

# Family Based Genome-Wide Association Study of Autism Spectrum Disorder in Middle Eastern Families

Yasser Al-Sarraj<sup>1,2</sup>, Eman Al-Dous<sup>1,2</sup>, Rowaida Z. Taha<sup>2</sup>, Dina Ahram<sup>2,3</sup>, Fouad Alshaban<sup>2</sup>, Mohammed Tolfat<sup>4</sup>, Hatem El-Shanti<sup>2,5</sup> and Omar M.E. Albagha<sup>1,2,6\*</sup>

<sup>1</sup> College of Health and Life Sciences, Hamad Bin Khalifa University, Doha 34110, Qatar; yalsarraj@qf.org.qa (Y.A.-S.); eyaldous@hbku.edu.qa (E.A.-D.)

<sup>2</sup> Qatar Biomedical Research Institute (QBRI), Hamad Bin Khalifa University, Doha 34110, Qatar; rtaha@hbku.edu.qa (R.Z.T.); da2747@cumc.columbia.edu (D.A.); falshaban@hbku.edu.qa (F.A.); hatem-el-shanti@uiowa.edu (H.E.-S.)

<sup>3</sup> Division of Nephrology, Columbia University Medical Center, New York, NY 10032, USA

<sup>4</sup> The Shafallah Center for Children with Special Needs, Doha 33123, Qatar; Mohamed.Tolfat@shafallah.org.qa

<sup>5</sup> Department of Pediatrics, Carver College of Medicine, University of Iowa, Iowa City, IA 52242, USA

<sup>6</sup> Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK

\* Correspondence: oalbagha@hbku.edu.qa

## Supplementary Materials

**Supplementary Table S1.** Family structure of the study subjects

Number of Families	Unaffected parents	Child with ASD	Unaffected Child	Total number of subjects
103	2	1	0	309
1	2	2	1	5
22	2	0	1	66
28	1	1	0	56
6	1	0	1	12
1	2	1	1	4
1	2	0	2	4
1	1	2	0	3
1	1	3	0	4
7	0	1	0	7
4	2	0	0	8
9	1	0	0	9

**Supplementary Table S2.** Replication signals for the lead SNP rs10099100\*.

Proxy SNP	Bp to lead SNP	iPSYCH Data		LD from EUR 1000 Genome		Our Data	
		P-value	OR	D'	R <sup>2</sup>	P-value	OR
rs6601521	-52172	0.001	0.95	0.93	0.30	0.013	0.95
rs7826189	-51085	0.024	1.03	0.53	0.14	0.04	1.04
rs733396	-45206	0.01	1.03	0.53	0.14	0.003	1.06
rs6601518	-28459	0.01	1.03	0.58	0.21	0.038	1.04
rs7827623	35844	0.007	0.95	0.90	0.37	0.03	0.95

\*We searched for close proxies (within 100 kb) in Linkage disequilibrium (LD) with the lead SNP based on LD data from European population from 1000 Genome (since the previously identified lead SNPs were found in studies of European populations). We then checked if the proxies were genotyped in our data set and compared results.

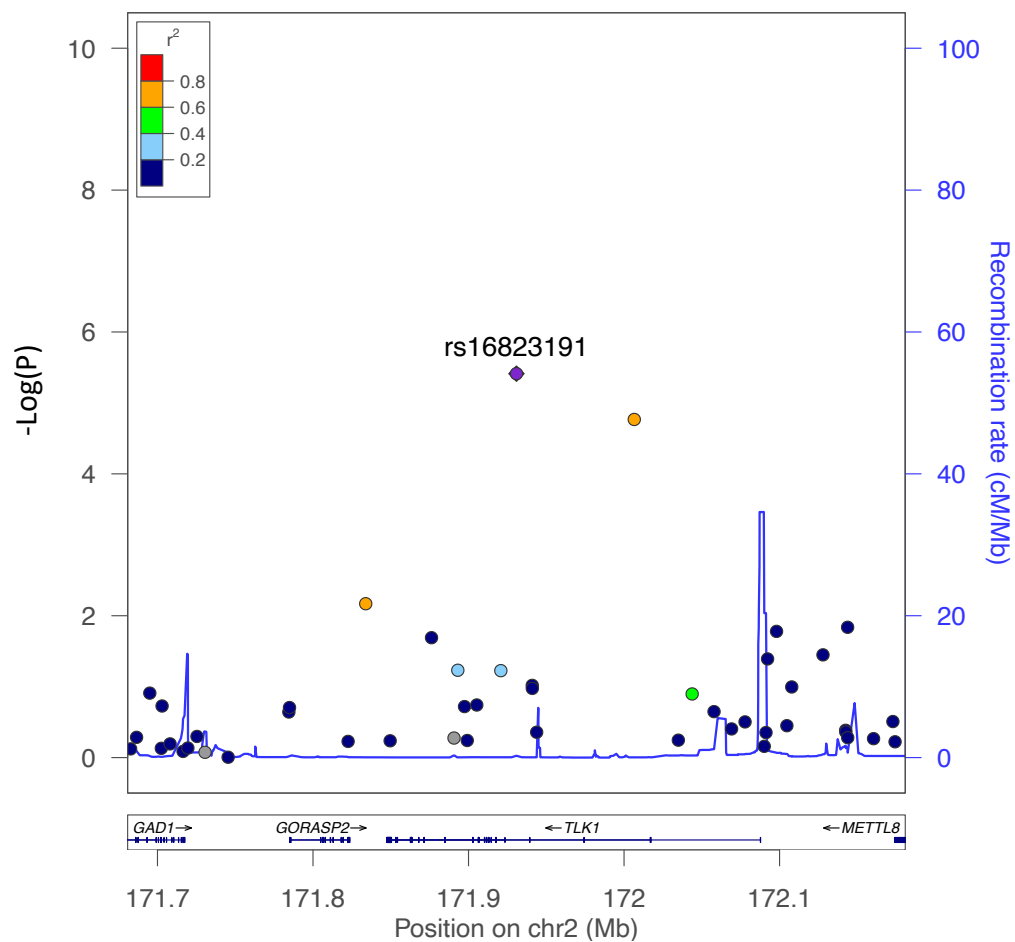
**Supplementary Table S3.** Replication analysis for SNPs with  $P < 1 \times 10^{-5}$  from the GWAS catalog\*.

