Population stratification analysis

After scree plot analysis to identify significant principal components (1), we identified principal components C1 and C2 as accounting for the most observed variance in population structure within our sample. Initially, we included C1 and C2 as covariates within our association analyses. However, after covariate analysis revealed that ancestry did not account for significant variance in remission, percentage change of MADRS score or time-to remission, we omitted C1 and C2 to avoid over correcting our models. There was little to no effect of C1 and C2 on the association of SLC6A2 rs2242446 across all three phenotypes of interest (see Supplementary Table S5).

Reference

1. Ledesma RD, Valero-Mora P. Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. Pract Assess Res Eval. 2007 Feb;12(2):1-1.

FIGURE S1. Flow diagram depicting: sample selection process, including exclusion after standard quality checks.

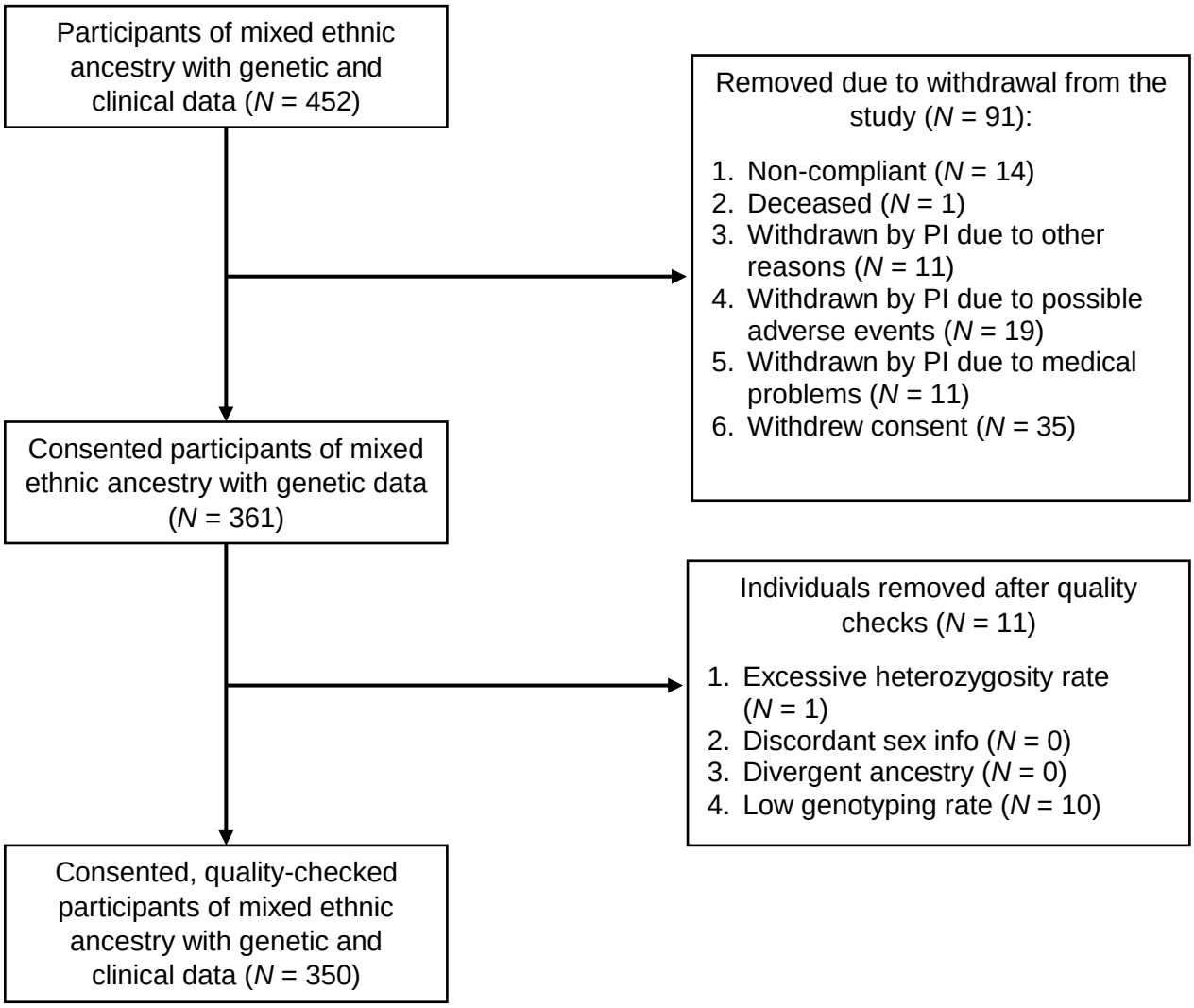
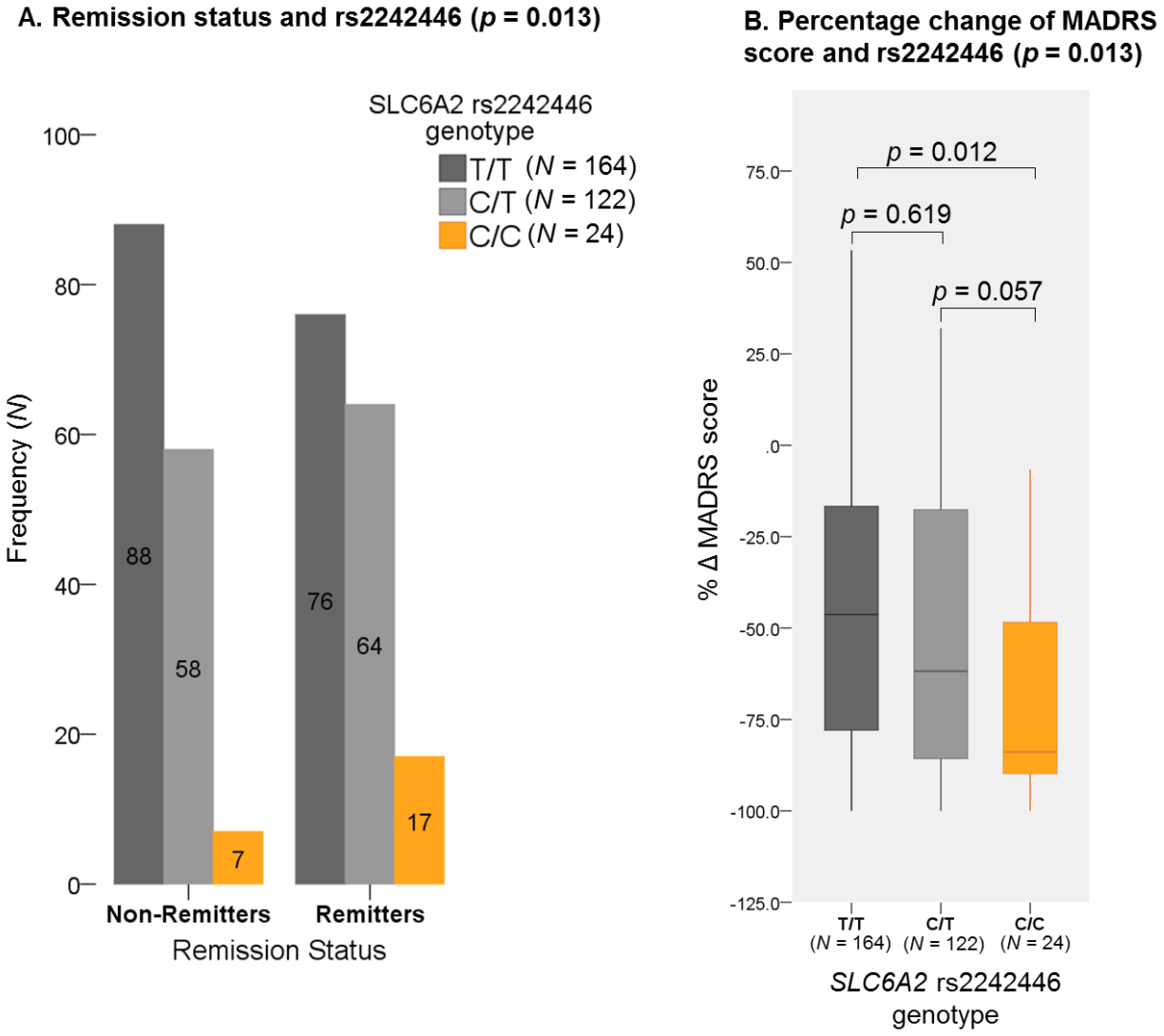


FIGURE S2. Association of *SCL6A2* variant rs2242446 with remission, percentage MADRS score change from baseline and time-to remission in the total, European-ancestry sample.



