

## Supplementary Information

Increased Nigral *SLC6A3* Activity in Schizophrenia Patients: Findings From the Toronto-McLean Cohorts

Zhicheng Lin <sup>#1,3</sup>, Nian Xiong <sup>1,3,8</sup>, Jinlong Yu <sup>1,3</sup>, Clement C. Zai <sup>4,5,6</sup>, Jennie G. Pouget <sup>4,5,6</sup>,  
Eden Martin<sup>7</sup>, Deborah L. Levy <sup>\*2,3</sup>, James L. Kennedy <sup>\*4,5,6</sup>

<sup>1</sup> Laboratory of Psychiatric Neurogenomics, McLean Hospital, Belmont, Massachusetts 02478 USA;

<sup>2</sup> Psychology Research Laboratory, McLean Hospital, Belmont, Massachusetts 02478 USA;

<sup>3</sup> Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, USA;

<sup>4</sup> Neurogenetics Section, Neuroscience Research Department, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, and

<sup>5</sup> Department of Psychiatry, and

<sup>6</sup> Institute of Medical Science, University of Toronto, Toronto, Ontario M5S 3E3 Canada;

<sup>7</sup> Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL 33136, USA;

<sup>8</sup> Department of Neurology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, 430074 Hubei, China.

## **Methods**

### *Subjects*

Postmortem midbrain tissue of 20 controls and 20 cases was provided by the Harvard Brain Tissue Resource Center at McLean Hospital, Belmont, MA, USA. Each postmortem sample contained tissue from the midbrain (coronal blocks for capture of DA neurons) and the cerebellum (for isolation of genomic DNA and genotyping).

Genomic DNA from post-mortem cerebellum tissue was extracted using the Qiagen RNeasy Lipid Tissue Mini Kit. These postmortem samples were used to examine the effects of *SLC6A3* alleles and genotypes on expression differences between SCZ cases and controls.

DNA extracted from blood was obtained from 861 controls and 819 patients with a diagnosis of a psychotic disorder (all controls and 610 cases came from the Neurogenetics Laboratory at the Centre for Addiction and Mental Health (CAMH) in Toronto, Ontario, Canada and 209 SCZ cases came from the Psychology Research Laboratory at McLean Hospital).

### *Laser capture microdissection (LCM)*

Individual dopamine (DA) neurons were isolated randomly from sections of postmortem midbrain blocks by LCM. Briefly, frozen blocks were sectioned at 10 µm thickness using a Leica CM 3050 cryostat and mounted on RNase-free slides. Fast immunostaining of the sections was used to identify DA neurons using the following procedures. Sections were warmed briefly to room temperature and fixed in 70% alcohol for a half min and then in acetone for another half min. The sections were rehydrated with PBS for 2 min, then incubated with an anti-tyrosine hydroxylase (TH) antibody (Vector Laboratories, Inc., Burlingame, CA) diluted at 1:100 for 5 min, ABC reagent for another 5 min, and then DAB (3,3'-diaminobenzidine tetrahydrochloride)

solution for 2 min. Sections were washed with PBS twice for one min, then dehydrated for a half min in 75% ethanol, another half min in 95% ethanol, twice for 1 min in 100% ethanol, and for 5 min in xylene, followed by air-drying and loading into a PixCell II LCM instrument (Arcturus, Mountain View, CA, USA). TH-positive DA neurons were captured onto CapSure™ HS caps covered with thermoplastic film. The harvested neurons (approximately 1,000 per sample) were dissolved from the film with 20 µl extraction buffer (included in Arcturus Pico Pure™ RNA isolation kit) for a half hour at 42°C, followed by -80°C storage.

*Quantitative real-time polymerase chain reaction (qRT-PCR) analysis of SLC6A3 mRNA levels*

Extracted RNA was converted into first-strand cDNA in 20µl by using a Verso cDNA Synthesis Kit (Thermo Scientific, Surrey, UK) in a cycling program of 42°C for 30 min followed by 95°C for 2 min. qRT-PCR was carried out as described elsewhere<sup>1</sup>. qRT-PCR used Invitrogen's One-Step qRT-PCR with SYBR® Green kit, 5'-cacagaagcacagtgtgccat-3' & 5'-tgagcagcatgatgaagaagacca-3' (amplicon 134bp) as PCR primers, and the CFX Connect™ Real-Time PCR Detection System (Bio-Rad, Life Science Research, 2000 Alfred Nobel Drive, Hercules, CA 94547, USA). Relative transcript concentration was inferred by normalizing against *β-actin* (5'-actgtgccatctacag-3' & 5'-ggtggtgaagctgtagcc-3'). *β-actin* was chosen as an internal control because *β-actin* was chosen as an internal control because cases and controls did not differ in its expression and this internal control could address postmortem tissue-associated RNA quality issue in qRT-PCR.<sup>2,3</sup> Phenotypic differences in genotypic expression were assessed by two-way ANOVA using GraphPad v5.03 (GraphPad Software, Inc., La Jolla, CA 92037 USA).

## *Genotyping*

Extraction of DNA samples from blood was performed using a standard high-salt method.<sup>4</sup> Genotyping of all markers was performed at the Neurogenetics Laboratory of CAMH. Two commercially available and six custom TaqMan® SNP Genotyping Assays were genotyped in all three sample sets (LifeTechnologies, Burlington, ON, Canada). Custom assay primers and probes, listed in Supplementary Table 1, were designed using the Custom TaqMan® Assay Design Tool on the LifeTechnologies website (<https://www.lifetechnologies.com/order/custom-genomic-products/tools/genotyping/>). For each reaction, 20 ng of genomic DNA was scaled to a total volume of 10 µL in an Applied Biosystems (AB) 2720 thermal cycler and amplified per manufacturer's directions. Post-amplification products were analyzed on the AB ViiA7 Real-Time PCR System and genotype calls were determined manually by comparison to six No Template Controls.