	GWAS Catalog			Our Data			Effect	
SNP	Effect allele	OR	P	Effect Allele	OR	P	Same Direction	PMID
rs910805	NA	NA	1.00E-10	G	1.00	0.88	NA	32606422
rs8453	G	NA	3.00E-08	G	0.92	0.08	NA	24189344
rs6508210	NA	NA	3.00E-08	C	1.00	0.95	NA	32606422
rs6537825	A	1.4	3.00E-08	A	0.98	0.58	No	24189344
rs3827735	C	NA	6.00E-08	C	1.04	0.13	NA	24189344
rs6537825	A	NA	6.00E-08	A	0.98	0.58	NA	24189344
rs11585926	T	NA	6.00E-08	T	0.97	0.19	NA	24189344
rs11587400	C	NA	7.00E-08	C	1.00	0.93	NA	24189344
rs10858047	T	NA	8.00E-08	T	1.02	0.64	NA	24189344
rs10489525	G	NA	9.00E-08	G	1.02	0.47	NA	24189344
rs8453	G	NA	9.00E-08	G	0.92	0.08	NA	24189344
rs4150167	NA	1.96	3.00E-07	A	1.03	0.43	NA	22843504
rs910805	G	1.08	3.00E-07	G	1.00	0.88	Yes	30804558
rs2391769	G	1.07	5.00E-07	G	1.01	0.59	Yes	30804558
rs2635182	T	1.07	5.00E-07	T	1.02	0.31	Yes	30804558
rs1452075	T	1.07	6.00E-07	T	0.98	0.49	No	30804558
rs7711337	NA	1.22	8.00E-07	A	1.04	0.07	NA	22843504
rs7834018	NA	1.56	8.00E-07	G	1.00	0.98	NA	22843504
rs11185408	G	1.06	1.00E-06	G	0.98	0.29	No	30804558
rs6538761	A	1.24	2.00E-06	A	1.00	0.82	Yes	24189344
rs10038113	T	1.33	3.00E-06	T	0.97	0.21	No	19456320
rs4916723	C	1.06	3.00E-06	C	1.00	0.93	Yes	30804558
rs7142002	NA	1.56	3.00E-06	G	1.05	0.16	NA	20663923
rs10742763	T	1.31	4.00E-06	T	1.14	0.06	Yes	32747698
rs6453278	NA	1.25	5.00E-06	C	1.01	0.82	NA	19812673
rs7627815	T	1.38	8.00E-06	T	0.95	0.36	No	32747698
rs2280840	G	NA	9.00E-06	G	1.00	0.99	NA	32996047
rs6599528	A	NA	9.00E-06	A	1.00	0.83	NA	32996047
rs2447097	A	1.53	9.00E-06	A	1.00	0.82	Yes	26398136

