allergic disease (pleiotropy); Blood protein levels; Lung function (FEV1); Lung function (FEV1/FVC); Inflammatory

eQTL rs ID	Gene symbol	confirmation	GWAS Catalog phenotypes
			bowel disease; Medication use (drugs used in diabetes)
rs9273841	<i>XXbac.BPG254F23.5</i>	novel	Immunoglobulin A; Multiple sclerosis; Ulcerative colitis or Crohn's disease; Systemic lupus erythematosus; Lymphoma; Ulcerative colitis; Multiple sclerosis (OCB status); Appendicular lean mass; FEV1; Lung function (FEV1/FVC); Peak expiratory flow; Follicular lymphoma; Hypothyroidism; Blood protein levels; Hepatitis C (spontaneous viral clearance); Random C-peptide levels in type I diabetes; C-peptide levels in type I diabetes; Hepatocellular carcinoma in hepatitis C infection; Urinary tract infection frequency; Hepatitis B; Itch intensity from mosquito bite adjusted by bite size; Mosquito bite size; Perceived unattractiveness to mosquitoes; Itch intensity from mosquito bite; Crohn's disease; IgA nephropathy; Lupus nephritis in systemic lupus erythematosus; C-reactive protein levels or total cholesterol levels (pleiotropy); Oral cavity and pharyngeal cancer; Oropharynx cancer; Oral cavity cancer
rs13238996	<i>GTF2IRD2B</i>	novel	Red cell distribution width; Cardiovascular disease; Lung function (FEV1/FVC); Body mass index; Menarche (age at onset); Chronic obstructive pulmonary disease or high blood pressure (pleiotropy); Diastolic blood pressure
rs7543453	<i>RP11.305E17.6</i>	confirmed	Lung cancer in ever smokers
rs760077	<i>THBS3</i>	confirmed	Renal function-related traits (BUN); Magnesium levels; Esophageal cancer and gastric cancer; Urate levels; Occipital lobe volume; General risk tolerance (MTAG); Red blood cell count; Serum urea levels; Lung function (FVC); Serum cancer antigen 15.3 levels; PR interval; Diastolic blood pressure; Mean arterial pressure; Systolic blood pressure; Dentures; Brain region volumes; Heel bone mineral density; Number of decayed; missing and filled tooth surfaces or use of dentures; Gastric cancer; Serum uric acid levels; Blood urea nitrogen levels; Estimated glomerular filtration rate; Non-cardia gastric cancer; Urinary albumin-to-creatinine ratio in non-diabetics; Urinary albumin-to-creatinine ratio; Gastric adenocarcinoma (histologically verified); Cancer (pleiotropy); Cancer; Hematocrit; Chronic inflammatory diseases (ankylosing spondylitis; Crohn's disease; psoriasis; primary sclerosing cholangitis; ulcerative colitis) (pleiotropy); Hemoglobin concentration; Urea levels; Serum magnesium levels; Gout; Microalbuminuria
rs6480781	<i>C10orf11</i>	confirmed	Lung function (FEV1/FVC)
rs2342607	<i>TMEM254.AS1</i>	confirmed	Chronic obstructive pulmonary disease-related biomarkers
rs2368238	<i>MPZL2</i>	confirmed	Blood protein levels; Lung adenocarcinoma; Non-small cell lung cancer; Lung cancer
rs11231152	<i>RP11.371A22.1</i>	confirmed	Waist-hip ratio; Heel bone mineral density; Waist-to-hip ratio adjusted for BMI; Lung function (FVC); FEV1; Peak expiratory flow; Blood protein levels; Waist-to-hip ratio adjusted for body mass index; Lung function (FEV1); Waist-to-hip ratio adjusted for BMI (joint analysis for main effect and physical activity interaction); Waist-to-hip ratio adjusted for BMI in active individuals; Waist-to-hip ratio adjusted for BMI x sex x age interaction (4df test)
rs61705659	<i>ARHGEF17</i>	confirmed	Pulmonary function (smoking interaction); FEV1; Lung function (FEV1/FVC); Peak expiratory flow; GIP levels in response to oral glucose tolerance test (120 minutes)
rs10876864	<i>RPS26</i>	confirmed	Type 1 diabetes; Vitiligo; Polycystic ovary syndrome; Appendicular lean mass; Smoking status (ever vs never smokers); Eosinophil counts; Red blood cell count; Respiratory diseases; Autoimmune traits; General cognitive ability; Lung function (FEV1/FVC); FEV1; Asthma; Eczema; Hypothyroidism; Menarche (age at onset); Hay fever and/or eczema; Household income (MTAG); Nasal polyps; Cognitive ability; years of educational attainment or schizophrenia (pleiotropy); Smoking status; Inflammatory skin disease; Alopecia areata; Lung function (FEV1);

