Supplementary Information

Synaptic scaffold evolution generated components of vertebrate cognitive complexity

J. Nithianantharajah[1,2], N.H. Komiyama[1,2], A. McKechanie[3], M. Johnstone[3], D. H. Blackwood[3], D. St Clair[4], R.D. Emes[5], L. N. van de Lagemaat[1,2], L.M. Saksida[6], T.J. Bussey[6*] & S.G.N. Grant[1,2,6*]

[1]Genes to Cognition Programme, Centre for Clinical Brain Sciences and Centre for Neuroregeneration, The University of Edinburgh, Chancellors Building, 47 Little France Crescent, Edinburgh UK EH16 4SB

[2]Genes to Cognition Programme, The Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire UK CB10 1SA

[3]Division of Psychiatry, The University of Edinburgh, Royal Edinburgh Hospital, Edinburgh UK EH105HF

[4]Institute of Medical Sciences, University of Aberdeen, Aberdeen UK AB252ZD

[5]School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus UK LE12 5RD

[6]Department of Experimental Psychology, University of Cambridge & The MRC and The Wellcome Trust Behavioral and Clinical Neuroscience Institute, University of Cambridge, Cambridge UK CB2 3EB

*These authors co-supervised the project

a

	DLG1 Hs	*DLG2* Hs	*DLG3* Hs	*DLG4* Hs
DLG1 Hs	100.0			
DLG2 Hs	80.3	100.0		
DLG3 Hs	76.0	76.9	100.0	
DLG4 Hs	74.1	72.0	69.9	100.0

b

	Dlg1 Mm	*Dlg2* Mm	*Dlg3* Mm	*Dlg4* Mm
Dlg1 Mm	100.0			
Dlg2 Mm	80.2	100.0		
Dlg3 Mm	76.3	77.4	100.0	
Dlg4 Mm	73.5	72.1	70.1	100.0

c

	DLG1 Hs	*DLG2* Hs	*DLG3* Hs	*DLG4* Hs
Dlg1 Mm	95.80			
Dlg2 Mm	80.80	96.60		
Dlg3 Mm	76.10	76.80	99.30	
Dlg4 Mm	73.90	71.50	69.90	99.20

d

		Dlg1		***Dlg2***		***Dlg3***		***Dlg4***	
Region	**Abbrev**	**Mouse**	**Human**	**Mouse**	**Human**	**Mouse**	**Human**	**Mouse**	**Human**
Dentate Gyrus	DG	11.0	7.9	10.0	8.1	3.8	9.7	27.1	7.5
Cornu Ammonis	CA	3.4	7.7	8.6	8.5	2.7	9.3	23.5	7.1
Subiculum	SUB	9.3	8.6	8.3	8.0	2.2	9.2	24.3	6.9
Cortex	CTX	13.0	8.2	9.2	7.9	2.0	9.0	26.8	6.7
Cortical Plate Amygdala	COA	17.7	8.3	7.6	7.8	3.5	8.9	28.7	6.5
Caudo-putamen	CP	10.8	8.5	7.9	7.7	1.8	9.0	29.5	7.2
Nucleus Accumbens	ACB	11.2	8.1	9.9	7.3	1.1	8.4	26.4	6.5
Substantia Nigra Pars Compacta	SNc	4.1	8.0	1.0	7.8	1.1	8.8	15.9	6.5
Substantia Nigra Pars Reticulata	SNr	3.4	8.5	1.1	7.4	0.8	7.7	10.1	6.2
Thalamus	TH	3.3	8.2	0.3	7.4	1.6	8.7	19.4	6.4
Zona Incerta	ZI	5.1	8.5	1.3	7.4	0.9	8.1	19.8	5.9
Hypothalamus	HY	6.0	7.9	1.3	7.3	0.3	8.3	14.9	6.2
Substantia Innominata	SI	5.2	8.4	2.9	7.8	1.0	8.4	17.6	6.6
Midbrain	MB	5.0	8.2	1.2	7.6	0.4	8.7	13.8	6.4
Myencephalon	MY	4.5	8.1	1.5	7.8	0.7	8.7	7.3	6.5
Pons	P	4.2	8.2	0.7	7.7	0.4	8.9	7.4	6.7
Cerebellar Cortex	CBX	5.6	8.2	3.3	7.6	0.2	8.7	13.0	6.7
Pearson R (Mouse vs Human)		**0.16**		**0.71**		**0.68**		**0.53**	
P-value		**0.557**		**0.002**		**0.004**		**0.035**	