For the VNTR markers, 50 ng total genomic DNA was combined with 1X MBI Fermentas PCR buffer containing KCl, 1.5 mM MgCl<sub>2</sub> (MBI Fermentas), 0.13 µg each primer (Supplementary Table 1; forward primer labeled with 5' fluorescent modifier), 0.16 mM each dNTP (MBI Fermentas) and 1 U Taq polymerase (MBI Fermentas) to a total volume of 25 µL. The PCR reactions were subjected to an initial denaturation at 95°C for 5 min, followed by 30 cycles of amplification in an AB 2720 thermal cycler: denaturing at 95°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds, and a final extension at 72°C for 10 min. One microliter of the PCR product was electrophoresed on an AB 3130-*Avant* Genetic Analyzer per manufacturer directions, and product sizes were determined by comparison to GeneScan 1200 (rs70957367) or 500 (rs3836790) LIZ size standard using

GeneMapper (version 4.0). Genotyping of 10% of the samples from each run was replicated for quality control purposes for each marker.

Independent genotype verification for rs67175440 in 40 brain samples was conducted at the Laboratory of Psychiatric Neurogenomics, McLean Hospital. DNA was amplified in a 10  $\mu$ L volume using the Maxima Hot Start PCR Master Mix (Fermentas). The primer sequences for both actin beta and *SLC6A3* have been described elsewhere.<sup>1</sup> The amplifications were performed in a thermocycler (Applied Biosystem, Life Technologies, Grand Island, NY 14072, USA) with initial denaturation at 95°C for 4 min, denaturation at 95°C for 30 seconds, annealing temperature of 63°C, extension at 72°C for 30 seconds, for 40 cycles, followed by final extension at 72°C for 7min. PCR product of 174 bp was subjected to restriction enzymatic digestion by *Bse*RI, followed by fragment sizing on 1% agarose gel and visualized under UV. Allele G was sensitive to the digestion for 127- and 47-bp bands.

#### *Ancestry Ascertainment*

A total of 64 ancestry-informative markers (AIMs)<sup>5</sup> were genotyped for all case and control subjects to ascertain ancestry. AIMs were genotyped using the TaqMan<sup>®</sup> OpenArray<sup>®</sup> method (LifeTechnologies<sup>®</sup>, CA, USA) per manufacturer's directions on the QuantStudio<sup>™</sup> 12K Flex Real-Time PCR System (LifeTechnologies<sup>®</sup>, CA, USA). Briefly, 2.5  $\mu$ L of DNA at a concentration of 20 ng/ $\mu$ L and 2.5  $\mu$ L of 40X TaqMan<sup>®</sup> OpenArray<sup>®</sup> Master Mix were combined manually in 384-well plates and then loaded onto the OpenArray<sup>®</sup> genotyping plates using the AccuFill System (LifeTechnologies<sup>®</sup>, CA, USA). Prepared arrays were amplified, visualized and analyzed on the QuantStudio<sup>™</sup> 12K Flex system. Genotypes were imported into the TaqMan<sup>®</sup> Genotyper software v1.3 and confirmed manually by two independent lab personnel. Six SNPs

were immediately excluded due to poor resolution of the genotype groups (rs1040404, rs7421394, rs7657799, rs870347, rs772262, rs3907047). Subjects with a missing genotype rate >10% were excluded from all analyses. Quality control filters for SNPs included a minor allele frequency >5% and a missing genotype rate of <5%. Of the 58 AIMs with adequate allelic discrimination, two SNPs had missing rates >10% (rs1408801 and rs3745099). Replication was performed for a random sample of 10% of study subjects for quality control, with concordance >97% for all 56 SNPs passing quality control filters. These 56 SNPs were used to calculate the first five components by multidimensional scaling analysis using PLINK.<sup>6</sup> Of the 56 SNPs passing quality control, those genotyped in five Phase 3 HapMap populations were analyzed: Utah residents with Northern and Western European ancestry from the CEPH collection (CEU), African ancestry in Southwest USA (ASW), Yoruban in Ibadan, Nigeria (YRI), Japanese in Tokyo, Japan (JPT), and Han Chinese in Beijing, China (CHB). (International HapMap Consortium)<sup>7</sup>

### *Association analysis*

PLINK was used for both data management and allelic association analysis.<sup>6</sup> Permutation procedures with 20,000 iterations (for EMP2 used) or SNPSpD<sup>8</sup> were used to estimate corrected *P*-values for multiple testing (adjusted *P* values). Linkage disequilibrium (LD) was analyzed and graphically displayed by Haploview.<sup>9</sup> Two GWAS datasets for US Caucasian SCZ in dbGaP (phs000021.v3.p2 and phs000167.v1.p1) were consulted for meta-analysis in PLINK. Standard quality control procedures were used to clean the datasets.<sup>10</sup> Caucasians and African Americans were separated from each other before the analyses. All *P* values < 0.05 were considered as statistically significant, either uncorrected or adjusted.

### *Imputation analysis*

Shapeit/2.727<sup>11</sup> ([https://mathgen.stats.ox.ac.uk/genetics\\_software/shapeit/shapeit.html](https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html)) was used to estimate haplotypes, impute2/2.3.0<sup>12</sup> ([https://mathgen.stats.ox.ac.uk/impute/impute\\_v2.html](https://mathgen.stats.ox.ac.uk/impute/impute_v2.html)) for high quality imputation (threshold of 0.9 using GTOOL version 0.7.5) from 1.3Mb to 1.5Mb of chromosome 5 (chr5) based on 1000G\_phase1integrated\_v3\_impute\_macGT1; only diallelic markers were retained, of which only SNPs with genotype rates  $\geq 5\%$  and MAFs of  $\geq 1\%$  were retained for the association analyses.

