SUPPLEMENTARY MATERIAL: Genome-wide association studies of cognitive and motor progression in Parkinson's disease

Supplementary Figures

Supplementary Figure 1. First two genetic principal components plotted for each cohort, from the final genetic PCA (after removing outliers in two passes). 3

Supplementary Figure 2. First two principal components, generated from a PCA merged with HapMap data. 3

Supplementary Figure 3. Scree plot and plot showing the proportion of variance explained in the motor progression principal component analysis. 4

Supplementary Figure 4. Correlation matrix between the random slopes for the motor progression variables. 4

Supplementary Figure 5. Scree plot and plot showing the proportion of variance explained in the cognitive progression principal component analysis. 5

Supplementary Figure 6.Correlation matrix between the random slopes for the cognitive progression variables. 5

Supplementary Figure 7. Scree plot and plot showing the proportion of variance explained in the composite (cross-domain - both motor and cognitive variables) progression principal component analysis. 6

Supplementary Figure 8. Composite progression GWAS: Manhattan plot for the gene-based test from MAGMA. 7

Supplementary Figure 9. QQ plots for the GWAS of composite progression (variant-based and gene-based tests). 7

Supplementary Figure 10. Manhattan plot of the gene-based test for motor progression from MAGMA. 8

Supplementary Figure 11.QQ plots for the GWAS of motor progression (variant-based and gene-based tests). 8

Supplementary Figure 12.Manhattan plot of the gene-based test for cognitive progression from MAGMA. 9

Supplementary Figure 13. QQ plots for the GWAS of cognitive progression (variant-based and gene-based tests). 9

Supplementary Figure 14. Forest plots for the top SNPs for motor and cognitive progression, showing effect sizes of the top SNPs in different cohorts. 10

Supplementary Figure 15. Heatmap of the PD GWAS risk loci and their association with composite, motor, or cognitive progression. 11

Supplementary Figure 16. Heatmap of candidate variants and their association with composite, motor, or cognitive progression. 12

Supplementary Tables

Supplementary Table 1. Correlation between motor progression principal components and cognitive progression principal components. 13

Supplementary Table 2. Correlation between first principal component from combined progression PCA and random slopes from individual measures. 13

Supplementary Table 3. Marginal R2 from mixed effects models showing the relationship between the raw scales (before percentage transformation and standardisation) and the composite progression score, in each cohort separately. 13

Supplementary Table 4. Top 10 independent SNPs from the GWAS of composite progression. 14

Supplementary Table 5. Top 10 independent SNPs from the GWAS of motor progression. 15

Supplementary Table 6. Motor progression GWAS performed in each cohort separately. ...16

Supplementary Table 7. Motor progression GWAS performed for each scale separately. ...16

Supplementary Table 8. Top 10 independent SNPs from the GWAS of cognitive progression. 17

Supplementary Table 9. Cognitive progression GWAS performed in each cohort separately. 18

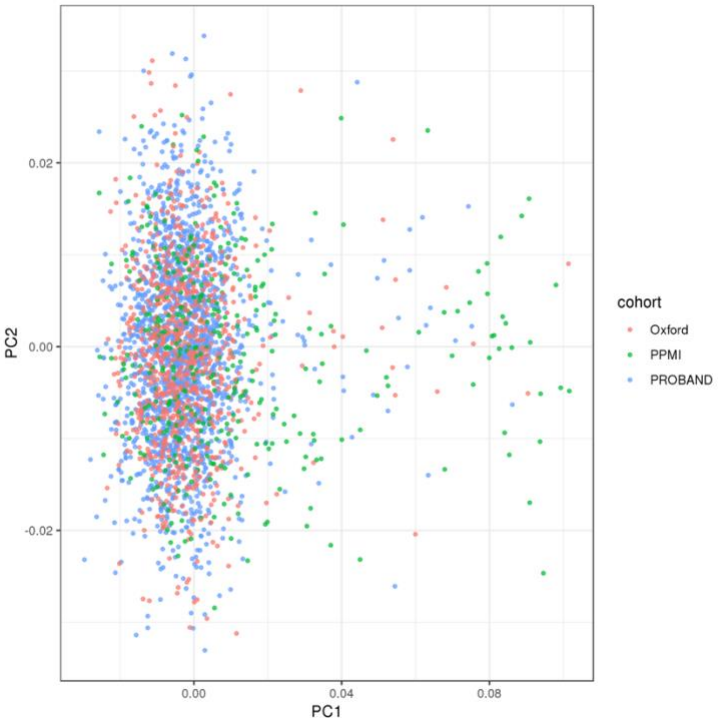
Supplementary Table 10. Cognitive progression GWAS performed for each scale separately. 18

Supplementary Table 11. PD risk SNPs missing for calculation of the genetic risk score, and whether a proxy was identified. 19

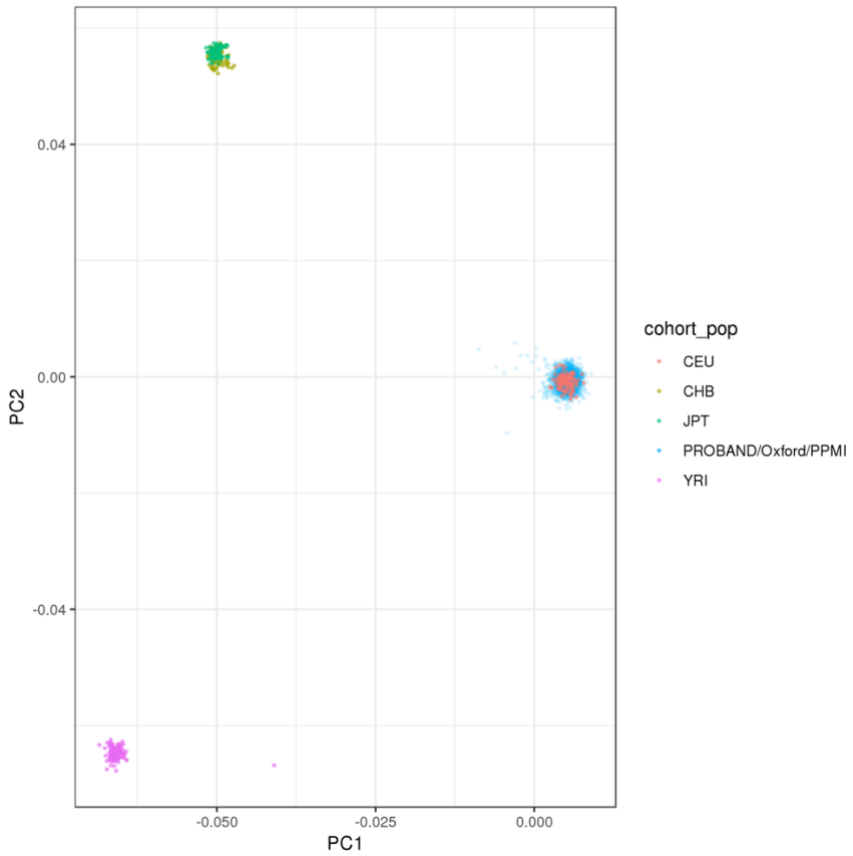
Supplementary Table 12. Sensitivity analysis excluding PD cases with less than 90% diagnostic certainty. 20