eQTL rs ID	Gene symbol	confirmation	GWAS Catalog phenotypes
			Body mass index; Rheumatoid arthritis; Heel bone mineral density; Age of smoking initiation (MTAG); Asthma (childhood onset); Intelligence (MTAG); Cognitive ability (MTAG); Educational attainment (years of education); Smoking initiation (ever regular vs never regular); Gastroesophageal reflux disease; Asthma or allergic disease (pleiotropy); Pediatric autoimmune diseases; Nonsyndromic cleft lip with cleft palate; Eosinophil percentage of white cells; Educational attainment (college completion); Eosinophil percentage of granulocytes; Sum eosinophil basophil counts; Neutrophil percentage of granulocytes; Allergic disease (asthma; hay fever or eczema); Asthma (adult onset); Medication use (drugs for peptic ulcer and gastro-oesophageal reflux disease); Medication use (glucocorticoids); Medication use (adrenergics; inhalants); Medication use (thyroid preparations); Asthma (moderate or severe); Educational attainment (MTAG); Self-reported math ability (MTAG); Highest math class taken; Cognitive performance; Cognitive performance (MTAG)
rs1131017	AC004057.1	confirmed	Type 1 diabetes; Vitiligo; Polycystic ovary syndrome; Appendicular lean mass; Smoking status (ever vs never smokers); Red blood cell count; Respiratory diseases; General cognitive ability; Lung function (FEV1/FVC); FEV1; Asthma; Eczema; Hypothyroidism; Menarche (age at onset); Hay fever and/or eczema; Allergic rhinitis; Household income (MTAG); Nasal polyps; Cognitive ability; years of educational attainment or schizophrenia (pleiotropy); Smoking status; Inflammatory skin disease; Alopecia areata; Lung function (FEV1); Body mass index; Rheumatoid arthritis; Heel bone mineral density; Brain region volumes; Lifetime smoking index; Age of smoking initiation (MTAG); Smoking initiation (ever regular vs never regular); Asthma (childhood onset); Cognitive function; Intelligence (MTAG); Cognitive ability (MTAG); Anorexia nervosa; Educational attainment (years of education); Smoking initiation; Allergic disease (asthma; hay fever or eczema); Asthma or allergic disease (pleiotropy); Pediatric autoimmune diseases; Eosinophil counts; Nonsyndromic cleft lip with cleft palate; Eosinophil percentage of white cells; Educational attainment (college completion); Eosinophil percentage of granulocytes; Neutrophil percentage of granulocytes; Asthma (adult onset); Smoking initiation (ever regular vs never regular) (MTAG); Medication use (glucocorticoids); Medication use (adrenergics; inhalants); Medication use (thyroid preparations); Asthma (moderate or severe); Educational attainment (MTAG); Self-reported math ability (MTAG); Highest math class taken; Cognitive performance; Cognitive performance (MTAG); Self-reported math ability; Highest math class taken (MTAG)
rs7147394	HAUS4	confirmed	Hypertension; Lung function (FEV1/FVC); Household income (MTAG); Cognitive ability; years of educational attainment or schizophrenia (pleiotropy); Cognitive ability (MTAG); Educational attainment (years of education); Educational attainment (MTAG); Highest math class taken; Cognitive performance (MTAG); Self-reported math ability; Highest math class taken (MTAG)
rs7183203	FBN1	confirmed	Breast cancer; Lung function (FEV1/FVC); Heel bone mineral density; Intraocular pressure
rs1628955	ANKDD1A	confirmed	Lung function (FEV1/FVC); Pulse pressure
rs34443712	TSPAN3	confirmed	Height; Red blood cell count; Lung function (FEV1/FVC)
rs2668714	RP11.156P1.2	confirmed	Lung function in never smokers (low FEV1 vs high FEV1); Loneliness; White matter microstructure (radial diusivities); White matter microstructure (axial diusivities); White matter microstructure (mean diusivities); White matter microstructure (fractional anisotropy); General cognitive ability; Neuroticism; Feeling guilty; Feeling miserable; Experiencing mood swings; Red cell distribution width; Mood instability; Breast cancer
rs2696506	RP11.44K6.4	confirmed	Lung function in never smokers (low FEV1 vs high FEV1); Loneliness; White matter microstructure (radial

