SUPPLEMENTARY INFORMATION

A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci

Nature Genetics: doi:10.1038/ng.3955

SUPPLEMENTARY NOTE

Enrichment of PDGene associations in PDWBS data

Top associations in PDGene were tested for enrichment of smaller p-values in the PDWBS data. All variants (of the 9,830 top PDGene variants) within 1Mb of previously associated loci were removed, and the remaining SNPs were LD-pruned with an r^2 threshold of 0.01 (using 1000 Genomes data for reference LD information). After the previous filtering step, 448 SNPs remained. Enrichment of variants observed at each significance threshold in the PDWBS data was tested in the remaining SNPs using a chi-square test. Results are shown in Supplementary Table 3. We observe significant enrichment of smaller p-values in the top PDGene associations in the PDWBS data.

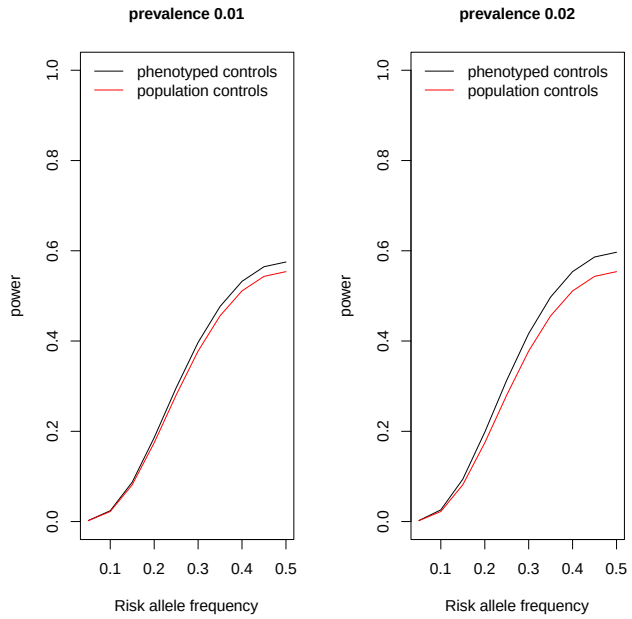
Power calculations for phenotyped versus population controls

The self-reported samples used as controls could result in a reduction of power in PDWBS. We investigated whether sample misdiagnosis or the presence of samples that may develop Parkinson's in the future within controls led to a reduction in power. Assuming that the misdiagnosis rate is unlikely to exceed the population prevalence of PD, we modified the method for computing power of a GWAS (CaTs)[1] by adjusting the controls disease allele frequency to account for misdiagnosis (based on prevalence) using the equation:

$$control_{adj}DAF = controlDAF*(1\text{-}prevalance) + caseDAF*(prevalence)$$

where *caseDAF* and *controlDAF* are unchanged from the original CaTs calculation.

We compared the unadjusted power calculation ("*phenotyped controls*") to the power calculation adjusted for a misdiagnosis rate equal to the population prevalence of PD (0.01) ("*population controls*"). Assuming a significance threshold alpha=$5x10^{-8}$ and genotype relative risk of 1.1, we observe minimal loss in power (see plot below). For example, at a disease allele frequency of 0.3, misdiagnosis of controls decreased the power from 0.397 to 0.378. Increasing the population prevalence of PD to 0.02 did not substantially decrease power (see below). Given these power calculations, we conclude that the use of self-reported controls has a modest impact on the power to detect effects within PDWBS.



PD genetic risk score

We calculated genetic risk scores (GRS) as described in previous studies to test whether the new PD risk loci significantly improved the GRS[2-4]. The GRS for previous loci used the 28 previously replicated GWAS variants (the 24 genome-wide significant loci in Table 1 in the main text, and four secondary signals in the *GBA*, *TMEM175*, *SNCA* and *HLA-DQB1* locus [3]) with the addition of rarer variants N370S and E326K in the *GBA* gene and G2019S in *LRRK2*[5] (Bras *et al* under review) (31 total variants). We compared this to a GRS calculated with an additional 16 novel loci present in the NeuroX data (see Table 2 in main text) (47 total variants). All GRS analyses were weighted using combined discovery effect estimates applied to the NeuroX allele counts, adjusting for age (onset in cases or last exam in controls), sex and eigenvectors 1-5 within a logistic regression model. Comparing areas under the curve of logistic regression models, we see a slight but significant increase in classification accuracy from the 31 variant GRS (AUC = 0.6439, 95% CI: 0.634-0.6537) to the 47 variant GRS (AUC = 0.6518, 95% CI: 0.6419-0.6616) quantified by DeLong's test ($|Z| = 5.4187$, $P = 6.002 \times 10^{-8}$)[6]. A significant increase in accuracy was still observed if we removed the three loci that did not validate at a one-sided $P<0.05$ in the NeuroX dataset ($|Z| = 5.342$, $P = 9.192 \times 10^{-8}$). While slightly higher AUC was detected in the 47 variant GRS than in the 44 variant GRS (AUC = 0.6515, 95% CI: 0.6417-0.6614), the increase in AUC was not significant ($P > 0.05$).

Differential expression between PD and healthy controls

Total RNA was extracted from the striatum and frontal cortex of subjects with a histopathological diagnosis of Parkinson's disease and aged matched healthy controls for sequencing. 0.5 µg of total RNA was used as an input material for library preparation using TruSeq RNA Sample Preparation Kit v2 (Illumina). Size of the libraries was confirmed using Fragment Analyzer (Advanced Analytical Technologies).

Library concentrations were determined by a qPCR-based method using a library quantification kit (KAPA). The libraries were multiplexed and then sequenced on Illumina HiSeq2500 to generate 25 million single-end 50-base reads per library. HTSeqGenie was used to perform feature counting, filtering & alignment. Only reads with unique genomic alignments were analyzed. Reads with alignments overlapping exons of the gene model were counted towards that gene. Gene models used in this step were from the IGIS gene model package on Bioconductor (TxDb.Hsapiens.BioMart.igis).

Genes with mean expression across all samples and tissues < 2 RPKM were removed. Differential expression between Parkinson's patients and healthy controls was assessed separately in the frontal cortex and in the striatum tissue. Differential expression was tested with DESeq2 [7]. Genes with an adjusted p-value < 0.05 in either tissue were considered as suggestive of differential expression. Because gene expression is derived from bulk tissue, it is equally possible that differential expression between the two disease states is due to differential cell-type composition of the tissue sample. Without expression data at cell-level resolution it is difficult to distinguish between a scenario with differential cell-type composition, differential gene expression or a mixture of the two. We therefore acknowledge that significant genes in our differential expression analysis are merely suggestive of differential gene expression between disease states.

Colocalization of eQTL and PD-associations

Colocalization is a formal way to test for overlap between the causal variant underlying a GWAS association and an eQTL association. One method typically used to carry this out on summary statistics is encoded by the R package, *coloc*[8]. Applying this method to GWAS and eQTL data requires making some key assumptions, including the assumption of a single causal variant and that the causal variant is included in the input summary statistics (see Giambortolomei *et al.* 2013 for a more detailed discussion). As the meta-analysis was carried out on a subset of SNPs provided publicly by PDGene, we did not have dense genotyping of variants in the PD-associated loci with summary statistics from the meta-analysis. Thus, we could not carry out a full colocalization analysis on all of the PD-associated loci discussed in this study.

As an alternative, we carried out colocalization only on the PDWBS dataset at the previously reported PD-loci that were genome-wide significant in the PDWBS cohort alone and for which we report an eQTL (7 SNP-gene pairings, see below). Colocalization between causal variants in PD associated loci and eQTLs was carried out with the coloc.abf function in the *coloc* R package[8] with prior probablities *p1* and *p2* parameters set to $1x10^{-5}$. eQTL summary statistics as pre-computed by GTEx and summary statistics from PDWBS were used as input for *coloc*. The colocalization analysis was run on all SNPs within 250kb of the index SNP that overlapped both the GTEx and the PDWBS GWAS data for each PD-associated locus. Of the eQTLs report, we find that all have moderate to high evidence of colocalization (posterior probability > 0.5).

SNP	Gene	GTEx tissue	Posterior probability of shared causal variant
rs11060180	*OGFOD2*	Brain_Anterior_cingulate_cortex_BA24	0.523
rs12637471	*MCCC1*	Muscle_Skeletal	0.554
rs1474055	*STK39*	Brain_Putamen_basal_ganglia	0.593
rs1474055	*STK39*	Testis	0.578
rs34311866	*DGKQ*	Brain_Cerebellum	0.514
rs34311866	*DGKQ*	Lung	0.858
rs356182	*SNCA*	Brain_Cerebellar_Hemisphere	0.548
rs823118	*NUCKS1*	Brain_Caudate_basal_ganglia	0.63
rs823118	*NUCKS1*	Skin_Sun_Exposed_Lower_leg	0.929
rs823118	*SLC41A1*	Brain_Cerebellum	0.521
rs823118	*SLC41A1*	Pancreas	0.626

Gene sets used for targeted gene-set enrichment analyses

Lysosomal, autophagy and mitochondrial gene-sets mapping to the background list of genes (see Methods section in the main paper) are listed below.

Lysosomal genes

ABCA2, ABCA3, ABCB9, ABCC10, ACP2, ACP5, ACPP, AGA, AHNAK, ANKFY1, ANPEP, AP1G1, AP3B1, AP3D1, AP3M1, ARL8A, ARL8B, ARSA, ARSB, ARSG, ASAH1, ATP11A, ATP11C, ATP13A2, ATP6AP1, ATP6V0A1, ATP6V0A2, ATP6V0A4, ATP6V0C, ATP6V0D1, ATP6V1A, ATP6V1B1, ATP6V1B2, ATP6V1C1, ATP6V1C2, ATP6V1D, ATP6V1E1, ATP6V1F, ATP6V1G1, ATP6V1H, BTD, C18orf8, C1orf85, CCZ1, CD164, CD63, CD68, CECR1, CLCN5, CLCN6, CLCN7, CLN3, CLN5, CLU, COL6A1, CP, CPVL, CREG1, CST7, CTBS, CTNS, CTSA, CTSB, CTSC, CTSD, CTSF, CTSH, CTSK, CTSO, CTSS, CTSZ, DAGLB, DEPDC5, DNAJC13, DNAJC5, DNASE1, DNASE2, DPP4, DPP7, ECE1, EGF, ELANE, ENPEP, ENPP1, ENTPD4, EPDR1, FLOT1, FLOT2, FUCA1, FUCA2, GAA, GALC, GALNS, GBA, GDAP2, GGH, GLA, GLB1, GM2A, GNA11, GNAI1, GNAI2, GNAI3, GNAQ, GNB1, GNB2, GNB4, GNS, GPLD1, GPR137, GRN, GUSB, HEXA, HEXB, HGSNAT, HPSE, HYAL1, HYAL2, IDS, IDUA, IFI30, IL4I1, LAMP1, LAMP2, LAMP3, LAMTOR2, LAPTM4A, LAPTM5, LGMN, LIPA, LITAF, LMBRD1, LNPEP, LRP1, LRP2, MAN2B1, MAN2B2, MANBA, MCOLN1, MCOLN2, MCOLN3, MFSD8, MIOS, MON1B, MPO, MTOR, MYLPF, NAAA, NAGA, NAGLU, NAPA, NAPG, NCSTN, NEU1, NEU4, NPC1, NPC2, NSF, OSTM1, P2RX4, PCYOX1, PI4K2A, PLA2G15, PLBD1, PLBD2, PLD1, PLOD1, PPT1, PPT2, PRCP, PRTN3, PSAP, PSEN1, PTGDS, RAB14, RAB2A, RAB5C, RDH14, RNASE1, RNASE6, RNASET2, RNF13, RPTOR, RRAGA, RRAGB, SCARB1, SCARB2, SCPEP1, SGSH, SIAE, SIDT2, SLC11A1, SLC11A2, SLC12A4, SLC15A3, SLC17A5, SLC26A11, SLC29A3, SLC2A13, SLC2A8, SLC30A2, SLC36A1, SLC44A2, SMCR8, SMPD1, SMPD4, SMPDL3A, SPG11, SPHK2, SPNS1, SPPL2A, STARD3, STARD3NL, STX7, TCIRG1, TLR3, TM9SF1, TMBIM1, TMEM175, TMEM192, TMEM55A, TMEM55B, TMEM63A, TMEM74, TMEM8A, TMEM9, TMEM92, TPP1, UBA52, VAMP7, VASN, VPS11, VPS16, VPS18, VPS33A, VPS35, VPS39, VPS41, WDR11, WDR41, ZFYVE26

Autophagy genes

AAMP, AFF4, ALDOA, AMBRA1, APOL1, ARNT, ARSA, ARSB, ATF4, ATF6, ATG10, ATG12, ATG16L1, ATG16L2, ATG2A, ATG2B, ATG3, ATG4A, ATG4B, ATG4C, ATG4D, ATG5, ATG7, ATG9A, ATG9B, ATIC, BAG1, BAG3, BAK1, BAX, BCL2, BCL2L1, BECN1, BID, BIRC5, BIRC6, BNIP1, BNIP3, BNIP3L, C12orf44, CALCOCO2, CAMKK2, CANX, CAPN1, CAPN10, CAPN2, CAPNS1, CASP1, CASP3, CASP4, CASP8, CCL2, CCR2, CD46, CDKN1A, CDKN1B, CDKN2A, CFLAR, CHMP2B, CHMP4B, CLDN9, CLN3, CSNK1A1, CTSB, CTSD, CX3CL1, CXCR4, DAPK1, DAPK2, DDIT3, DIRAS3, DLC1, DNAJB1, DNAJB9, DRAM1, EDEM1, EEF2, EEF2K, EGFR, EIF2AK2, EIF2AK3, EIF2S1, EIF4EBP1, EIF4G1, ERBB2, ERN1, ERO1L, FADD, FAS, FKBP1A, FKBP1B, FOS, FOXO1, FOXO3, GAA, GABARAP, GABARAPL1, GABARAPL2, GAPDH, GNAI3, GNB2L1, GOPC, GRID1, GRID2, HDAC1, HDAC6, HGS, HIF1A, HSP90AB1, HSPA5, HSPA8, HSPB8, IFNG, IKBKB, IKBKE, IL24, ILF2, INTS7, IRGM, ITGA3, ITGA6, ITGB1, ITGB4, ITPR1, KAT8, KIAA0226, KIF5B, KLHL24, LAMP1, LAMP2, MAP1LC3A, MAP1LC3B, MAP1LC3C, MAP2K7, MAPK1, MAPK3, MAPK8, MAPK8IP1, MAPK9, MBTPS2, MLST8, MTMR14, MTOR, MYC, NAF1, NAMPT, NBR1, NCKAP1, NFE2L2, NFKB1, NKX2-3, NLRC4, NPC1, NRG1, NRG2, NRG3, P4HB, PARK2, PARP1, PEA15, PELP1, PEX14, PEX3, PIK3C3, PIK3R4, PINK1, PPP1R15A, PPP1R15B, PRKAB1, PRKAR1A, PRKCD, PRKCQ, PTEN, PTK6, RAB11A, RAB1A, RAB24, RAB33B, RAB5A, RAB7A, RAC1, RAF1, RB1, RB1CC1, RBX1, RELA, RGS19, RHEB, RPS6KB1, RPTOR, SAR1A, SERPINA1, SESN2, SH3GLB1, SIRT1, SIRT2, SPHK1, SPNS1, SQSTM1, ST13, STK11, TBK1, TM9SF1, TMEM74, TNFSF10, TP53,