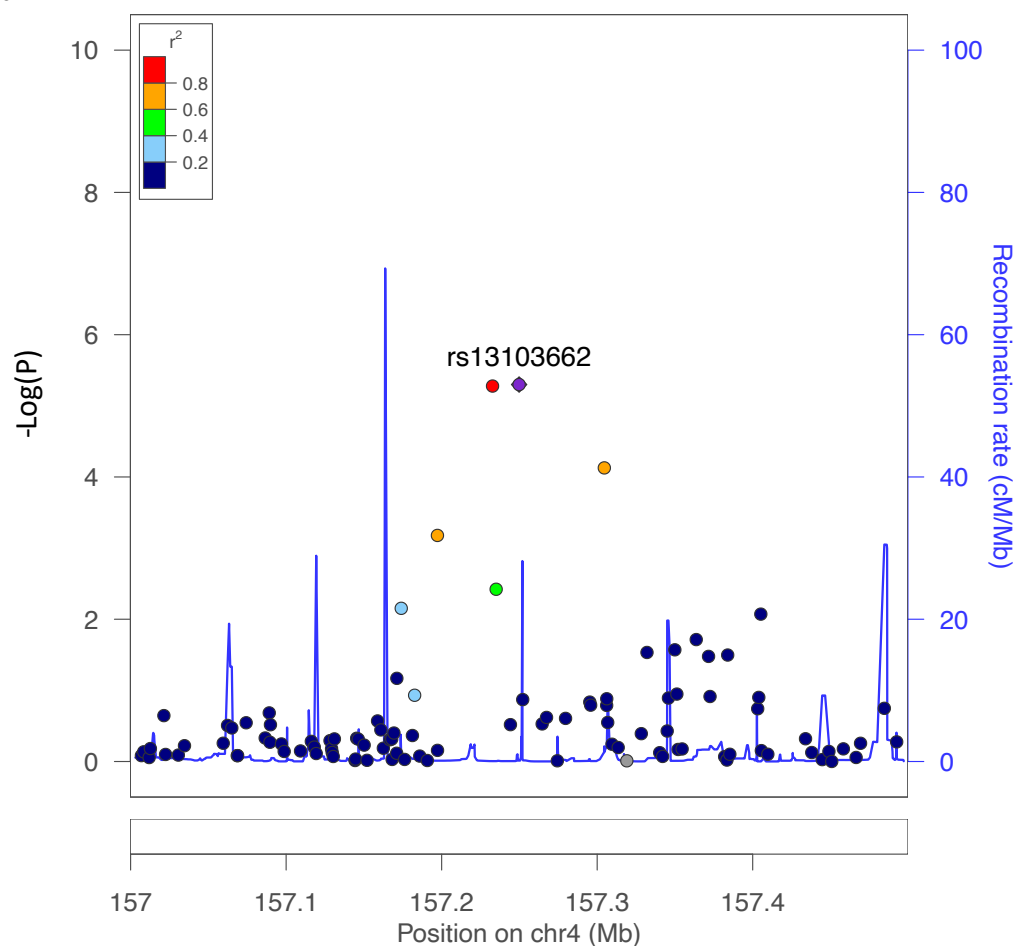
\*GWAS catalog data were retrieved from <https://www.ebi.ac.uk/gwas/home> using autism or autism spectrum disorder as phenotype. OR; odds ratio for effect allele, PMID; PubMed ID.