### *Plasmid construction for luciferase (Luc) reporting of rs1478435 allelic regulation of 2.5 kb SLC6A3 promoter activity*

The SV40 promoter in Promega's pGL3 Promoter Vector was replaced by a 2.5 kb core promoter of *SLC6A3* through the *Bgl*III/*Hind*III sites. The 2.5 kb fragment was PCR amplified using *PfuUltra* II Fusion HS DNA Polymerase (Agilent Technologies, Santa Clara, CA, USA) with the two primers, 5'-tttagatctttcacggagcattcctgtgt-3' and 5'-ttttaagcttcagcgctgggcggtctca-3', resulting in pSLC6A3-2.5kb. There was a *Kpn*I site upstream of the *Bgl*III site so that *Kpn*I/*Bgl*III double digestion of pSLC6A3-2.5kb allowed insertions of two rs1478435 alleles, C and T. Each allele was prepared by annealing two DNA oligonucleotides f (forward) and r (reverse), as listed below:

Oligo name	Sequence 5' to 3'
rs1478435Tf	cat atg aag aaa agt gaa tgg ggg cca gac gca gtg gct cat gcc t
rs1478435Tr	gat cag gca tga gcc act gcg tct ggc ccc cat tca ctt ttc ttc ata tgg tac
rs1478435Cf	cat atg aag aaa agt gaa tgg ggg ccg gac gca gtg gct cat gcc t

rs1478435Cr    gat cag gca tga gcc act gcg tcc ggc ccc cat tca ctt ttc ttc ata tgg tac

The insertions yielded two allelic expression plasmids, pSLC6A3-2.5kb-T and pSLC6A3-2.5kb-C, where two alleles regulated the 2.5 kb *SLC6A3* promoter, which was located upstream of firefly luciferase gene *luc* in the Promega's vector.

#### *Promoter activity analysis by Luc reporting*

Procedures for analyzing regulation of the *SLC6A3* promoter activity by rs1478435 alleles have been described elsewhere. Briefly, we used at least three independent preparations of plasmid DNA, each with the same quality (optical density 260/280 ratio of approximately 1.8), for a mixture of the same genotype in order to rule out any DNA quality-related artifact; we used two wells in each experiment and repeated the experiment 4-5 times in different plates and cell passages for the same treatment/condition, in order to rule out well/plate-related artifact. SK-N-AS of human origin or SN4741 of mouse origin cells, maintained in Dulbecco's Modified Eagle Medium with 10% fetal bovine serum and 1x penicillin/streptomycin at 32°C, were split one day before transfection by using Qiagen's SuperFect Transfection Reagent (2.4 µL and 0.4 µg DNA for each of 24 wells per plate; Valencia, CA, USA) per manufacturer's protocol. Transfected cells were always grown at 37°C for *luc* expression. A half mL of fresh medium was added to each well one day after transfection, followed by harvesting one day later for protein quantification (Bio-Rad Protein Assay Kit I, Bio-Rad, Hercules, CA, USA). Ten percent of the cell lysate in each well was used for Luc activity measurements in Promega's Luciferase Assay System with Bio-Tek/Gen5 (Winooski, VT, USA). Data are all presented as fold of vector



pSLC6A3-2.5kb activity in mean  $\pm$  standard error. ANOVA or *Student's t*-test results with *P* values of  $< 0.05$  were considered as statistically significant.

## References

1. Zhou Y, Michelhaugh SK, Schmidt CJ, Liu JS, Bannon MJ, Lin Z. Ventral midbrain correlation between genetic variation and expression of the dopamine transporter gene in cocaine-abusing versus non-abusing subjects. *Addict Biol* Jan 2014;19(1):122-131.
2. Hashimoto R, Straub RE, Weickert CS, Hyde TM, Kleinman JE, Weinberger DR. Expression analysis of neuregulin-1 in the dorsolateral prefrontal cortex in schizophrenia. *Molecular psychiatry* Mar 2004;9(3):299-307.
3. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* Apr-Jun 2006;27(2-3):126-139.
4. Lahiri DK, Nurnberger JI, Jr. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic acids research* Oct 11 1991;19(19):5444.
5. Kosoy R, Nassir R, Tian C, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Human mutation* Jan 2009;30(1):69-78.
6. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* Sep 2007;81(3):559-575.
7. The International HapMap Project. *Nature* Dec 18 2003;426(6968):789-796.
8. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *American journal of human genetics* Apr 2004;74(4):765-769.
9. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics (Oxford, England)* Jan 15 2005;21(2):263-265.
10. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc* Sep 2010;5(9):1564-1573.
11. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods* Feb 2012;9(2):179-181.
12. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics* Jun 2009;5(6):e1000529.

**Supplementary Table 1.** DNA oligos used for genotyping.

Variant	Assay ID/ Primer-probe sequences
rs11564750	C__31406264_10
rs11564751	F: TGCCACAGAAGGGATCTCTCTAC R: CCAAGCCCTGGCTCACT VIC: ACAGGTTTCAATGAAAAG FAM: AGGTTTCAATAAAAAG
rs12652860	F: CCACCCCTCCTCTCAACCA R: ACGAATCCAAATTCTAGAGATGAAAACCA VIC: TGCATTTTGCATCTGAG FAM: TGCATTTTGCCTCTGAG
rs2455391	F: GGCATTCACGGGCATTCC R: GTGCCCCAGCTGTACCAT VIC: CGTTCTCAGCGCCTG FAM: CGTTCTCAGCACCTGA
rs2550948	C__3284858_10
rs3756450	C__27477615_10
rs67175440	F: CATGGAGGCCTCAAGACAGA R: ACAGCCTTGCACTTTGATAGGA VIC: CACTCTGGTTTGCTCCT FAM: ACTCTGGTTTGCCTCT
rs6860992	F: GGAATGCAATGGCACCATATCAG R: TGACGTGGGAGAATCTCTTGAAC

	VIC: CACTGCAACCTCCACC FAM: CACTGCAACTTCCACC
rs70957367	F: 5'- NED-GCTGAAAGCCACTGAAAAGG R: 5'- GCCGCATTCGTCTTTAGTCT
rs3836790	F: 5'- 6FAM-CTGGTCCTGCCCTTCATC R: 5'- TGAGTTTGAATGCACAAATGAG