TP53INP2, TP63, TP73, TSC1, TSC2, TUSC1, ULK1, ULK2, ULK3, USP10, UVRAG, VAMP3, VAMP7, VEGFA, WDFY3, WDR45, WIPI1, WIPI2, ZFYVE1

Mitochondrial genes

AADAT, AARS2, AASS, ABAT, ABCA13, ABCA9, ABCB10, ABCB6, ABCB7, ABCB8, ABCB9, ABCD1, ABCD2, ABCD3, ABCF2, ABHD10, ABHD11, ACAA1, ACAA2, ACACA, ACACB, ACAD10, ACAD11, ACAD8, ACAD9, ACADL, ACADM, ACADS, ACADSB, ACADVL, ACAT1, ACCS, ACLY, ACN9, ACO1, ACO2, ACOT13, ACOT2, ACOT7, ACOT9, ACOX1, ACOX3, ACP6, ACSF2, ACSF3, ACSL1, ACSL4, ACSL6, ACSM1, ACSM2A, ACSM3, ACSM5, ACSS1, ACSS3, ACYP2, ADCK1, ADCK2, ADCK3, ADCK4, ADCK5, ADHFE1, AFG3L2, AGK, AGMAT, AGPAT5, AGR2, AGXT, AGXT2, AHCYL1, AIFM1, AIFM2, AIFM3, AK2, AK3, AK4, AKAP1, AKAP10, AKR1B10, AKR7A2, ALAS1, ALAS2, ALDH18A1, ALDH1B1, ALDH1L1, ALDH1L2, ALDH2, ALDH3A2, ALDH4A1, ALDH5A1, ALDH6A1, ALDH7A1, ALDH9A1, ALKBH1, ALKBH3, ALKBH7, AMACR, AMT, ANGEL2, APEX2, APOA1BP, APOO, APOOL, APOPT1, ARF5, ARG2, ARL2, ARMC10, ARMS2, ASAH2, ATAD1, ATAD3A, ATAD3B, ATIC, ATP10D, ATP5A1, ATP5B, ATP5C1, ATP5D, ATP5E, ATP5F1, ATP5G1, ATP5G2, ATP5G3, ATP5H, ATP5I, ATP5J, ATP5J2, ATP5J2-PTCD1, ATP5L, ATP5O, ATP5S, ATP5SL, ATPAF1, ATPAF2, ATPIF1, ATXN2, AUH, AURKAIP1, BAD, BAK1, BAX, BCAT2, BCKDHA, BCKDHB, BCKDK, BCL2, BCL2L1, BCL2L13, BCL2L2, BCS1L, BDH1, BID, BLOC1S1, BNIP3, BNIP3L, BOLA1, BOLA3, BPHL, C10orf10, C10orf2, C12orf10, C12orf65, C14orf159, C14orf2, C15orf40, C15orf48, C15orf61, C16orf91, C17orf89, C19orf52, C19orf70, C1QBP, C20orf24, C21orf33, C2orf47, C2orf69, C3orf33, C5orf63, C6orf136, C6orf203, C6orf57, C7orf55, C8orf82, CA5A, CA5B, CARKD, CARS2, CASP8, CAT, CBR3, CBR4, CCBL2, CCDC109B, CCDC127, CCDC51, CCDC58, CCDC90B, CCT7, CDC25C, CECR5, CEP89, CHCHD1, CHCHD10, CHCHD2, CHCHD3, CHCHD4, CHCHD5, CHCHD6, CHCHD7, CHDH, CHPT1, CISD1, CISD2, CISD3, CKMT1A, CKMT1B, CKMT2, CLIC4, CLPB, CLPP, CLPX, CLYBL, CMC1, CMC2, CMC4, CMPK2, COA1, COA3, COA4, COA5, COA6, COA7, COASY, COMT, COMTD1, COQ10A, COQ10B, COQ2, COQ3, COQ4, COQ5, COQ6, COQ7, COQ9, COX10, COX11, COX14, COX15, COX16, COX17, COX18, COX19, COX20, COX4I1, COX4I2, COX5A, COX5B, COX6A1, COX6A2, COX6B1, COX6B2, COX6C, COX7A1, COX7A2, COX7A2L, COX7B, COX7C, COX8A, COX8C, CPOX, CPS1, CPT1A, CPT1B, CPT1C, CPT2, CRAT, CRLS1, CROT, CRY1, CRYZ, CS, CYB5A, CYB5B, CYB5R2, CYB5R3, CYC1, CYCS, CYP11A1, CYP11B2, CYP24A1, CYP27A1, CYP27B1, D2HGDH, DAP3, DARS2, DBI, DBT, DCAKD, DCXR, DDAH1, DDX28, DECR1, DGUOK, DHCR24, DHODH, DHRS1, DHRS4, DHRS7B, DHRSX, DHTKD1, DHX30, DIABLO, DLAT, DLD, DLST, DMGDH, DMPK, DNA2, DNAJA3, DNAJC11, DNAJC15, DNAJC19, DNAJC28, DNAJC30, DNAJC4, DNLZ, DNM1L, DTYMK, DUS2, DUSP26, DUT, EARS2, ECH1, ECHDC1, ECHDC2, ECHDC3, ECHS1, ECI1, ECI2, ECSIT, EEFSEC, EFHD1, EHHADH, ELAC2, EMC2, ENDOG, EPHX2, ERAL1, ETFA, ETFB, ETFDH, ETHE1, EXOG, FABP1, FAHD1, FAHD2A, FAM136A, FAM162A, FAM185A, FAM195A, FAM210A, FAM210B, FAM213A, FARS2, FASN, FASTK, FASTKD1, FASTKD2, FASTKD3, FASTKD5, FBXL4, FDPS, FDX1, FDX1L, FDXR, FECH, FH, FHIT, FIS1, FKBP10, FKBP8, FLAD1, FOXRED1, FPGS, FTH1, FTMT, FTSJ2, FUNDC1, FUNDC2, FXN, GADD45GIP1, GAPDH, GARS, GATC, GATM, GBAS, GCAT, GCDH, GCSH, GDAP1, GFER, GFM1, GFM2, GHITM, GK, GLDC, GLOD4, GLRX2, GLRX5, GLS, GLS2, GLUD1, GLYAT, GLYCTK, GNG5, GOLPH3, GOT2, GPAM, GPD2, GPI, GPT2, GPX1, GPX4, GRHPR, GRPEL1, GRPEL2, GRSF1, GSR, GSTK1, GSTO1, GSTZ1, GTPBP10, GTPBP3,

GTPBP6, GUF1, GUK1, HADH, HADHA, HADHB, HAGH, HAO2, HARS2, HCCS, HDHD3, HEBP1, HEMK1, HIBADH, HIBCH, HIGD1A, HIGD2A, HINT1, HINT2, HINT3, HK1, HK2, HMBS, HMGCL, HMGCS2, HOGA1, HRSP12, HSCB, HSD17B10, HSD17B4, HSD17B8, HSDL1, HSDL2, HSPA9, HSPB7, HSPD1, HSPE1, HTATIP2, HTRA2, IARS2, IBA57, ICT1, IDE, IDH1, IDH2, IDH3A, IDH3B, IDH3G, IDI1, IFI27, IMMP1L, IMMP2L, IMMT, ISCA1, ISCA2, ISCU, ISOC2, IVD, KARS, KIAA0100, KIAA0141, KIAA0391, KIF1B, KMO, KRT5, L2HGDH, LACE1, LACTB, LACTB2, LAMC1, LAP3, LARS2, LDHAL6B, LDHB, LDHD, LETM1, LETM2, LETMD1, LIAS, LIPT1, LIPT2, LONP1, LONP2, LRPPRC, LYPLA1, LYPLAL1, LYRM1, LYRM2, LYRM4, LYRM5, LYRM7, LYRM9, MACROD1, MALSU1, MAOA, MAOB, MARC1, MARC2, MARCH5, MARS2, MAVS, MCAT, MCCC1, MCCC2, MCEE, MCU, MCUR1, MDH1, MDH2, ME1, ME2, ME3, MECR, METAP1D, METTL15, METTL17, METTL5, METTL8, MFF, MFN1, MFN2, MGARP, MGME1, MGST1, MGST3, MICU1, MICU2, MIEF1, MINOS1, MIPEP, MLH1, MLYCD, MMAB, MMACHC, MMADHC, MOCS1, MPC1, MPC2, MPST, MPV17, MPV17L, MPV17L2, MRM1, MRPL1, MRPL10, MRPL11, MRPL12, MRPL13, MRPL14, MRPL15, MRPL16, MRPL17, MRPL18, MRPL19, MRPL2, MRPL20, MRPL21, MRPL22, MRPL23, MRPL24, MRPL27, MRPL28, MRPL3, MRPL30, MRPL32, MRPL33, MRPL34, MRPL35, MRPL36, MRPL37, MRPL38, MRPL39, MRPL4, MRPL40, MRPL41, MRPL42, MRPL43, MRPL44, MRPL45, MRPL46, MRPL47, MRPL48, MRPL49, MRPL50, MRPL51, MRPL52, MRPL53, MRPL54, MRPL55, MRPL57, MRPL9, MRPS10, MRPS11, MRPS12, MRPS14, MRPS15, MRPS16, MRPS17, MRPS18A, MRPS18B, MRPS18C, MRPS2, MRPS21, MRPS22, MRPS23, MRPS24, MRPS25, MRPS26, MRPS27, MRPS28, MRPS30, MRPS31, MRPS33, MRPS34, MRPS35, MRPS36, MRPS5, MRPS6, MRPS7, MRPS9, MRRF, MRS2, MSRA, MSRB2, MSRB3, MTCH1, MTCH2, MTCP1, MTERF, MTERFD1, MTERFD2, MTERFD3, MTFMT, MTFP1, MTFR1, MTFR1L, MTG1, MTG2, MTHFD1, MTHFD1L, MTHFD2, MTHFD2L, MTHFS, MTIF2, MTIF3, MTO1, MTPAP, MTRF1, MTRF1L, MTX1, MTX2, MUL1, MUT, MUTYH, NADK2, NAGS, NARS2, NBR1, NCEH1, NCOA4, NDUFA1, NDUFA10, NDUFA11, NDUFA12, NDUFA13, NDUFA2, NDUFA3, NDUFA4, NDUFA5, NDUFA6, NDUFA7, NDUFA8, NDUFA9, NDUFAB1, NDUFAF1, NDUFAF2, NDUFAF3, NDUFAF4, NDUFAF5, NDUFAF6, NDUFAF7, NDUFB1, NDUFB10, NDUFB11, NDUFB2, NDUFB3, NDUFB4, NDUFB5, NDUFB6, NDUFB7, NDUFB8, NDUFB9, NDUFC1, NDUFC2, NDUFS1, NDUFS2, NDUFS3, NDUFS4, NDUFS5, NDUFS6, NDUFS7, NDUFS8, NDUFV1, NDUFV2, NDUFV3, NEU4, NFS1, NFU1, NGRN, NIF3L1, NIPSNAP1, NIPSNAP3A, NIPSNAP3B, NIT1, NIT2, NLN, NLRX1, NME3, NME4, NME6, NMNAT3, NNT, NOA1, NRD1, NSUN3, NSUN4, NT5C, NT5DC2, NT5DC3, NT5M, NTHL1, NUBPL, NUCB2, NUDT13, NUDT19, NUDT2, NUDT5, NUDT6, NUDT8, NUDT9, OAT, OBSCN, OCIAD1, OCIAD2, OGDH, OGDHL, OGG1, OMA1, OPA1, OPA3, OSBPL1A, OSGEPL1, OTC, OXA1L, OXCT1, OXLD1, OXNAD1, OXR1, OXSM, P4HB, PACSIN2, PAICS, PAK7, PAM16, PANK2, PARK7, PARL, PARS2, PC, PCBD2, PCCA, PCCB, PCK2, PDE12, PDF, PDHA1, PDHA2, PDHB, PDHX, PDK1, PDK2, PDK3, PDK4, PDP1, PDP2, PDPR, PDSS1, PDSS2, PET100, PET112, PET117, PEX11B, PGAM5, PGS1, PHB, PHB2, PHYH, PI4KA, PICK1, PIF1, PINK1, PISD, PITRM1, PKLR, PLGRKT, PMAIP1, PMPCA, PMPCB, PNPLA8, PNPO, PNPT1, POLDIP2, POLG, POLG2, POLRMT, PPA2, PPIF, PPM1K, PPOX, PPTC7, PPWD1, PRDX2, PRDX3, PRDX4, PRDX5, PRDX6, PRELID1, PRELID2, PREPL, PRODH, PRODH2, PROSC, PRSS35, PSMA6, PSTK, PTCD1, PTCD2, PTCD3, PTGES2, PTPMT1, PTPN4, PTRH1, PTRH2, PTS, PUS1, PUSL1, PXMP2, PXMP4, PYCR1, PYCR2, PYURF, QDPR, QRSL1, QTRT1, RAB11FIP5, RAB24, RAB32, RAB35, RARS, RARS2, RBFA,

RCN2, RDH11, RDH13, RDH14, RECQL4, REXO2, RFK, RHOT1, RHOT2, RMDN1, RMDN3, RMND1, RNASEH1, RNMTL1, ROMO1, RPIA, RPL10A, RPL34, RPL35A, RPS14, RPS15A, RPS18, RPUSD3, RPUSD4, RSAD1, RTN4IP1, SAMM50, SARDH, SARS2, SCCPDH, SCO1, SCO2, SCP2, SDHA, SDHAF1, SDHAF2, SDHB, SDHC, SDHD, SDR39U1, SDSL, SECISBP2, SELO, SEPT4, SERAC1, SERHL2, SETD9, SFXN1, SFXN2, SFXN3, SFXN4, SFXN5, SHMT1, SHMT2, SIRT3, SIRT4, SIRT5, SLC16A1, SLC16A11, SLC16A7, SLC22A4, SLC25A1, SLC25A10, SLC25A11, SLC25A12, SLC25A13, SLC25A14, SLC25A15, SLC25A16, SLC25A17, SLC25A18, SLC25A19, SLC25A20, SLC25A21, SLC25A22, SLC25A23, SLC25A24, SLC25A25, SLC25A26, SLC25A27, SLC25A28, SLC25A29, SLC25A3, SLC25A30, SLC25A31, SLC25A32, SLC25A33, SLC25A34, SLC25A35, SLC25A36, SLC25A37, SLC25A38, SLC25A39, SLC25A4, SLC25A40, SLC25A41, SLC25A42, SLC25A43, SLC25A44, SLC25A45, SLC25A46, SLC25A47, SLC25A48, SLC25A5, SLC25A51, SLC25A53, SLC25A6, SLC30A6, SLC30A9, SLC37A4, SLIRP, SLMO1, SLMO2, SMDT1, SNAP29, SND1, SOD1, SOD2, SPATA19, SPATA20, SPG7, SPR, SPRYD4, SPTLC2, SQRDL, SSBP1, STAR, STARD7, STOM, STOML1, STOML2, STX17, SUCLA2, SUCLG1, SUCLG2, SUOX, SUPV3L1, SURF1, SYNJ2BP, TACO1, TAMM41, TARS, TARS2, TBRG4, TCAIM, TCHP, TCIRG1, TDRKH, TEFM, TFAM, TFB1M, TFB2M, THEM4, THG1L, THNSL1, TIMM10, TIMM10B, TIMM13, TIMM17A, TIMM17B, TIMM21, TIMM22, TIMM23, TIMM44, TIMM50, TIMM8A, TIMM8B, TIMM9, TIMMDC1, TK2, TKT, TMBIM4, TMEM11, TMEM126A, TMEM126B, TMEM143, TMEM14C, TMEM177, TMEM186, TMEM205, TMEM65, TMEM70, TMLHE, TOMM20, TOMM22, TOMM34, TOMM40, TOMM40L, TOMM5, TOMM6, TOMM7, TOMM70A, TOP1MT, TOP3A, TPI1, TRAP1, TRIAP1, TRIT1, TRMT1, TRMT10C, TRMT11, TRMT2B, TRMT61B, TRMU, TRNT1, TRUB2, TSFM, TSPO, TST, TSTD1, TSTD3, TTC19, TUBB3, TUFM, TXN2, TXNDC12, TXNRD1, TXNRD2, TYSND1, UCP1, UCP2, UCP3, UNG, UQCC1, UQCC2, UQCR10, UQCR11, UQCRB, UQCRC1, UQCRC2, UQCRFS1, UQCRH, UQCRQ, USMG5, VARS2, VDAC1, VDAC2, VDAC3, VWA8, WARS2, WBSCR16, WDR81, XPNPEP3, XRCC6BP1, YARS2, YBEY, YME1L1, ZADH2