Supplementary Figure 1. Conservation of Dlg paralogs in mouse and human

a-c. Sequence comparison of Dlgs in mouse and human. Percent identity of human (Hs) and Mouse (Mm) aligned Dlg genes.

a. Comparison of human paralogs
b. Comparison of mouse paralogs
c. Comparison of human-mouse paralogs.

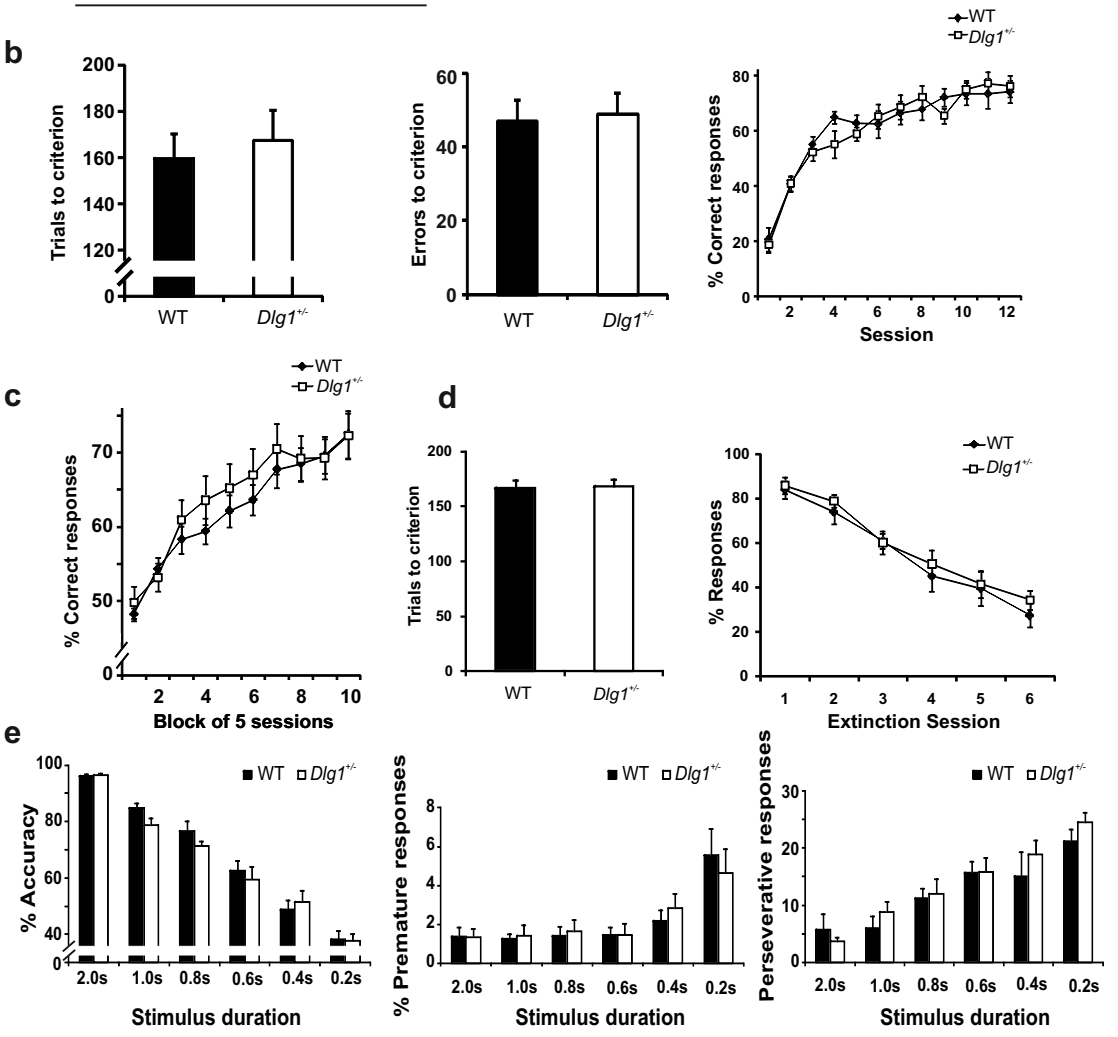
Peptide sequences were aligned using Prank (Löytynoja & Goldman, BMC Bioinformatics 11, 579; 2010). Percent identities between pairs of peptide sequences were calculated using Bio3D (Grant et al., Bioinformatics 22, 2695-2696; 2006).