**Supplementary Figure S1.** Regional association plot for chr 2q31.1 (a) and 4q32.1 (b) regions

**a**

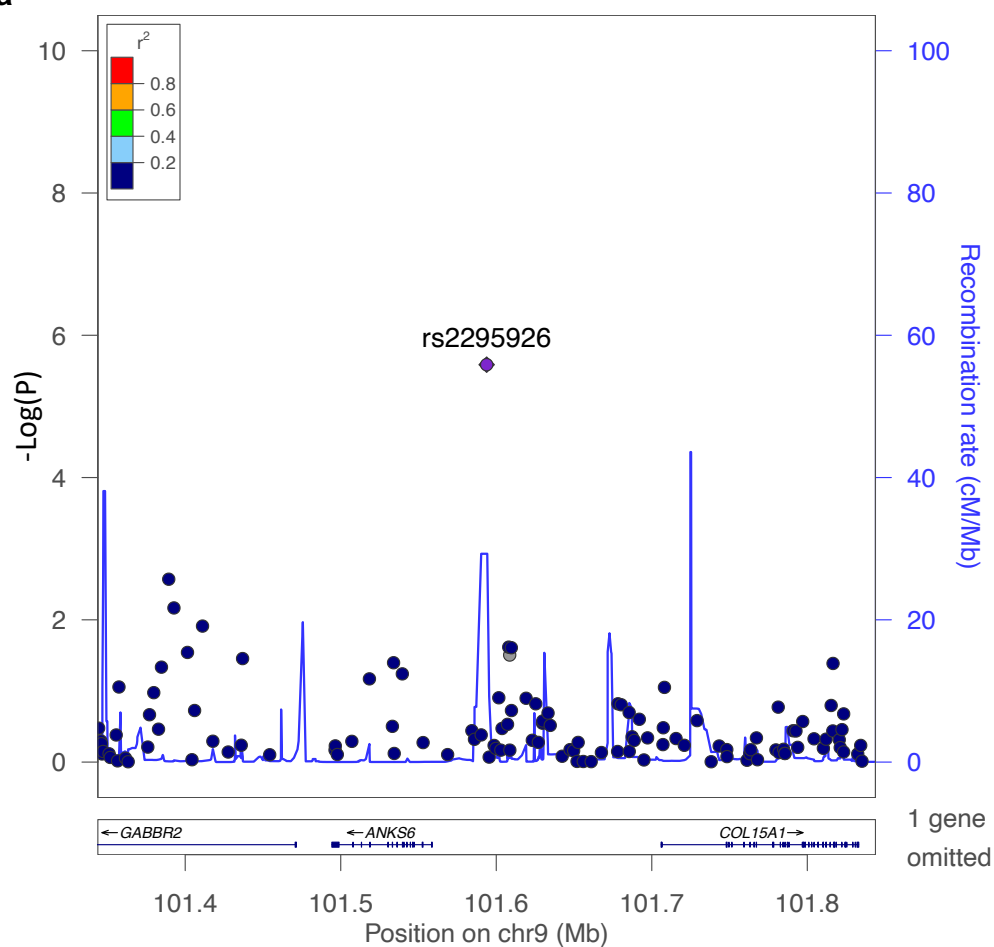


**b**

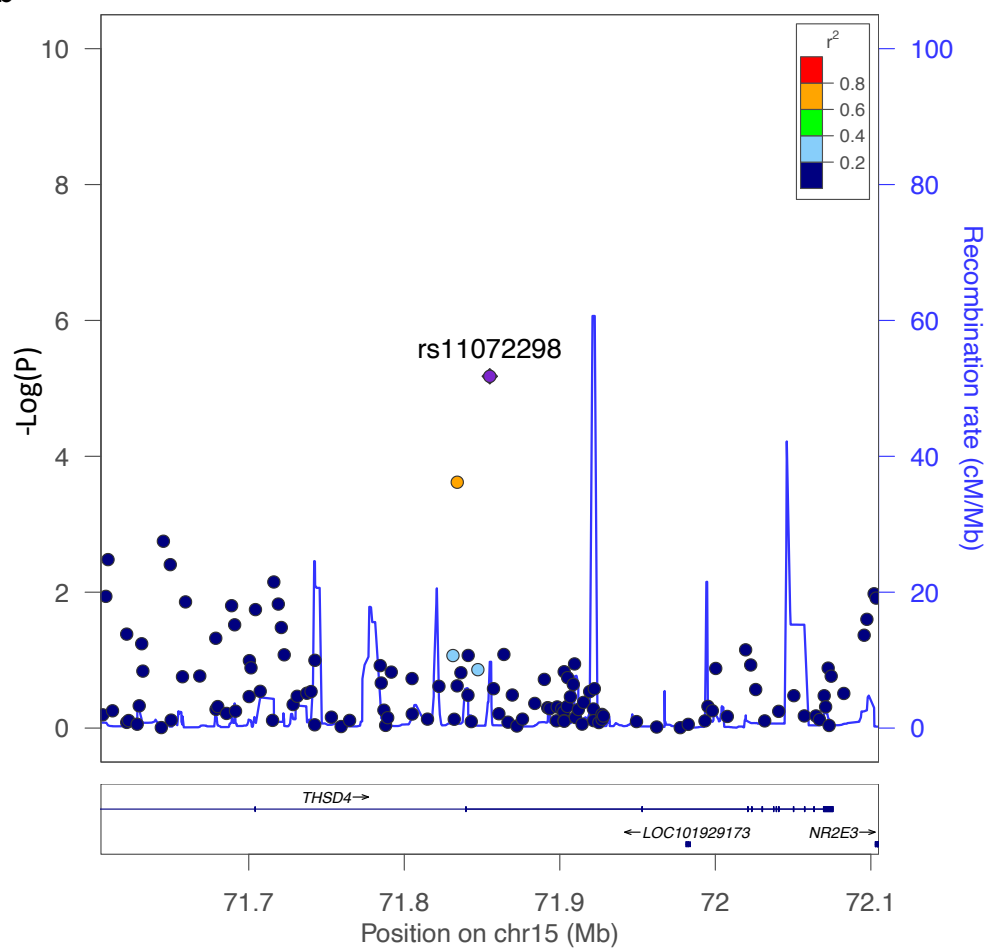