Non-targeted gene-set enrichment analysis of PD-associated regions

Overall, we found the PD associated regions to be enriched for genes in the lysosomal KEGG pathway ($P_{adjusted}$=0.02) (Supplementary Table 12). Of the 115 genes mapping to this pathway, 12 mapped to 8 of the PD associated intervals (*SCARB2, PSAPL1, ATP6V0A1, NEU1, PPT2, NAGLU, IDUA, LAMP3, GBA, GALC, ABCB9,* and *CTSB*). Our candidate gene pipeline had nominated four of these twelve genes (*ATP6V0A1, GBA, GALC* AND *CTSB*). We note that both *NAGLU* and *ATP6V0A1* were within 250kb of the index SNP rs601999, though only *ATP6V0A1* was nominated by our candidate pipeline as a candidate gene.

23andMe Research Team

Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, David A. Hinds, Bethann S.

Hromatka, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Vladimir Vacic, and Catherine H. Wilson

IPDGC consortium members and affiliations:

Alastair J Noyce (Department of Molecular Neuroscience, UCL, London, UK), Alexis Brice (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Anamika Giri (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Angelika Oehmig (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Arianna Tucci (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Aude Nicolas (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Claudia Schulte (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research), Mark R Cookson (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA), Cornelis Blauwendraat (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA), Kathrin Brockmann (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Demis Kia (UCL Genetics Institute; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Fabrice Danjou (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Faraz Faghri (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA; Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA), Gavin Charlesworth (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), J Raphael Gibbs (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Huw R Morris (National Hospital for Neurology and Neurosurgery, University College London, London, UK), Helene Plun-Favreau (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Dena G Hernandez (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Peter Holmans (Biostatistics &

Bioinformatics Unit, Institute of Psychological Medicine and Clinical Neuroscience, MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff, UK), Huw R Morris (National Hospital for Neurology and Neurosurgery, University College London, London, UK), Iris Jansen (VU University Medical Center, Amsterdam, Netherlands), John Hardy (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Javier Simón-Sánchez (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Jose M Bras (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Joshua M. Shulman (Baylor College of Medicine, Houston, Texas, USA), John Quinn (Institute of Translational Medicine, University of Liverpool, Liverpool, UK), Juan A. Botía (Universidad de Murcia, Murcia, Spain), Kin Y Mok (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Kimberley Billingsley (Institute of Translational Medicine, University of Liverpool, Liverpool, UK), Lasse Pihlstrom (Department of Neurology, Oslo University Hospital, Oslo, Norway), Lea R'Bibo (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Codrin Lungu (National Institutes of Health Parkinson Clinic, NINDS, National Institutes of Health, Bethesda, MD, USA), Manu Sharma (Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tubingen and Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen Germany), Maria Martinez (INSERM UMR 1220; and Paul Sabatier University, Toulouse, France), Mina Ryten (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Valentina Escott-Price (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, UK), Niccolo E. Mencacci (Department of Molecular Neuroscience, UCL, London, UK), Mike A. Nalls (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA; Contractor/consultant with Data Tecnica, MD, USA), Nicholas W Wood (UCL Genetics Institute; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Patrick Lewis (University of Reading, Reading, UK), Paul Denny (University College London, London, UK), Peter Heutink (DZNE, German Center for Neurodegenerative Diseases and Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany), Patrizia Rizzu (DZNE, German Center for Neurodegenerative Diseases), Pille Taba (Department of Neurology and Neurosurgery, University of Tartu, Tartu, Estonia), Rita Guerreiro (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Ruth Lovering (University College London, London, UK), Raquel Duran Ogalla (University College London, London, UK), Rebecca Foulger (University College London, London, UK), Laurie Robak (Baylor College of Medicine, Houston, Texas, USA), Steven Lubbe (Ken and Ruth Davee Department of Neurology, Northwestern

University Feinberg School of Medicine, Chicago, IL, USA), Steven Finkbeiner (Departments of Neurology and Physiology, University of California, San Francisco; Gladstone Institute of Neurological Disease; Taube/Koret Center for Neurodegenerative Disease Research, San Francisco, CA, USA), Sigurlaug Sveinbjörnsdóttir (Department of Neurology, Landspítali University Hospital, Reykjavík, Iceland; Department of Neurology, MEHT Broomfield Hospital, Chelmsford, Essex, UK; and Queen Mary College, University of London, London, UK), Andrew B Singleton (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Sonja Scholz (Neurodegenerative Diseases Research Unit, National Institute of Neurological Disorders and Stroke, Bethesda, MD, USA), Sulev Koks (Department of Pathophysiology, University of Tartu, Tartu, Estonia), Suzanne Lesage (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Jean-Christophe Corvol (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, Centre d'Investigation Clinique Pitié Neurosciences CIC-1422, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Thomas Foltynie (UCL Institute of Neurology, London, UK), Thomas Gasser (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), T. Ryan Price (University California Irvine, Irvine, CA, USA), Una-Marie Sheerin (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Nigel Williams (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff, UK), Xylena Reed (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA).

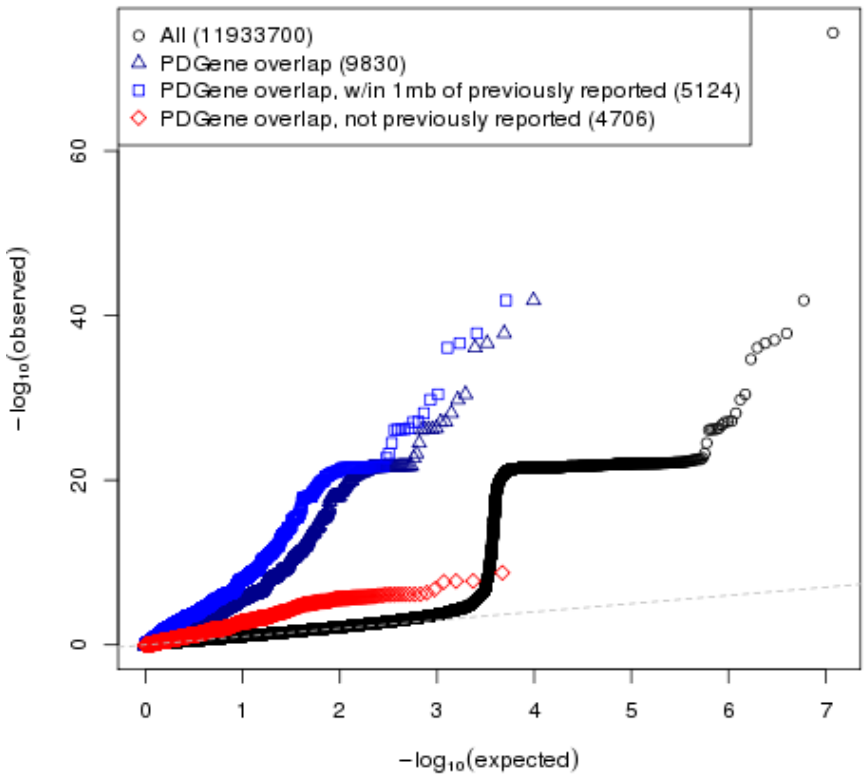
Acknowledgements

We thank the UK Parkinson's Disease Society Brain Bank (UKPDSBB). Mike A. Nalls' participation is supported by a consulting contract between Data Tecnica International and the National Institute on Aging, NIH, Bethesda, MD, USA. This work was supported in part by the Intramural Research Programs of the National Institute of Neurological Disorders and Stroke (NINDS), the National Institute on Aging (NIA), and the National Institute of Environmental Health Sciences both part of the National Institutes of Health, Department of Health and Human Services; project numbers Z01-AG000949-02 and Z01-ES101986. In addition this work was supported by the Department of Defense (award W81XWH-09-2-0128), and The Michael J Fox Foundation for Parkinson's Research. This work was supported by National Institutes of Health grants R01NS037167, R01CA141668, P50NS071674, American Parkinson Disease Association (APDA); Barnes Jewish Hospital

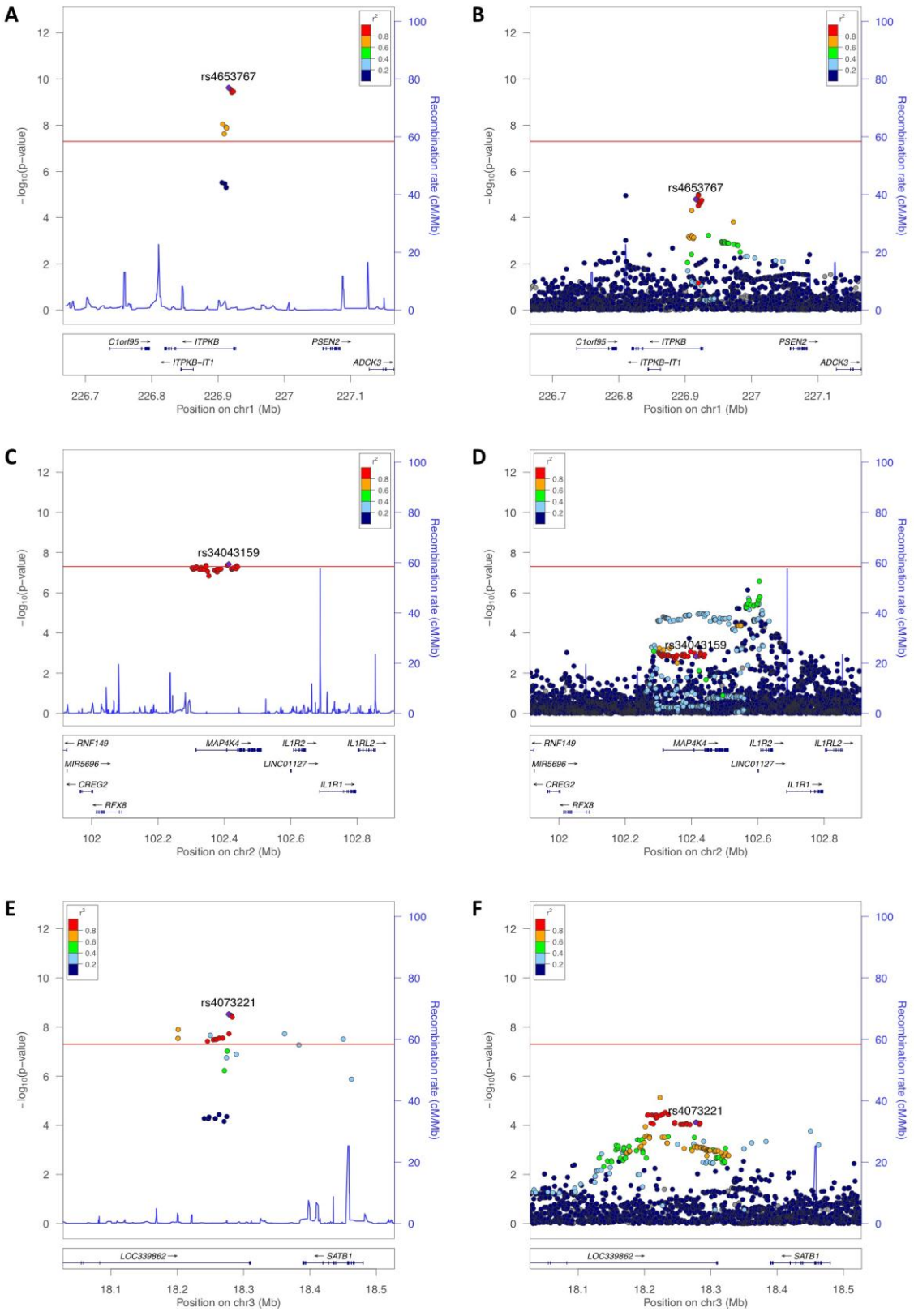
Foundation; Greater St Louis Chapter of the APDA; Hersenstichting Nederland; the Prinses Beatrix Fonds. The KORA (Cooperative Research in the Region of Augsburg) research platform was started and financed by the Forschungszentrum für Umwelt und Gesundheit, which is funded by the German Federal Ministry of Education, Science, Research, and Technology and by the State of Bavaria. This study was also funded by the German National Genome Network (NGFNplus number 01GS08134, German Ministry for Education and Research); by the German Federal Ministry of Education and Research (NGFN 01GR0468, PopGen); and 01EW0908 in the frame of ERA-NET NEURON and Helmholtz Alliance Mental Health in an Ageing Society (HA-215), which was funded by the Initiative and Networking Fund of the Helmholtz Association. The French GWAS work was supported by the French National Agency of Research (ANR-08-MNP-012). This study was also funded by France-Parkinson Association, the French program "Investissements d'avenir" funding (ANR-10-IAIHU-06) and a grant from Assistance Publique-Hôpitaux de Paris (PHRC, AOR-08010) for the French clinical data. This study was also sponsored by the Landspitali University Hospital Research Fund (grant to SSv); Icelandic Research Council (grant to SSv); and European Community Framework Programme 7, People Programme, and IAPP on novel genetic and phenotypic markers of Parkinson's disease and Essential Tremor (MarkMD), contract number PIAP-GA-2008-230596 MarkMD (to HP and JHu). This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (http://biowulf.nih.gov), and DNA panels, samples, and clinical data from the National Institute of Neurological Disorders and Stroke Human Genetics Resource Center DNA and Cell Line Repository. People who contributed samples are acknowledged in descriptions of every panel on the repository website. We thank the French Parkinson's Disease Genetics Study Group and the Drug Interaction with genes (DIGPD) study group: Y Agid, M Anheim, A-M Bonnet, M Borg, A Brice, E Broussolle, J-C Corvol, P Damier, A Destée, A Dürr, F Durif, A Elbaz, D Grabil, S Klebe, P. Krack, E Lohmann, L. Lacomblez, M Martinez, V Mesnage, P Pollak, O Rascol, F Tison, C Tranchant, M Vérin, F Viallet, and M Vidailhet. We also thank the members of the French 3C Consortium: A Alpérovitch, C Berr, C Tzourio, and P Amouyel for allowing us to use part of the 3C cohort, and D Zelenika for support in generating the genome-wide molecular data. We thank P Tienari (Molecular Neurology Programme, Biomedicum, University of Helsinki), T Peuralinna (Department of Neurology, Helsinki University Central Hospital), L Myllykangas (Folkhalsan Institute of Genetics and Department of Pathology, University of Helsinki), and R Sulkava (Department of Public Health and General Practice Division of Geriatrics, University of Eastern Finland) for the Finnish controls (Vantaa85+ GWAS data). We used genome-wide association data generated by the Wellcome Trust Case-Control Consortium 2 (WTCCC2) from UK patients with Parkinson's disease and UK control