**Supplementary Table 2.** Discordance between AIM and Self-report on "Caucasians".

Sample	AIM only	Self-report only	AIM and self-report	% Discordance
Controls (CTL)	817	857	811	0.7
Toronto (TS)	433	427	388	10.4
McLean (LS)	173	187	167	3.5
Brain tissue	33	39	33	0.0
Total	1456	1510	1399	3.9

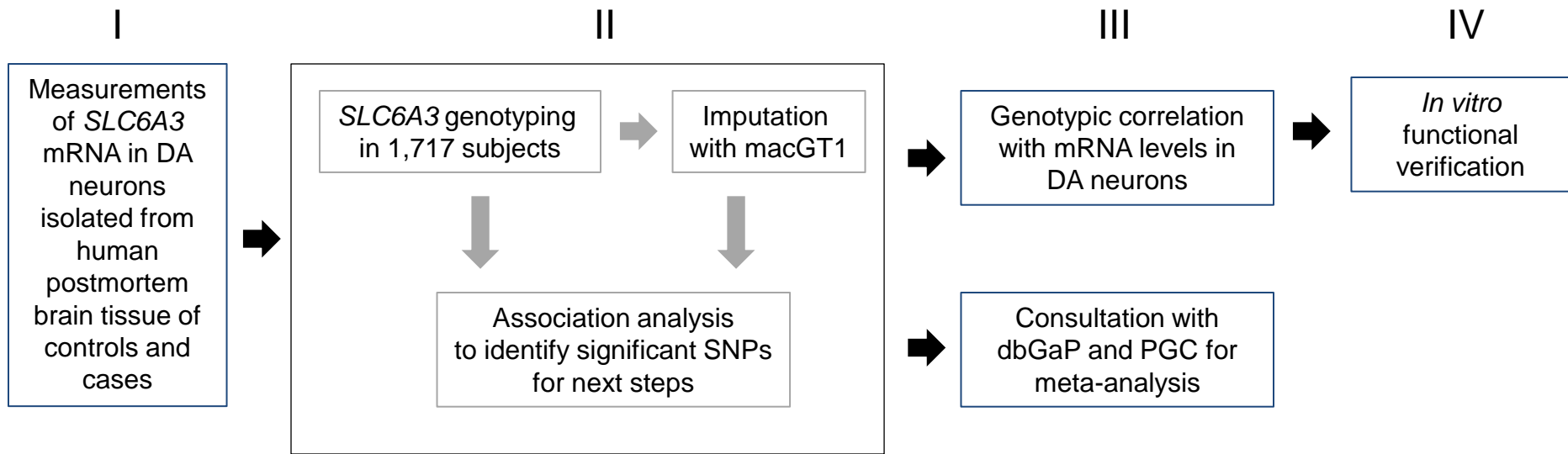
Note: % discordance =  $100 \times (\text{AIM only} - \text{AIM and self-report}) / \text{AIM only}$ .

**Supplementary Table 3.** Meta-analysis of four cohorts with ancestry and phenotype QC.

BP	SNP	A1	A2	N	P	OR	
1378700	rs55989583	T	C		2	0.6452	0.9747
1378772	rs55754205	T	C		3	0.5982	0.9762
1379762	rs37017	T	C		3	0.2661	0.9367
1379826	rs2113328	A	G		3	0.7625	0.9867
1391161	rs11133762	T	C		3	0.6296	0.9634
1395077	rs40184	T	C		3	0.715	1.0124
1396813	rs11564773	C	T		2	0.3197	1.086
1398007	rs11564772	T	C		3	0.9187	1.0061
1398797	rs11564771	A	G		3	0.804	1.0151
1398806	rs11564770	G	A		3	0.804	1.0151
1404466	rs11564768	G	A		2	0.932	0.9895
1404548	rs6869645	T	C		3	0.8372	0.9875
1404726	rs11564766	T	C		2	0.4623	1.0769
1404922	rs11564765	C	T		2	0.6931	1.0417
1405806	rs11564764	T	C		2	0.4828	1.0754
1406036	rs6876225	A	C		3	0.8631	1.0112
1406663	rs11564763	A	G		2	0.2642	1.127
1409719	rs2245660	T	C		2	0.9417	0.9921
1423905	rs464049	A	G		3	0.2257	0.9595
1424364	rs62331112	G	A		2	0.114	0.9139
1425049	rs10475005	G	A		2	0.1116	0.9139
1425159	rs250686	T	C		3	0.1338	0.9499
1425166	rs57212133	C	T		2	0.1997	0.9266
1425235	rs10475006	G	A		2	0.1054	0.9124
1425313	rs61696543	G	T		2	0.1054	0.9124
1426079	rs250684	T	C		2	0.04769	1.1021
1426151	rs250683	C	T		2	0.01129	1.129
1426166	rs10051340	G	C		2	0.1054	0.9124
1426267	rs72715534	G	T		3	0.8142	0.976
1426416	rs418265	A	G		3	0.2134	0.9583
1426959	rs2173947	G	A		2	0.05757	0.9034
1427803	rs250682	C	G		2	0.01437	1.1231
1428011	rs250681	C	T		3	0.2276	0.9595
1428111	rs10052016	G	A		2	0.048	0.9007
1428135	rs10053602	C	T		2	0.1358	0.9222
1428514	rs393795	T	G		2	0.0358	1.1082
1428883	rs365663	A	G		2	0.5489	0.9742
1429187	rs427284	C	G		2	0.01424	1.1241
1430093	rs458334	A	G		3	0.19	0.9555
1430244	rs464061	A	G		2	0.01303	1.1271
1430616	rs11737901	T	C		3	0.2964	0.9584
1430775	rs410209	G	A		2	0.01391	1.126
1430834	rs409588	T	G		2	0.07999	1.0867