Supplementary Table 13. Sensitivity analysis excluding fastest and slowest progressing cases (top and bottom 5% of each distribution) 21

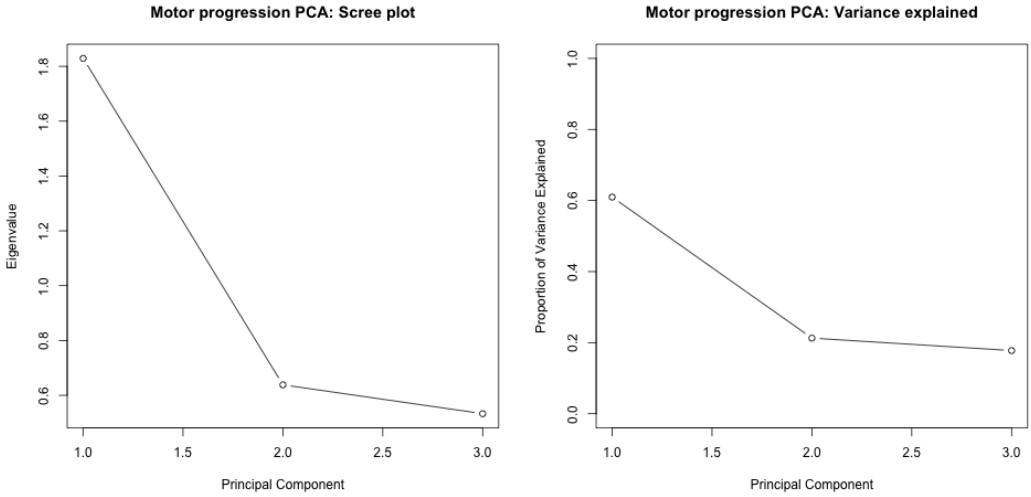
Supplementary Figure 1. First two genetic principal components plotted for each cohort, from the final genetic PCA (after removing outliers in two passes).



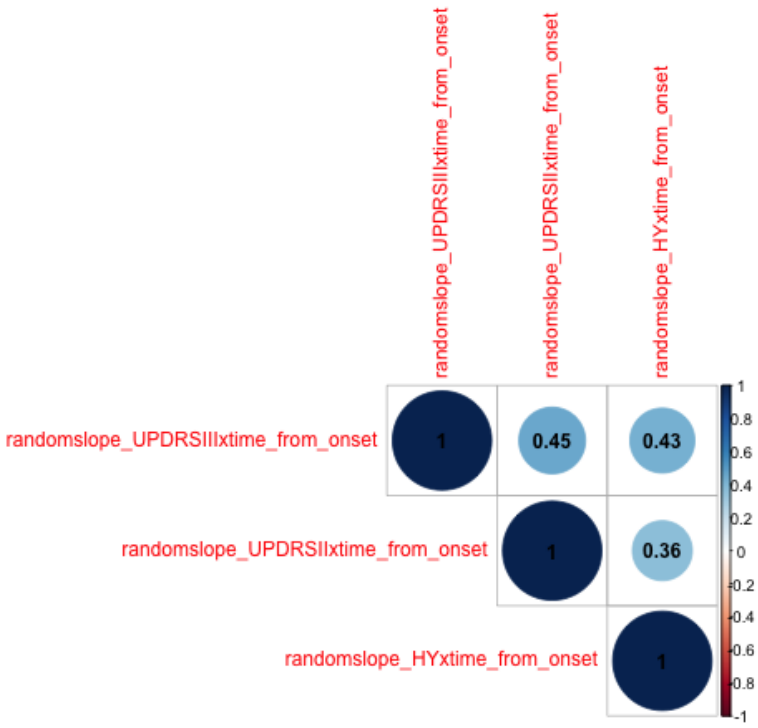
Supplementary Figure 2. First two principal components, generated from a PCA merged with HapMap data.



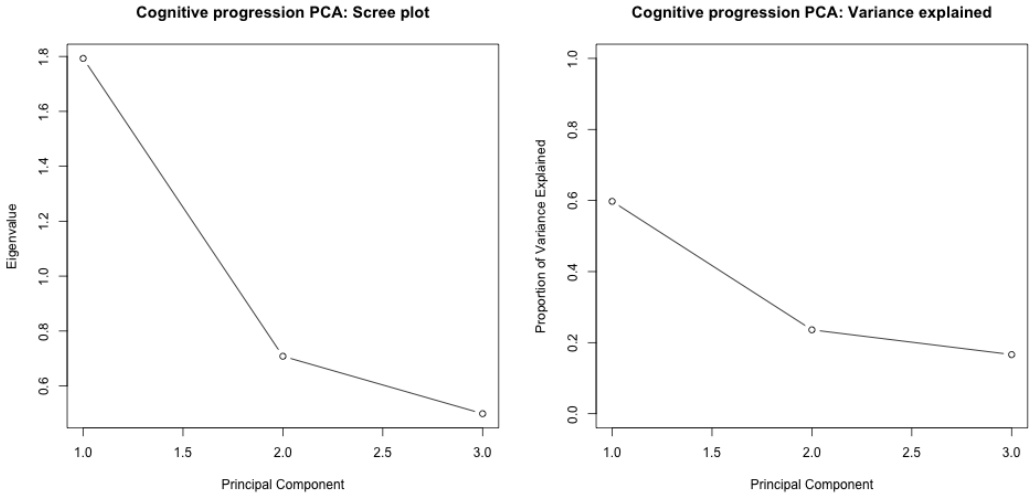
Supplementary Figure 3. Scree plot and plot showing the proportion of variance explained in the motor progression principal component analysis.