individuals from the 1958 Birth Cohort and National Blood Service. Genotyping of UK replication cases on ImmunoChip was part of the WTCCC2 project, which was funded by the Wellcome Trust (083948/Z/07/Z). UK population control data was made available through WTCCC1. This study was supported by the Medical Research Council and Wellcome Trust disease centre (grant WT089698/Z/09/Z to NW, JHa, and ASc). As with previous IPDGC efforts, this study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of the Data are available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113, 085475 and 090355. This study was also supported by Parkinson's UK (grants 8047, J-0804, and G-1307) and the Medical Research Council (G0700943 and G1100643). We thank Jeffrey Barrett and Jason Downing for assistance with the design of the ImmunoChip and NeuroX arrays. DNA extraction work that was done in the UK was undertaken at University College London Hospitals, University College London, who received a proportion of funding from the Department of Health's National Institute for Health Research Biomedical Research Centres funding. This study was supported in part by the Wellcome Trust/Medical Research Council Joint Call in Neurodegeneration award (WT089698) to the Parkinson's Disease Consortium (UKPDC), whose members are from the UCL Institute of Neurology, University of Sheffield, and the Medical Research Council Protein Phosphorylation Unit at the University of Dundee. Faraz Faghri's contribution to this work has been supported in part through the RDA/US Data Share grant from the Alfred P. Sloan Foundation #G-2014-13746 and supported by grant 1U54GM114838 awarded by NIGMS through funds provided by the trans-NIH big data to Knowledge (BD2K) initiative (www.bd2k.nih.gov) and a collaborative award from the National Science Foundation (NSF award numbers CNS-1329686, CNS-1329737, CNS-1330142, and CNS-1330491). This work was supported by the Medical Research Council grant MR/N026004/1.

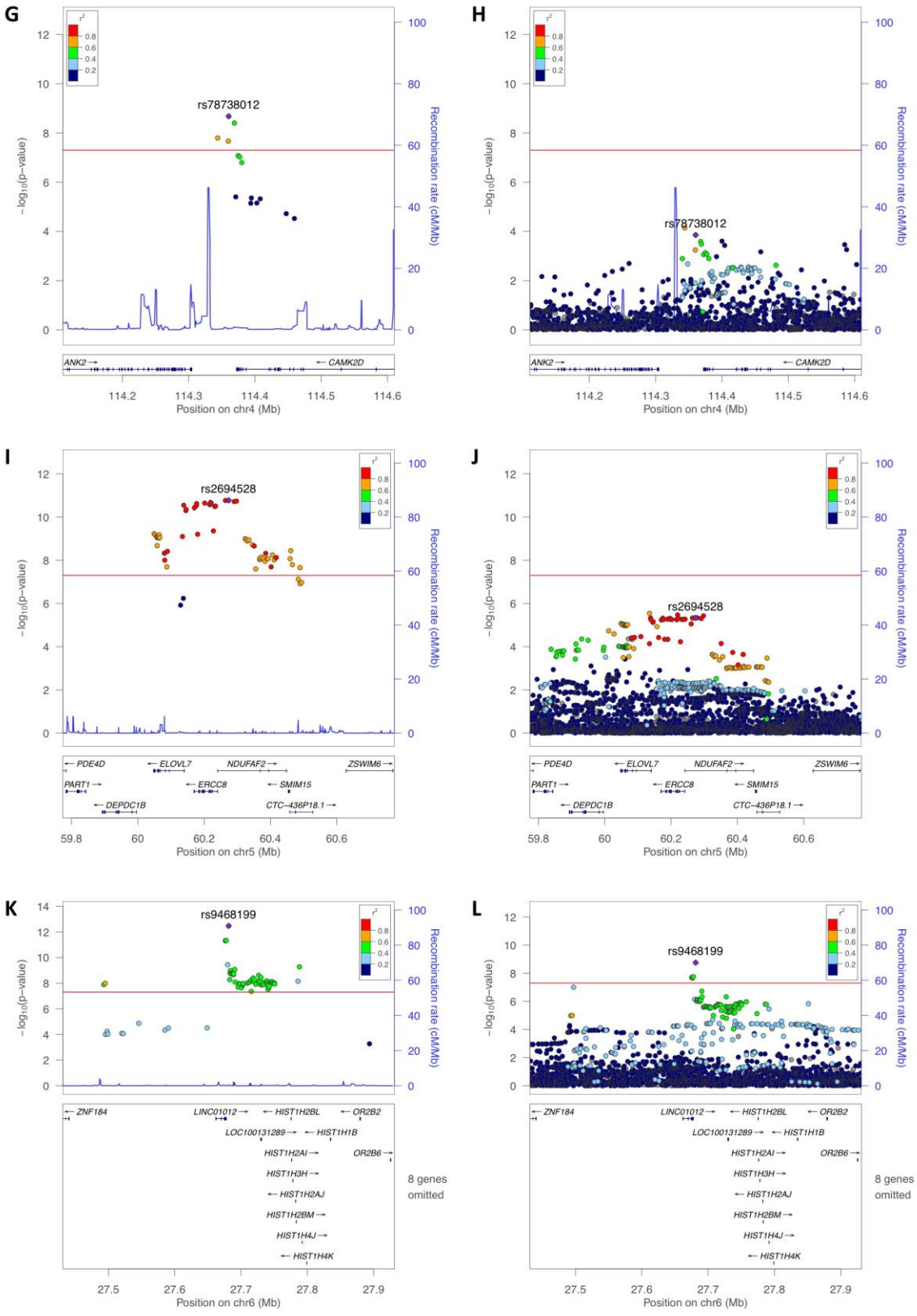
SUPPLEMENTARY FIGURES

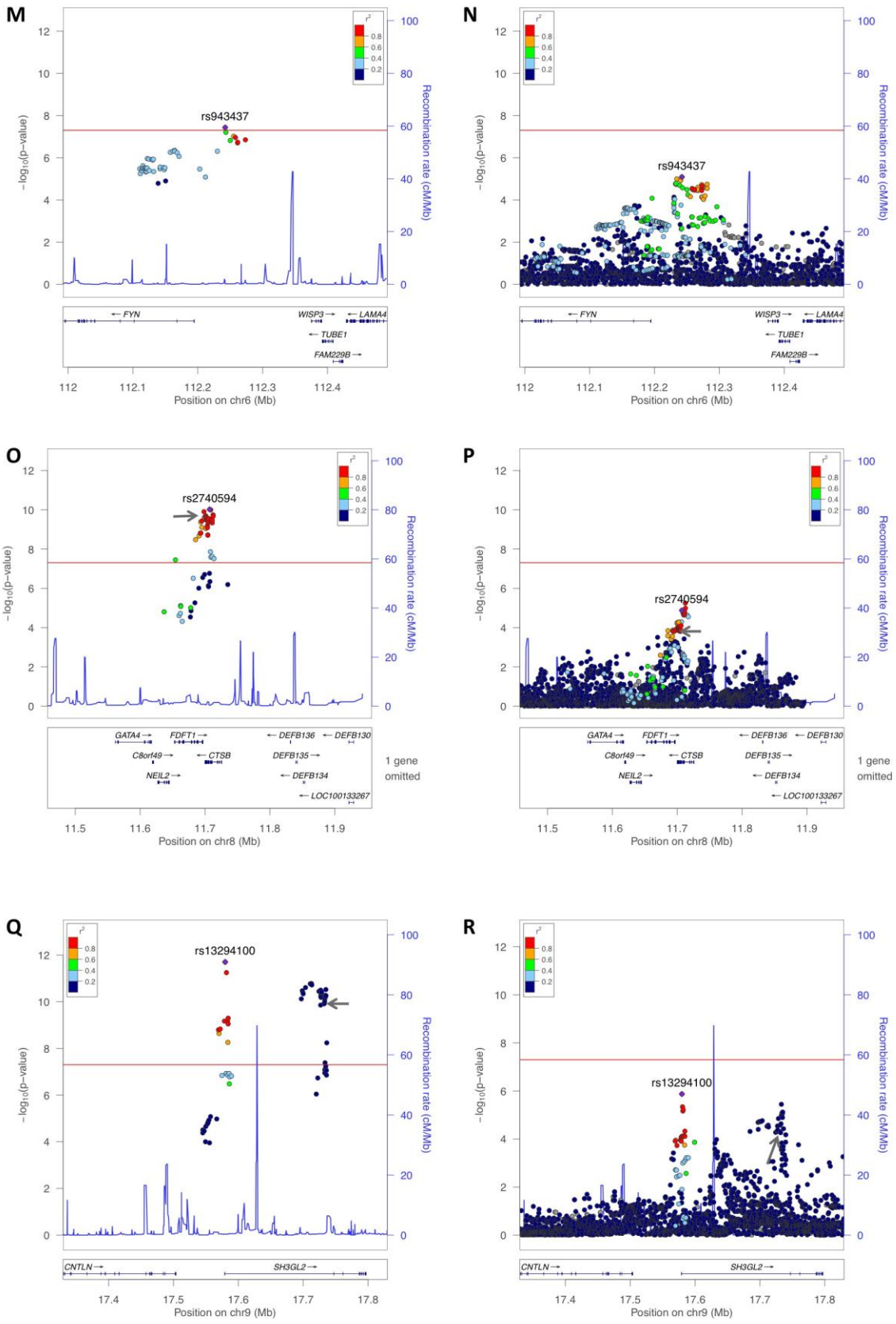


Supplementary Figure 1. QQ-plot of PDWBS p-values. Genome-wide p-values for variants in PDWBS are shown in black circles ("All", 11,933,700 SNPs). Association p-values for the 9,830 variants analyzed in this meta-analysis are shown in dark blue triangles ("PDGene overlap"). Of this subset, those within 1Mb of previously associated regions are displayed in blue squares ("PDGene overlap, w/in 1mb of previously reported"). The 4,706 variants outside of these regions are displayed in red diamonds ("PDGene overlap, not previously reported").

