Smith SM et al.

Polymorphisms in the choline transporter SLC44A1 are associated with reduced cognitive performance in both normotypic and prenatal alcohol-exposed children.

Supplementary Tables 1-3
Supplementary Figure 1 and legend

Supplementary Table 1. Variants in *SLC44A1* that were evaluated in the study participants.

POSTION GRCh38	ID	REF	ALT	Location
105,244,650-904	**Exon 1**			Exon 1
105,244,667	rs390529	T	C	Exon 1 (5'UTR)
105,244,725	rs555397265	C	CCAGCCGCCGCTG	Exon 1 (5'UTR)
105,245,799	rs167867	G	A	Intron 1
105,245,878	rs327960	A	C	Intron 1
105,248,163	rs170582	G	A	Intron 1
105,252,720	rs75106836	T	C	Intron 1
105,254,308	rs116350994	G	C	Intron 1
105,255,270	rs79085380	G	C	Intron 1
105,255,507	rs140200024	GTTTAA	G	Intron 1
105,256,650	rs116747281	T	A	Intron 1
105,260,869	rs76992248	C	T	Intron 1
105,262,345	rs327994	C	T	Intron 1
105,264,780	rs76992248	T	G	Intron 1
105,268,318	rs78837910	C	T	Intron 1
105,270,070	rs79632593	T	C	Intron 1
105,273,898	rs34668343	T	C	Intron 1
105,274,748	rs79971185	G	A	Intron 1
105,276,512	rs74847242	G	T	Intron 1
105,278,539	rs150185127	C	T	Intron 1
105,278,601	rs111877278	G	A	Intron 1
105,279,809	rs79223354	A	G	Intron 1
105,280,580	rs71494513	A	G	Intron 1
105,280,614	rs35293517	A	G	Intron 1
105,281,632	rs76054992	A	G	Intron 1
105,283,226	rs115995379	A	G	Intron 1
105,283,721	rs115467043	A	G	Intron 1
105,284,839	rs76066704	A	G	Intron 1
105,285,292	rs74773787	T	C	Intron 1
105,285,891	rs143438338	A	G	Intron 1
105,286,849	rs138013654	AACT	A	Intron 1
105,289,223	rs10991608	C	T	Intron 1
105,289,232	rs115049292	T	C	Intron 1
105,290,750	rs327982	C	T	Intron 1
105,291,402	rs59370172	C	T	Intron 1
105,291,831	rs75690841	C	T	Intron 1

105,292,685	rs327984	G	A	Intron 1
105,294,585	rs114676441	C	T	Intron 1
105,295,803	rs114897292	C	T	Intron 1
105,296,546	rs327991	C	T	Intron 1
105,297,413	rs2089807	C	T	Intron 1
105,297,774	rs12347364	T	A	Intron 1
105,298,026	rs115441783	A	G	Intron 1
105,298,131	rs75575818	G	T	Intron 1
105,298,426	rs149357410	C	T	Intron 1
105,299,219-307	**Exon 2**			Exon 2
105,299,786	rs116754351	C	T	Intron 2
105,299,805	rs10991609	C	G	Intron 2
105,302,321	rs114365914	G	A	Intron 2
105,305,895	rs9644967	G	A	Intron 2
105,306,276	rs76025089	G	A	Intron 2
105,306,727	rs35455036	A	AC	Intron 2
105,309,586	rs113544435	G	A	Intron 2
105,309,724-865	**Exon 3**			Exon 3
105,310,529	rs12376222	A	G	Intron 3
105,313,914	rs112461270	C	T	Intron 3
105,315,716	rs10991619	C	T	Intron 3
105,317,659	rs76402817	G	T	Intron 3
105,320,075	rs79600643	T	C	Intron 3
105,320,589	rs142832243	T	A	Intron 3
105,321,204	rs12349416	G	A	Intron 3
105,321,443	rs77139204	A	G	Intron 3
105,321,461	rs112445608	CAAATT	C	Intron 3
105,324,016	rs138909285	G	T	Intron 3
105,326,849	rs75252146	A	G	Intron 3
105,327,673	rs12379026	T	A	Intron 3
105,328,053	rs10991625	C	A	Intron 3
105,328,967	rs73512035	G	A	Intron 3
105,331,975	rs10991629	C	T	Intron 3
105,332,351	rs13298325	C	A	Intron 3
105,334,467	rs113909391	A	G	Intron 3
105,335,000	rs10991630	G	A	Intron 3
105,335,391	rs73512040	A	G	Intron 3
105,335,562-696	**Exon 4**			Exon 4
105,338,369	rs77491366	G	A	Intron 4
105,338,437	rs10991632	C	A	Intron 4
105,338,890	rs75638188	T	C	Intron 4
105,339,231	rs111867599	AAAAAAAA	N=11-13 As	Intron 4
105,342,624	rs116567682	C	T	Intron 4