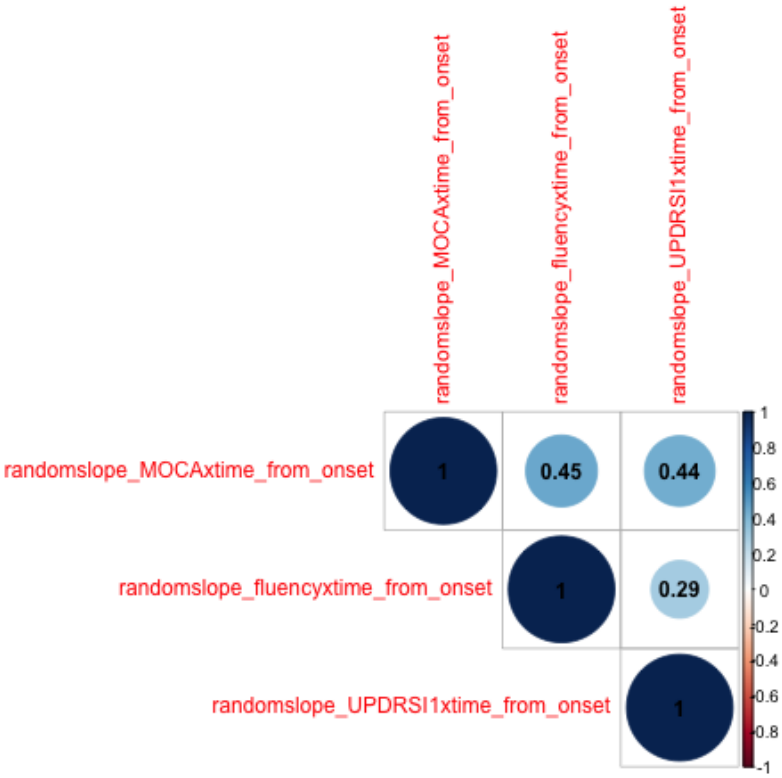
Supplementary Figure 4. Correlation matrix between the random slopes for the motor progression variables.



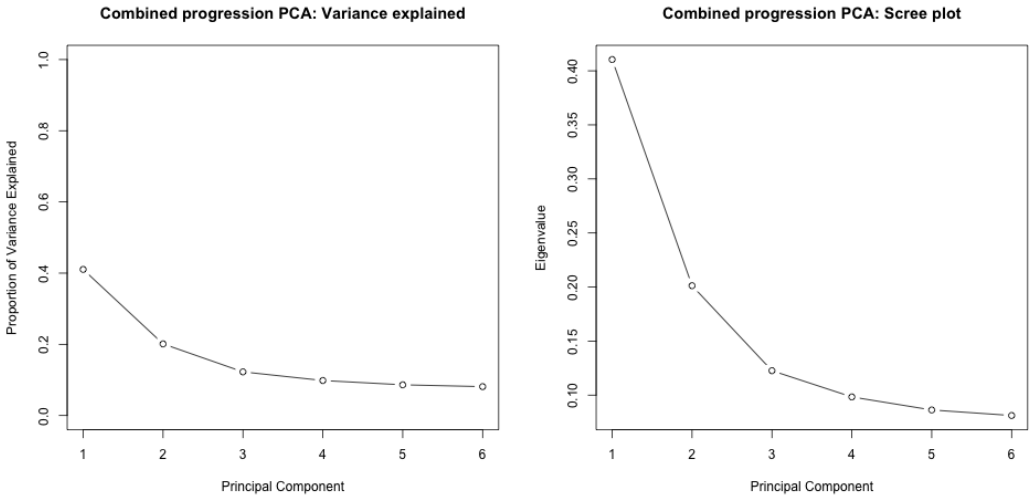
Supplementary Figure 5. Scree plot and plot showing the proportion of variance explained in the cognitive progression principal component analysis.



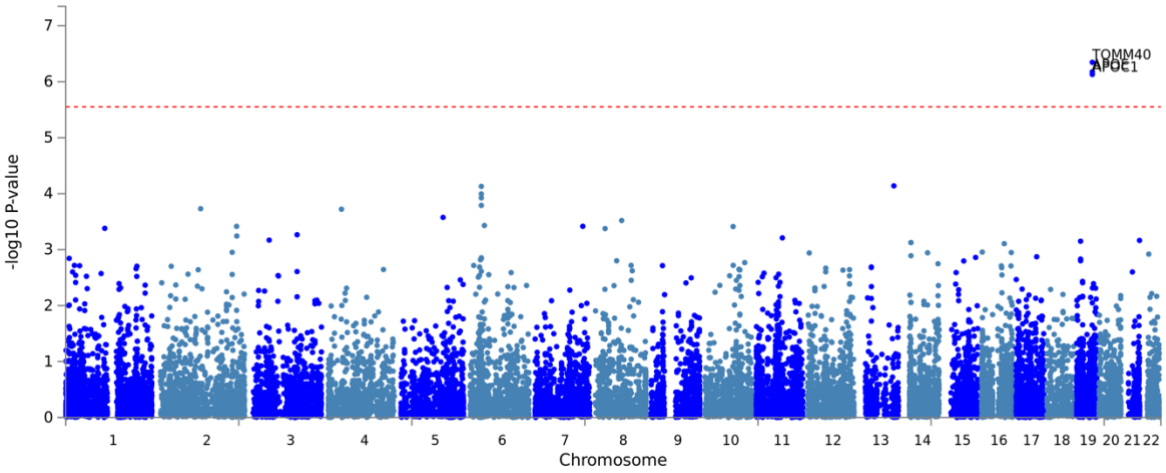
Supplementary Figure 6.Correlation matrix between the random slopes for the cognitive progression variables.



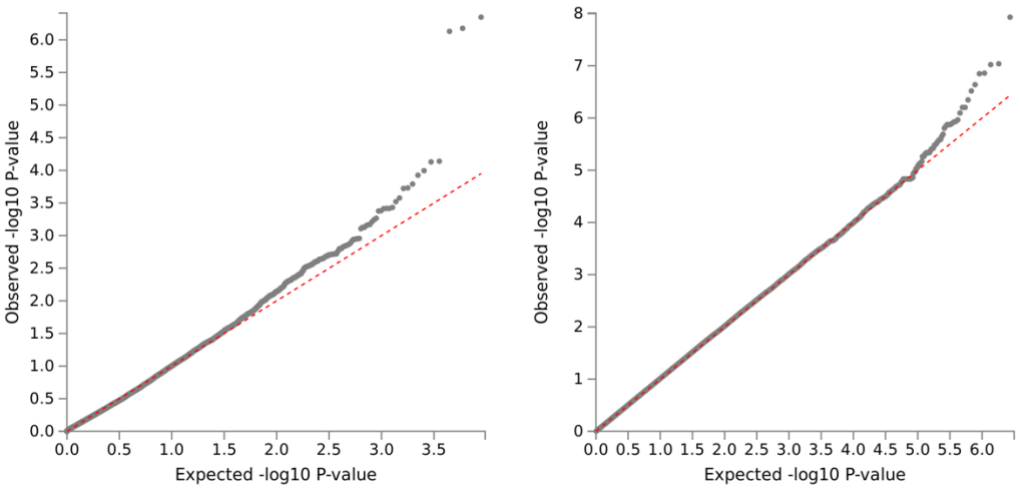
Supplementary Figure 7. Scree plot and plot showing the proportion of variance explained in the composite (cross-domain - both motor and cognitive variables) progression principal component analysis.



