Supplemental Data

Profiling allele specific gene expression in brains from individuals with autism spectrum disorder reveals preferential minor allele usage

Author

Changhoon Lee[1,6], Eun Yong Kang[2], Michael J. Gandal[1], Eleazar Eskin[2,3] & Daniel H. Geschwind[1,3-5]

Supplementary Tables

Supplementary table 1. The detailed information of brain samples that used for ASE study. It contains sample ID, brain ID, region, diagnosis, detailed diagnosis, primary cause of death, secondary cause of death, age, sex, RIN, PMI, brain bass, brain bank, sequence batch, SNP annotation, ancestry, AT dropout from Picard, and GC dropout from Picard.

Supplementary table 2. Results of ASE for BA9, BA41, vermis, and cortex. For each tissue, the table contains SNP information, the output of linear mixed model, and gene information. SNP information contains chromosome, location, reference allele, alternative allele, and SNP ID. For the output of linear mixed model, beta value, standard error (SE), p-values are provided for allele, age, sex, sequencing batch, RIN, brain bank (bank), and ancestry (ethnicity). Gene info provides SNP location at all known transcripts which contains ENSG, HGNC, biotype, ENST,

and exon/intron number or promoter location. There are also the gene lists showing ASE (EMSG and HGNC genes). The sample size of BA9, BA41, vermis, and cortex were 67, 64, 64, and 131, respectively.

Supplementary table 3. Results of ASE for BA9, BA41, vermis, and cortex in control and idiopathic ASD. For each group, the table contains SNP information, the output of linear mixed model, and gene information as described at Supplementary table 2. The ASE gene lists are also provided. The sample size of BA9, BA41, vermis, and cortex in control were 37, 32, 32, and 69, respectively. In idiopathic ASD, they were 30, 32, 32, and 62, respectively.

Supplementary table 4. Additional data of ASE GO analyses in cortex. REViGO[49] generated the p-values and other results from the GO analysis. The table contains its raw GO analysis data for common ASE genes between control and idiopathic ASD groups for above interactive graph of Fig. 3d, control-specific ASE genes (**Supplementary Fig. 3c**), and idiopathic ASD-specific ASE genes (**Supplementary Fig. 3d**).

Supplementary table 5. MAE SNP numbers per chromosome (Chr). Major and minor allele MAE SNPs were counted in control, ASD, and dup15q.

Control			ASD			Dup15q		
Chr	Major allele MAE	Minor allele MAE	Chr	Major allele MAE	Minor allele MAE	Chr	Major allele MAE	Minor allele MAE
1	435	9	1	279	6	1	143	4
2	432	12	2	283	13	2	171	17
3	375	5	3	304	18	3	108	16
4	335	8	4	201	8	4	82	4
5	309	12	5	193	4	5	70	5
6	321	5	6	233	9	6	96	6
7	248	5	7	247	18	7	75	4
8	276	7	8	163	9	8	78	13
9	235	6	9	117	3	9	64	3
10	266	4	10	145	2	10	59	4
11	290	13	11	192	15	11	83	16
12	252	3	12	159	5	12	84	4
13	189	6	13	132	9	13	50	10
14	220	34	14	145	25	14	75	23
15	404	67	15	151	187	15	123	36
16	179	7	16	119	7	16	65	5
17	156	1	17	87	3	17	64	3
18	113	6	18	82	3	18	38	3
19	161	9	19	103	0	19	59	6
20	122	3	20	62	6	20	35	1
21	57	2	21	41	1	21	19	2
22	101	2	22	64	5	22	40	2
X	33	1	X	18	1	X	26	5
Total	5,509	227	Total	3,520	357	Total	1,707	192

Supplementary table 6. Gene lists used for gene set enrichment and GO studies. The lists contain ASD risk genes (SFARI[36]; Methods), known imprinted genes (Methods), PSD[32], FMRP target[28], HuR target[29], and RBFOX1 target genes[30], cell marker genes[31] (neuron, astrocyte, oligodendrocyte, microglia, and endothelial), up- and down-regulated genes in ASD cortex[4], and genes containing risk variants in psychiatric disease datasets from *de novo* variant data (SCZ[47], ID[47], ASD1[47], and ASD2[45]; Methods). Brain expressed genes (ENSG and HGNC genes) were used as background genes for the gene set enrichment and GO analyses.