eQTL rs ID	Gene symbol	confirmation	GWAS Catalog phenotypes
			diusivities); White matter microstructure (axial diusivities); White matter microstructure (mean diusivities); White matter microstructure (fractional anisotropy); General cognitive ability; Neuroticism; Feeling guilty; Feeling miserable; Experiencing mood swings; Red cell distribution width; Mood instability; Breast cancer
rs113680823	<i>DCXR</i>	confirmed	Lung function (FVC); FEV1
rs62113084	<i>ZFP82</i>	confirmed	Lung function (FVC)
rs406133	<i>SHKBP1</i>	confirmed	Age at loss of ambulation in Duchenne muscular dystrophy; Lung function (FVC); Smoking initiation (ever regular vs never regular) (MTAG)
rs4569473	<i>RP11.301O19.1</i>	confirmed	Lung function (FEV1/FVC); Pulse pressure; Diastolic blood pressure; Systolic blood pressure
rs2735313	<i>LCA5L</i>	confirmed	Lung function (FEV1/FVC); Platelet count
rs13051949	<i>WRB</i>	confirmed	Lung function (FEV1/FVC); Esotropia; Non-accommodative esotropia; Platelet count; Plateletcrit
rs1800818	<i>PDGFB</i>	confirmed	Lung function (FEV1/FVC)
rs6775611	<i>ANAPC13</i>	confirmed	Lung function (FVC); Height; Neurocognitive impairment in HIV-1 infection (dichotomous)
rs9852745	<i>RSRC1</i>	confirmed	Lung function in heavy smokers (high FEV1 vs average FEV1); Waist-hip ratio; Depression; Body mass index
rs11724788	<i>CYTL1</i>	confirmed	Lung function (FVC); Height
rs10064943	<i>RP11.43F13.1</i>	confirmed	Lung adenocarcinoma
rs12658646	<i>PDE8B</i>	confirmed	Caudate nucleus volume; Lung function (FEV1/FVC); Estimated glomerular filtration rate
rs4869923	<i>GINM1</i>	confirmed	Brainstem volume; Lung cancer
rs2393593	<i>HIST1H4H</i>	confirmed	Parental longevity (father's age at death or father's attained age); Offspring birth weight; Birth weight; Hand grip strength; Lifetime smoking index; Chronotype; Smoking cessation (MTAG); Lung cancer; Height; Cisplatin-induced ototoxicity; Educational attainment; Educational attainment (years of education); Educational attainment (MTAG)
rs67509210	<i>BTN3A2</i>	confirmed	Urinary metabolite modules (eigenmetabolites) in chronic kidney disease; General cognitive ability; Life satisfaction; Neuroticism; Hip circumference variance; Hip circumference; Intelligence (MTAG); Urinary tract infection frequency; Squamous cell lung carcinoma; Lung cancer in ever smokers; Headache; Urate levels
rs67859638	<i>BTN3A2</i>	confirmed	Urinary tract infection frequency; Squamous cell lung carcinoma
rs192804591	<i>HIST1H4L</i>	confirmed	Schizophrenia; White matter microstructure (radial diusivities); White matter microstructure (mean diusivities); Schizophrenia; schizoaffective disorder or bipolar disorder; Pulse pressure; Schizophrenia or bipolar disorder; Lung cancer; Urate levels
rs1811359	<i>HLA.DPB2</i>	confirmed	Hepatitis B; Aspirin exacerbated respiratory disease in asthmatics; Hepatitis B (viral clearance); Severe aplastic anemia; Chronic hepatitis B infection; Hepatitis B vaccine response; Wegener's granulomatosis; Antineutrophil cytoplasmic antibody-associated vasculitis; Mouth ulcers; Response to hepatitis B vaccine; Knee osteoarthritis; Osteoarthritis of the hip or knee; Osteoarthritis; White coat effect (clinic diastolic blood pressure minus ambulatory diastolic blood pressure); Thyroid peroxidase antibody levels; Acute graft versus host disease in bone marrow transplantation (recipient effect); Susceptibility to persistent hepatitis B virus infection; EGFR mutation-positive lung adenocarcinoma; Sjögren's syndrome

eQTL rs ID	Gene symbol	confirmation	GWAS Catalog phenotypes
rs7797295	GNA12	confirmed	Height; Ulcerative colitis; Birth weight; Body fat distribution (trunk fat ratio); Body fat distribution (leg fat ratio); Heel bone mineral density; White blood cell count; Chronic obstructive pulmonary disease; Lung function (FEV1/FVC); White matter microstructure (mode of anisotropy); Brain region volumes; Hip circumference; Hip circumference adjusted for BMI; Waist circumference adjusted for body mass index; Waist circumference adjusted for BMI in non-smokers; Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction); Inflammatory bowel disease; Waist circumference adjusted for BMI in active individuals; Waist circumference adjusted for BMI (adjusted for smoking behaviour); Waist circumference adjusted for BMI (joint analysis main effects and smoking interaction)