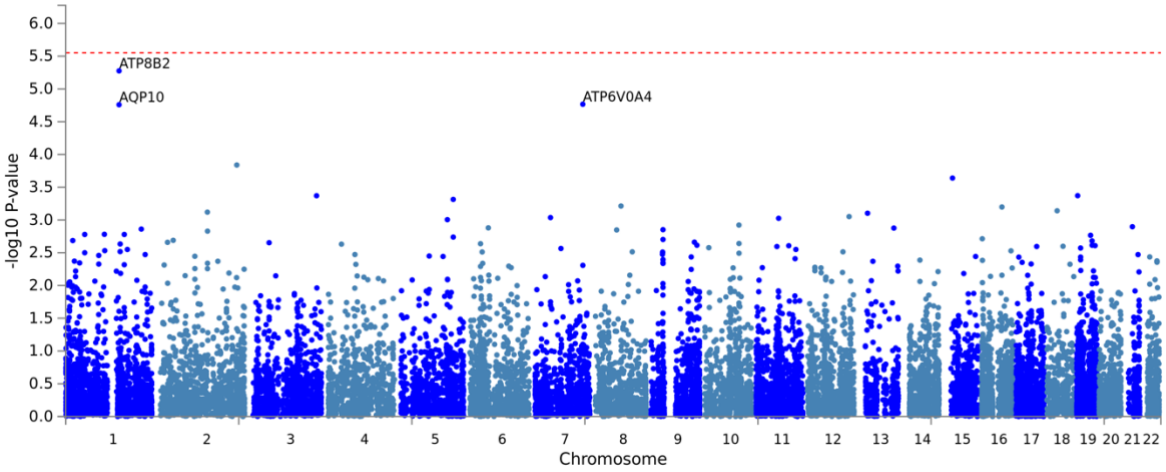
Supplementary Figure 8. Composite progression GWAS: Manhattan plot for the gene-based test from MAGMA.Genome-wide significance indicated by the red dashed line is defined as p = 0.05/17802 (the number of mapped protein coding genes) = 2.81 x 10_{-6}.



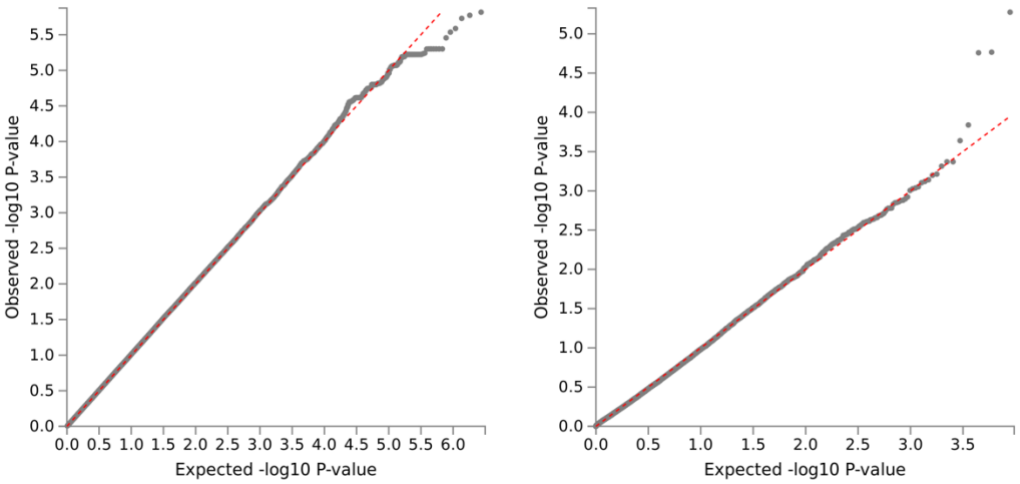
Supplementary Figure 9. QQ plots for the GWAS of composite progression (variant-based and gene-based tests).



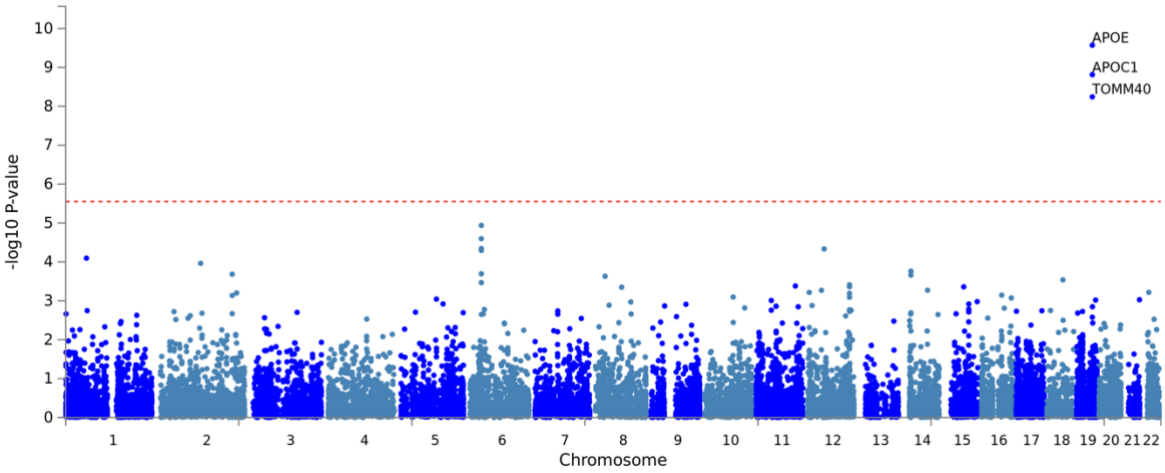
Supplementary Figure 10. Manhattan plot of the gene-based test for motor progression from MAGMA. Genome-wide significance was defined at p = 0.05/17802 (the number of mapped protein coding genes) = 2.81 x 10^{-6}.



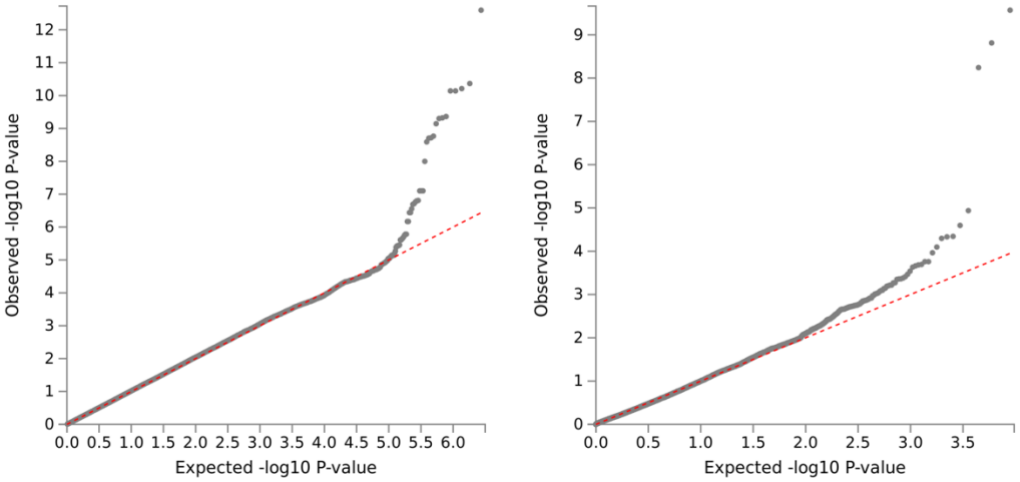
Supplementary Figure 11.QQ plots for the GWAS of motor progression (variant-based and gene-based tests).