d. Gene expression correlation of Dlg1-4 in the brain in mouse and human.
Mouse values are median voxel expression intensity values based on in situ hybridisation data from the Allen Institute mouse brain atlas (Lein et al, Nature 445, 168-76; 2007). Voxel expression data were matched to regions using the provided 200µm resolution atlas and the relevant brain region ontology. Median expression values in mouse brain regions were computed from the voxel expression distribution of each gene in each brain region.
Human values are median log2 transformed expression values from microarray analysis of two brains from the Allen Institute human brain atlas (Hawrylycz et al, Nature 489, 391–399; 2012). After quantile normalization of each sample, samples were selected from each region using the provided brain region ontology. Median expression values in human brain regions were computed as the median of relevant samples in the brain region.

a

Operant conditioning:
Pre-training phases in the touchscreen

Sessions to criterion	WT	*Dlg1*$^{+/-}$
Phase 1	2.0 ± 0.0	2.0 ± 0.0
Phase 2	1.0 ± 0.0	1.0 ± 0.0
Phase 3	2.5 ± 0.2	2.4 ± 0.3
Phase 4	1.0 ± 0.0	1.0 ± 0.0
Phase 5	2.0 ± 0.0	2.1 ± 0.1



Supplementary Figure 2. Dlg1$^{+/-}$ mice show normal performance on all tasks within the touchscreen cognitive battery

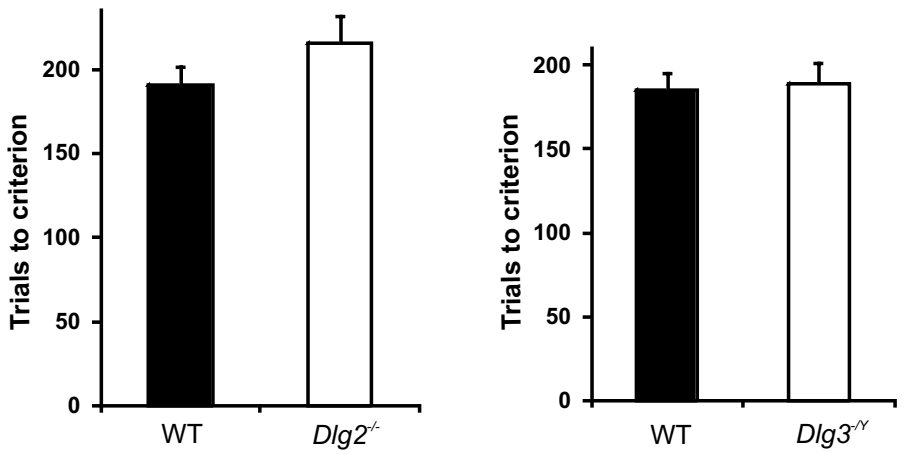
a. Mice were trained through several phases to nose-poke a stimulus displayed on the touchscreen to attain a reward (operant conditioning). Animals were required to successfully complete and reach the set criterion at each phase before advancing to the next phase. Phase 1: animals were acclimated for 20min on 2 days to the operant chamber and required to consume reward pellets freely available from the magazine. Phase 2: a single visual stimulus was displayed on the touchscreen after which, the disappearance of the stimuli coincided with delivery of a food reward, presentation of a tone and illumination of the pellet magazine. Phase 3: animals were required to nose poke a visual stimulus that appeared on the touchscreen in order to obtain a reward. Phase 4: animals were additionally required to initiate the commencement of a new trial with a head entry into the pellet magazine. Phase 5: In addition to that described for previous phases, responses at a blank part of the screen during stimulus presentation now produced a 5s time-out and were not rewarded. See methods for further details.