C. Time-to Remission and *SLC6A2* rs2242446 (p = 0.020)

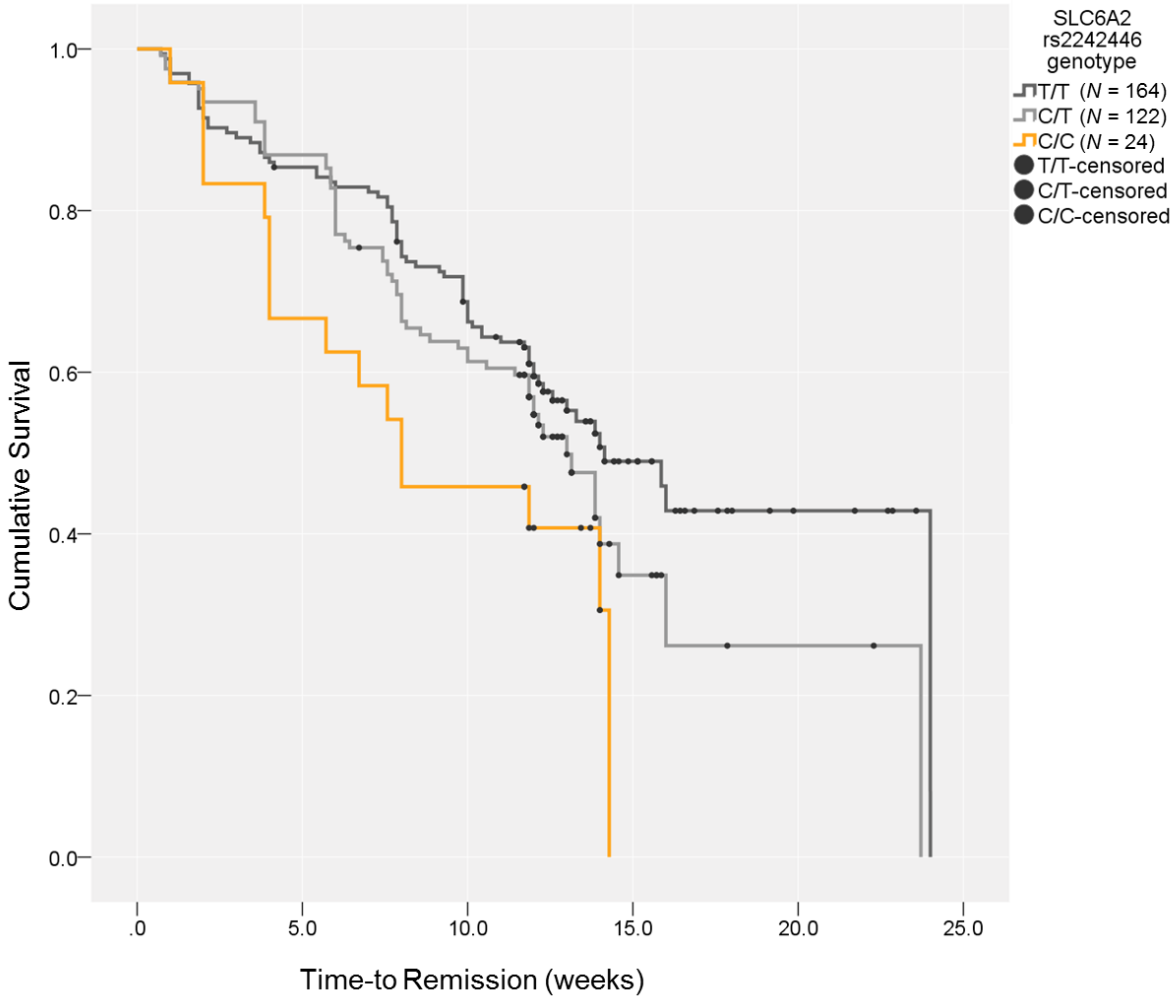
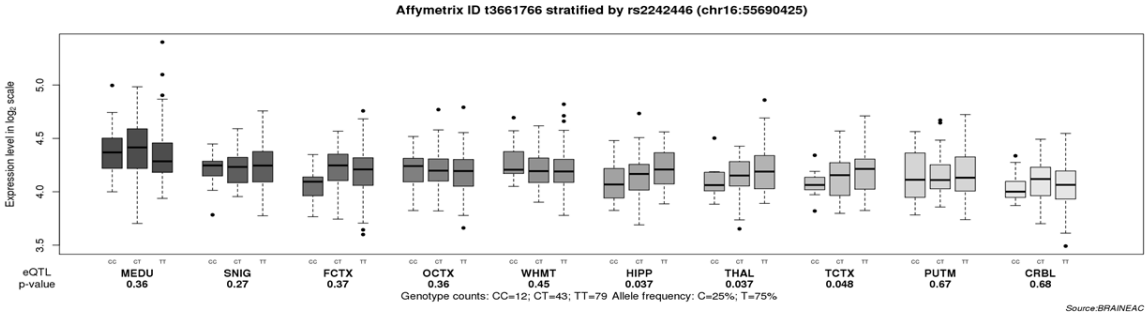
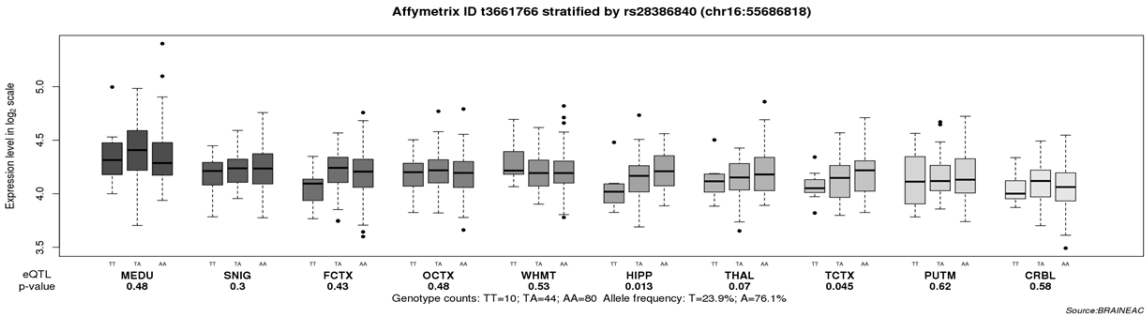