Supplementary Figure 12.Manhattan plot of the gene-based test for cognitive progression from MAGMA. Genome-wide significance was defined at P = 0.05/17802 (the number of mapped protein coding genes) = 2.81 x 10^{-6}.

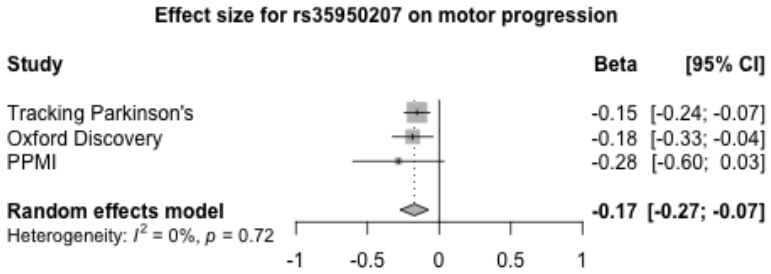


Supplementary Figure 13. QQ plots for the GWAS of cognitive progression (variant-based and gene-based tests).

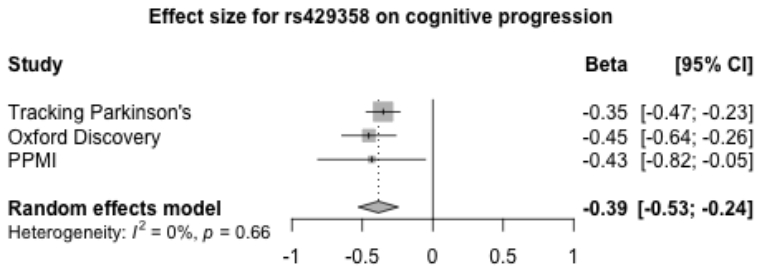


Supplementary Figure 14. Forest plots for the top SNPs for motor and cognitive progression, showing effect sizes of the top SNPs in different cohorts.

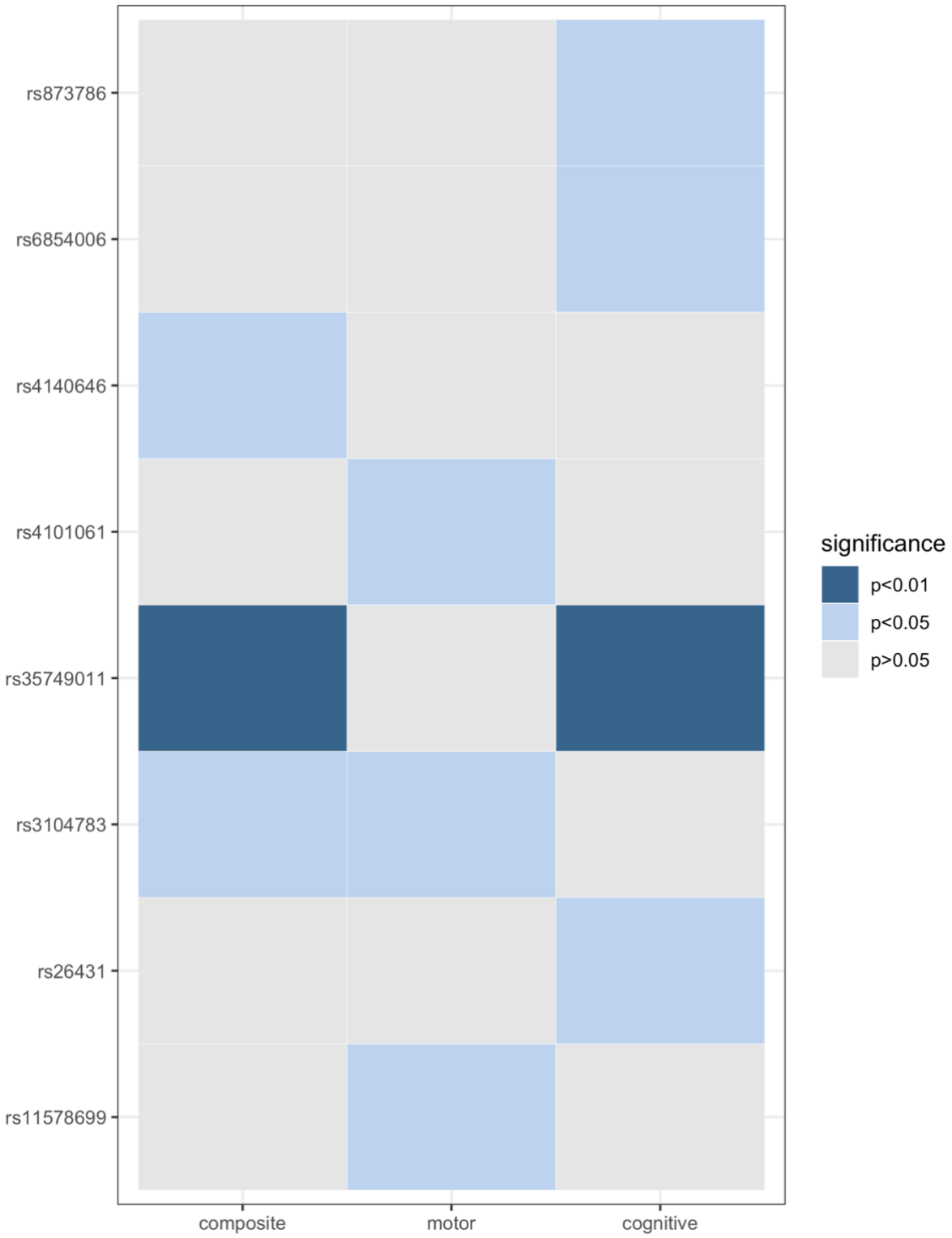
A)