105,343,266	rs116835488	G	C	Intron 4
105,343,460	rs10991633	A	G	Intron 4
105,343,745	rs112970775	TCAGATA	T	Intron 4
105,344,449	rs12377753	A	G	Intron 4
105,344,478	rs12377755	A	G	Intron 4
105,345,430	rs10991634	C	T	Intron 4
105,346,125	rs77157131	C	T	Intron 4
105,346,555	rs112651128	G	A	Intron 4
105,346,871	rs12380580	C	T	Intron 4
105,348,355-449	**Exon 5**			Exon 5
105,349,196	rs78132990	C	A	Intron 5
105,350,842	rs10991636	C	G	Intron 5
105,351,432	rs111442823	G	C	Intron 5
105,353,266	rs71494522	T	A	Intron 5
105,354,844	rs10991638	G	A	Intron 5
105,356,208-379	**Exon 6**			Exon 6
105,356,772	rs10991640	C	T	Intron 6
105,357,909	rs35533765	A	G	Intron 6
105,358,344-434	**Exon 7**			Exon 7
105,360,791	rs10991641	C	T	Intron 7
105,361,038	rs150451876	C	A	Intron 7
105,361,187-328	**Exon 8**			Exon 8
105,361,698	rs76795767	T	C	Intron 8
105,362,820-363,007	**Exon 9**			Exon 9
105,363,170	rs35157236	C	CTT	Intron 9
105,363,248	rs34253461	C	T	Intron 9
105,363,621	rs113670898	T	A	Intron 9
105,363,656	rs80234385	G	A	Intron 9
105,363,789	rs79144680	C	G	Intron 9
105,364,554-720	**Exon 10**			Exon 10
105,365,193	rs4538947	G	A	Intron 10
105,365,483-639	**Exon 11**			Exon 11
105,365,767	rs7860321	C	T	Intron 11
105,366,343-429	**Exon 12**			Exon 12
105,367,199	rs11506819	C	A	Intron 12
105,368,865	rs12379533	C	T	Intron 12
105,369,272	rs71494524	A	G	Intron 12
105,371,588	rs10991643	G	A	Intron 12
105,372,094	rs71364089	TGACAA	T	Intron 12
105,373,018	rs74344739	T	C	Intron 12
105,373,364	rs117555942	T	C	Intron 12
105,373,837	rs111516986	T	G	Intron 12

105,374,595-736	**Exon 13**			Exon 13
105,379,395	rs10991644	A	G	Intron 13
105,380,358	rs549490262	CT	C	Intron 13
105,380,558	rs13297419	T	C	Intron 13
105,381,737	rs71494526	C	G	Intron 13
105,381,743	rs10991645	C	A	Intron 13
105,383,122-359	**Exon 14**			Exon 14
105,384,640	rs71494527	A	G	Intron 14
105,385,418-502	**Exon 15 (short)**			
105,385,418-713	**Exon 15 (long)**			
105,385,482	rs3199966	T	G	Exon 15
105,389,033-397,056	**Exon 16 (short)**			
105,389,918	rs2771040	G	A	Exon 16 (3'UTR)
105,399,766	rs74472811	G	C	
105,400,131	rs35322610	G	A	
105,424,107	rs117432520	T	A	

Short variant of Exon 15 + Exon 16 encodes the plasma membrane variant and its 5'UTR.
Long variant of Exon 15 encodes the mitochondrial variant and its 3'UTR.