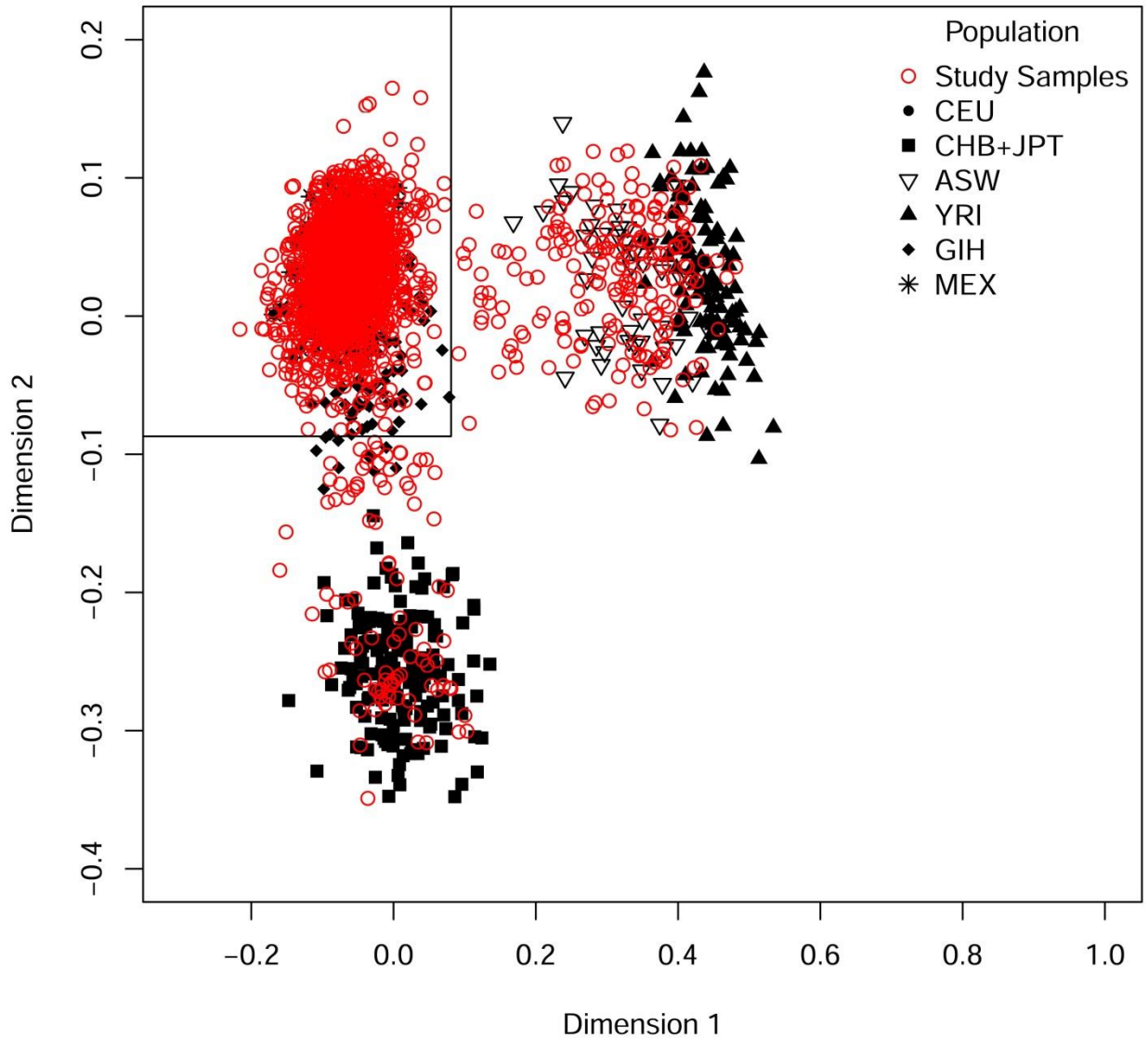
1432825	rs460000	T	G	2	0.08989	1.0834
1432876	rs465130	A	G	2	0.07466	1.0882
1433272	rs62331115	A	C	2	0.1694	0.923
1433401	rs4975646	A	G	3	0.3329	0.9615
1438105	rs1048955	C	T	2	0.6095	0.9735
1438174	rs1048953	A	G	2	0.4405	0.96
1438354	rs403636	A	C	3	0.9066	0.9951
1443728	rs2937639	C	T	3	0.8361	0.9929
1445616	rs2975226	A	T	2	0.4357	1.0327
1446389	rs2652511	A	G	2	0.4628	1.0307
1446719	rs2617596	C	G	2	0.7639	0.9878
1447027	rs6413429	A	C	2	0.5728	1.0539
1447860	rs2652510	T	C	2	0.791	0.9876
1448148	rs3756450	G	A	4	0.1126	1.0696
1448246	rs2617595	T	C	2	0.942	0.9966
1448288	rs2617594	G	A	2	0.7412	0.9848
1450444	rs2550948	C	T	3	0.2696	0.9559
1450461	rs10079467	C	T	2	0.2859	1.1065
1450506	rs2550947	T	C	3	0.3083	0.9592
1450513	rs2550946	A	G	3	0.2791	0.9567
1453772	rs12652860	A	C	3	0.07371	0.9292
1454004	rs12654851	T	G	2	0.8761	1.0114
1454612	rs1478435	T	C	3	0.06009	0.9197
1455250	rs10063727	A	G	2	0.3958	1.0835
1455516	rs4639276	C	T	3	0.5217	0.9737
1455799	rs72717506	A	G	2	0.3958	1.0835
1455946	rs6860992	C	T	3	0.4692	0.9707
1456803	rs10060889	A	G	3	0.09892	0.9332
1457554	rs748209	A	C	3	0.06827	0.9221
1457986	rs905201	C	T	2	0.5716	0.9745
1458018	rs2937650	A	G	3	0.06174	0.9205
1461167	rs7737692	G	A	2	0.9501	1.0027
1461389	rs2292023	A	C	2	0.8375	1.0133
1462098	rs9728	G	A	2	0.7444	1.0215
1462802	rs2036391	A	G	2	0.3246	1.0949
1463589	rs56264807	C	T	2	0.2199	1.1226
1465730	rs10068779	G	A	2	0.6003	1.0493
1465732	rs10068782	G	A	2	0.5956	1.0498