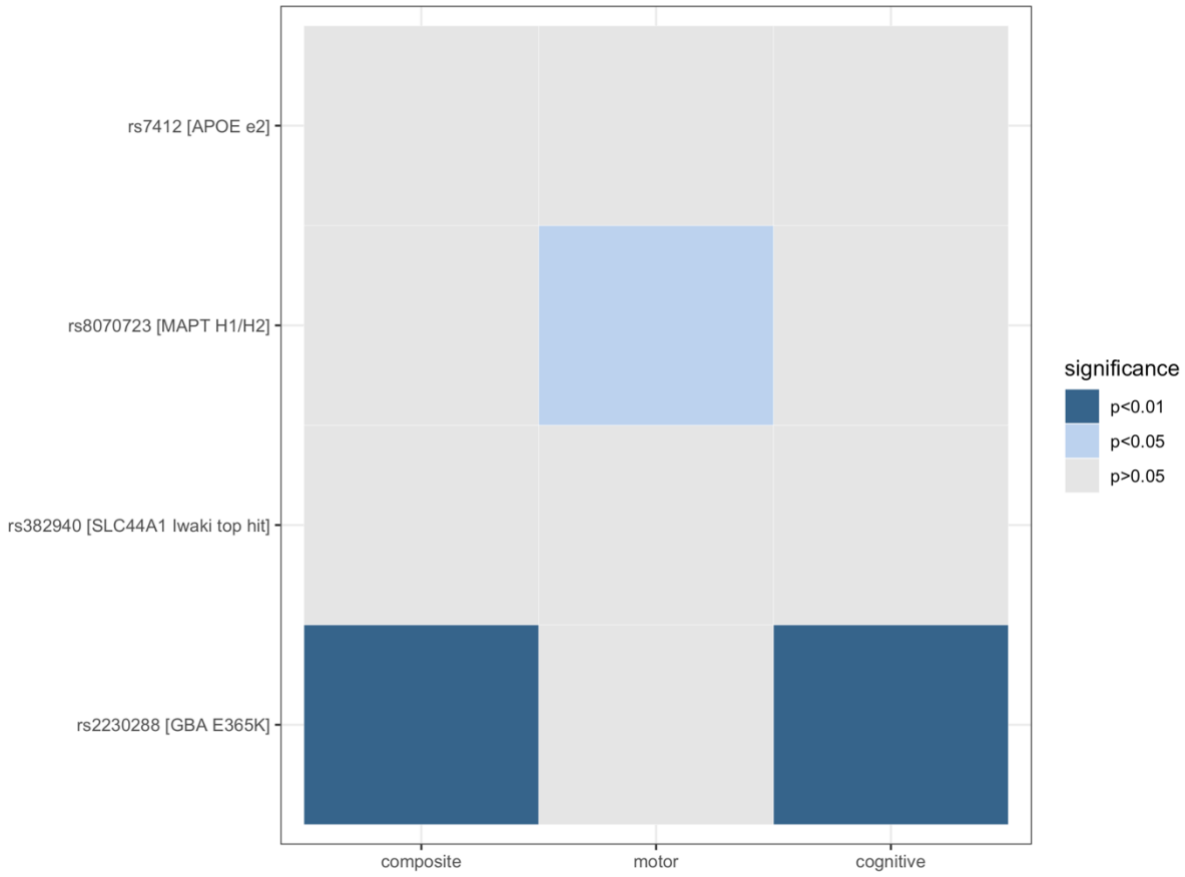
B)



Supplementary Figure 15. Heatmap of the PD GWAS risk loci and their association with composite, motor, or cognitive progression. Only variants with at least one association < 0.05 are shown in the heatmap.



Supplementary Figure 16. Heatmap of candidate variants and their association with composite, motor, or cognitive progression.



Supplementary Table 1. Correlation between motor progression principal components and cognitive progression principal components. Pearson's r, r2, and p values are reported.

	motor_PC1	motor_PC2	motor_PC3
cog_PC1	r = -0.35 r_2 = 0.12 p < 2.2 x 10^{-16}	r = 0.14 r_2 = 0.02 p = 6.2 x 10^{-15}	r = 0.06 r_2 = 0.004 p = 4.5 x 10^{-4}
cog_PC2	r = -0.11 r_2 = 0.01 p = 1.9 x 10^{-9}	r = 0.09 r_2 = 0.007 p = 1.7 x 10^{-6}	r = 0.06 r_2 = 0.004 p = 3.0 x 10^{-4}
cog_PC3	r = -0.01 r_2 = 1.8 x 10^{-4} p = 0.05	r = 0.05 r_2 = 0.003 p = 3.5 x 10^{-3}	r = 0.005 r_2 = 2.5 x 10^{-5} p = 0.78

Supplementary Table 2. Correlation between first principal component from combined progression PCA and random slopes from individual measures. Pearson's r is reported.

	PC1	PC2	PC3	PC4	PC5	PC6
MDS-UPDRSIII	0.65	-0.48	0.12	-0.33	-0.46	0.11
MDS-UPDRSII	0.72	-0.28	-0.29	-0.30	0.40	-0.27
Hoehn and Yahr	0.57	-0.54	0.30	0.51	0.17	0.08
MoCA total	0.67	0.48	0.07	0.22	-0.26	-0.45
Semantic fluency	0.56	0.53	0.50	-0.23	0.22	0.23
MDS-UPDRS 1.1	0.66	0.30	-0.54	0.18	-0.05	0.37
% of variance explained	*41.0%*	*20.1%*	*12.3%*	*9.9%*	*8.6%*	*8.1%*

Supplementary Table 3. Marginal R2 from mixed effects models showing the relationship between the raw scales (before percentage transformation and standardisation) and the composite progression score, in each cohort separately.

	UPDRS_III	UPDRS_II	H&Y	MoCA	Semantic fluency	UPDRS1.1
Tracking Parkinson's	0.12	0.15	0.10	0.18	0.13	0.18
Oxford	0.10	0.16	0.07	0.23	0.15	0.21
PPMI	0.03	0.07	0.04	0.04	0.09	0.11

Supplementary Table 4. Top 10 independent SNPs from the GWAS of composite progression.

Chr	Position (GRCh38)	SNP	Effect allele (minor)	Ref allele	Effect allele freq	Nearest gene	Distance to gene (kb)	Beta	SE	p value original	p value conditional (COJO)
19	44908684	rs429358	C	T	0.14	APOE	0	0.35	0.06	1.17E-08	1.07E-08
10	33942102	rs224750	T	C	0.37	PARD3	167458	-0.21	0.04	1.09E-06	1.20E-06
15	94318611	rs11634227	C	T	0.41	MCTP2	0	-0.21	0.04	1.19E-06	1.32E-06
19	50760039	rs4802739	C	A	0.40	GPR32	10425	0.20	0.04	1.27E-06	1.05E-06
6	119112570	rs79987229	T	A	0.01	FAM184A	0	0.85	0.18	2.57E-06	1.21E-06
15	45744252	rs17554587	C	G	0.22	SQRDL	52958	0.24	0.05	3.11E-06	3.39E-06
5	4699328	rs62343939	T	C	0.05	ADAMTS16	441002	0.43	0.09	3.25E-06	3.52E-06
5	122191027	rs17367669	T	G	0.22	LOC100505841	8364	0.23	0.05	3.31E-06	3.59E-06
7	17673826	rs10253857	T	C	0.22	SNX13	116935	-0.23	0.05	3.86E-06	4.19E-06
2	108292945	rs13424530	A	G	0.44	SULT1C2	0	0.20	0.04	4.06E-06	3.25E-06

Supplementary Table 5. Top 10 independent SNPs from the GWAS of motor progression.