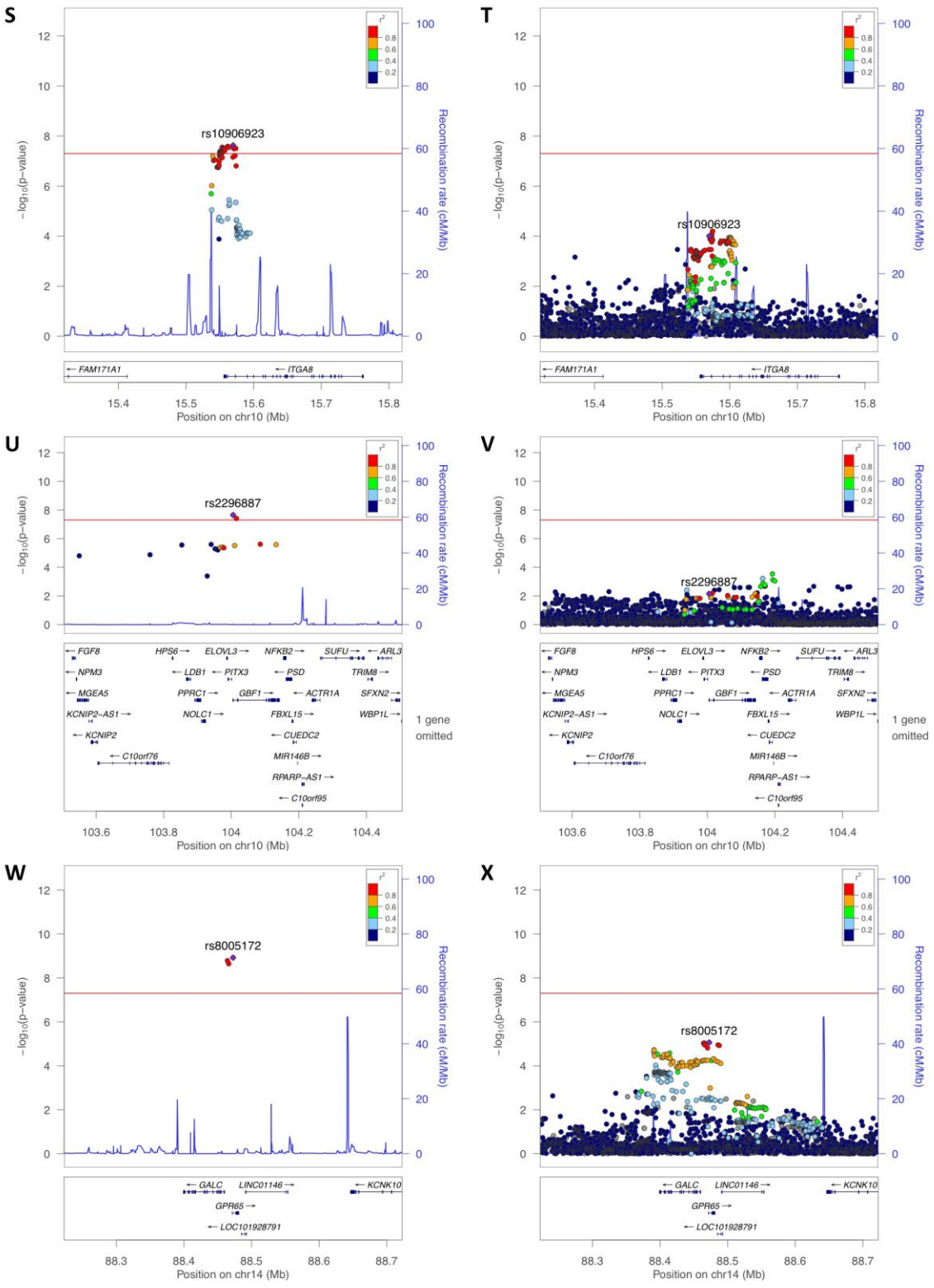


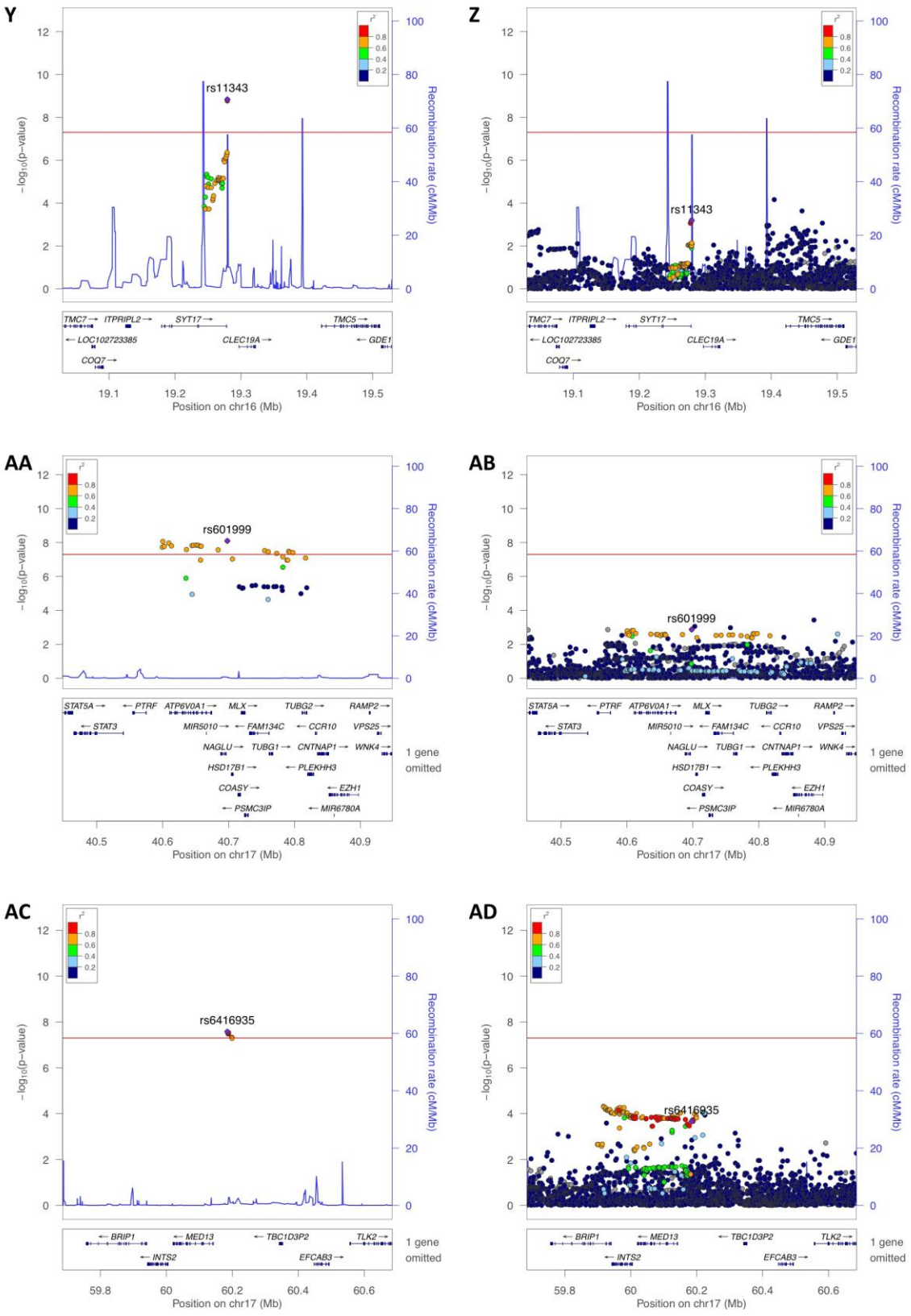
Nature Genetics: doi:10.1038/ng.3955


G
H
I
J
K
L
rs78738012
rs2694528
rs9468199
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Position on chr4 (Mb)
Position on chr5 (Mb)
Position on chr6 (Mb)
ANK2
CAMK2D
PDE4D
ELOVL7
NDUFAF2
ZSWIM6
PART1
ERCC8
SMIM15
DEPDC1B
CTC-436P18.1
ZNF184
LINC01012
HIST1H2BL
OR2B2
LOC100131289
HIST1H1B
HIST1H2AI
OR2B6
HIST1H3H
HIST1H2AJ
HIST1H2BM
HIST1H4J
HIST1H4K
8 genes omitted



Nature Genetics: doi:10.1038/ng.3955

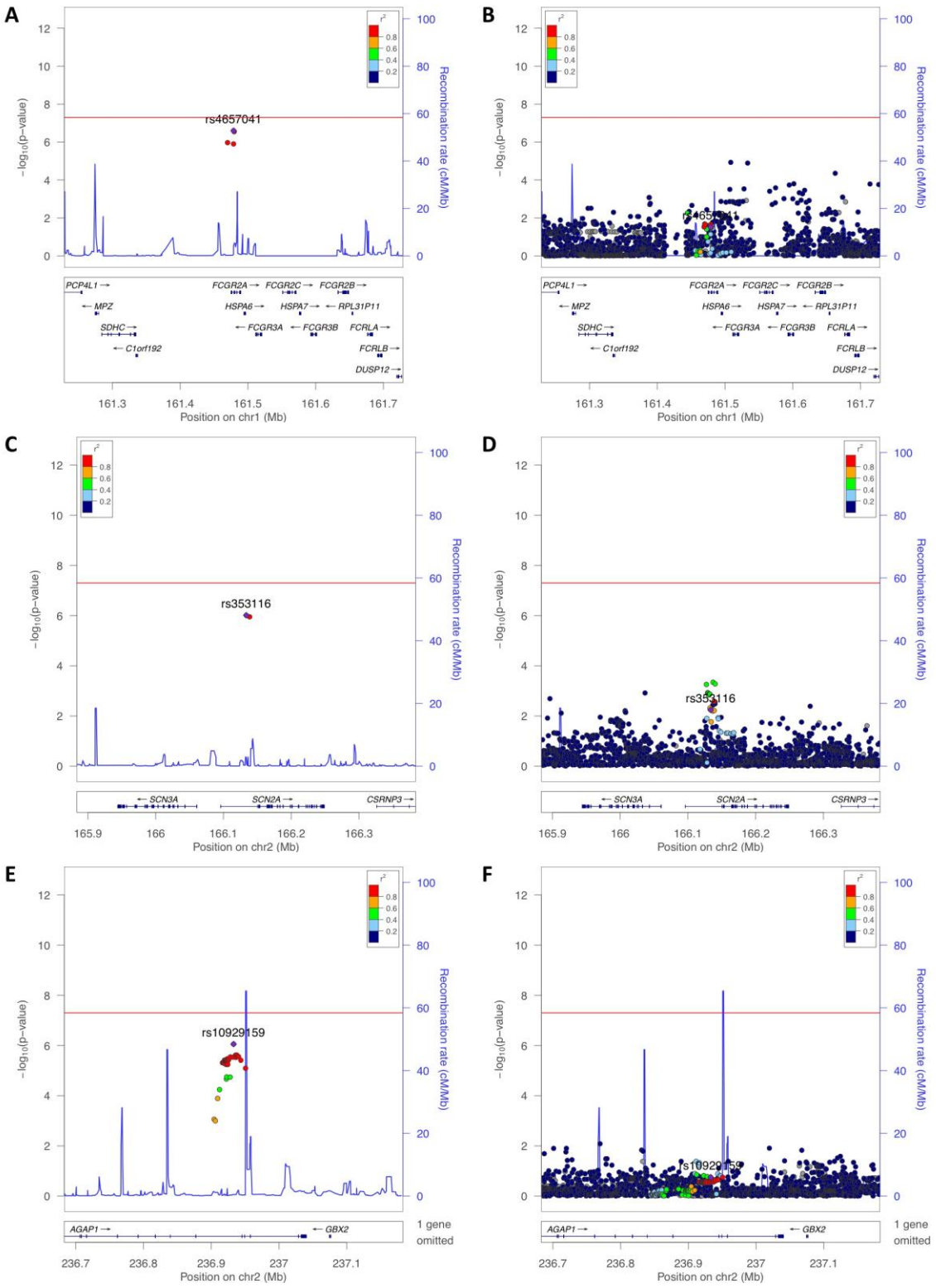

S
T
U
V
W
X
rs10906923
rs2296887
rs8005172
−log10(p−value)
Recombination rate (cM/Mb)
r²
0.8
0.6
0.4
0.2
FAM171A1
ITGA8
Position on chr10 (Mb)
FGF8
HPS6
ELOVL3
NFKB2
SUFU
ARL3
NPM3
LDB1
PITX3
PSD
TRIM8
MGEA5
PPRC1
GBF1
ACTR1A
SFXN2
KCNIP2-AS1
NOLC1
FBXL15
WBP1L
KCNIP2
CUEDC2
C10orf76
MIR146B
RPARP-AS1
C10orf95
1 gene omitted
GALC
LINC01146
KCNK10
GPR65
LOC101928791
Position on chr14 (Mb)

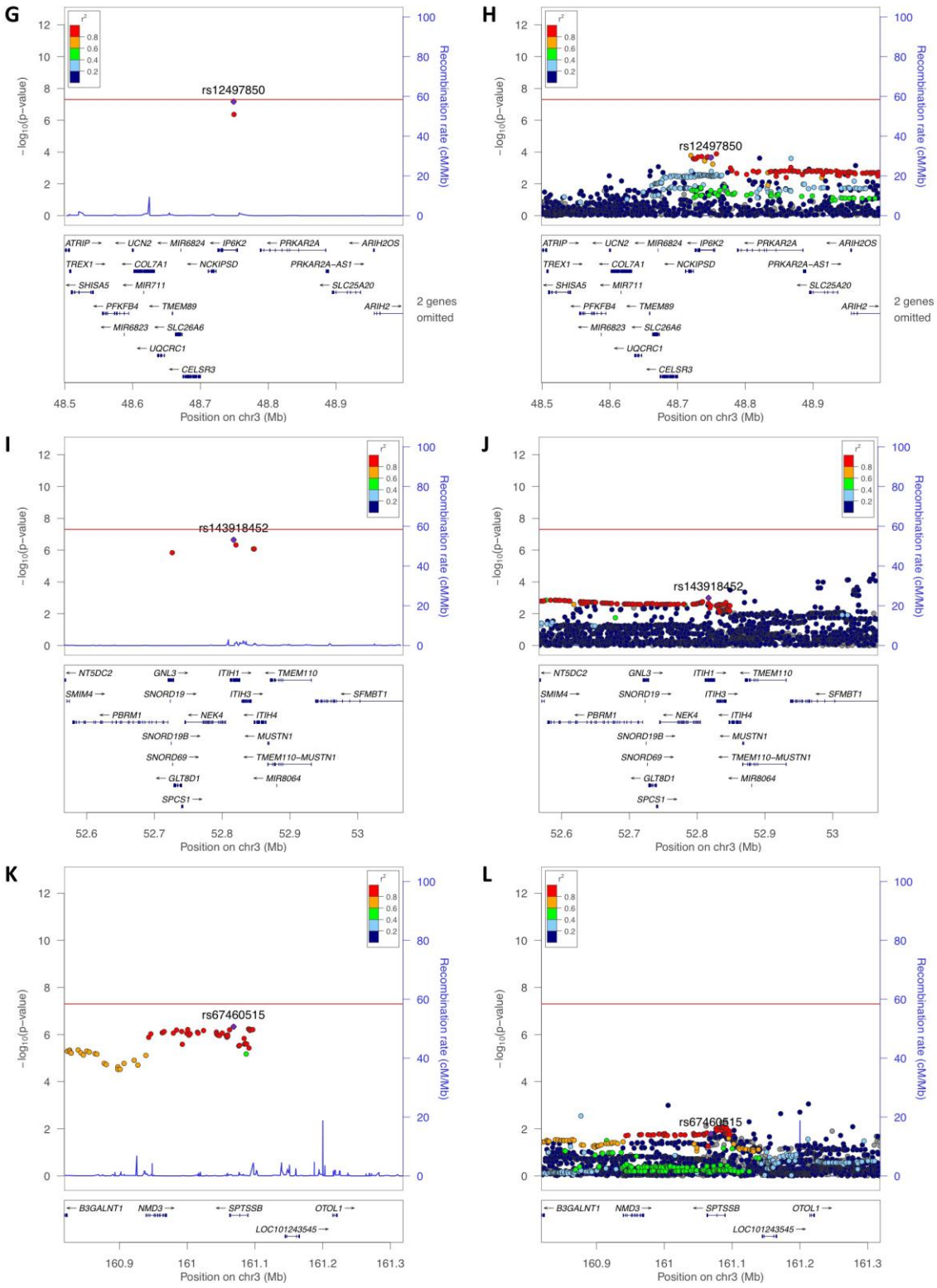


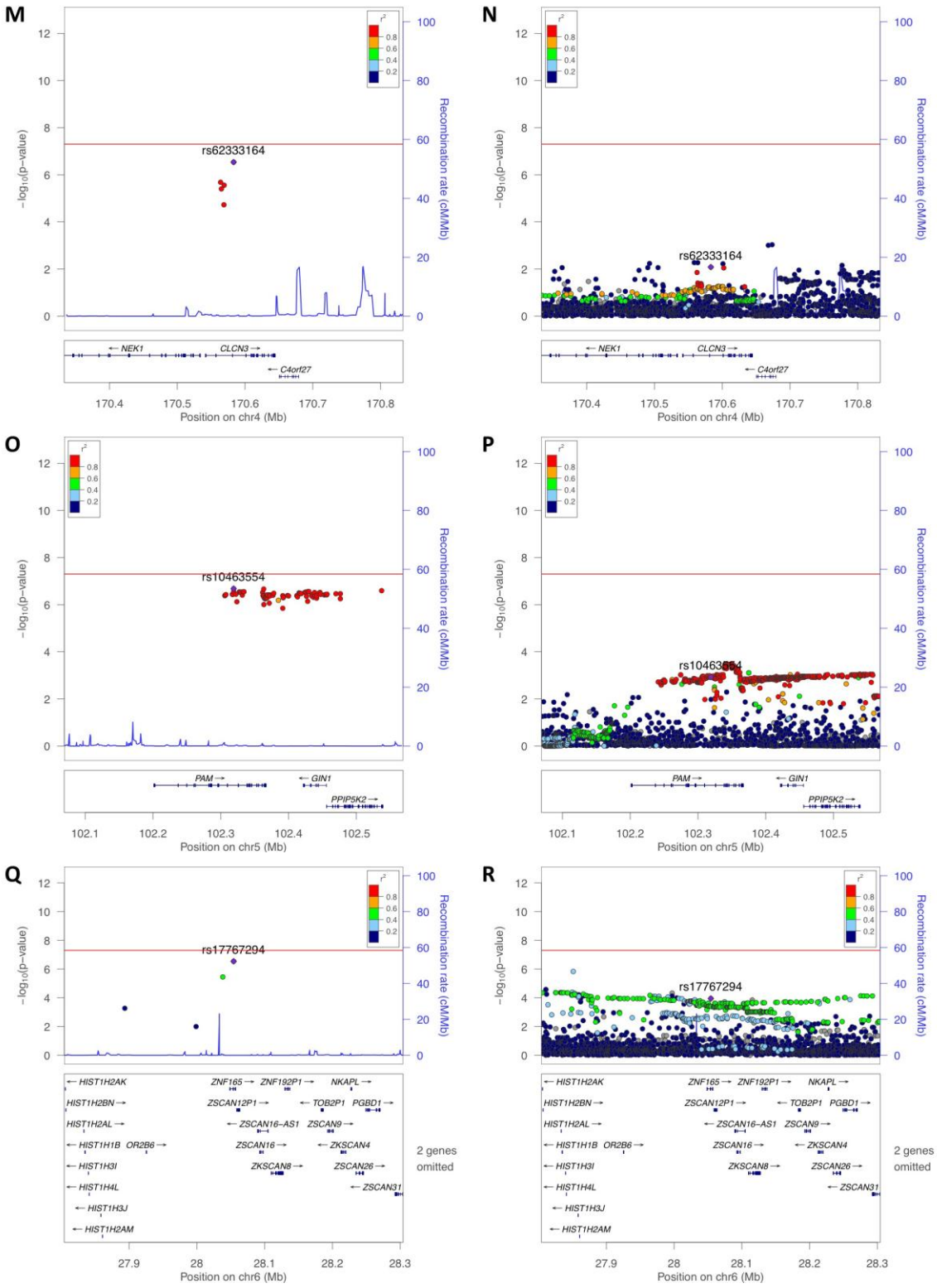
Nature Genetics: doi:10.1038/ng.3955

Supplementary Figure 2. Regional association plots for loci with $P<5x10^{-8}$ in the discovery phase meta-analysis. For the 15 novel loci, regional association plots are displayed for p-values in the meta-analysis (left column), as well as in the PDWBS data alone (right column). The index SNP for each region is labeled and plotted as a purple diamond. Colors for all other variants refer to the LD (r^2) with the index SNP as measured in 1000 Genomes data. Grey refers to missing LD information for this SNP. Recombination rates for the regions are displayed in blue. For each association, all variants within 250kb of the index SNP are shown, with the exception of 4 regions (rs34043159, rs2694528, rs2296887, rs6416935) where variants within 500kb of the index SNP are shown. Grey arrows in (O-R) denote associations previously reported at non-genome-wide significant thresholds (rs1296028 and rs1536076). In two regions, the PDWBS GWAS indicate an alternate top signal. For rs34043159 (C-D), rs11679611 is the top association signal in the PDWBS data and for rs2296887 (U-V), rs9665587 is the top association signal in the PDWBS data. Plots were generated with Locus Zoom [9].

