Supplementary material

Cognitive tests in kindred

Intellectual level

The Vocabulary, Digit Span, and Similarities subtests from the WAIS-R were administered and the scores were pro-rated to obtain an estimate of VIQ (ref 19, main text).

Executive Functions

1).Verbal fluency: The subject had to produce as many words beginning with the letter S (phonemic fluency) and as many animal names (semantic fluency) both in a minute. Performance on fluency tests, particularly for the phonemic category (letter "s"), is a sensitive indicator of frontal lobe functioning.[31]

2). The Hayling Test is a response suppression task. The subject has to complete two series of 15 sentences each missing the last word. In the first section a sensible completion is required and in the second a nonsensical completion. The test yields two measures of mental processing speed and an error score for the second series. Performance on this measure has been shown to involve frontal brain regions in healthy individuals[32] and to be adversely affected by frontal lobe pathology.[33] 3) Cognitive estimates. This is a semantic reasoning task that requires the subject to provide a reasonable estimation to ten questions that have no exact answer, based on their available semantic knowledge. The questions are of the format 'How fast do race horses gallop?'. Penalties are awarded for inaccurate responses and the higher the score the poorer the reasoning demonstrated. This test has been shown to be sensitive to frontal lobe pathology.[19, 34]

Memory.

1). Verbal Recall. This was assessed using the Story Recall subtest from the Adult Memory and Information Processing Battery.[35] The subject is read a short story and then has to recall as many details as possible immediately following presentation and again following a delay of 30 minutes. Performance measures used were the immediate recall score and the % retained score (delayed recall/immediate recall x 100). 2). Verbal Learning. The List Learning test from the AMIPB was employed. The subject is presented with a list of 15 words on five occasions and following each presentation has to recall as many of the words as possible. A second list of words is then presented and following one attempt at recall is required to recall as many words from the first list (delayed recall). The total number of words remembered in the learning phase (verbal learning trails) and in the delayed recall condition (verbal learning delay) were recorded.

MRI in kindred

Coronal T1-weighted images were acquired at 1.5T: acquisition parameters were TE= 4.2, TI= 450, TR= 15, NEX= 1, flip angle= 20, acquisition matrix 256 x 128, field of view 24cm, producing 124 contiguous slices, voxel dimension 0.9375mm x 0.9375mm x 1.5mm. Data were reformatted in multiple planes to allow careful examination of regions of interest. Additionally T2 and FastFLAIR sequences were obtained (T2 and PD sequence: TE1= 30, TE2= 120, TR= 2000, NEX= 1, acquisition matrix 256 x 128, field of view 24x18cm, slice thickness 5mm contiguous; FastFLAIR sequence TE1= 152, TE2= 2200, TR= 10002, NEX= 1, acquisition matrix 256 x 128, field of view 24cm, slice thickness 5mm contiguous).

Cognitive testing in the LBC1921 cohort

Mini-Mental State Examination (MMSE): MMSE[36] was as used as a screen for dementia. The maximum score is 30. A score of less than 24 was used as an indicator of possible dementia.

Moray House Test (MHT):This general mental ability test was previously described in detail.[18, 37] It mainly assesses verbal reasoning skills. Subjects took this test at the age of 11 and again at about age 79.

Raven's Standard Progressive Matrices: Non-verbal reasoning was examined at age 79 using Raven's Standard Progressive Matrices.[38]

Verbal fluency: Prefrontal executive function was examined at age 79 using the Verbal Fluency test. Subjects named as many words as possible in one minute for each of the letters C, F, L.[39,40]

Logical Memory: Verbal declarative memory was measured at age 79 using the Logical Memory test, which is a sub-test from Wechsler Memory Scale-Revised.[41]

National Adult Reading Test (NART): NART was used at age 79 to assess prior cognitive ability.[42-44]

g factor: A g factor, as a measure of general intelligence, was created by principal component analysis of the age 79 MHT, Raven's Standard Progressive Matrices, Verbal Fluency and Logical Memory scores. A single component accounted for 53.5% of the total variance.