Step:



**Supplemental Fig. 1.** Four-step design for *SLC6A3* activity in SCZ.

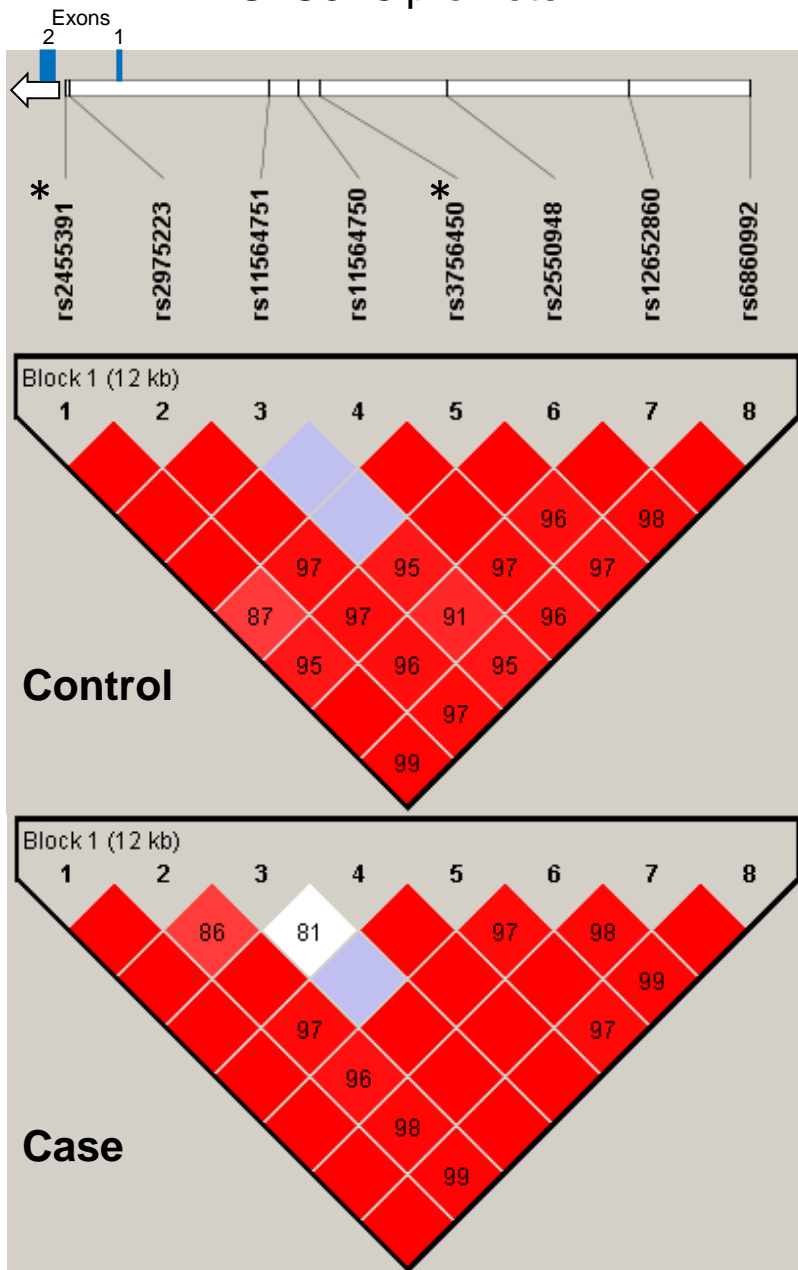
## Ancestry Clustering



**Supplementary Fig. 2.** PCA plot of subjects ethnicities (red circles), comparing to HapMap standards. AIM defined “Caucasians” are indicated by the rectangle in upper left corner.