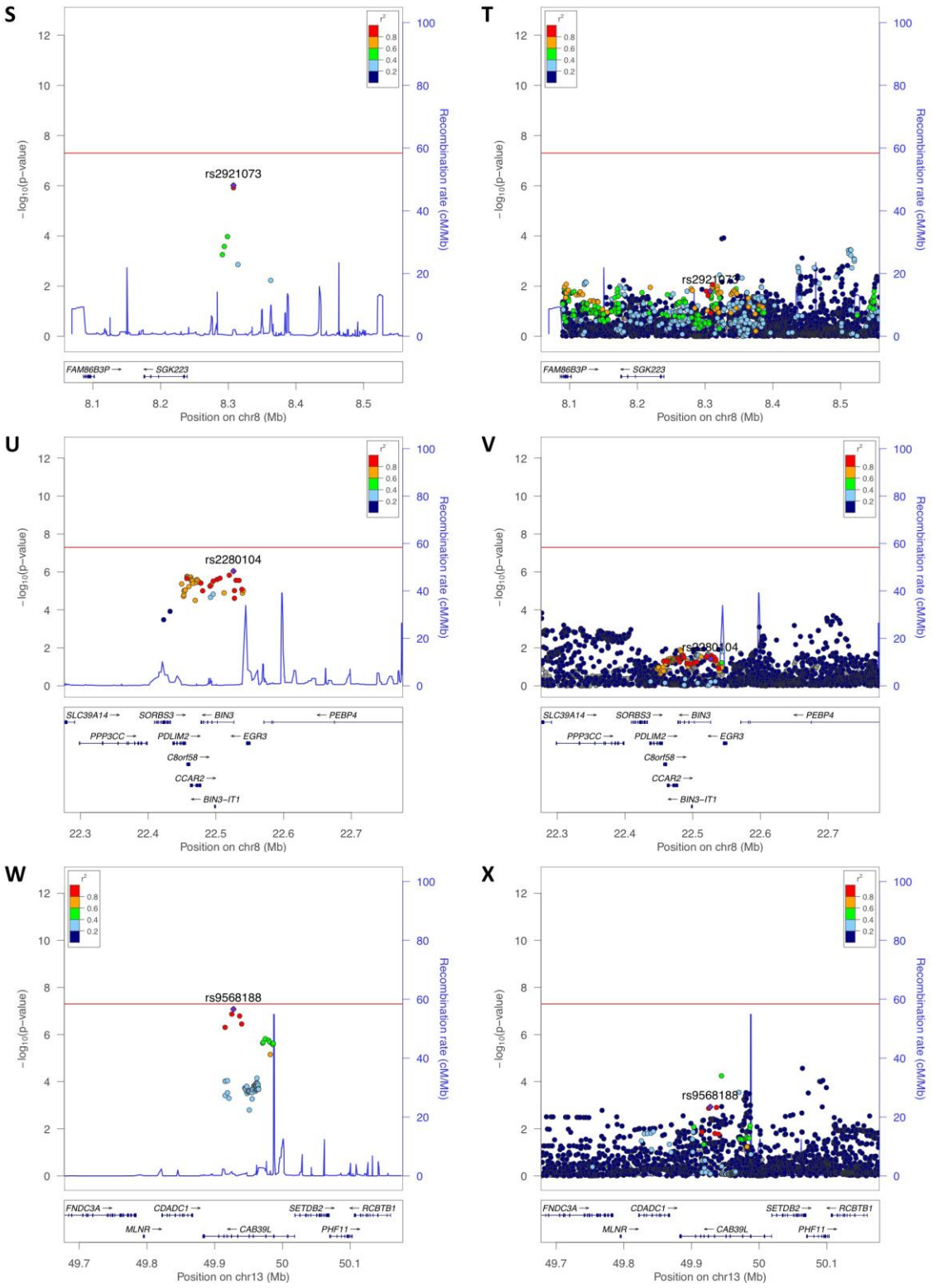
Nature Genetics: doi:10.1038/ng.3955

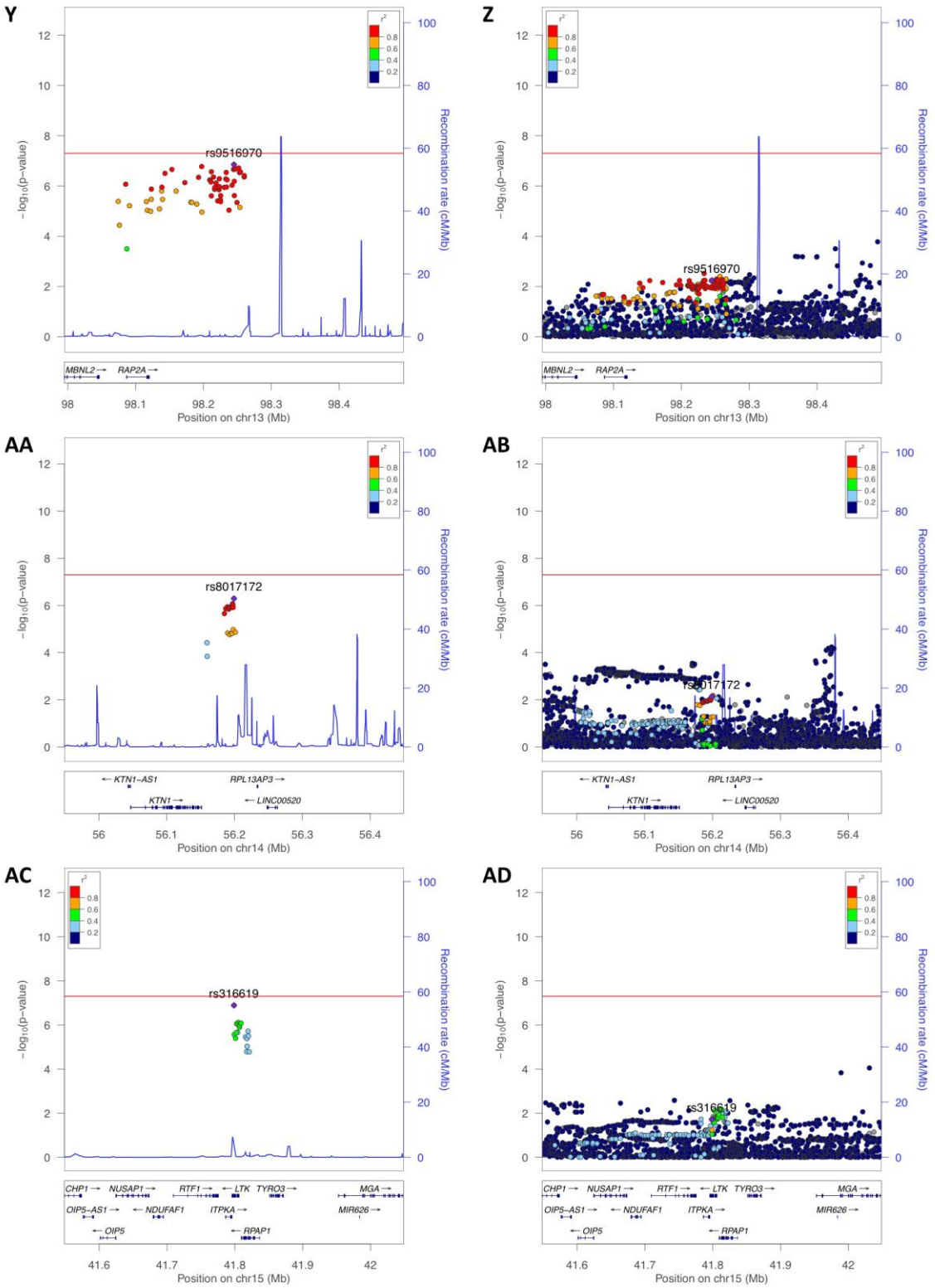

A
B
C
D
E
F
rs4657041
rs353116
rs10929159
−log10(p−value)
Recombination rate (cM/Mb)
Position on chr1 (Mb)
Position on chr2 (Mb)
PCP4L1
MPZ
SDHC
C1orf192
FCGR2A
FCGR2C
FCGR2B
HSPA6
HSPA7
RPL31P11
FCGR3A
FCGR3B
FCRLA
FCRLB
DUSP12
SCN3A
SCN2A
CSRNP3
AGAP1
GBX2
1 gene omitted

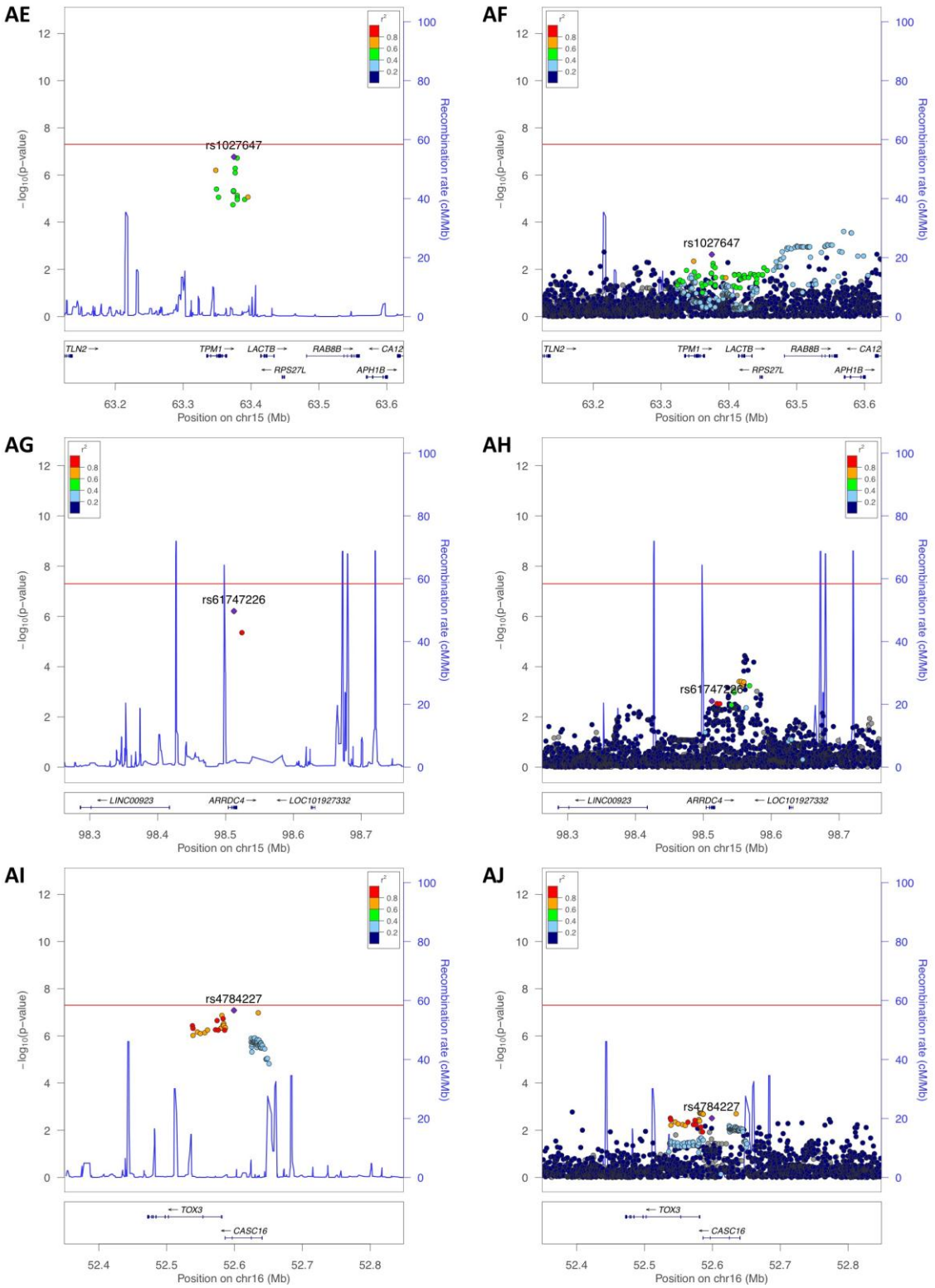

G
H
I
J
K
L
rs12497850
rs143918452
rs67460515
−log10(p−value)
Recombination rate (cM/Mb)
r²
0.8
0.6
0.4
0.2
ATRIP
UCN2
MIR8824
IP6K2
PRKAR2A
ARIH2OS
TREX1
COL7A1
NCKIPSD
PRKAR2A−AS1
SHISA5
MIR711
SLC25A20
PFKFB4
TMEM89
ARIH2
MIR6823
SLC26A6
UQCRC1
CELSR3
2 genes omitted
NT5DC2
GNL3
ITIH1
TMEM110
SMIM4
SNORD19
ITIH3
SFMBT1
PBRM1
NEK4
ITIH4
SNORD19B
MUSTN1
SNORD69
TMEM110−MUSTN1
GLT8D1
MIR8064
SPCS1
B3GALNT1
NMD3
SPTSSB
OTOL1
LOC101243545
Position on chr3 (Mb)