b. Visual discrimination and reversal learning. Total number of trials (left graph) and errors (middle graph) to reach learning criterion on visual discrimination. Percentage of correct responses on reversal learning across sessions (right graph).

c. Object-location paired-associates learning. Percentage of correct responses across training sessions.

d. Extinction learning. Number of trials to reach acquisition criterion (left graph) and percentage of responses made during extinction learning across sessions (right graph).

e. Performance on the 5-CSRTT. Percentage accuracy (% correct responses)(left graph), percentage of premature responses (middle graph) and number of perseverative responses (right graph).

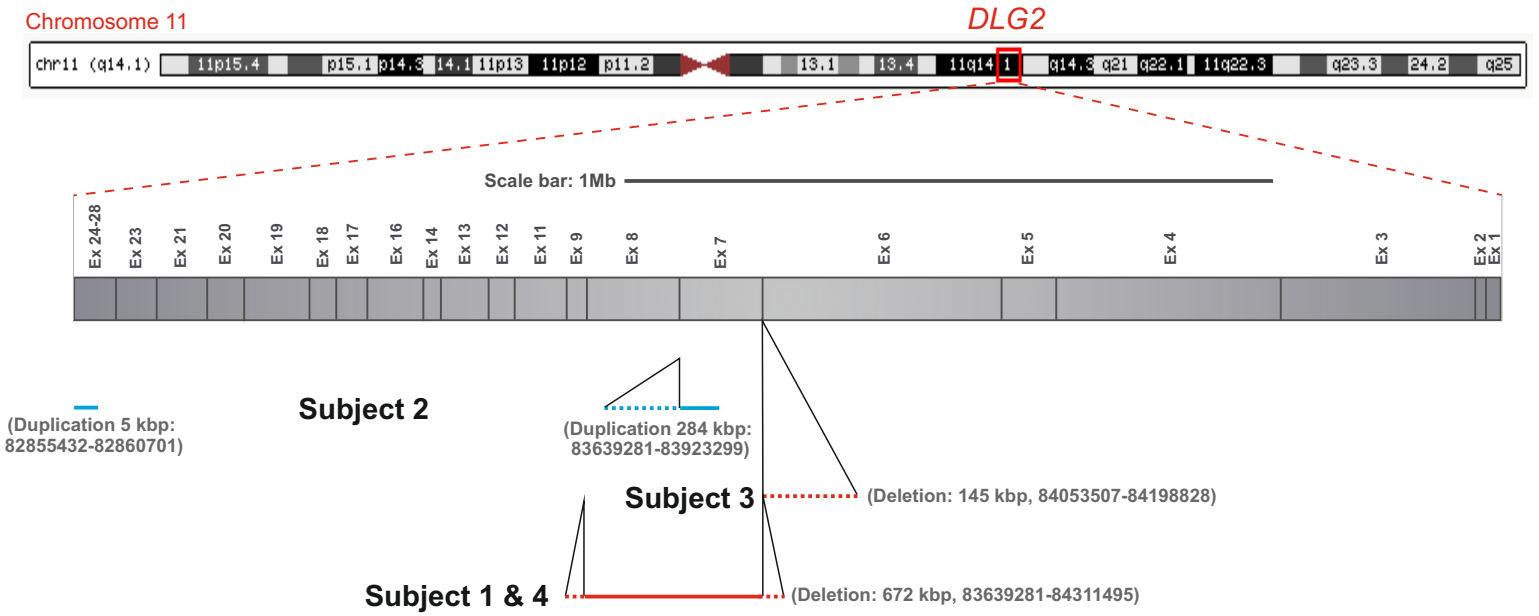


Supplementary Figure 3. Acquisition of instrumental response

Number of trials to reach acquisition criterion prior to testing extinction.