FIGURE S3. Neural expression profiles for rs28386840 and rs2242446 across specific brain regions, analyzed in 134 brains of individuals without neurodegenerative disorders (retrieved from *BRAINEAC.org, UK Brain Consortium*).

A. eQTLs for *SLC6A2* rs2242446 across brain regions.



B. eQTLs for *SLC6A2* rs28386840 across brain regions.



Abbreviations: CRBL, cerebellum; eQTL, expression quantitative trait loci; FCTX, frontal cortex; HIPP, hippocampus; MEDU, medulla; OCTX, occipital cortex; PUTM, putamen; SNIG, substantia nigra; TCTX, temporal cortex; WHMT, white matter.

TABLE S1. Genetic variants genotyped.

Gene	Variant (SNP)	CHR	Location (BP)	Type	Minor Allele (Ref > Alt)	Global / Sample MAF	Functional Annotations
Primary							
SLC6A2	rs2242446	16	55656513	5' UTR	C > T	0.25 / 0.26	8 PHMs; 10 EHMs; 12 DHS; 1 TFBS; 4 Δ motifs; 4 eQTLs
	rs5569	16	55697923	SYN	G > A	0.23 / 0.30	2 EHMs; 1 DHS; 1 Δ motifs
SLC6A4	rs25531	17	30237328	UPS	T > C	0.14 / 0.11	4 PHMs; 19 EHMs; 8 Δ motifs
	5-HTTLPR	17	-	PRM	L / S	-	-
	STin2 VNTR	17	-	INT	12, 10, 9, 7	-	-
Secondary							
HTR1A	rs6295	5	63962738	INT	C > G	0.45 / 0.48	4 PHMs; 1 EHMs; 2 Δ motifs; 1 eQTLs
HTR1B	rs11568817	6	77463665	UPS	A > C	0.30 / 0.42	18 PHMs; 7 EHMs; 19 DHS
	rs130058	6	77463564	UPS	T > A	0.18 / 0.26	18 PHMs; 7 EHMs; 25 DHS; 9 Δ motifs
	rs6296	6	77462543	SYN	C > G	0.34 / 0.26	15 PHMs; 6 EHMs; 16 DHS; 3 TFBS
HTR2A	rs2274639	13	46856128	EXN	C > G	0.18 / 0.13	2 EHMs; 1 DHS; 2 Δ motifs
	rs6311	13	46897343	UPS	C > T	0.44 / 0.45	11 PHMs; 8 EHMs; 7 DHS; 2 Δ motifs
	rs9567746	13	46882413	INT	A > G	0.14 / 0.18	3 Δ motifs
HTR2C	rs17260600	23	114786945	INT	A > C	0.03† / 0.03	1 PHMs; 4 Δ motifs
	rs1801412	23	114908141	3'UTR	T > G	0.06 / 0.05	2 EHMs; 2 Δ motifs
	rs3813929	23	114584047	UPS	C > T	0.12 / 0.17	7 PHMs; 11 EHMs; 7 DHS; 1 TFBS
	rs518147	23	114584109	5'UTR	G > C	0.17 / 0.36	7 PHMs; 11 EHMs; 7 DHS; 1 TFBS
	rs6318	23	114731326	MIS	G > C	0.17 / 0.18	5 Δ motifs
	rs6644093	23	114829460	INT	G > T	0.13 / 0.14	3 eQTLs
TPH1	rs1800532	11	18026269	INT	G > T	0.32 / 0.38	3 GRASP QTLs; 13 eQTLs
TPH2	rs11178997	12	71938373	UPS	T > A	0.15 / 0.10	4 Δ motifs
	rs11178998	12	71938935	UPS	A > G	0.08 / 0.08	5 PHMs; 13 EHMs; 43 DHS; 25 TFBS; 3 Δ motifs
	rs4570625	12	71946293	UPS	G > T	0.35 / 0.26	4 Δ motifs

†Variant excluded from analysis due to minor allele frequency less than 5% in the whole mixed-ancestry sample.