Genotypic analysis in LBC1921 cohort

Genotyping of the tagging SNPs was carried out by TaqMan fluorescence based allelic discrimination, with primers designed using the Applied Biosystems (ABI) Assay by Design tool (Applied Biosystems, Foster City, CA). Polymerase chain reactions were carried out according to the standard ABI protocol for 5μl reaction volume on 384-well plates. SDS version 2.1 software (Applied Biosystems, Foster City, CA) was used for analysis. Primer sequences are available on request.

Statistical analysis

Multiple regression analyses were implemented in STATISTICA (StatSoft Inc, Tulsa, USA), with the respective cognitive measure as the dependent variable, and genotype scores as the independent predictors. Genotypes were scored using an unrestricted model allowing arbitrary effects of the genotypes at each locus. Regression analyses were performed both with and without including sex as an independent variable. Haplotype analyses were implemented using the haplo.stats software[45], in the R console. Those individuals with data missing from more than two loci or present at less than two loci in the section were excluded from the haplotype analysis.

Association was undertaken in 469 individuals for the available cognitive measures that most clearly paralleled those enhanced in the *RIMS1* affected subjects: these were the National Adult Reading Test (NART), Wechsler Logical Memory, and verbal fluency.

Nineteen tagging SNPs were selected, giving average-locus haplotype r^2 values for the five sections within the gene of 0.81, 0.81, 0.81, 0.81 and 0.86, and an average-locus haplotype r^2 value of 0.82 within the gene as a whole. Average-locus r^2 values for the flanking 100kb sections were 0.89 and 0.67 for the upstream and downstream sections respectively. No significant violations of Hardy-Weinberg equilibrium were observed for any of the tagging SNPs. Each individual SNP was checked for association with the cognitive measures (Supplementary Table 2), including both sex

and genotype as predictors. There were no sex-by-genotype interactions, and none of the individual SNPs was associated with any of the cognitive measures. Sex was a significant predictor of score for Raven's matrices (P=0.0005), and g (P=0.03; P values uncorrected for multiple testing). Additionally, the haplotypes generated by the tagging SNPs from each section showed no association with any of the cognitive measures (Supplementary Table 3).

Generation of *RIMS1* riboprobes and in situ hybridisation.

The cDNA sequence (nucleotides 590-995 numbering from the start codon) was amplified using the polymerase chain reaction from human brain cDNA (Clontech) using KOD DNA polymerase (Novagen). This was A-tailed and then inserted into the vector pGEM-T Easy (Promega). The construct was then sequenced using vector primers to check on the orientation of the insert. Antisense and sense probes were generated by in vitro transcription using T7 and SP6 RNA polymerase under standard procedures. Digoxigenin-dUTP was incorporated into riboprobes during in vitro transcription by using the DIG RNA labelling mix (Roche) according to the manufacturer's instructions. In situ hybridization was performed as reported in Lai et al.[46]

***RIMS1* evolution analysis results**

Preliminary Analysis – Human-Chimp and Rat-Mouse Alignment:

Primate Lineage:

Nonsynonymous Changes = 4
Synonymous Changes = 10

Ka = NS changes/nonsynonymous site
= 4 / 3637.83 = 0.0010

Ks = S changes/synonymous site
= 10 / 1383.17 = 0.0072

Ka/Ks = .0010/.0072 = 0.152

Rodent Lineage:

Nonsynonymous Changes = 18
Synonymous Changes = 138

Ka = nonsynonymous change/nonsynonymous site
= 18 / 3002.67 = 0.0060

Ks = synonymous change/synonymous site
= 138 / 1116.33 = 0.124

Ka/Ks = 0.0060/0.124 = 0.0484

Statistical Significance:

Dorus et al.[5] tested for significance with a one-tailed Fisher's exact test using the numbers of nonsynonymous and synonymous substitutions as the text values. These values are too large, therefore Chi-squared is more appropriate.

	Rodent	Primate
Nonsynonymous	18	4
Synonymous	*138*	*10*

$X^2 = 3.32$

P = 0.065

There is a marginally significant disparity between substitution rates in primates and rodents. However, the Human-Chimp alignment is expected to reduce statistical power in detecting evolutionary signatures due to the sequence similarities.