Supplementary Table 2

Pairwise Linkage Disequilibrium Values for SLC44A1 Variants Associated with Any Neurobehavioral Outcome in Participants with Heavy PAE. *

Reference SNP	Location	SNP #2	D'
rs75106836	Intron 1	rs143438338	**0.9968**
		rs59370172	**0.9968**
		rs12347364	0.1298
		rs10991629	**0.8826**
		rs3199966	**0.9962**
		rs2771040	**0.9961**
rs143438338	Intron 1	rs59370172	**0.9968**
		rs12347364	0.2766
		rs10991629	**0.8911**
		rs3199966	**0.9962**
		rs2771040	**0.9961**
rs59370172	Intron 1	rs12347364	0.0029
		rs10991629	**0.9962**
		rs3199966	**0.9962**
		rs2771040	**0.9961**
rs12347364	Intron 1	rs10991629	**0.9947**
		rs3199966	**0.9948**
		rs2771040	**0.9952**
rs10991629	Intron 3	rs3199966	0.7769
		rs2771040	**0.9030**
rs3199966	Exon 15	rs2771040	**0.9997**

* D', disequilibrium value for the comparison between the reference SNP and SNP #2. **Boldface** indicates pairwise associations ≥0.80.

Supplementary Table 3

Additional SNPs in *SLC44A1* that are Associated with Neurobehavioral Measures in the Genotype × Exposure Association[1]