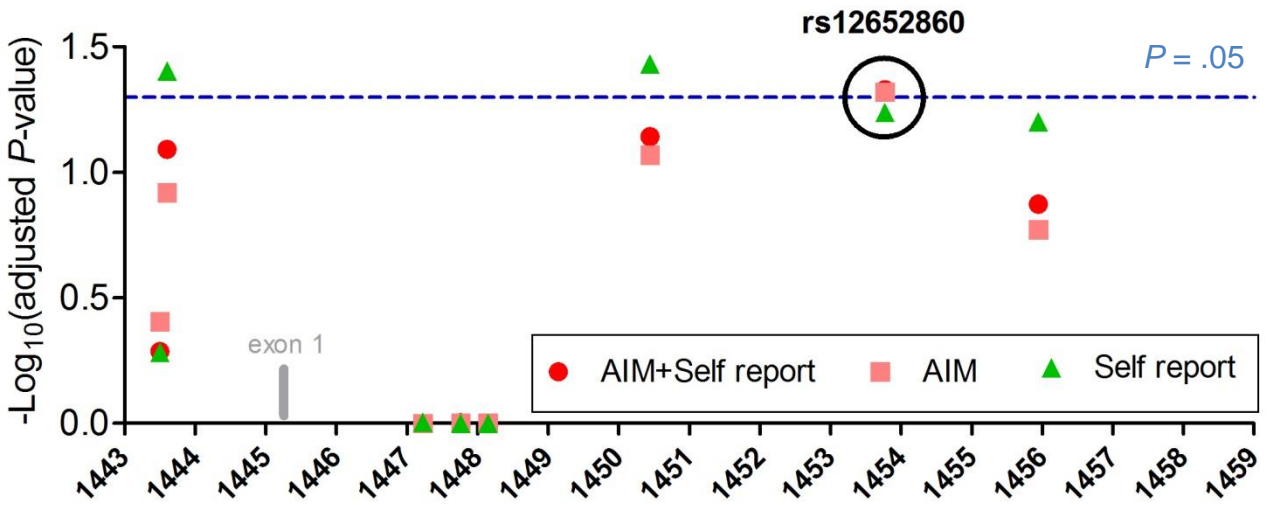


## SLC6A3 promoter

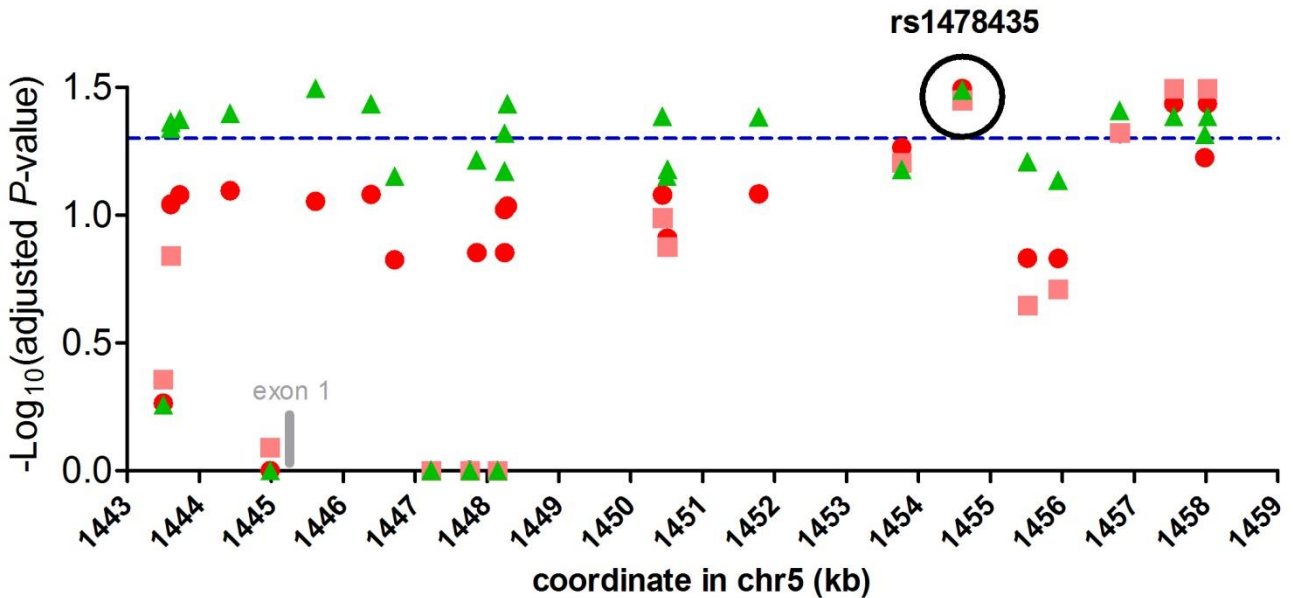


**Supplementary Fig. 3.** Linkage disequilibrium (LD) of the *SLC6A3* promoter regions in controls (upper) versus patients with SCZ (case, lower panel). Location of eight genotyped promoter SNPs is indicated on top. Two are located in Intron 1 and the other four in upstream regions. \*, previously reported positive association with SCZ.