**Supplementary Figure S2.** Regional association plot for chr 9q22.33 (a) and 15q23 (b) regions

**a**

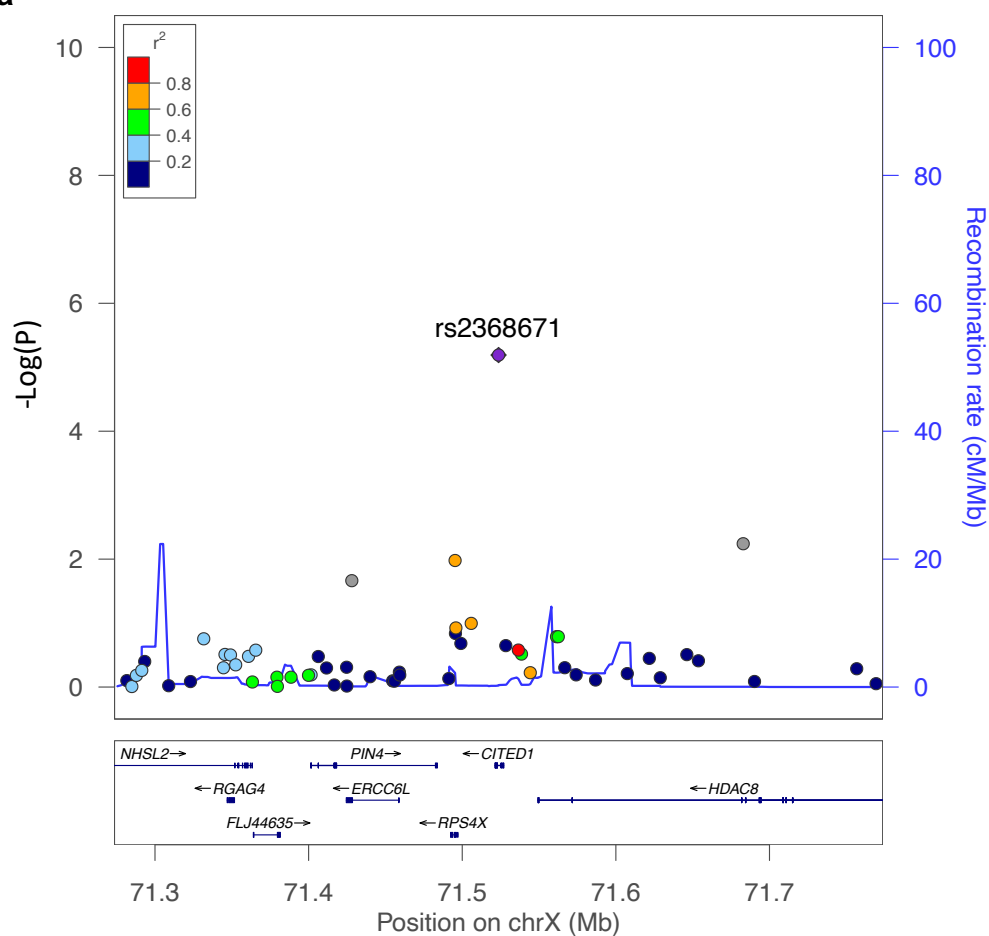


**b**

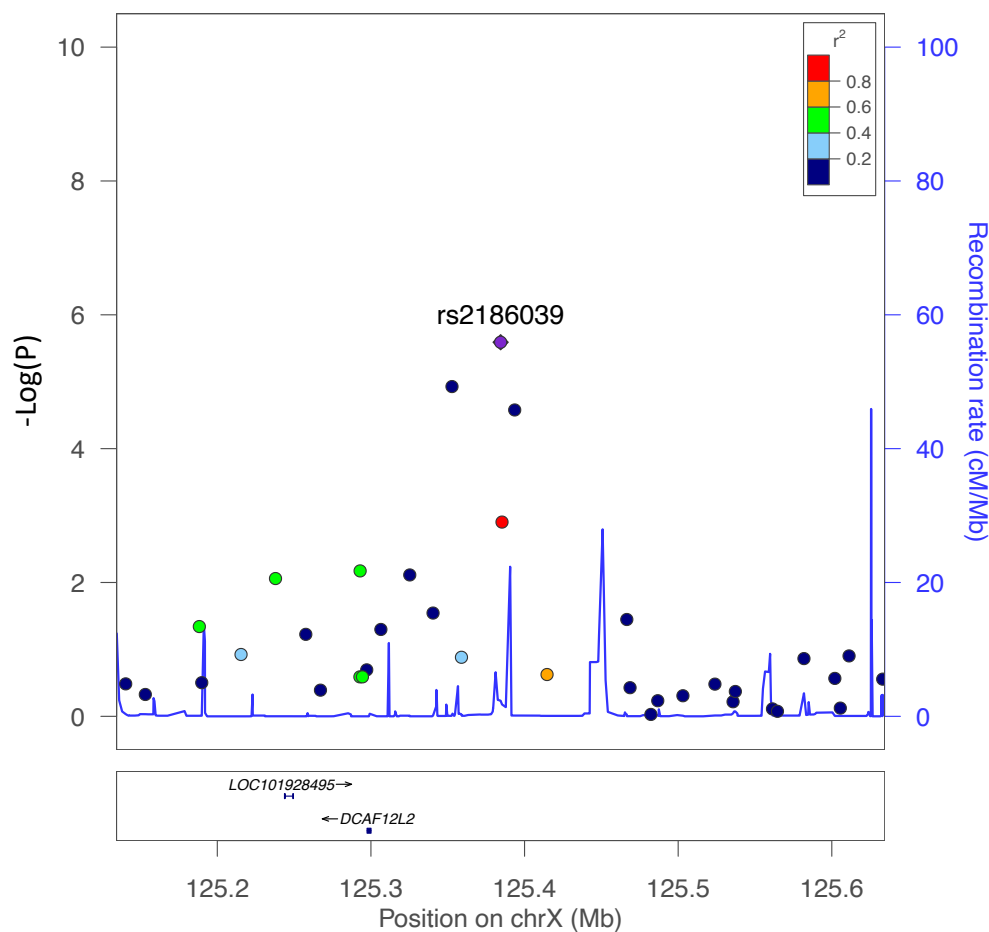


**Supplementary Figure S3.