Abbreviations. Δ, change; Alt, alternative allele; BP, base pair; CHR, chromosome; DHS, DNAse hypersensitivity site; EHM, enhancer histone mark; eQTL, expression quantitative trail loci; EXN, exonic; INT, intronic; MAF, minor allele frequency; Ref, reference allele; SNP, single nucleotide polymorphism; TFBS, transcription factor binding site.

TABLE S2. Functional annotations for genotyped ancestry informative markers after quality control (*N* = 46).

Gene	Variant (SNP)	CHR	Location (BP)	Type	Minor Allele (Ref > Alt)	Total Sample MAF	Global/ Sample EUR MAF	Global/ Sample AFR MAF	HWE (*p*)	Functional Annotations
HIVEP3	rs1325502	1	41894599	INT	G > A	0.21	0.16 / 0.15	0.73 / 0.30	0.002	3 PHMs; 12 EHMs; 5 DHS; 1 GRASP QTL
USP24	rs12130799	1	55197699	INT	A > G	0.06	0.06 / 0.07	0.01 / 0.00	1.000	1 PHM; 8 EHMs; 1 DHS; 2 hits
RP4-575N6.5	rs3737576	1	101244007	DWN	T > C	0.05	0.06 / 0.05	0.01 / 0.03	0.077	3 EHMs; 1 Δ motif; 1 GRASP QTL
SEMA6C	rs7554936	1	151150013	UPS	C > T	0.40	0.68 / 0.36	0.04 / 0.14	0.321	3 Δ motifs; 7 GRASP QTLs; 24 eQTLs
TIPRL	rs1040404	1	168190652	INT	G > A	0.34	0.32 / 0.29	0.84 / 0.21	0.159	2 DHS; 1 TFBS; 1 Δ motif; 8 GRASP QTLs; 23 eQTLs
HMCN1	rs1407434	1	186179900	INT	G > A	0.10	0.10 / 0.09	0.24 / 0.20	0.401	3 EHMs; 1 eQTL
AC007463.2	rs798443	2	7828144	DWN	G > A	0.28	0.82 / 0.22	0.07 / 0.15	0.001	5 Δ motifs; 8 GRASP QTLs
AC011897.2	rs7421394	2	14616225	UPS	A > G	0.36	0.28 / 0.30	0.93 / 0.12	0.002	2 EHMs; 1 GRASP QTL; 1 eQTL
ALK	rs4666200	2	29315545	INT	G > A	0.32	0.72 / 0.28	0.13 / 0.30	0.466	3 PHMs; 10 EHMs; 5 DHS; 1 TFBS; 3 Δ motifs; 1 GRASP QTL; 1 eQTL
CDC42EP3	rs4670767	2	37714253	INT	G > T	0.12	0.11 / 0.12	0.04 / 0.05	0.027	1 PHM; 3 EHMs; 4 Δ motifs; 1 GRASP QTL
CTNNA2	rs13400937	2	79637797	INT	T > G	0.32	0.24 / 0.25	0.86 / 0.14	0.007	4 Δ motifs
EDAR	rs260690	2	108963282	INT	C > A	0.14	0.92 / 0.09	0.35 / 0.45	0.000	3 EHMs; 1 DHS; 2 TFBS; 2 Δ motifs
AC074093.1	rs10496971	2	145012376	INT	T > G	0.10	0.07 / 0.09	0.05 / 0.12	0.131	10 DHS; 5 TFBS; 5 Δ motifs; 1 GRASP QTL
GORASP1	rs9809104	3	39104938	INT	T > C	0.26	0.21 / 0.20	0.86 / 0.18	0.008	2 PHMs; 16 EHMs; 5 GRASP QTLs; 11 eQTLs
ROBO1	rs6548616	3	79350425	INT	T > C	0.34	0.30 / 0.29	0.91 / 0.20	0.054	3 Δ motifs
GTF2E1	rs12629908	3	120803869	DWN	G > A	0.06	0.05 / 0.05	0.11 / 0.14	0.111	4 Δ motifs
LPP	rs1513181	3	188857208	INT	G > A	0.14	0.17 / 0.14	0.19 / 0.09	0.670	1 EHM; 7 Δ motifs; 1 GRASP QTL
LIMCH1	rs10007810	4	41552347	INT	G > A	0.28	0.20 / 0.22	0.93 / 0.14	0.002	2 DHS; 1 Δ motif; 1 GRASP QTL
Y_RNA	rs316598	5	2364512	DWN	T > C	0.33	0.26 / 0.28	0.95 / 0.14	0.006	-