rs80233585	PPP1R16A	confirmed	General cognitive ability; Reaction time; Lung function (FVC); Cognitive ability; years of educational attainment or schizophrenia (pleiotropy); Height; Blood metabolite levels; Age at first birth; Mean corpuscular hemoglobin concentration; Educational attainment (years of education); Smoking initiation (ever regular vs never regular) (MTAG); Self-reported math ability (MTAG); Self-reported math ability; Highest math class taken; Educational attainment (MTAG); Cognitive performance (MTAG); Highest math class taken (MTAG)
rs1545837	KIAA1967	confirmed	Parkinson's disease or first degree relation to individual with Parkinson's disease; Lung function (FVC); Triglycerides; Parkinson's disease; Exploratory eye movement dysfunction in schizophrenia (responsive search score); Exploratory eye movement dysfunction in schizophrenia (cognitive search score)
rs11778927	RP11.51J9.5	confirmed	Lung function (FEV1/FVC); Migraine
rs3213849	FGFR1	confirmed	Lung function (FEV1/FVC); PR interval; Body mass index; Age of smoking initiation (MTAG); Smoking initiation (ever regular vs never regular) (MTAG); Response to cognitive-behavioural therapy in anxiety disorder; Waist-to-hip ratio adjusted for BMI; Waist-hip ratio
rs10870201	CARD9	confirmed	Crohn's disease; Ulcerative colitis; Red cell distribution width; Lung function (FVC); FEV1; Type 2 diabetes; Glucose homeostasis traits; Ankylosing spondylitis; Inflammatory bowel disease; Peak insulin response (insulin secretion adjusted); Pediatric autoimmune diseases; IgA nephropathy; Granulocyte percentage of myeloid white cells; Chronic inflammatory diseases (ankylosing spondylitis; Crohn's disease; psoriasis; primary sclerosing cholangitis; ulcerative colitis) (pleiotropy)
rs10870160	SDCCAG3	confirmed	Crohn's disease; Ulcerative colitis; Red cell distribution width; White blood cell count; Lung function (FVC); Morningness; Morning person; Glucose homeostasis traits; Ankylosing spondylitis; Inflammatory bowel disease; Chronotype; Basal cell carcinoma; Squamous cell carcinoma; Cutaneous squamous cell carcinoma; Pediatric autoimmune diseases; IgA nephropathy; Eosinophil counts; Chronic inflammatory diseases (ankylosing spondylitis; Crohn's disease; psoriasis; primary sclerosing cholangitis; ulcerative colitis) (pleiotropy); Myeloid white cell count; Sum neutrophil eosinophil counts; Granulocyte count; Sum basophil neutrophil counts; Neutrophil count

<i>HLA-G</i>	major histocompatibility complex, class I, G	6:31242089 (A/T)	rs2524080	trans	intergenic	8.51x10 ⁻¹⁶
<i>HLA-G</i>	major histocompatibility complex, class I, G	6:29936065 (C/T)	rs17180100	cis	intergenic	1.59x10 ⁻²⁸

NIHR BioResource Collaborators

NIHR BioResource – Rare Disease Consortium¹, Julian Adlard², Munaza Ahmed³, Tim Aitman^{4,5}, Hana Alachkar⁶, David Allsup⁷, Jeff Almeida-King⁸, Philip Ancliff⁹, Richard Antrobus¹⁰, Ruth Armstrong^{11,12,13}, Gavin Arno^{14,15}, Sofie Ashford^{1,16}, William Astle^{1,16,17}, Anthony Attwood^{1,16}, Chris Babbs^{18,19}, Tamam Bakchoul²⁰, Tadbir Bariana^{21,22}, Julian Barwell^{23,24}, David Bennett²⁵, David Bentley²⁶, Agnieszka Bierzynska²⁷, Tina Biss²⁸, Marta Bleda²⁹, Harm Bogaard³⁰, Christian Bourne²⁶, Sara Boyce³¹, John Bradley¹, Gerome Breen^{32,33}, Paul Brennan^{34,35}, Carole Brewer³⁶, Matthew Brown^{1,16}, Michael Browning³⁷, Rachel Buchan^{38,39}, Matthew Buckland⁴⁰, Teofila Bueser^{41,42,43}, Siobhan Burns⁴⁰, Oliver Burren²⁹, Paul Calleja⁴⁴, Gerald Carr-White⁴², Keren Carss^{1,16}, Ruth Casey^{11,12,13}, Mark Caulfield⁴⁵, John Chambers^{46,47}, Jennifer Chambers^{48,49}, Floria Cheng⁴⁹, Patrick F Chinnery^{1,50,51}, Martin Christian⁵², Colin Church⁵³, Naomi Clements Brod^{1,16}, Gerry