Human-Macaque and Rat-Mouse Alignment:

<u>Primate Lineage:</u>

Nonsynonymous Changes = 11

Synonymous Changes = 28

Ka = NS changes/nonsynonymous site

= 14 / 3563.3 = 0.0039

Ks = S changes/synonymous site

= 52 / 1353.67 = 0.038

Ka/Ks = .0039/.038 = 0.102

<u>Rodent Lineage:</u>

As Above:

Ka/Ks = 0.0060/0.124 = 0.048

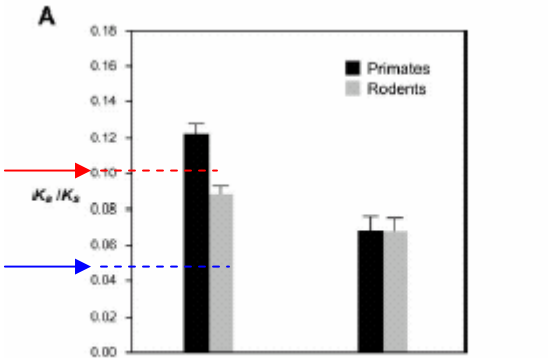

Figure 4: Placement of Primate (red) and Rodent (blue) Ka/Ks values for RIM-1 with respect to the average Ka/Ks obtained for 214 CNS genes.

Statistical Significance:

	Rodent	Primate
Nonsynonymous	18	14
Synonymous	*138*	*52*

$X^2 = 3.54$

$0.5 \leq P \leq 0.1$

Ka/Ks for Human-Macaque is slightly less significant than that of Human-Chimp. The summary of all sequences and Ka/Ks calculations is in the Tables below, A-C.

Table A: Number of Sequence Changes Between Species:

	Non-Synonymous	Synonymous	Non-Synonymous	Synonymous	Non-Synonymous	Synonymous
	Chimpanzee		Macaque		Squirrel Monkey	
Human	4	10	14	52	22	43
Chimp	----		18	57	25	49
Macaque	----		----		19	48

Table B: Chi Square Values of each primate alignment compared to the Rat-Mouse alignment

	Chimpanzee	Macaque	Squirrel Monkey
Human	3.31	3.52	15.40*
Chimpanzee		5.98*	16.34*
Macaque			9.58*

Chi Square tests of all primate comparisons vs. the rodent comparison show significance (*) in human lineage only when compared with the squirrel monkey (significance = P>3.84). This comparison does not represent a similar divergence time (as the Human-Macaque relationship does) and therefore is probably not as significant as it appears. The Chimp-Macaque Ka/Ks is significantly different from the rodent Ka/Ks, however.

Table C: Pairwise Primate Ka/Ks Calculations.

	Chimpanzee	Macaque	Squirrel Monkey
Human	0.151919821	0.10227747	0.19195082
Chimpanzee	--	0.120054536	0.191678276
Macaque	--	--	0.14858768

Ka/Ks values based on the alignments of species gene sequences.

Figure: Identification of Inter-species Sequence Changes in the Phylogenetic Tree