** Regional association plot for chr Xq13.1 (a) and Xq25 (b) regions

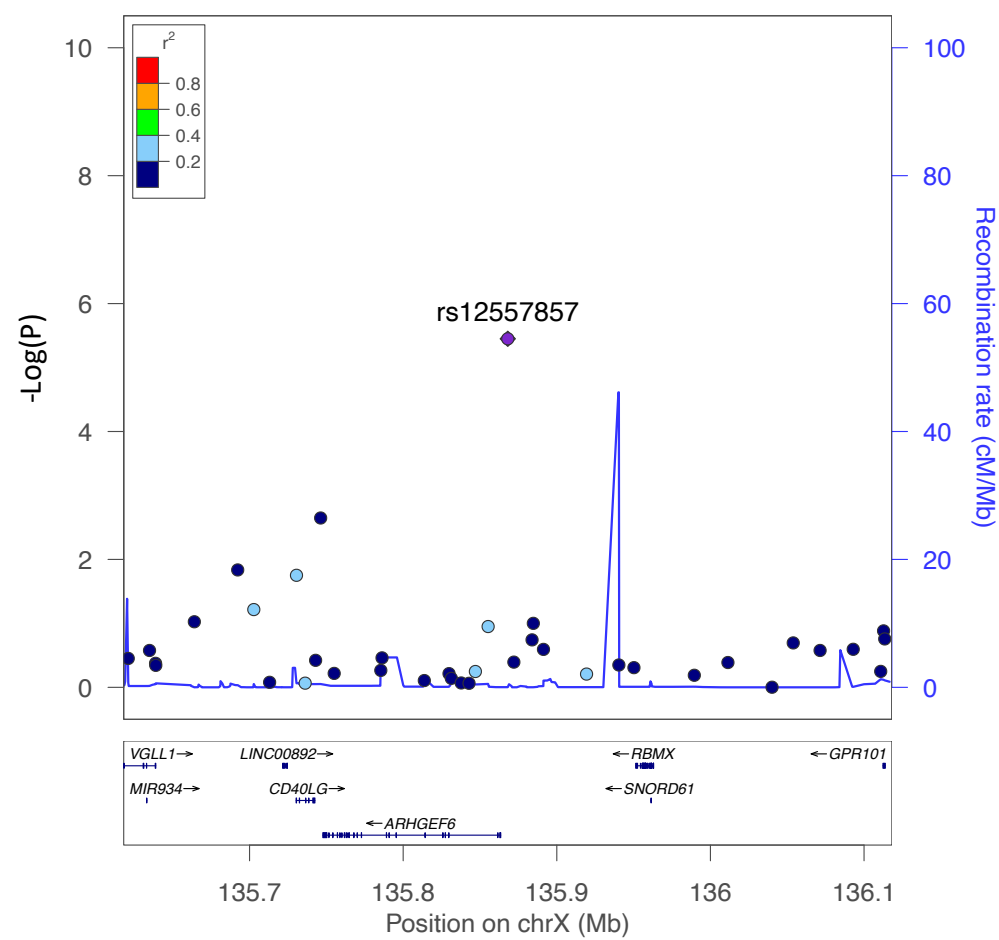
**a**



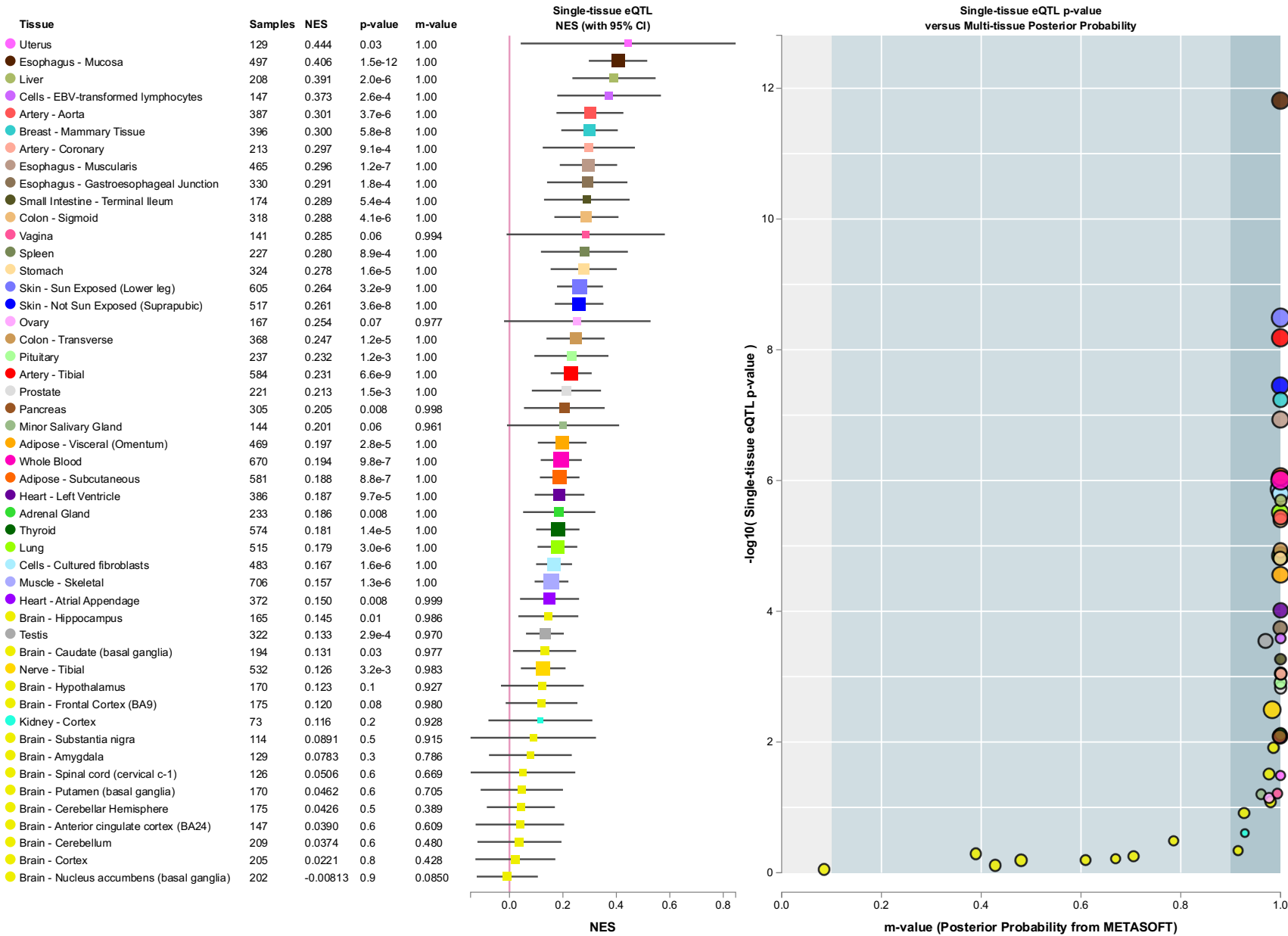
**b**



**Supplementary Figure S4.** Regional association plot for chromosome Xq26.3 region.