Coghlan⁴⁰, Elizabeth Colby²⁷, Trevor Cole⁵⁴, Janine Collins⁵⁵, Peter Collins⁵⁶, Camilla Colombo²⁶, Robin Condliffe⁵⁷, Stuart Cook^{38,58,59,60}, Terry Cook⁶¹, Nichola Cooper⁶², Paul Corris^{63,64}, Abigail Crisp-Hihn^{1,16}, Nicola Curry⁶⁵, Cesare Danesino⁶⁶, Matthew Daniels^{67,68}, Louise Daugherty^{1,16}, John Davis^{1,16}, Sri V V Deevi^{1,16}, Timothy Dent⁶⁸, Eleanor Dewhurst^{1,16}, Peter Dixon⁴⁸, Kate Downes^{1,16}, Anna Drazyk⁶⁹, Elizabeth Drewe⁷⁰, Tina Dutt⁷¹, David Edgar⁷², Karen Edwards^{1,16}, William Egner⁷³, Wendy Erber⁷⁴, Marie Erwood^{1,16}, Maria C Estiu⁷⁵, Gillian Evans⁷⁶, Dafydd Gareth Evans⁷⁷, Tamara Everington⁷⁸, Mélanie Eyries⁷⁹, Remi Favier^{80,81,82}, Debra Fletcher^{1,16}, James Fox^{1,16}, Amy Frary^{1,16}, Courtney French⁸³, Kathleen Freson⁸⁴, Mattia Frontini^{1,16}, Daniel Gale⁸⁵, Henning Gall⁸⁶, Claire Geoghegan²⁶, Terry Gerightly²⁶, Stefano Ghio⁸⁷, Hossein-Ardeschir Ghofrani^{62,86}, Simon Gibbs³⁸, Kimberley Gilmour⁸⁸, Barbara Girerd^{89,90,91}, Sarah Goddard⁹², Keith Gomez^{21,22}, Pavels Gordins⁹³, David Gosal⁶, Stefan Gräf^{1,16,29}, Luigi Grassi^{1,16}, Daniel Greene^{1,16,17}, Lynn Greenhalgh⁹⁴, Andreas Greinacher⁹⁵, Paolo Gresele⁹⁶, Philip Griffiths^{97,98}, Sofia Grigoriadou⁹⁹, Russell Grocock²⁶, Detelina Grozeva¹¹, Scott Hackett¹⁰⁰, Charaka Hadinnapola²⁹, William Hague¹⁰¹, Matthias Haime^{1,16,29}, Matthew Hall⁷⁰, Helen Hanson⁹⁴, Kirsty Harkness¹⁰², Andrew Harper^{38,67,103}, Claire Harris⁶⁴, Daniel Hart⁵⁵, Ahamad Hassan¹⁰⁴, Grant Hayman¹⁰⁵, Alex Henderson¹⁰⁶, Jonathan Hoffmann⁵⁴, Rita Horvath^{107,108}, Arjan Houweling³⁰, Luke Howard³⁸, Fengyuan Hu^{1,16}, Gavin Hudson¹⁰⁷, Joseph Hughes²⁶, Aarnoud Huissoon¹⁰⁰, Marc Humbert^{89,90,91}, Sean Humphray²⁶, Sarah Hunter²⁶, Matthew Hurles¹⁰⁹, Louise Izatt¹¹⁰, Roger James^{1,16}, Sally Johnson¹¹¹, Stephen Jolles^{112,113}, Jennifer Jolley^{1,16}, Neringa Jurkute^{14,22}, Mary Kasanicki¹¹⁴, Hanadi Kazkaz¹¹⁵, Rashid Kazmi³¹, Peter Kelleher³⁹, David Kiely⁵⁷, Nathalie Kingston¹, Robert Klima⁴⁴, Myrto Kostadima^{1,16}, Gabor Kovacs^{116,117}, Ania Koziell^{118,119}, Roman Kreuzhuber^{1,16}, Taco Kuijpers^{120,121}, Ajith Kumar³, Dinakantha Kumararatne¹¹⁴, Manju Kurian^{122,123}, Michael Laffan^{124,125}, Fiona Laloo⁷⁷, Michele Lambert^{126,127}, Hana Lango Allen^{1,16}, Allan Lawrie¹²⁸, Mark Layton¹²⁴, Claire Lentaigne^{124,125}, Adam Levine⁸⁵, Rachel Linger^{1,16}, Hilary Longhurst⁹⁹, Eleni Louka^{18,19}, Robert MacKenzie Ross¹²⁹, Bella Madan¹³⁰, Eamonn Maher^{11,131}, Jesmeen Maimaris⁸⁸, Sarah Mangles¹³², Rutendo Mapeta^{1,16}, Kevin Marchbank⁶⁴, Stephen Marks⁹, Hugh S Markus⁶⁹, Hanns-Ulrich Marschall¹³³, Andrew Marshall^{134,135,136}, Jennifer Martin^{1,16,29}, Mary Mathias¹³⁷, Emma Matthews^{22,138}, Heather Maxwell¹³⁹, Paul McAlinden⁶⁴, Mark McCarthy^{19,103,140}, Stuart Meacham^{1,16}, Adam Mead¹⁴¹, Karyn Megy^{1,16}, Sarju Mehta¹⁴², Michel Michaelides¹⁴, Carolyn Millar^{124,125}, Shahin Moledina⁹, David Montani^{89,90,91}, Tony Moore^{14,15}, Nicholas Morrell^{1,29}, Monika Mozere⁸⁵, MPGN/C3 Glomerulopathy Rare Renal Disease group¹⁴³, Keith Muir¹⁴⁴, Andrew Mumford^{145,146}, Michael Newnham²⁹, Jennifer O'Sullivan¹³⁰, Samya Obaji⁵⁶, Steven Okoli^{18,19}, Andrea Olschewski¹¹⁶, Horst Olschewski^{116,117}, Kai Ren Ong⁵⁴, Elizabeth Ormondroyd^{67,68}, Willem Ouwehand^{1,16}, Sofia Papadia^{1,16}, Soo-Mi Park^{12,13,147}, David Parry⁵, Joan Paterson^{11,12,13}, Andrew Peacock⁵³, John Peden²⁶, Kathelijne Peerlinck⁸⁴, Christopher Penkett^{1,16}, Joanna Pepke-Zaba¹⁴⁸, Romina Petersen^{1,16}, Angela Pyle¹⁰⁷, Stuart Rankin⁴⁴, Anupama Rao⁹, F Lucy Raymond^{1,11}, Paula Rayner-Matthews^{1,16}, Christine Rees²⁶, Augusto Rendon⁴⁵, Tara Renton⁴³, Andrew Rice^{149,150}, Sylvia Richardson¹⁷, Alex Richter¹⁰, Irene Roberts^{18,19,151}, Catherine Roughley⁷⁶, Noemi Roy^{18,19,151}, Omid Sadeghi-Alavijeh⁸⁵, Moin Saleem²⁷, Nilesh Samani¹⁵², Alba Sanchis-Juan^{1,16}, Ravishankar Sargur⁷³, Simon Satchell²⁷, Sinisa Savic¹⁵³, Laura Scelsi⁸⁷, Sol Schulman¹⁵⁴, Marie Scully¹¹⁵, Claire Searle¹⁵⁵, Werner Seeger⁸⁶, Carrock Sewell¹⁵⁶, Denis Seyres^{1,16}, Susie Shapiro⁶⁵, Olga Sharmardina^{1,16}, Rakefet