7SK	rs870347	5	6844922	DWN	A > C	0.05	0.08 / 0.05	0.07 / 0.06	0.259	1 GRASP QTL; 1 eQTL
NNT	rs6451722	5	43711276	DWN	G > A	0.28	0.20 / 0.23	0.84 / 0.23	0.003	1 DHS; 3 Δ motifs; ; 18 eQTLs
SGCD	rs6556352	5	156044704	INT	C > T	0.35	0.34 / 0.29	0.95 / 0.11	0.015	12 Δ motifs
CDYL	rs1040045	6	4746925	INT	G > A	0.32	0.72 / 0.27	0.19 / 0.18	0.052	1 Δ motif; 1 eQTL
RP11-125M16.1	rs2504853	6	12534879	DWN	T > C	0.33	0.27 / 0.26	0.92 / 0.05	0.002	3 EHMs; 1 DHS; 1 Δ motif;
PKHD1	rs2397060	6	51746672	INT	T > C	0.20	0.14 / 0.15	0.74 / 0.30	4×10^{-4}	2 Δ motifs
UTRN	rs4463276	6	144734195	INT	G > A	0.31	0.23 / 0.25	0.90 / 0.16	1.000	5 Δ motifs; 1 eQTL
SCIN	rs731257	7	12629626	INT	G > A	0.14	0.16 / 0.15	0.03 / 0.02	0.589	1 DHS; 2 Δ motifs; 1 GRASP QTL
ELN	rs4717865	7	74039869	INT	G > A	0.08	0.09 / 0.08	0.01 / 0.00	0.019	1 PHM; 12 EHMs; 3 DHS; 1 Δ motif; 1 GRASP QTL
DLC1	rs3943253	8	13501991	INT	A > G	0.07	0.06 / 0.05	0.36 / 0.19	1.000	6 PHMs; 1 EHM; 4 DHS; 2 Δ motifs
KIF13B	rs1471939	8	29083788	INT	C > T	0.26	0.80 / 0.24	0.72 / 0.41	1.000	1 Δ motif; 2 GRASP QTLs; 2 eQTLs
TYRP1	rs1408801	9	12672320	UPS	A > G	0.15	0.11 / 0.10	0.51 / 0.48	0.055	4 Δ motifs; 1 eQTL
ASTN2	rs10513300	9	117367927	INT	T > C	0.05	0.06 / 0.05	0.00 / 0.03	0.204	-
USP54	rs4746136	10	73541236	INT	G > A	0.12	0.18 / 0.13	0.02 / 0.06	0.448	1 EHM; 4 Δ motifs; 1 GRASP QTL
HABP2	rs4918842	10	113557053	INT	T > C	0.11	0.11 / 0.11	0.09 / 0.08	0.054	2 PHMs; 13 EHMs; 1 DHS; 2 Δ motifs
U6	rs2946788	11	23988984	DWN	G > T	0.33	0.75 / 0.28	0.14 / 0.20	0.012	2 PHMs; 3 EHMs; 2 DHS; 2 Δ motifs; 1 eQTL
USP12	rs9319336	13	27050219	DWN	T > C	0.08	0.06 / 0.06	0.18 / 0.15	0.003	1 EHM; 1 DHS
LINC00457	rs7997709	13	34273600	DWN	C > T	0.07	0.94 / 0.06	0.81 / 0.18	0.410	1 EHM; 2 Δ motifs
TBC1D4	rs9530435	13	75419751	INT	T > C	0.23	0.84 / 0.18	0.14 / 0.26	4×10^{-4}	8 EHMs; 2 DHS; ; 3 Δ motifs; 1 GRASP QTL
BRF1	rs3784230	14	105212718	INT	A > G	0.46	0.38 / 0.41	0.96 / 0.12	0.086	5 EHMs; 1 DHS; ; 5 Δ motifs; 1 GRASP QTL; 6 eQTLs
RP11-661P17.1	rs8035124	15	91562478	INT	A > C	0.27	0.16 / 0.21	0.70 / 0.23	0.001	12 EHMs; 3 Δ motifs
LA16c-313D11.12	rs4984913	16	690466	INT	A > G	0.34	0.32 / 0.28	0.84 / 0.15	0.007	23 PHMs; 5 EHMs; 53 DHS; 46 TFBs; 3 Δ motifs; 5 GRASP QTLs; 90 eQTLs
RP11-256I9.3	rs818386	16	65372805	INT	T > C	0.17	0.80 / 0.19	0.99 / 0.00	0.125	7 EHMs; 8 Δ motifs