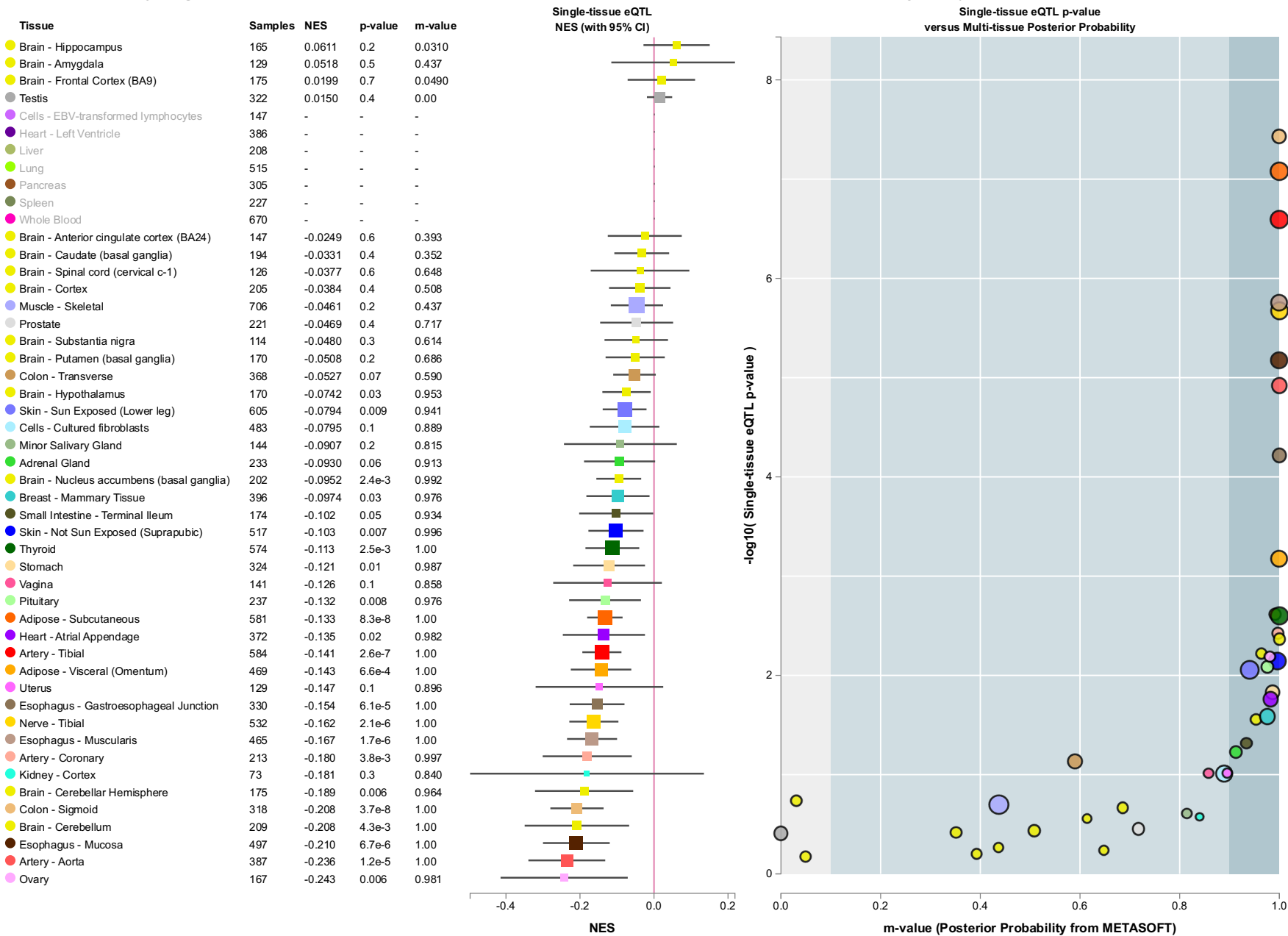


**Supplementary Figure S5.** *PIN4* expression in multiple tissues in relation to rs2368671 genotype.

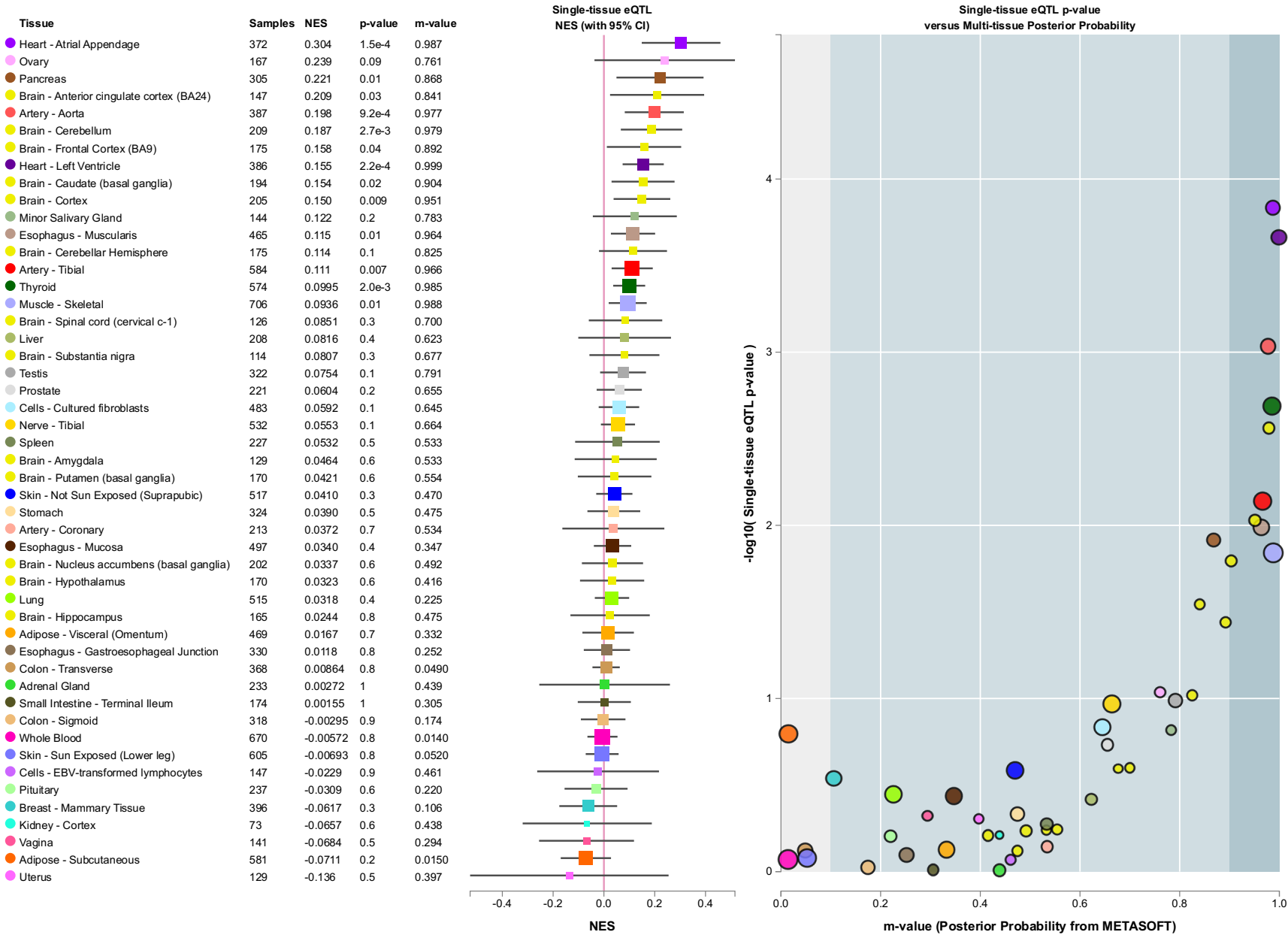




Supplementary Figure S6. DCAF12L2 expression in multiple tissues in relation to rs2186039 genotype



Supplementary Figure S7. *ARHGEF6* expression in multiple tissues in relation to rs12557857 genotype



Adapted from GTEX portal (<https://gtexportal.org/home/>).