Shtoyerman¹⁵⁷, Keith Sibson¹³⁷, Lucy Side³, Ilenia Simeoni^{1,16}, Michael Simpson¹⁵⁸, Suthesh Sivapalaratnam⁵⁵, Anne-Bine Skytte¹⁵⁹, Katherine Smith⁴⁵, Kenneth G C Smith^{29,160}, Katie Snape¹⁶¹, Florent Soubrier⁷⁹, Simon Staines^{1,16}, Emily Staples²⁹, Hannah Stark^{1,16}, Jonathan Stephens^{1,16}, Kathleen Stirrups^{1,16}, Sophie Stock^{1,16}, Jay Suntharalingam¹²⁹, Emilia Swietlik²⁹, R Campbell Tait¹⁶², Kate Talks²⁸, Rhea Tan⁶⁹, James Thaventhiran²⁹, Andreas Themistocleous²⁵, Moira Thomas¹⁶³, Kate Thomson^{67,68}, Adrian Thrasher⁹, Chantal Thys⁸⁴, Marc Tischkowitz¹⁶⁴, Catherine Titterton^{1,16}, Cheng-Hock Toh⁷¹, Mark Toshner²⁹, Matthew Traylor⁶⁹, Carmen Treacy^{29,148}, Richard Trembath⁴¹, Salih Tuna^{1,16}, Wojciech Turek⁴⁴, Ernest Turro^{1,16,17}, Tom Vale²⁵, Chris Van Geet⁸⁴, Natalie Van Zuydam²⁵, Marta Vazquez-Lopez⁴⁹, Julie von Ziegenweidt^{1,16}, Anton Vonk Noordegraaf³⁰, Quintin Waisfisz³⁰, Suellen Walker⁹, James Ware^{38,39,58}, Hugh Watkins^{67,68,103}, Christopher Watt^{1,16}, Andrew Webster^{14,15}, Wei Wei⁵⁰, Steven Welch¹⁰⁰, Julie Wessels⁹², Sarah Westbury^{145,146}, John-Paul Westwood¹¹⁵, John Wharton⁶², Deborah Whitehorn^{1,16}, James Whitworth^{11,12,13}, Martin R Wilkins⁶², Catherine Williamson^{48,165}, Edwin Wong⁹⁸, Nicholas Wood^{166,167}, Yvette Wood^{1,16}, Geoff Woods^{11,114}, Emma Woodward⁷⁷, Stephen Wort^{39,41}, Austen Worth⁹, Katherine Yates^{1,16,29}, Patrick Yong¹⁶⁸, Tim Young^{1,16}, Ping Yu^{1,16}, Patrick Yu-Wai-Man⁵⁰

Affiliations

¹NIHR BioResource, Cambridge University Hospitals NHS Foundation, Cambridge Biomedical Campus, Cambridge, UK. ²Chapel Allerton Hospital, Leeds Teaching Hospitals NHS Trust, Leeds, UK. ³North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. ⁴MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, London, UK. ⁵Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. ⁶Salford Royal NHS Foundation Trust, Salford, UK. ⁷Queens Centre for Haematology and Oncology, Castle Hill Hospital, Hull and East Yorkshire NHS Trust, Cottingham, UK. ⁸European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, UK. ⁹Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. ¹⁰University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. ¹¹Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. ¹²Cancer Research UK Cambridge Centre, Cambridge Biomedical Campus, Cambridge, UK. ¹³NIHR Cambridge Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK. ¹⁴Moorfields Eye Hospital NHS Foundation Trust, London, UK. ¹⁵UCL Institute of Ophthalmology, University College London, London, UK. ¹⁶Department of Haematology, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. ¹⁷MRC Biostatistics Unit, Cambridge Institute of Public Health, University of Cambridge, Cambridge, UK. ¹⁸MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. ¹⁹NIHR Oxford Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, UK. ²⁰Center for Clinical Transfusion Medicine, University Hospital of Tübingen, Tübingen, Germany. ²¹The Katharine Dormandy Haemophilia Centre and Thrombosis Unit, Royal Free London NHS Foundation Trust, London, UK. ²²University College London, London, UK. ²³Department of Clinical Genetics, Leicester Royal Infirmary, University Hospitals of Leicester, Leicester, UK. ²⁴University of Leicester, Leicester, UK. ²⁵The Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK. ²⁶Illumina Limited, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, UK. ²⁷Bristol Renal, University of Bristol, Bristol, UK. ²⁸Haematology Department, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. ²⁹Department of Medicine, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. ³⁰Department of Pulmonary Medicine, VU University Medical Centre, Amsterdam, The Netherlands. ³¹Southampton General Hospital, University Hospital Southampton NHS Foundation Trust, Southampton, UK. ³²MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London,

London, UK.³³NIHR Biomedical Research Centre for Mental Health, Maudsley Hospital, London, UK.