WDR88	rs8113143	19	33161341	INT	C > A	0.37	0.34 / 0.31	0.96 / 0.09	2×10^{-4}	7 EHMs; 1 DHS; 3 Δ motifs; 2 GRASP QTLs; 12 eQTLs
RP5-839B4.7	rs6104567	20	10214785	INT	T > G	0.28	0.28 / 0.28	0.13 / 0.23	1.000	27 Δ motifs; 2 GRASP QTLs
RP5-1010E17.2	rs3907047	20	55384376	DWN	T > C	0.04	0.05 / 0.04	0.01 / 0.00	1.000	1 Δ motif
RP11-191L9.4	rs5768007	22	47812123	INT	C > T	0.12	0.12 / 0.12	0.01 / .03	0.199	2 Δ motifs

Abbreviations. Δ, change; Alt, alternative allele; BP, base pair; CHR, chromosome; DHS, DNAse hypersensitivity site; EHM, enhancer histone mark; eQTL, expression quantitative trail loci; EXN, exonic; INT, intronic; MAF, minor allele frequency; Ref, reference allele; SNP, single nucleotide polymorphism; TFBS, transcription factor binding site.

TABLE S3. Baseline covariates with venlafaxine remission, percentage MADRS score change post-treatment and time-to remission.

Potential Covariates	% Δ MADRS score			Remitters vs. Non-remitters			Time-to Remission (weeks)		
	M	r_s	*95% C.I.*	*M / Frequency (N)*	*OR*	*95% C.I.*	*M*	r_s	*95% C.I.*
Sex*	M_F = -55.1±35.9 M_M = -41.8±38.6	-	-	N_F = 128/95 N_M = 51/76	2.00	1.28, 3.13	M_F = 10.6±3.9 M_M = 10.0±4.9	-	-
Age (years)*	-	-0.18	-0.28, 0.08	M_{Rem} = 69.9±7.6 $M_{Non\text{-}rem}$ = 67.4±6.1	-	-	-	-0.08	-0.18, 0.03
Ethnic ancestry	M_{EUR} = -49.6±37.1 M_{AFR} = -56.5±36.2 M_{ALL} = -50.2±37.2	-	-	N_{EUR} = 158/153 N_{AFR} = 18/15 N_{ALL} = 179/171	-	-	M_{EUR} = 10.4±4.6 M_{AFR} = 9.8±4.8 M_{ALL} = 10.3±4.6	-	-
Site (A, B or C)*	M_A = -58.4±33.9 M_B = -45.8±41.0 M_C = -43.1±36.2	-	-	N_A = 87/59 N_B = 47/46 N_C = 45/63	-	-	M_A = 9.7±4.7 M_B = 9.7±4.0 M_C = 11.5±5.0	-	-
Baseline MADRS score*	-	0.21	0.11, 0.31	M_{Rem} = 24.8±5.3 $M_{Non\text{-}R}$ = 28.4±5.4	-	-	-	0.29	0.19, 0.38
Treatment length (days)*	-	-0.11	-0.21, -0.005	M_{Rem} = 97.0±18.8 $M_{Non\text{-}rem}$ = 92.4±18.4	-	-	-	0.42	0.33, 0.50
Age of onset (years)	-	-0.08	-0.18, 0.03	M_{Rem} = 43.0±21.9 $M_{Non\text{-}rem}$ = 39.0±21.3	-	-	-	-0.12	-0.22, 0.02
Current MDE duration (weeks)*	-	0.12	0.01, 0.22	M_{Rem} = 227.7±540.5 $M_{Non\text{-}rem}$ = 430.4±782.1	-	-	-	0.24	0.13, 0.33
MDE type (single vs. recurrent)	M_S = -45.6±38.0 M_R = -51.9±36.9	-	-	N_S = 44/51 N_R = 135/120	1.30	0.81, 2.09	M_S = 10.6±4.8 M_R = 10.2±4.8	-	-

Note. All associations were derived from bivariate analyses (without covariates).
Abbreviations: Δ, change; 95% C.I., 95% confidence interval; AFR, African ancestry; *d*, Cohen's *d*; EUR, European ancestry; *F*, female; *M*, male; M, mean; MDE, major depressive episode; OR, odds ratio; R, recurrent; Rem, remission; r_s, Spearman rank correlation coefficient; S, single; VEN, venlafaxine.
Significance levels: $*p \leq 0.05$

TABLE S4. Multivariate logistic regression results examining the association between remission status and *SLC6A2* rs2242446 genotype in the total, mixed ancestry sample.