Dlg2$^{-/-}$mice (left graph) and *Dlg3*$^{-/Y}$ mice (right graph).

a



b

					CANTAB TESTS Standard Z-score		
	Age	**Gender**	***DLG2* CNV**	**Diagnosis**	**IED Total errors (adjusted)**	**PAL Total errors (adjusted)**	**RVP A'**
Subject 1	**67**	**F**	**Deletion**	**Schizophrenia**	**-4.99**	**-1.41**	**-3.82**
Subject 2	**33**	**M**	**Duplication**	**Schizophrenia**	**0.22**	**-1.34**	**-2.09**
Subject 3	**61**	**F**	**Deletion**	**Schizophrenia**	**-2.72**	**-2.29**	**-1.61**
Subject 4	**24**	**F**	**Deletion**	**Nil**	**0.57**	**0.19**	**-1.03**

Supplementary Figure 4. Human mutations in *DLG2*.

a. Schematic diagram showing location of CNVs (illustrated by red (deletion) or blue (duplication) horizontal bars; exonic sequences illustrated by solid bars and intronic sequences by broken bars) within relative positions of exons (Ex) within *DLG2* for subjects 1-4. For each subject, size of CNV (kilo base pairs, kbp) and genome positions indicated in parenthesis. Subjects 1 and 4 CNV deletions originate within an intronic sequence between exons 8-9 and exons 5-6 resulting in the complete deletion of exons 7 & 8. Subject 2 has two CNV duplication sites: 1) an intronic sequence between exons 7 & 8 and the first 13 amino acids of exon 7 and 2) within exons 26 and 27. Subject 3 CNV deletion spans an intronic sequence between exons 6 & 7. Co-ordinates for Subjects 1, 2 & 4 were defined by MAQ assays; co-ordinates for Subject 3 from GWAS (this region als includes intronic sequences between exons 6 & 7). All numbering is based on UCSC Genome Browser Mar. 2006; hg18.

b. Summary CANTAB test results for subjects 1-4. Age, Gender, *DLG2* CNV type and Diagnosis is indicated for each subject. Individual standard scores (z-scores) for the Intra/Extradimensional set-shift (IED) task, Paired Associates Learning (PAL) task and Rapid Visual Information Processing (RVP) task are shown (corresponding to group histograms in Figure 6).

a SESSIONS TO CRITERION (BASELINE PERFORMANCE)

	4s	2s
WT	12.00±1.38	11.55±1.27
Dlg1$^{+/-}$	13.00±1.24	11.45±1.30
WT	14.45±1.40	12.45±1.97
Dlg2$^{-/-}$	21.78±1.39*	32.33±4.48**
WT	12.86±1.20	11.79±1.62
Dlg3$^{-/Y}$	11.80±1.09	11.70±2.30

b % OMISSIONS

	2.0s	1.0s	0.8s	0.6s	0.4s	0.2s
WT	12.00±1.52	25.94±0.95	29.60±2.53	39.55±3.13	41.37±3.97	41.82±3.61
Dlg1$^{+/-}$	13.79±1.20	23.91±1.30	32.27±2.98	33.93±4.73	36.77±4.16	39.48±4.08
WT	15.39±1.55	25.70±1.59	33.09±2.76	32.91±3.37	36.33±3.69	40.61±3.01
Dlg2$^{-/-}$	14.81±1.39	29.56±1.65	31.44±1.37	37.22±1.75	44.44±3.22	44.07±3.78
WT	16.25±0.74	32.56±2.44	37.39±2.19	36.92±2.33	39.44±2.65	42.00±3.55
Dlg3$^{-/Y}$	16.00±1.03	32.19±2.48	34.37±2.51	39.11±2.27	44.44±3.70	44.96±2.18