³⁴Newcastle University, Newcastle upon Tyne, UK.³⁵Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.³⁶Department of Clinical Genetics, Royal Devon & Exeter Hospital, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK.³⁷Department of Immunology, Leicester Royal Infirmary, Leicester, UK.³⁸National Heart and Lung Institute, Imperial College London, Royal Brompton Hospital, London, UK.³⁹Royal Brompton Hospital, Royal Brompton and Harefield NHS Foundation Trust, London, UK.⁴⁰Royal Free London NHS Foundation Trust, London, UK.⁴¹King's College London, London, UK.⁴²Guy's and St Thomas' Hospital, Guy's and St Thomas' NHS Foundation Trust, London, UK.⁴³King's College Hospital NHS Foundation Trust, London, UK.⁴⁴High Performance Computing Service, University of Cambridge, Cambridge, UK.⁴⁵Genomics England Ltd, London, UK.⁴⁶Epidemiology and Biostatistics, Imperial College London, London, UK.⁴⁷Imperial College Healthcare NHS Trust, London, UK.⁴⁸Division of Women's Health, King's College London, London, UK.⁴⁹Women's Health Research Centre, Surgery and Cancer, Faculty of Medicine, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK.⁵⁰Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK.⁵¹Medical Research Council Mitochondrial Biology Unit, Cambridge Biomedical Campus, Cambridge, UK.⁵²Children's Renal and Urology Unit, Nottingham Children's Hospital, QMC, Nottingham University Hospitals NHS Trust, Nottingham, UK.⁵³Golden Jubilee National Hospital, Glasgow, UK.⁵⁴West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK.⁵⁵The Royal London Hospital, Barts Health NHS Foundation Trust, London, UK.⁵⁶The Arthur Bloom Haemophilia Centre, University Hospital of Wales, Cardiff, UK.⁵⁷Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital NHS Foundation Trust, Sheffield, UK.⁵⁸MRC London Institute of Medical Sciences, Imperial College London, London, UK.⁵⁹National Heart Research Institute Singapore, National Heart Centre Singapore, Singapore, Singapore.⁶⁰Division of Cardiovascular & Metabolic Disorders, Duke-National University of Singapore, Singapore, Singapore.⁶¹Imperial College Renal and Transplant Centre, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK.⁶²Department of Medicine, Imperial College London, London, UK.⁶³National Pulmonary Hypertension Service (Newcastle), The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.⁶⁴Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.⁶⁵Oxford Haemophilia and Thrombosis Centre, The Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK.⁶⁶Department of Molecular Medicine, General Biology, and Medical Genetics Unit, University of Pavia, Pavia, Italy.⁶⁷Department of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.⁶⁸Oxford University Hospitals NHS Foundation Trust, Oxford, UK.⁶⁹Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK.⁷⁰Nottingham University Hospitals NHS Trust, Nottingham, UK.⁷¹The Roald Dahl Haemostasis and Thrombosis Centre, The Royal Liverpool Hospital, Liverpool, UK.⁷²Regional Immunology Service, Kelvin Building, Royal Victoria Hospital, Belfast, UK.⁷³Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK.⁷⁴Pathology and Laboratory Medicine, University of Western Australia, Crawley, Australia.⁷⁵Ramón Sardá Mother's and Children's Hospital, Buenos Aires, Argentina.⁷⁶Haemophilia Centre, Kent & Canterbury Hospital, East Kent Hospitals University Foundation Trust, Canterbury, UK.⁷⁷Manchester Centre for Genomic Medicine, Saint Mary's Hospital, Manchester, UK.⁷⁸Salisbury District Hospital, Salisbury NHS Foundation Trust, Salisbury, UK.⁷⁹Departement de genetique, Hopital Pitie-Salpetriere, Paris, France.⁸⁰Service d'Hématologie Biologique, Hôpital d'enfants Armand Trousseau, Paris, France, Paris, France.⁸¹Inserm U1170, Villejuif, France.⁸²Assistance Publique-Hôpitaux de Paris, Département d'Hématologie, Hôpital Armand Trousseau, Paris, France.⁸³Department of Paediatrics, School of Clinical Medicine, University of Cambridge , Cambridge Biomedical Campus, Cambridge, UK.⁸⁴Department of Cardiovascular Sciences, Center for Molecular and Vascular Biology, University of Leuven, Leuven, Belgium.⁸⁵UCL Centre for Nephrology, University College London, London, UK.⁸⁶University of Giessen and Marburg Lung Center (UGMLC), Giessen, Germany.⁸⁷Division of Cardiology, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy.⁸⁸UCL Great Ormond

Street Institute of Child Health, London, UK. ⁸⁹Université Paris-Sud, Le Kremlin-Bicêtre, France. ⁹⁰Service de Pneumologie, DHU Thorax Innovation, Hôpital Bicêtre, Le Kremlin-Bicêtre, France. ⁹¹INSERM U999, LabEx LERMIT, Centre Chirurgical Marie Lannelongue, Le Plessis Robinson, France. ⁹²University Hospitals of North Midlands NHS Trust, Stoke-on-Trent, UK. ⁹³East Yorkshire Regional Adult Immunology and Allergy Unit, Hull Royal Infirmary, Hull & East Yorkshire Hospitals NHS Trust, Hull, UK. ⁹⁴Department of Clinical Genetics, Liverpool Women's NHS Foundation, Liverpool, UK. ⁹⁵Institute for Immunology and Transfusion Medicine, University of Greifswald, Greifswald, Germany. ⁹⁶Section of Internal and Cardiovascular Medicine, University of Perugia, Perugia, Italy. ⁹⁷Mitochondrial Research Group, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. ⁹⁸Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. ⁹⁹Barts Health NHS Foundation Trust, London, UK. ¹⁰⁰Birmingham Heartlands Hospital, Heart of England NHS Foundation Trust, Birmingham, UK. ¹⁰¹ARCH: Australian Research Centre for Health of Women and Babies, Robinson Research Institute, Discipline of Obstetrics and Gynaecology, The University of Adelaide, Women's and Children's Hospital, Adelaide, Australia. ¹⁰²Department of Neurology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. ¹⁰³Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. ¹⁰⁴Department