Chr	Position (GRCh38)	SNP	Effect allele (minor)	Ref allele	Effect allele freq	Nearest gene	Distance to gene (kb)	Beta	SE	p value original	p value conditional
5	122193658	rs5870994	C	CTT	0.23	LOC100505841	10995	0.21	0.04	1.36E-06	1.49E-06
9	8454921	rs7870456	T	C	0.22	PTPRD	0	0.21	0.04	1.53E-06	1.68E-06
15	94320087	rs72767442	A	T	0.41	MCTP2	0	-0.18	0.04	1.69E-06	1.85E-06
2	23493673	rs6741991	G	A	0.26	KLHL29	0	0.20	0.04	2.91E-06	3.17E-06
1	154319482	rs35950207	T	C	0.31	AQP10	1585	-0.18	0.04	5.01E-06	5.40E-06
6	119067987		T	TAAAC	0.01	FAM184A	0	0.70	0.15	5.03E-06	5.40E-06
12	5829410	rs74709761	C	G	0.04	ANO2	0	-0.41	0.09	6.42E-06	8.72E-06
11	114821560	rs4436579	T	C	0.29	NXPE2	114443	0.18	0.04	7.47E-06	8.02E-06
12	12677103	rs12813102	C	A	0.04	GPR19	0	0.43	0.10	7.70E-06	1.05E-05
15	71520619	rs4128840	A	G	0.41	THSD4	0	-0.17	0.04	7.95E-06	8.53E-06

Supplementary Table 6. Motor progression GWAS performed in each cohort separately. Progression scores were created in the merged cohort. The results for the top 5 independent hits from the combined motor progression GWAS are shown here. These show that the effects and allele frequencies are consistent across all three cohorts.

SNP	**Nearest gene**	**Combined**			**Tracking Parkinson's**			**Oxford**			**PPMI**		
		Beta	*p*	*MAF*	*Beta*	*p*	*MAF*	*Beta*	*p*	*MAF*	*Beta*	*p*	*MAF*
rs5870994	LOC100505841	0.21	1.36E-06	0.23	0.20	5.31e-05	0.23	0.17	0.04	0.21	0.30	0.064	0.26
rs7870456	PTPRD	0.21	1.53E-06	0.22	0.12	0.016	0.23	0.32	3.81e-05	0.23	0.36	0.052	0.21
rs72767442	MCTP2	-0.18	1.69E-06	0.41	-0.14	0.001	0.41	-0.15	0.026	0.40	-0.40	0.005	0.41
rs6741991	KLHL29	0.20	2.91E-06	0.26	0.10	0.040	0.25	0.17	0.022	0.26	0.65	7.19e-05	0.27
rs35950207	AQP10	-0.18	5.01E-06	0.31	-0.15	0.0007	0.31	-0.18	0.010	0.30	-0.28	0.080	0.32

Supplementary Table 7. Motor progression GWAS performed for each scale separately. The results for the top 5 independent hits from the combined motor progression GWAS are shown here. The random slope from the mixed effects model for each scale was used as the progression measure. These results show that the effects are consistent across each of the different motor scales.

SNP	**Nearest gene**	**Combined**		**MDS-UPDRSIII random slope**		**MDS-UPDRSII random slope**		**Hoehn and Yahr random slope**	
		Beta	*p*	*Beta*	*p*	*Beta*	*p*	*Beta*	*p*
rs5870994	LOC100505841	0.21	1.36E-06	0.013	2.32e-05	0.012	9.43e-05	0.006	0.0008
rs7870456	PTPRD	0.21	1.53E-06	0.008	0.008	0.008	0.012	0.010	4.74e-10
rs72767442	MCTP2	-0.18	1.69E-06	-0.013	1.42e-06	-0.008	0.003	-0.004	0.006
rs6741991	KLHL29	0.20	2.91E-06	0.011	0.0004	0.010	0.002	0.007	1.08e-05
rs35950207	AQP10	-0.18	5.01E-06	-0.014	3.21e-06	-0.005	0.106	-0.006	2.79e-05

Supplementary Table 8. Top 10 independent SNPs from the GWAS of cognitive progression.

Chr	Position (GRCh38)	SNP	Effect allele (minor)	Ref allele	Effect allele freq	Nearest gene	Distance to gene (kb)	Beta	SE	p value original	p value conditional
19	44908684	rs429358	C	T	0.14	APOE	0	-0.38	0.05	2.53E-13	4.20E-13
12	20812884	rs143371462	G	A	0.02	SLCO1B3	0	-0.64	0.13	6.76E-07	7.53E-07
3	23951314	rs113730632	G	A	0.05	NR1D2	0	0.41	0.09	1.65E-06	6.59E-07
12	125083207	rs6488987	C	T	0.36	AACS	0	0.18	0.04	1.65E-06	1.95E-06
11	8882396	rs34105455	G	A	0.13	ST5	0	-0.25	0.05	3.64E-06	3.94E-06
8	74970819	rs2956605	A	C	0.40	CRISPLD1	13654	-0.17	0.04	3.70E-06	4.02E-06
11	107127349	rs17092224	C	G	0.11	CWF19L2	198996	-0.27	0.06	3.85E-06	4.62E-06
22	34970241	rs5755468	C	T	0.37	ISX-AS1	0	-0.17	0.04	4.16E-06	4.49E-06
9	84627637	rs148603475	T	C	0.08	NTRK2	40821	-0.31	0.07	6.08E-06	8.94E-06
20	18097908	rs1124933	A	G	0.42	PET117	39947	-0.16	0.04	7.94E-06	8.50E-06

Supplementary Table 9. Cognitive progression GWAS performed in each cohort separately. Progression scores were created in the merged cohort. The results for the top 5 independent hits from the combined cognitive progression GWAS are shown here. These show that the effects and allele frequencies are consistent across all three cohorts.