c CORRECT RESPONSE LATENCY

	2.0s	1.0s	0.8s	0.6s	0.4s	0.2s
WT	1.41±0.03	1.33±0.04	1.43±0.07	1.37±0.07	1.39±0.08	1.57±0.07
Dlg1$^{+/-}$	1.45±0.08	1.43±0.09	1.32±0.10	1.43±0.11	1.40±0.13	1.47±0.09
WT	1.55±0.05	1.33±0.06	1.30±0.08	1.21±0.09	1.16±0.10	1.18±0.04
Dlg2$^{-/-}$	1.53±0.03	1.26±0.05	1.29±0.04	1.20±0.05	1.17±0.05	1.25±0.06
WT	1.50±0.09	1.39±0.10	1.28±0.09	1.25±0.08	1.31±0.10	1.24±0.08
Dlg3$^{-/Y}$	1.44±0.06	1.29±0.05	1.25±0.07	1.23±0.05	1.25±0.06	1.07±0.04

d REINFORCER LATENCY

	2.0s	1.0s	0.8s	0.6s	0.4s	0.2s
WT	1.77±0.13	1.70±0.09	1.89±0.05	1.96±0.16	1.79±0.10	1.85±0.15
Dlg1$^{+/-}$	1.89±0.11	1.94±0.07	2.00±0.10	1.90±0.08	1.95±0.15	1.78±0.07
WT	1.88±0.16	1.74±0.07	1.91±0.18	1.71±0.11	1.68±0.07	1.62±0.10
Dlg2$^{-/-}$	2.02±0.22	1.92±0.23	1.83±0.14	1.81±0.10	1.83±0.20	1.76±0.15
WT	1.80±0.07	1.76±0.08	1.79±0.08	1.74±0.06	1.68±0.04	1.73±0.11
Dlg3$^{-/Y}$	1.70±0.08	1.94±0.12	1.91±0.06	1.87±0.11	1.69±0.11	1.83±0.13

Supplementary Table 1. Measures on the 5-CSRTT

a. Number of sessions to reach baseline performance criterion (see methods) at stimulus durations of 4s and 2s. Independent samples t-tests, * $p<0.005$, ** $p<0.001$.

b. Percentage of omissions (failure to respond to any window during stimulus display or the limited holding period).

c. Average correct response latency (time from when the stimulus is displayed and a nose-poke response made).

d. Reinforcer latency (time taken to collect reward) were analysed for correct responses only.

UCSC Browser position chr11:82846116-85020429

Amplicon	Forward	Reverse
DLG2_ex28	82843239	82843606
DLG2_ex26	82855432	82855521
DLG2_ex25	82857899	82858017
DLG2_ex24	82860619	82860701
DLG2_ex23	82872277	82872594
DLG2_ex21	82930545	82930648
DLG2_ex20	83021515	83021813
DLG2_ex19	83070391	83070618
DLG2_ex18	83175563	83175696
DLG2_ex17	83222813	83222887
DLG2_ex16	83262414	83262826
DLG2_ex14	83353917	83354130
DLG2_ex13	83369324	83369616
DLG2_ex12	83448002	83448265
DLG2_ex11	83487418	83487622
DLG2_ex09	83639281	83639432
DLG2_ex08	83662311	83662509
DLG2_ex07	83923205	83923299
DLG2_ex06	84311311	84311495
DLG2_ex05	84500109	84500266
DLG2_ex04	84543536	84543664
DLG2_ex03	84673572	84673829
DLG2_ex02	84987298	84987438
DLG2_ex01	85015427	85015739
cnv1_01	84242534	84242681
cnv1_02	84249728	84249852
cnv2_01	84388144	84388322
cnv2_02	84399144	84399309
cnv2_03	84440577	84440935
cnv2_04	84463125	84463398
cnv2_05	84493892	84494159
cnv3_01	84902373	84902765

Supplementary Table 2. Positions selected in UCSC browser for PCR primers encompassing the DLG2 gene