of Neurology, Leeds Teaching Hospital NHS Trust, Leeds, UK. ¹⁰⁵Epsom & St Helier University Hospitals NHS Trust, London, UK. ¹⁰⁶Northern Genetics Service, The Newcastle upon Tyne Hospitals NHS Foundation Trust, International Centre for Life, Newcastle upon Tyne, UK. ¹⁰⁷Wellcome Centre for Mitochondrial Research, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. ¹⁰⁸John Walton Muscular Dystrophy Research Centre, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. ¹⁰⁹Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. ¹¹⁰Department of Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, UK. ¹¹¹Department of Paediatric Nephrology, Great North Children's Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. ¹¹²University Hospital of Wales, Cardiff, UK. ¹¹³Cardiff & Vale University LHB, Cardiff, UK. ¹¹⁴Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ¹¹⁵University College London Hospitals NHS Foundation Trust, London, UK. ¹¹⁶Ludwig Boltzmann Institute for Lung Vascular Research, Graz, Austria. ¹¹⁷Dept of Internal Medicine, Division of Pulmonology, Medical University of Graz, Graz, Austria. ¹¹⁸Division of Transplantation Immunology and Mucosal Biology, Department of Experimental Immunobiology, Faculty of Life Sciences and Medicine, King's College London, London, UK. ¹¹⁹Department of Paediatric Nephrology, Evelina London Children's Hospital, Guy's & St Thomas' NHS Foundation Trust, London, UK. ¹²⁰Department of Pediatric Hematology, Immunology, Rheumatology and Infectious Diseases, Emma Children's Hospital, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, The Netherlands. ¹²¹Department of Clinical Genetics, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, The Netherlands. ¹²²Molecular Neurosciences, Developmental Neurosciences, UCL Great Ormond Street Institute of Child Health, London, UK. ¹²³Department of Neurology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. ¹²⁴Department of Haematology, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. ¹²⁵Department of Haematology, Imperial College London, London, UK. ¹²⁶Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, USA. ¹²⁷Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, USA. ¹²⁸Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Sheffield, UK. ¹²⁹Royal United Hospitals Bath NHS Foundation Trust, Bath, UK. ¹³⁰Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, UK. ¹³¹Cambridge NIHR Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK. ¹³²Haemophilia, Haemostasis and Thrombosis Centre, Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK. ¹³³Wallenberg Laboratory, Department of Molecular and Clinical Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ¹³⁴Faculty of Medical and Human Sciences, Centre for Endocrinology and Diabetes, Institute of Human Development, University of Manchester, Manchester, UK. ¹³⁵Department of Clinical Neurophysiology, Manchester Royal Infirmary, Central Manchester University Hospitals National Health Service Foundation Trust, Manchester Academic Health

Science Centre, Manchester, UK. ¹³⁶National Institute for Health Research/Wellcome Trust Clinical Research Facility, Manchester, UK. ¹³⁷Department of Haematology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. ¹³⁸The National Hospital for Neurology and Neurosurgery, University College London Hospitals NHS Foundation Trust, London, UK. ¹³⁹Royal Hospital for Children, NHS Greater Glasgow and Clyde, Glasgow, UK. ¹⁴⁰Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, UK. ¹⁴¹Centre for Haematology, Department of Medicine, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. ¹⁴²Department of Clinical Genetics, Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ¹⁴³MPGN/C3 Glomerulopathy Rare Renal Disease group, , UK. ¹⁴⁴Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK. ¹⁴⁵School of Cellular and Molecular Medicine, University of Bristol, Bristol, UK. ¹⁴⁶University Hospitals Bristol NHS Foundation Trust, Bristol, UK. ¹⁴⁷East Anglian Regional Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ¹⁴⁸Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK. ¹⁴⁹Pain Research, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK. ¹⁵⁰Chelsea and Westminster Hospital NHS Foundation Trust, London, UK. ¹⁵¹Department of Paediatrics, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. ¹⁵²Departments of Cardiovascular Sciences and NIHR Leicester Cardiovascular Biomedical Research Unit, University of Leicester, Leicester, UK. ¹⁵³The Leeds Teaching Hospitals NHS Trust, Leeds, UK. ¹⁵⁴Beth Israel Deaconess Medical Centre and Harvard Medical School, Boston, USA. ¹⁵⁵Department of Clinical Genetics, Nottingham University Hospitals NHS Trust, Nottingham, UK. ¹⁵⁶Scunthorpe General Hospital, Northern Lincolnshire and Goole NHS Foundation Trust, Scunthorpe, UK. ¹⁵⁷Clinical Genetics Institute, Kaplan Medical Center, Rehovot, Israel. ¹⁵⁸Genetics and Molecular Medicine, King's College London, London, UK. ¹⁵⁹Aarhus University Hospital, Aarhus, Denmark. ¹⁶⁰Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. ¹⁶¹Department of Clinical Genetics, St George's University Hospitals NHS Foundation Trust, London, UK. ¹⁶²Glasgow Royal Infirmary, NHS Greater Glasgow and Clyde, Glasgow, UK. ¹⁶³Gartnavel General Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK. ¹⁶⁴Addenbrooke's Treatment Centre, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ¹⁶⁵Institute of Reproductive and Developmental Biology, Surgery and Cancer, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. ¹⁶⁶Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK. ¹⁶⁷UCL Genetics Institute, London, UK. ¹⁶⁸Frimley Park Hospital, NHS Frimley Health Foundation Trust, Camberley, UK.