SNP	**Nearest gene**	**Combined**			**Tracking Parkinson's**			**Oxford**			**PPMI**		
		Beta	*p*	*MAF*	*Beta*	*p*	*MAF*	*Beta*	*p*	*MAF*	*Beta*	*p*	*MAF*
rs429358	APOE	-0.38	2.53E-13	0.14	-0.35	6.88e-09	0.14	-0.45	3.56e-06	0.14	-0.43	0.026	0.13
rs143371462	SLCO1B3	-0.64	6.76E-07	0.02	-0.54	0.0002	0.02	-0.73	0.003	0.02	-1.04	0.049	0.02
rs113730632	NR1D2	0.41	1.65E-06	0.05	0.36	0.0004	0.05	0.19	0.240	0.05	0.80	0.004	0.06
rs6488987	AACS	0.18	1.65E-06	0.36	0.21	1.66e-06	0.37	0.05	0.428	0.36	0.29	0.042	0.34
rs34105455	ST5	-0.25	3.64E-06	0.13	-0.15	0.018	0.13	0.10	0.0002	0.12	-0.39	0.047	0.14

Supplementary Table 10. Cognitive progression GWAS performed for each scale separately.The results for the top 5 independent hits from the combined cognitive progression GWAS are shown here. The random slope from the mixed effects model for each scale was used as the progression measure. These results show that the direction of effects and p values are consistent across each of the different cognitive scales.

SNP	**Nearest gene**	**Combined**		**MoCA random slope**		**Fluency random slope**		**MDS-UPDRS 1.1 random slope**	
		Beta	*p*	*Beta*	*p*	*Beta*	*p*	*Beta*	*p*
rs429358	APOE	-0.38	2.53E-13	0.02	6.84e-13	0.007	4.04e-06	0.02	1.04e-07
rs143371462	SLCO1B3	-0.64	6.76E-07	0.03	0.0001	0.02	0.0001	0.03	9.78e-05
rs113730632	NR1D2	0.41	1.65E-06	-0.02	0.0004	-0.01	1.595e-05	-0.01	0.006
rs6488987	AACS	0.18	1.65E-06	-0.007	0.002	-0.004	9.16e-05	-0.009	2.37e-05
rs34105455	ST5	-0.25	3.64E-06	0.01	2.64e-05	0.006	0.0008	0.009	0.002

Supplementary Table 11. PD risk SNPs missing for calculation of the genetic risk score, and whether a proxy was identified. All other SNPs from Nalls et al. (2019) were present in our final genotype dataset. In total, 75 SNPs were used for creation of the genetic risk score including 2 proxies.

SNP	chr_pos_GRCh38	Effect allele	Other allele	EAF	Proxy identified	R_2 for proxy	D' for proxy	Correlated alleles
rs114138760	1:154925709	c	g	0.0112	No			
rs76763715	1:155235843	t	c	0.9953	No			
rs6658353	1:161499264	c	g	0.5011	No			
rs73038319	3:18320267	a	c	0.9592	No			
rs13117519	4:113447909	t	c	0.1744	No			
rs75859381	6:132889222	t	c	0.9673	No			
rs76949143	7:66544864	a	t	0.0507	rs62469072	0.9717	1	T=C,A=T
rs117896735	10:119776815	a	g	0.0166	No			
rs34637584	12:40340400	a	g	0.0015	No			
rs10847864	12:122842051	t	g	0.364	No			
rs11610045	12:132487182	a	g	0.4896	No			
rs6497339	16:19266171	a	t	0.4536	No			
rs62053943	17:45666837	t	c	0.1552	No			
rs117615688	17:45720942	a	g	0.067	No			
rs11658976	17:46789439	a	g	0.5802	No			
rs61169879	17:61840005	t	c	0.1641	rs3744434	0.935	0.9777	C=A,T=C

Supplementary Table 12. Sensitivity analysis excluding PD cases with less than 90% diagnostic certainty. The top SNPs in the main analysis are shown, with the results from the sensitivity analysis for comparison. 5.2% (51/985) patients were removed from Oxford Discovery, 21.3% (419/1966) patients were removed from Tracking Parkinson's.

SNP	**Nearest gene**	**Results for top SNPs in full dataset**		**Results in PD cases with ≥ 90% diagnostic certainty**		
		Beta	*p*	*Beta*	*p*	*N*
Composite progression						
rs429358	APOE	0.35	1.17E-08	0.34	6.00e-07	2459
rs224750	PARD3	-0.21	1.09E-06	-0.20	5.36e-05	2459
rs11634227	MCTP2	-0.21	1.19E-06	-0.20	2.14e-05	2459
rs4802739	GPR32	0.20	1.27E-06	0.22	2.86e-06	2459
rs79987229	FAM184A	0.85	2.57E-06	0.99	3.64e-07	2459
Motor progression						
rs5870994	LOC100505841	0.21	1.36E-06	0.19	4.19e-05	2496
rs7870456	PTPRD	0.21	1.53E-06	0.19	0.0001	2496
rs72767442	MCTP2	-0.18	1.69E-06	-0.15	0.0001	2496
rs6741991	KLHL29	0.20	2.91E-06	0.19	2.27e-05	2496
rs35950207	AQP10	-0.18	5.01E-06	-0.18	2.13e-05	2496
Cognitive progression						
rs429358	APOE	-0.38	2.53E-13	-0.39	2.05e-12	2474
rs143371462	SLCO1B3	-0.64	6.76E-07	-0.64	3.22e-06	2474
rs113730632	NR1D2	0.41	1.65E-06	0.43	4.34e-06	2474
rs6488987	AACS	0.18	1.65E-06	0.19	1.69e-06	2474
rs34105455	ST5	-0.25	3.64E-06	-0.25	1.79e-05	2474

Supplementary Table 13. Sensitivity analysis excluding fastest and slowest progressing cases (top and bottom 5% of each distribution)

SNP	**Nearest gene**	**Results for top SNPs in full dataset**		**Results in PD cases excluding extreme 5%**		
		Beta	*p*	*Beta*	*p*	*N*
Composite progression						
rs429358	APOE	0.35	1.17E-08	0.17	0.0002	2483
rs224750	PARD3	-0.21	1.09E-06	-0.10	0.002	2483
rs11634227	MCTP2	-0.21	1.19E-06	-0.10	0.003	2483
rs4802739	GPR32	0.20	1.27E-06	0.11	0.004	2483
rs79987229	FAM184A	0.85	2.57E-06	0.42	0.002	2483
Motor progression						
rs5870994	LOC100505841	0.21	1.36E-06	0.12	0.0003	2570
rs7870456	PTPRD	0.21	1.53E-06	0.10	0.002	2570
rs72767442	MCTP2	-0.18	1.69E-06	-0.07	0.012	2570
rs6741991	KLHL29	0.20	2.91E-06	0.08	0.008	2570
rs35950207	AQP10	-0.18	5.01E-06	-0.06	0.029	2570
Cognitive progression						
rs429358	APOE	-0.38	2.53E-13	-0.17	1.28e-05	2511
rs143371462	SLCO1B3	-0.64	6.76E-07	-0.32	0.001	2511
rs113730632	NR1D2	0.41	1.65E-06	0.17	0.006	2511
rs6488987	AACS	0.18	1.65E-06	0.09	0.002	2511
rs34105455	ST5	-0.25	3.64E-06	-0.15	0.0002	2511