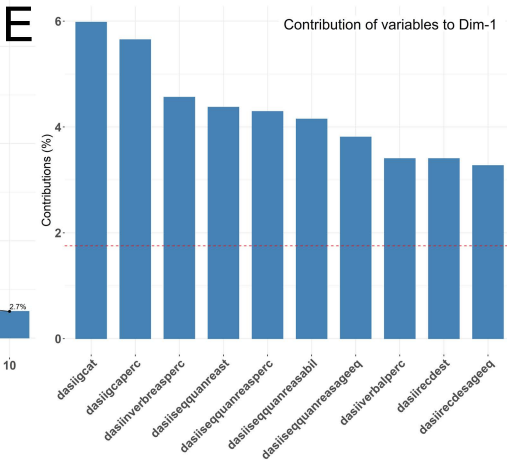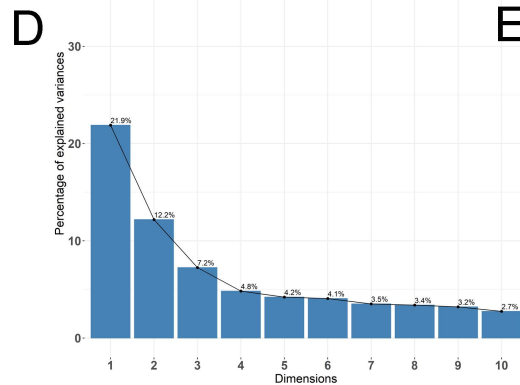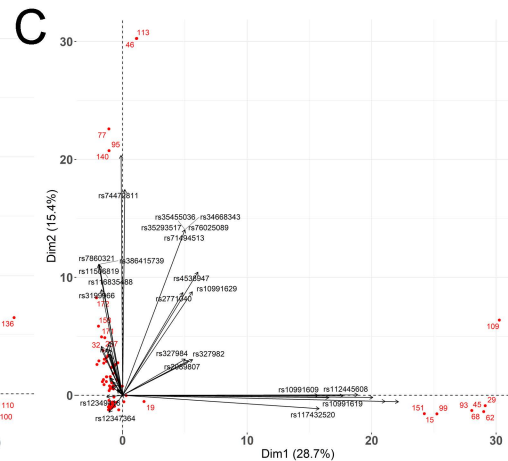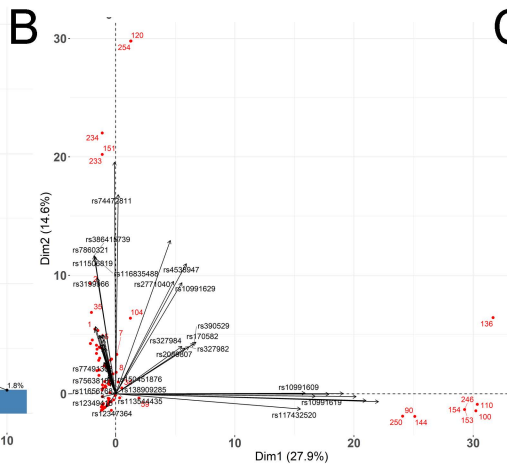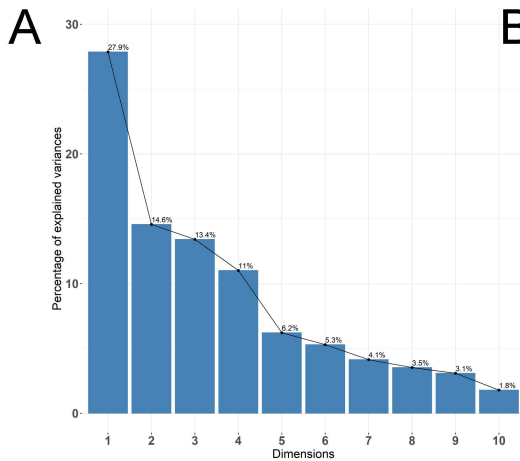| SNP	Position	LD with rs3199966 (D′)	Neurobehavioral Measurement	Padj	B ± SE	Effect Size	Effect Allele
rs75106836	Intron 1; T>C	0.9962	DAS-II, General cognitive abilities, T-score	0.033	1.33±0.62	2.15	C (T>C)
			DAS-II, Sequential quantitative reasoning, T-score	0.035	1.58±0.63	2.51	
rs143438338	Intron 1; C>T	0.9962	DAS-II, Sequential quantitative reasoning, T-score	0.035	1.59±0.63	2.52	T (C>T)
rs59370172	Intron 1; C>T	0.9962	DAS-II, Sequential quantitative reasoning, T-score	0.035	1.59±0.64	2.48	T (C>T)
rs12347364	Intron 1; T>A	0.9948	CVLT, Serial cluster ratio, Z-score	0.0031	-1.26±0.47	-2.68	A (T>A)
rs10991629	Intron 3; C>T	0.7769	DAS-II, Recall of designs, T-score	0.035	0.62±0.24	2.58	T (C>T)
			DAS-II, Sequential quantitative reasoning, T-score	0.010	0.71±0.23	3.09	