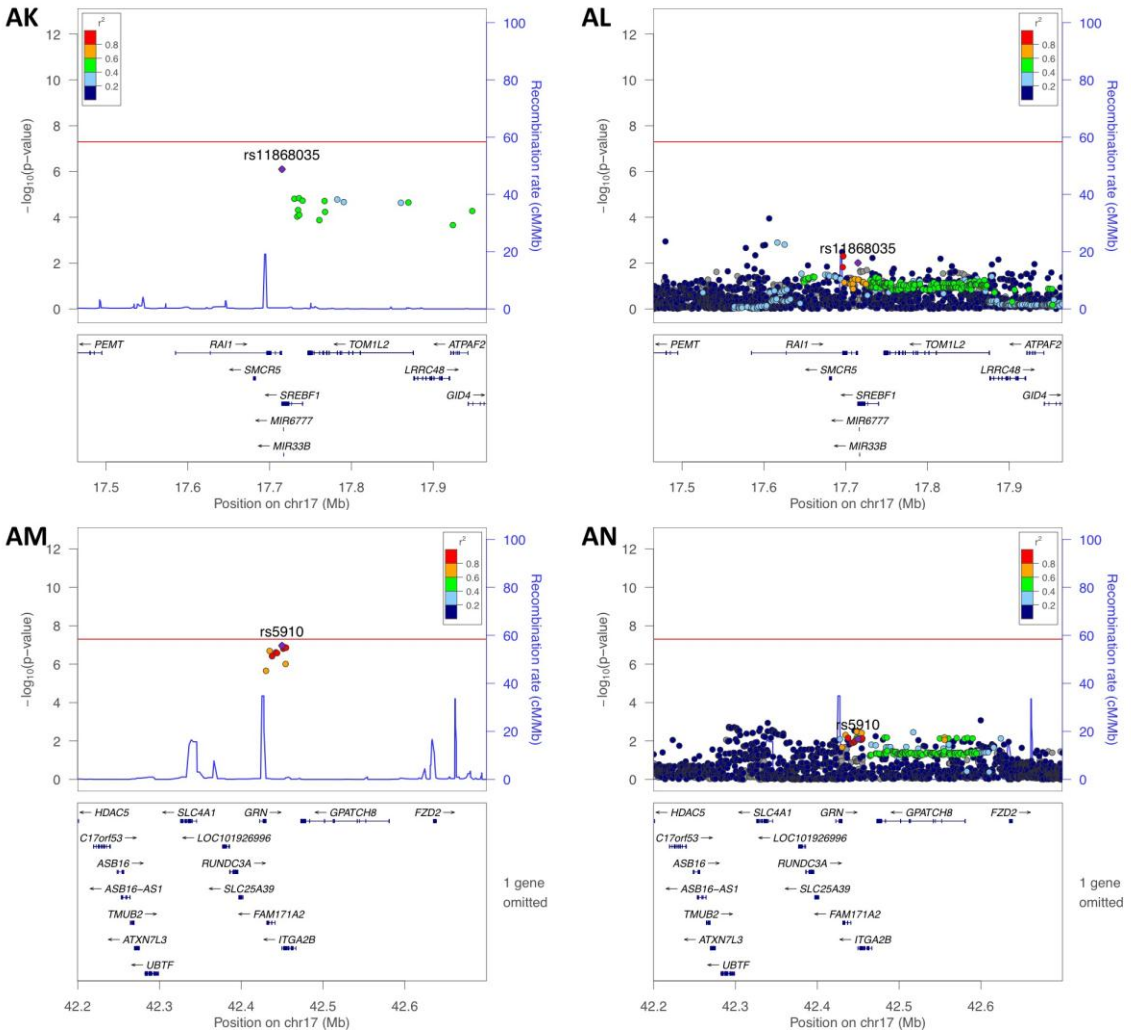


Nature Genetics: doi:10.1038/ng.3955

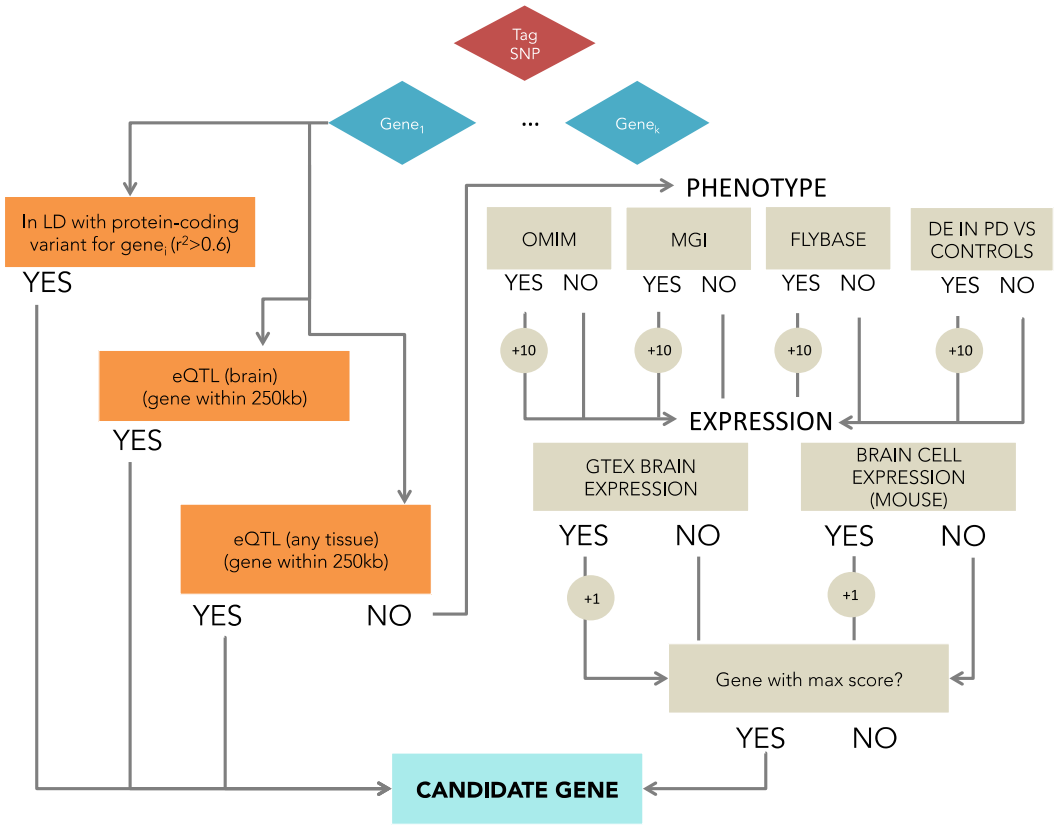

S
T
U
V
W
X
rs2921073
rs2280104
rs9568188
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Position on chr8 (Mb)
Position on chr13 (Mb)
FAM86B3P
SGK223
SLC39A14
SORBS3
BIN3
PEBP4
PPP3CC
PDLIM2
EGR3
C8orf58
CCAR2
BIN3-IT1
FNDC3A
CDADC1
SETDB2
RCBTB1
MLNR
CAB39L
PHF11


Y
Z
AA
AB
AC
AD
rs9516970
rs8017172
rs316619
MBNL2
RAP2A
KTN1-AS1
KTN1
RPL13AP3
LINC00520
CHP1
NUSAP1
RTF1
LTK
TYRO3
MGA
OIP5-AS1
NDUFAF1
ITPKA
MIR626
OIP5
RPAP1
Position on chr13 (Mb)
Position on chr14 (Mb)
Position on chr15 (Mb)
Recombination rate (cM/Mb)
−log10(p−value)

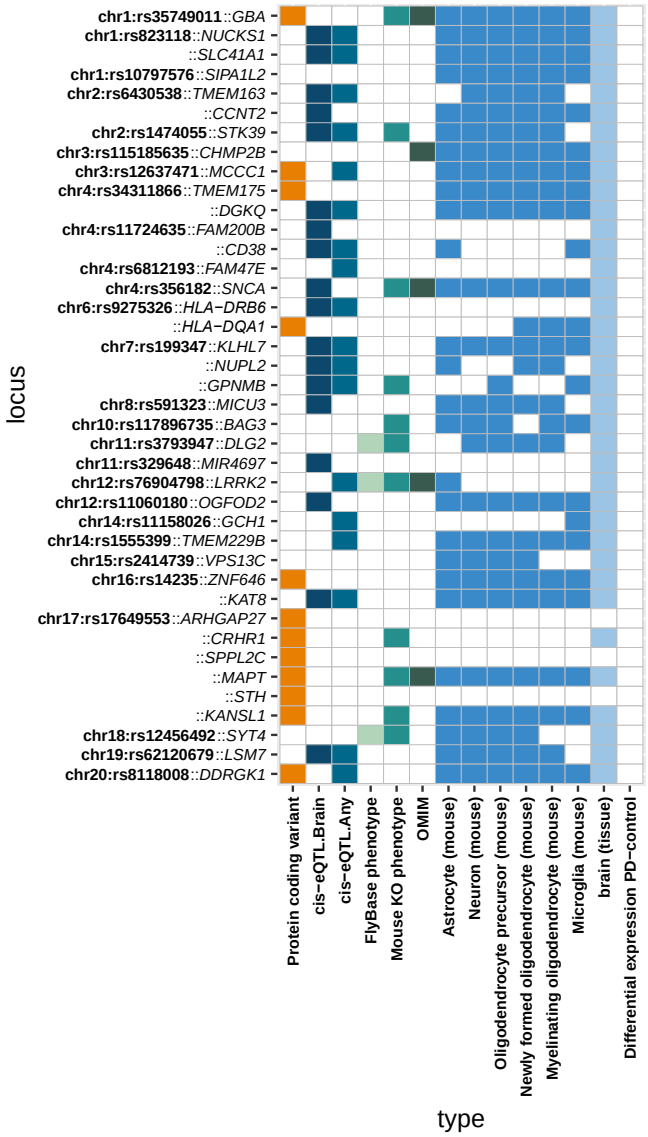

AE
AF
AG
AH
AI
AJ
rs1027647
rs61747226
rs4784227
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Position on chr15 (Mb)
Position on chr16 (Mb)
TLN2
TPM1
LACTB
RAB8B
CA12
RPS27L
APH1B
LINC00923
ARRDC4
LOC101927332
TOX3
CASC16



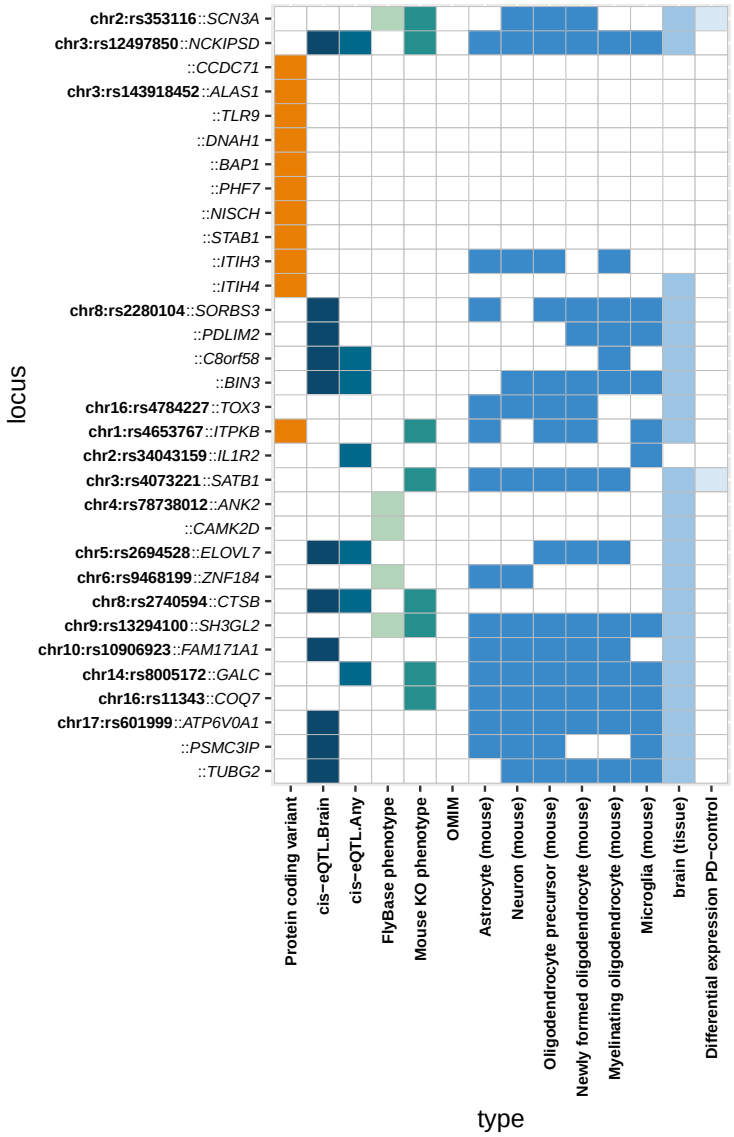
Supplementary Figure 3. Regional association plots for loci with $5x10^{-8}<P<1x10^{-6}$ in the discovery phase meta-analysis.



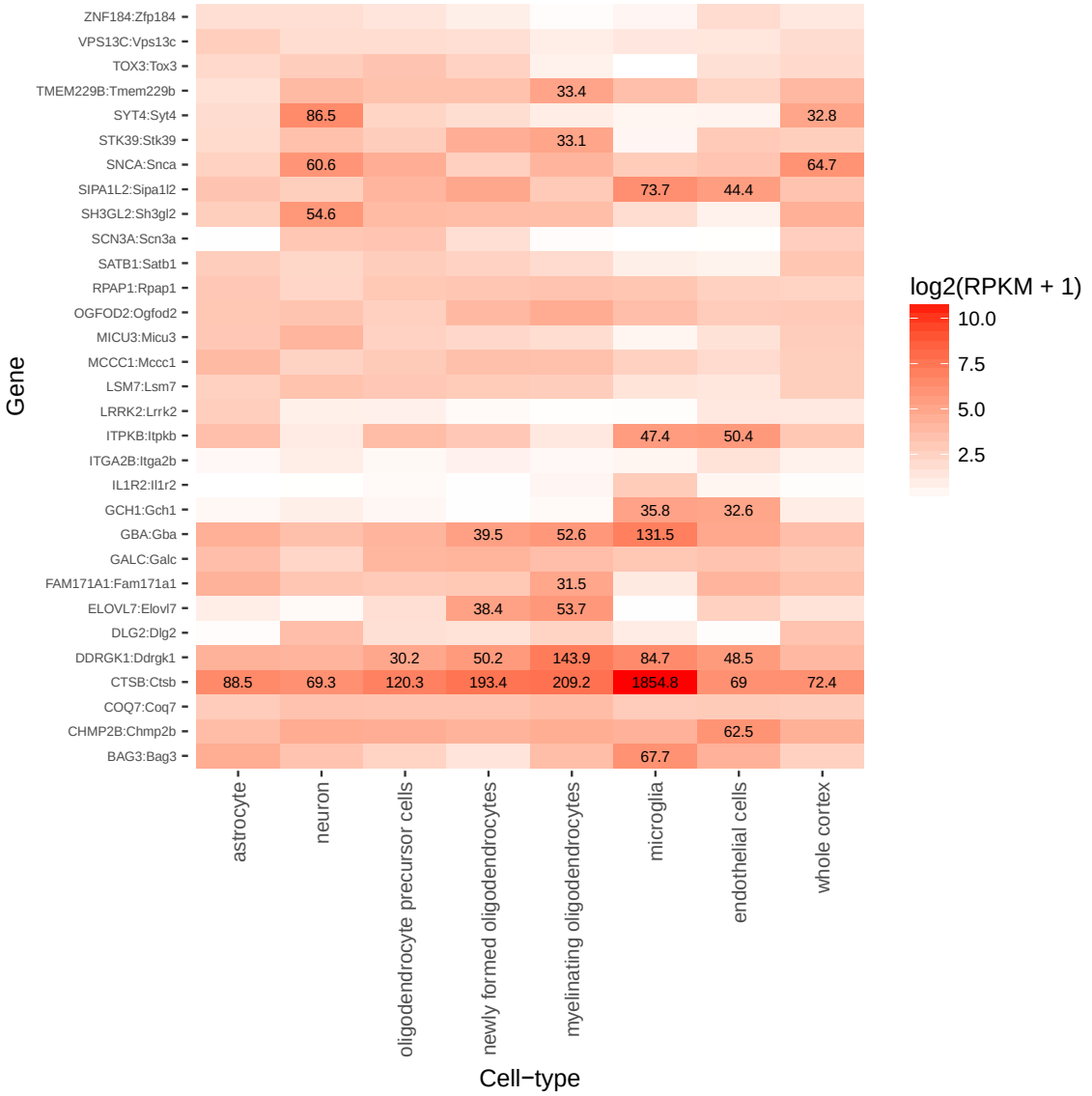
Supplementary Figure 4. Candidate gene nomination pipeline. A description of the candidate gene nomination procedure can be found in the Methods section in the main text.