						95% *C.I.*	
	β	S.E.	Wald	Sig. (*p*)	*OR*	Lower	Upper
Site	-0.19	.14	1.79	**.180**	0.83	0.62	1.09
Age	0.03	.02	3.61	**.057**	1.04	0.99	1.07
Sex	0.83	.25	10.76	**.001**	2.30	1.40	3.78
Baseline MADRS score	-.14	.02	31.80	**<.001**	.87	0.83	.92
Treatment length (days)	.28	.12	5.23	**.022**	1.32	1.04	1.67
SLC6A2 rs2242446 genotype	.51	.19	6.78	**.009**	1.67	1.13	2.42

Abbreviations: 95% C.I., 95% confidence interval; *β* , regression coeffecient; MDE, major depressive episode; OR, oods ratio; S.E., standard error; VEN, venlafaxine.

TABLE S5. Effect of the first two principal components on the association analyses between the three outcomes of interest and *SLC6A2* rs2242446 genotype.

	% Δ MADRS		Remission Status		Time-to Remission	
Parameters	*(η^2)*	*p*	*OR [95%C.I.]*	*p*	*(X^2)*	*p*
Original analysis	0.30	0.006	1.66 [1.13, 2.42]	0.009	9.47	0.009
Incl. C1	0.33	0.004	1.70 [1.16, 2.50]	0.006	N/A	N/A
Incl. C1 and C2	0.32	0.004	1.71 [1.16, 2.50]	0.006	N/A	N/A

Note. Each analysis is multivariate as described in the manuscript, already including covariates: age, sex, recruitment site, baseline MADRS score and duration in treatment.

TABLE S6. Genetic variant association with phenotypes in the African-ancestry subsample (*N* = 33).

Gene	Variant	% Δ MADRS (Π^2)	Remission *OR* [*95%C.I.*]	Time-to Remission (X^2)
SLC6A2	rs2242446	0.084	3.81 [0.54, 26.78]	4.84
	rs5569	0.006	2.24 [0.16, 29.78]	0.29
SLC6A4	rs25531-LPR	0.22	0.62 [0.35, 1.04]	6.91
	VNTR	0.084	0.30 [0.04, 2.30]	2.60
HTR1A	rs6295	0.049	0.99 [0.19, 5.06]	1.39
HTR1B	rs11568817	0.030	1.21 [0.24, 6.09]	0.97
	rs130058	0.024	3.38 [0.21, 53.18]	1.76
	rs6296	0.055	1.79 [0.31, 10.34]	1.09
HTR2A	rs2274639	0.049	1.25 [0.29, 5.50]	0.94
	rs6311	0.010	1.28 [0.27, 6.06]	0.40
	rs9567746	0.010	0.24 [0.01, 4.10]	1.22
HTR2C[b]	$rs3813929_F$	-[a]	-[a]	-[a]
	$rs3813929_M$	-[a]	-[a]	0.78
	$rs51814_F$	0.017	-[a]	0.32
	$rs51814_M$	-[a]	-[a]	2.14
	$rs6318_F$	0.018	-[a]	0.03
	$rs6318_M$	-[a]	-[a]	2.14
	$rs6644093_F$	-[a]	-[a]	-[a]
	$rs6644093_M$	-[a]	-[a]	-[a]
	$rs1801412_F$	-[a]	-[a]	-[a]
	$rs1801412_M$	-[a]	-[a]	-[a]
TPH1	rs1800532	0.10	3.10 [0.50, 19.18]	1.94
TPH2	rs11178997	0.018	0.78 [0.14, 4.16]	0.70
	rs11178998	0.017	-	9.04**
	rs4570625	0.062	1.78 [0.42, 7.51]	3.45

Abbreviations. Δ, change; F, female; M, male; X^2, Mantel-Cox chi-squared; Π^2, partial eta squared; *OR,* odds ratio

Note. Presented p-values are not adjusted for multiple testing. Bonferroni-adjust thresholds are 0.0125 and 0.0031 for primary and secondary analyses, respectively.

[a]Too few remitter/genotype cases to conduct analysis.

[b]Analyses conducted separately for males and females due to the location of *HTR2C* on the X-chromosome

Significance levels: $^{**}p \leq 0.01$, $^{*}p \leq 0.05$.