[1] SNPs in addition to rs3199966 and rs2771040 as presented in Table 4; D′=1.00 for rs3199966 and rs2771040. CVLD, California Verbal Learning Test; DAS-II, Differential Abilities Scale, 2nd Ed.; SE, Standard Error. T-scores and Z-scores reflect age-normalized values that are internal to the testing instrument. P-adjusted by Bonferroni correction.

Supplementary Table 4
ANOVA Reports for the Exposure × Genotype Regression Analysis of rs3199966 and rs2771040 Presented in Figure 4.*

Measure	Exposure	Genotype	Genotype × Exposure	Control TT vs. Con TG	Control TT vs. Con GG	Control TT vs. PAE TT	Control TT vs. PAE TG	Control TT vs. PAE GG
General Cognitive Abilities	$F_{(2,162)}=67.41$ $p=6.57 \times 10^{-14}$	$F_{(2,162)}=8.965$ $p=2.03 \times 10^{-4}$	$F_{(2,162)}=6.424$ $p=0.0021$	$p=1.33 \times 10^{-4}$	$p=0.041$	$p=1.0 \times 10^{-8}$	$p=1.0 \times 10^{-8}$	$p=0.0033$
Recall of Designs	$F_{(2,157)}=40.38$ $p=2.16 \times 10^{-9}$	$F_{(2,157)}=5.999$ $p=0.0031$	$F_{(2,157)}=4.146$ $p=0.0176$	$p=6.26 \times 10^{-4}$	$p=0.278$	$p=1.0 \times 10^{-8}$	$p=4.3 \times 10^{-6}$	$p=0.0346$
Nonverbal Reasoning	$F_{(2,162)}=35.99$ $p=1.25 \times 10^{-8}$	$F_{(2,162)}=3.931$ $p=0.0215$	$F_{(2,162)}=4.261$ $p=0.00157$	$p=0.0153$	$p=0.230$	$p=1.0 \times 10^{-8}$	$p=7.06 \times 10^{-5}$	$p=0.186$
Sequential Quantitative Reasoning	$F_{(2,158)}=50.56$ $p=3.77 \times 10^{-11}$	$F_{(2,158)}=8.965$ $p=2.03 \times 10^{-4}$	$F_{(2,158)}=6.675$ $p=0.00165$	$p=0.00156$	$p=0.090$	$p=1.0 \times 10^{-8}$	$p=5.0 \times 10^{-7}$	$p=0.276$
Executive Function Difficulties	$F_{(2,169)}=256.8$ $p=2.00 \times 10^{-16}$	$F_{(2,169)}=0.085$ $p=0.918$	$F_{(2,169)}=3.639$ $p=0.0283$	$p=0.758$	$p=1.00$	$p=1.0 \times 10^{-8}$	$p=1.0 \times 10^{-8}$	$p=3.48 \times 10^{-5}$
				Control AA vs. Con AG	Control AA vs. Con GG	Control AA vs. PAE AA	Control AA vs. PAE AG	Control AA vs. PAE GG
CVLT Serial Cluster Ratio	$F_{(2,169)}=0.005$ $p=0.944$	$F_{(2,169)}=0.001$ $p=0.999$	$F_{(2,169)}=1.334$ $p=0.266$	$p=0.997$	$p=0.913$	$p=0.970$	$p=1.00$	$p=0.950$

* All analyses were performed for rs3199966(T>G) except for CVLT Serial Cluster Ratio, which tested rs2771040(A>G). Analysis used the generalized linear regression package 'glm' in R and an Additive gene model, followed by two-way analysis of variance and Tukey multiple comparison of means.

Supplementary Figure 1

Legend for Supplementary Figure 1. Exploration of Principal Component Analysis (PCA) Findings of Participant Characteristics. The PCAs are presented in Figure 1. (A-C) PCA exploration for the comparison of genotype against participant characteristics (ancestry, sex, age, and PAE status) from Figures 1A-1D. (A) Loadings plot of PC dimensions 1 through 10 shows that only PC1 makes a substantial contribution (27.9%) to the variance between participants. (B) Loadings plot identifies individuals and SNPs within *SLC44A1* that contribute to the variance in both PC1 and PC2. (C) Removal of those outliers from PC1 does not appreciably change the loadings of PC1 and PC2, showing those individuals did not drive the variance. (D-F) PCA exploration of neurobehaviors against participant characteristics (ancestry, sex, age, and PAE status) from Figures 1E-1H. (D) Loadings plot of PCA dimensions 1 through 10 shows that PC1 and PC2 are the primary contributions to the variance in neurobehavioral outcomes for the participants. (E) Scree plot shows the top 10 drivers of PC1 are cognitive measures within the DAS-II instrument including general cognitive abilities T-score (dasiigcat) and %ile (dasiigcaperc); nonverbal reasoning %ile (dasiinverbreasperc); sequential quantitative reasoning T-score (dasiiseqquanreast), %ile (dasiiseqquanreastperc), ability score (dasiiseqquanreastabil), and age-equivalent (dasiiseqquanreastageeq); verbal ability cluster %ile (dasiiverbalperc); and recall of designs T-score (dasiirecdest) and age-equivalent (dasiirecdesageeq). (F) Scree plot shows the top 10 drivers of PC2 are mostly age-related measures including subject age at testing (briefpage, discage, ysrage); the item set administered in the word definition and verbal similarities tests (dasiiworddefset; dasiiverbsimset), standard error on the Connors hyperactivity measurement (conners_hyim_se), sequential quantitative reasoning ability score (dasiiseqquanreasabil), and the age-equivalent measures of sequential quantitative reasoning, recall of designs, and verbal similarities (dasiiseqeuanreasageeq, dasiirecdesageeq, dasiiverbsimageeq), all in DAS-II.