Supplementary Figure 5. Expanded candidate gene evidence plot for previously reported loci. Each locus is listed by chromosome followed by candidate genes. Loci with tied candidate genes are listed across several lines. Support for a gene with evidence from each category is indicated by a filled cell. Supporting evidence includes whether associated SNPs are cis-eQTLs; whether SNPs are in LD with protein-coding variants; whether genes have a neurological related phenotype in flies ("FlyBase"), mice ("Mouse KO") or humans ("OMIM"); whether genes were expressed in any brain-specific cell type in mouse, or in human brain samples ("GTEx"); and, finally, whether genes were differentially expressed between PD patients and healthy controls.



Supplementary Figure 6. Expanded candidate gene evidence plot for new PD associations. Similar to Supplementary Figure 5, we present here a more detailed view of support for candidate genes of newly discovered PD associations.



Supplementary Figure 7. Mouse brain cell-type expression of PD candidate genes. For PD candidate genes, expression of the mouse homolog (if available) is plotted in specific mouse brain cells, as well as the whole cortex. Colors represent the log2(RPKM+1) of the gene, while the actual RPKM values are displayed for genes with expression >30 RPKMs. No enrichment of PD candidate genes for expression in a specific cell-type is observed.

SUPPLEMENTARY TABLES

Supplementary Table 1. Enrichment statistics for PD heritability across 24 annotations. Data are provided in the corresponding file.

Nature Genetics: doi:10.1038/ng.3955

Cell-type	Proportion SNPs	Proportion h^2 (std. error)	Enrichment (std. error)	Enrichment P-value
Adrenal / Pancreas	0.094	0.535 (0.117)	5.716 (1.253)	4.370×10^{-5}
Cardiovascular	0.111	0.382 (0.140)	3.435 (1.262)	0.036
CNS	0.149	0.559 (0.111)	3.757 (0.746)	0.001
Connective Bone	0.115	0.333 (0.119)	2.897 (1.036)	0.044
GI	0.168	0.323 (0.140)	1.925 (0.837)	0.251
Immune	0.233	0.457 (0.146)	1.960 (0.624)	0.110
Kidney	0.043	0.142 (0.081)	3.332 (1.909)	0.198
Liver	0.072	0.267 (0.111)	3.704 (1.533)	0.068
Other	0.203	0.523 (0.150)	2.582 (0.738)	0.033
SkeletalMuscle	0.104	0.271 (0.126)	2.615 (1.212)	0.187

Supplementary Table 2. Enrichment statistics for PD heritability across 10 cell-type groups.

Significance threshold	Observed # SNPs	Expected # SNPs	P-value
0.1	157	45	7.05×10^{-19}
0.05	128	22	5.68×10^{-21}
0.001	17	0	8.93×10^{-5}
1.00×10^{-6}	1	0	1

Supplementary Table 3. Enrichment of PDGene associations in PDWBS. Within the 9830 SNPs analyzed in the meta-analysis, the number of SNPs observed at various significance thresholds in the PDWBS GWAS is more than expected by chance alone (with the exception of the 1×10^{-6} threshold). After removing SNPs within regions previously associated with PD, the number of SNPs observed at each significance threshold in the PDWBS study was compared to the expected number of SNPs under the null hypothesis of no association (uniform distribution of p-values). P-values are calculated from a chi-squared test.

Supplementary Table 4. Candidate PD associations with $P<1x10^{-6}$ in the discovery meta-analysis. Data are available in the corresponding file.

Supplementary Table 5. NeuroX and joint meta-analysis association statistics for loci with $P<1x10^{-6}$ in the discovery meta-analysis. Grayed out rows represent weaker proxies that were not included in the main manuscript. No suitable proxies were available for rs62333164 and rs4657041. (*) denotes that publicly available data from PDGene for this SNP includes overlapping NeuroX data. Data are available in the corresponding file.

Supplementary Table 6. Variants with regulomeDB scores > 3 that are in LD with novel PD-associated index SNPs. Data are available in the corresponding file.

CHR	BP	Condition SNP	Candidate gene	Most significant SNP after conditioning	$P_{original}$	$P_{condition}$
1	205723572	rs823118	*NUCKS1,SLC41A1*	rs11557080	$1.23x10^{-7}$	$3.99x10^{-4}$
4	951947	rs34311866	*TMEM175,DGKQ*	rs34884217*	$1.38x10^{-6}$	$3.7x10^{-5}$
4	15737101	rs11724635	*FAM200B,CD38*	rs7674301	0.152	$3.75x10^{-7}$
4	77198986	rs6812193	*FAM47E*	rs4859430*	$3.60x10^{-4}$	$5.31x10^{-9}$
4	90626111	rs356182	*SNCA*	rs7681154*	0.171	$2.38x10^{-13}$
6	32666660	rs9275326	*HLA-DRB6,HLA-DQA1*	rs532965	$7.71x10^{-7}$	$1.11x10^{-4}$
9	17579690	rs13294100	*SH3GL2*	rs4961588	$3.88x10^{-5}$	$2.27x10^{-4}$
12	40614434	rs76904798	*LRRK2*	rs28370830	$1.14x10^{-14}$	$4.08x10^{-15}$
12	123303586	rs11060180	*OGFOD2*	rs10847864	$3.31x10^{-13}$	$7.12x10^{-5}$
16	31121793	rs14235	*ZNF646,KAT8*	rs4889685	$4.22x10^{-5}$	$4.02x10^{-4}$

Supplementary Table 7. Loci with residual associations after conditional analysis. Loci with residual associations ($P_{condition} < 1x10^{-3}$) after conditional analysis in the PDWBS data are listed. The most significant SNP in the locus after conditioning is listed along with its original p-value and its p-value after conditioning on the "Condition SNP." (*) denotes the SNP listed is the independent variant reported in the Nalls *et al.* study.

Index	Coding variant	r^2	Gene	Type	Transcript ID	Change
rs9275326	rs9272785	0.61	*HLA-DQA1*	missense	ENST00000343139	p.Ala210Thr
rs8118008	rs2295547	0.66	*DDRGK1*	missense	ENST00000380201	p.Phe298Leu
rs35749011	rs2230288	0.89	*GBA*	missense	ENST00000327247	p.Glu365Lys
rs34311866	rs34311866	1.00	*TMEM175*	missense	ENST00000264771	p.Met393Thr
rs17649553	rs62063857	0.99	*STH*	missense	ENST00000537309	p.Gln7Arg
rs17649553	rs62621252	0.99	*SPPL2C*	missense	ENST00000329196	p.Ser224Pro
rs17649553	rs63750417	0.99	*MAPT*	missense	ENST00000344290	p.Pro202Leu
rs17649553	rs34579536	0.99	*KANSL1*	missense	ENST00000262419	p.Ile1085Thr
rs17649553	rs16940681	0.99	*CRHR1*	missense	ENST00000339069	p.Glu280Gln
rs17649553	rs12949256	0.62	*ARHGAP27*	missense	ENST00000428638	p.Ala117Thr
rs14235	rs749670	0.89	*ZNF646*	missense	ENST00000394979	p.Glu327Gly
rs12637471	rs2270968	0.69	*MCCC1*	missense	ENST00000492597	p.His355Pro
rs4653767	rs147889095	0.85	*ITPKB*	disruptive inframe deletion	ENST00000429204	p.Ser93_Ser95del
rs62333164	rs34755939	0.62	*CLCN3*	missense	ENST00000504131	p.Ala15Gly
rs61747226	rs61747226	1.00	*ARRDC4*	missense	ENST00000268042	p.Thr235Met
rs5910	rs5911	0.98	*ITGA2B*	missense	ENST00000262407	p.Ile874Ser
rs4657041	rs1801274	1.00	*FCGR2A*	missense	ENST00000271450	p.His167Arg
rs143918452	rs5743845	1.00	*TLR9*	missense	ENST00000360658	p.Arg863Gln
rs143918452	rs150311081	1.00	*STAB1*	missense	ENST00000321725	p.Arg1019His
rs143918452	rs201539522	1.00	*PHF7*	splice acceptor variant & intron variant	ENST00000327906	NA
rs143918452	rs61736838	1.00	*NISCH*	missense	ENST00000345716	p.Ala1028Val
rs143918452	rs151083454	1.00	*ITIH4*	missense	ENST00000266041	p.Gly893Ser
rs143918452	rs148156289	1.00	*ITIH3*	missense	ENST00000449956	p.Met757Ile
rs143918452	rs61729450	1.00	*DNAH1*	missense	ENST00000420323	p.Ile190Thr
rs143918452	rs146661777	1.00	*BAP1*	missense	ENST00000478368	p.Asp145His
rs143918452	rs35338461	1.00	*ALAS1*	missense	ENST00000394965	p.Arg13Gln
rs12497850	rs4955419	0.71	*CCDC71*	missense	ENST00000321895	p.Gln317Leu
rs11868035	rs11649804	0.94	*RAI1*	missense	ENST00000353383	p.Pro165Thr
rs10463554	rs34813	0.92	*GIN1*	missense	ENST00000399004	p.Thr239Met

Supplementary Table 8. Functional variants for PD associated loci. Functional variants are listed for PD associated loci, in addition to the r^2 between the index variant and the coding variant (measured in 1000 Genomes) the Ensembl transcript ID, as well as the corresponding amino acid change (if applicable) is noted. If the variant impacted several transcripts, only the longest transcript is listed. If the index variant is in LD with multiple coding variants for the same gene only the one with the largest r^2 is listed.

Supplementary Table 9. Cis-eQTLs for PD associated loci. For PD associated regions with significant cis-eQTLs (as displayed in Figure 3), the brain and non-brain tissue with the most significant p-value is listed for each PD loci. "Risk expression" considers the direction of expression the risk allele is associated with. Data are available in the corresponding file.

Supplementary Table 10. Genes within 250kb of PD associated loci with evidence of co-expression and/or protein-protein interactions. Data are available in the corresponding file.

Category	P-value	Obs.	Exp.	Candidate genes in these gene sets
Lysosome				
Previously reported	1.33×10^{-3}	2	0.22	*GBA , TMEM175*
Novel associations	3.64×10^{-5}	3	0.19	*CTSB, GALC, ATP6V0A1*
All	3.35×10^{-6}	5	0.40	
Autophagy				
Previously reported	1.09×10^{-3}	2	0.20	*BAG3, KAT8*
Novel associations	NA	1	0.18	*CTSB*
All	5.71×10^{-3}	3	0.38	
Mitochondrial				
Previously reported	NA	1	1.03	*MCCC1*
Novel associations	0.058	2	0.89	*COQ7, ALAS1*
All	0.13	3	1.92	

Supplementary Table 11. Targeted enrichment analysis of PD associations. PD associations are enriched for genes in the lysosomal and autophagy pathways. P-values were not calculated for gene-sets with less than 2 observed genes. Expected number of genes ("Exp.") is calculated as (#genes in the gene-set)/(# genes in the background set)*(# genes associated with PD).

Supplementary Table 12. General gene-set enrichment analysis of PD associations. Data are available in the corresponding file.

Nature Genetics: doi:10.1038/ng.3955

Supplementary Table 13. All protein-coding genes within 250 kb of PD-associated loci used as input for STRING analysis. Data are available in the corresponding file.

Supplementary Table 14. Minimum p-values for the 42 genes in the lysosomal pathway for which SNPs tested in this meta-analysis were within 250kb. Data are available in the corresponding file.

Nature Genetics: doi:10.1038/ng.3955

Supplementary Table 15. Minimum p-values for 117 genes in the mitochondrial pathway for which SNPs tested in this meta-analysis were within 250kb of. Data are available in the corresponding file.

Supplementary Table 16. Minimum p-values for 25 genes in the autophagy pathway for which SNPs tested in this meta-analysis were within 250kb of. Data are available in the corresponding file.

REFERENCES

1. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nature genetics. 2006;38(2):209-13. doi: 10.1038/ng1706. PubMed PMID: 16415888.
2. International Parkinson Disease Genomics C, Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet. 2011;377(9766):641-9. doi: 10.1016/S0140-6736(10)62345-8. PubMed PMID: 21292315; PubMed Central PMCID: PMCPMC3696507.
3. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nature genetics. 2014;46(9):989-93. doi: 10.1038/ng.3043. PubMed PMID: 25064009; PubMed Central PMCID: PMCPMC4146673.
4. International Parkinson's Disease Genomics C, Wellcome Trust Case Control C. A two-stage meta-analysis identifies several new loci for Parkinson's disease. PLoS genetics. 2011;7(6):e1002142. doi: 10.1371/journal.pgen.1002142. PubMed PMID: 21738488; PubMed Central PMCID: PMCPMC3128098.
5. Nalls MA, McLean CY, Rick J, Eberly S, Hutten SJ, Gwinn K, et al. Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: a population-based modelling study. Lancet Neurol. 2015;14(10):1002-9. doi: 10.1016/S1474-4422(15)00178-7. PubMed PMID: 26271532; PubMed Central PMCID: PMCPMC4575273.
6. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-45. PubMed PMID: 3203132.
7. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PubMed Central PMCID: PMCPMC4302049.
8. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS genetics. 2014;10(5):e1004383. doi: 10.1371/journal.pgen.1004383. PubMed PMID: 24830394; PubMed Central PMCID: PMCPMC4022491.
9. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336-7. doi: 10.1093/bioinformatics/btq419. PubMed PMID: 20634204; PubMed Central PMCID: PMCPMC2935401.