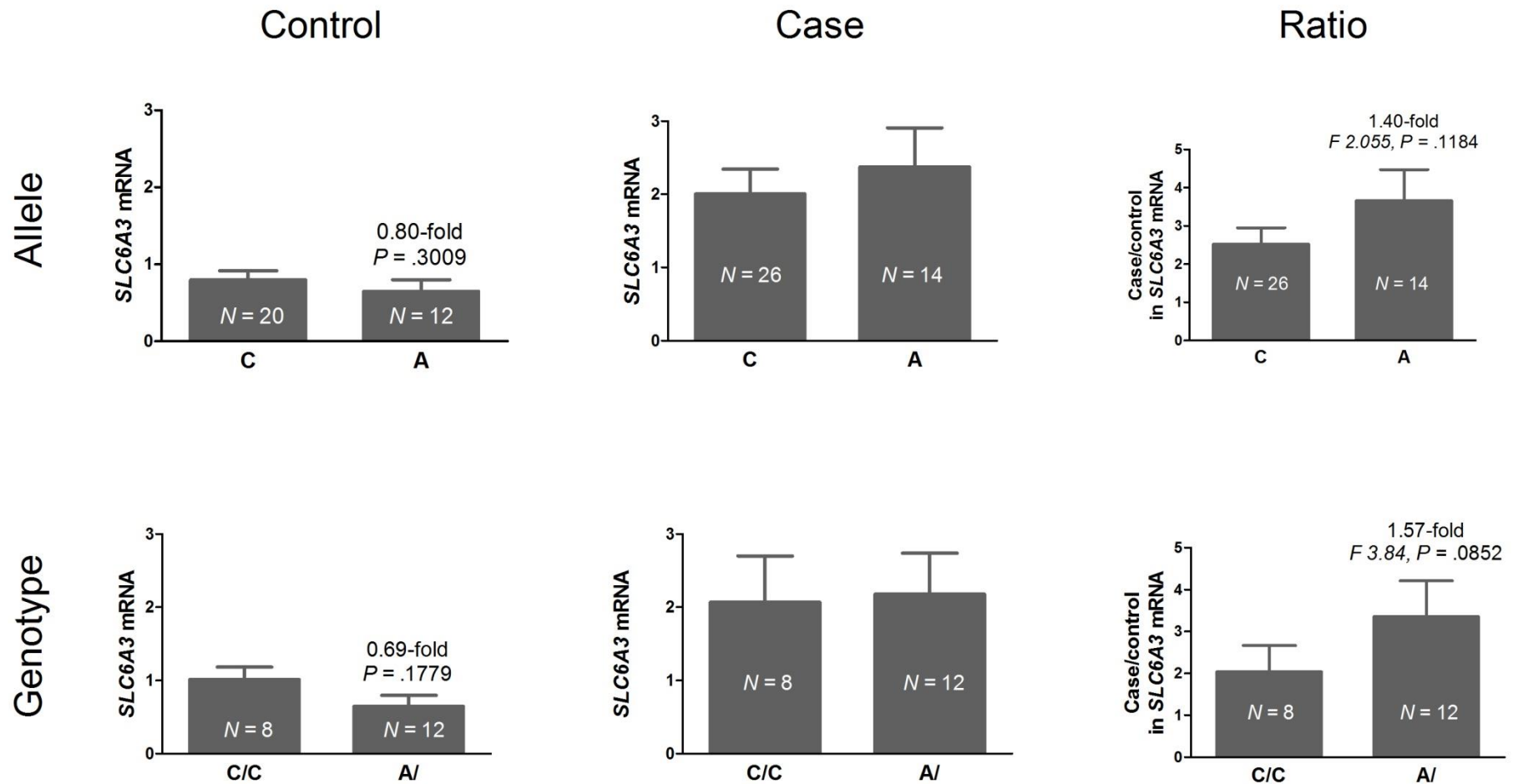
# A Genotyped



# B Imputed

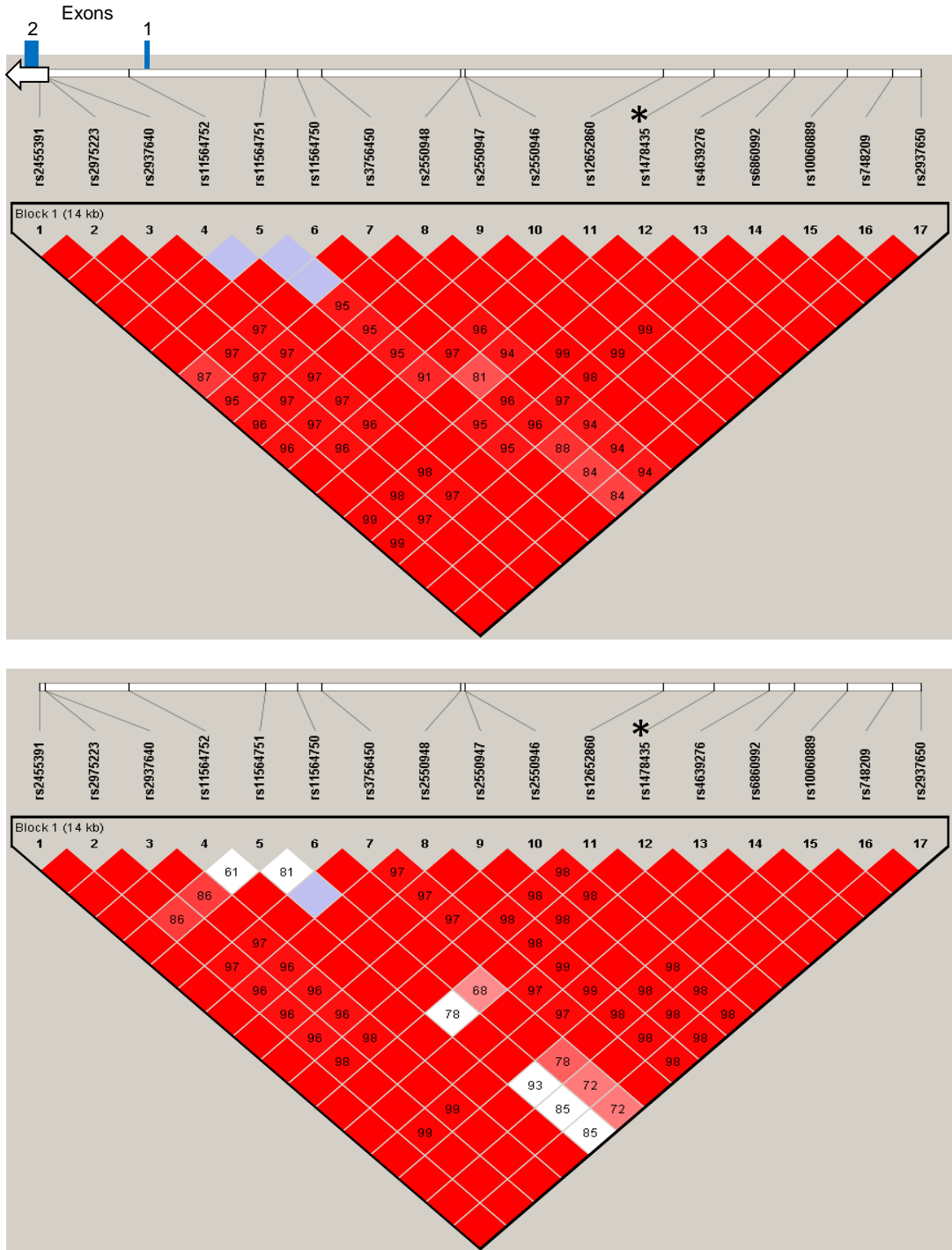


**Supplementary Fig. 4.** Association of *SLC6A3* genotype (A, also see Table 2) or imputed markers (B, also see Table 3) with SCZ by adjusted  $P$  values ( $-\text{Log}_{10}$  adjusted  $P$  values). Caucasians are defined in three categories, AIM+Self report, AIM and Self-report; blue dashed line,  $P < .05$ ; 5' promoter is on the right side of exon 1; circle, the most significant signals.



**Supplementary Fig. 5.** Correlation between *SLC6A3* mRNA levels in postmortem DA neurons and rs12652860 alleles (upper) or genotype (lower) in controls (left), cases (middle) or case/control ratio (right). Comparison was in *Student's t* tests.

# SLC6A3 promoter



**Supplementary Fig. 6.** LD of imputed genotype for controls (upper) and cases (lower). \*, rs1478435.