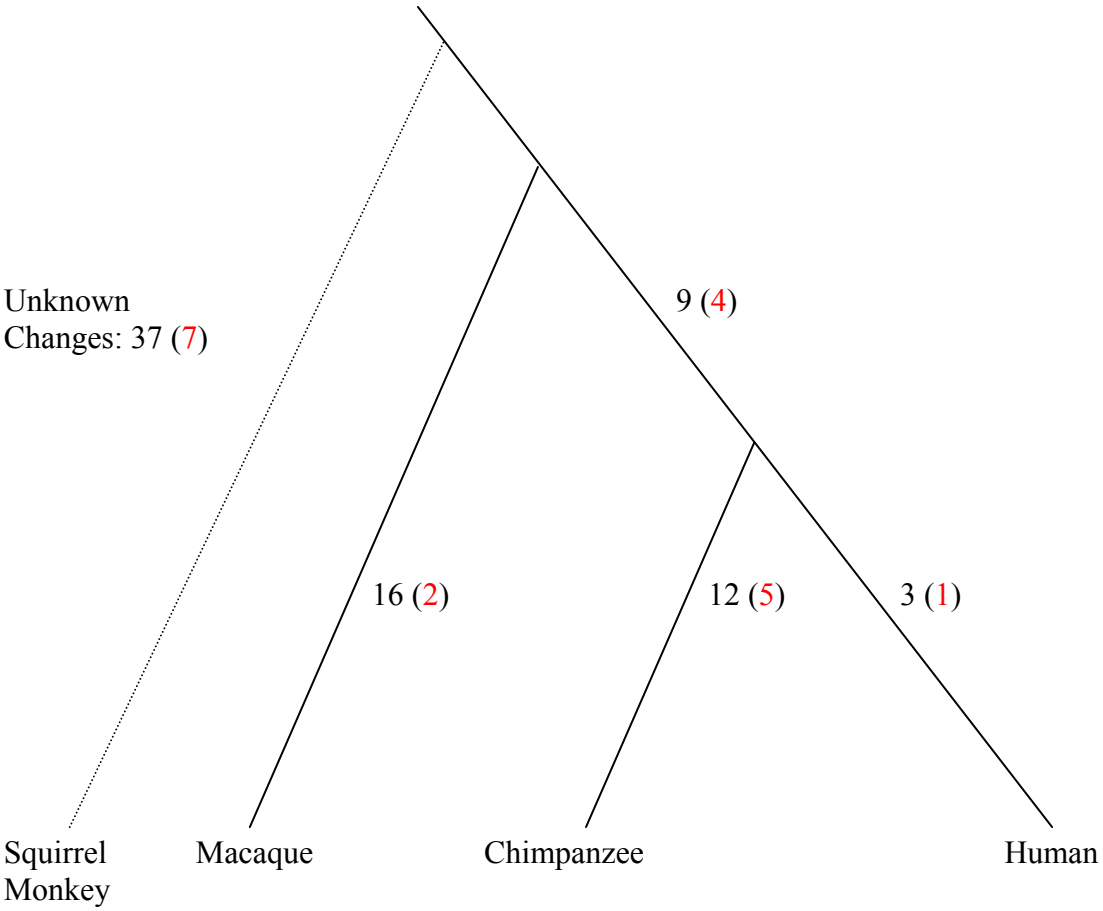

Figure. Sequence analysis using the squirrel monkey as an outlying reference allows the identification of the origin of sequence changes between species (nonsynonymous changes in red). Of a total 77 variants, only three of them are specific to the human lineage (two synonymous changes and one nonsynonymous). The chimpanzee has the most species-specific nonsynonymous changes (5), and the human-chimpanzee common ancestor has four. Due to limitations in the sequencing of squirrel monkey exons, 37 variants were unable to be assigned (seven of which are nonsynonymous changes, the majority of which (22) are located in exon 7, an exon that does not lie in a functional domain).

Supplementary Material References

31. Henry JD, Crawford JR. A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology* 2004;**18**:284-95.

32. Collette F, Van der Linden M, Delfiore G, Degueldre C, Salmon E. The functional anatomy of inhibition processes investigated with the Hayling Test. *Neuroimage* 2001;**14**:258-267.

33. Burgess P, Shallice T. The Hayling Island and Brixton Test Manual. Bury St Edmunds: Thames Valley Test Co; 1997.

34. Della Sala S, MacPherson SE, Phillips LH, Sacco L, Spinnler H. The role of semantic knowledge on the cognitive estimation task -- evidence from Alzheimer's disease and healthy adult aging. *J Neurol*. 2004;**251**:156-164.

35. Coughlan A, Hollows S. The Adult Memory and Information Processing Battery. Leeds: St James Hospital; 1986.

36. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr.Res.* 1975;**12**:189-198.

37. Scottish Council for Research in Education. The intelligence of Scottish children: A national survey of an age-group. London: University of London Press; 1933.

38. Raven JC, Court JH, Raven J. Manual for Raven's Progressive Matrices and Vocabulary Scales. London: H. K. Lewis; 1977.

39. Boone KB, Ponton MO, Gorsuch RL, Gonzalez JJ, Miller BL. Factor analysis of four measures of prefrontal lobe functioning. *Arch. Clin. Neuropsychol.* 1998;**13**:585-595.

40. Lezak M.. Neuropsychological testing. Oxford: Oxford University Press; 1995.

41. Wechsler D, Wechsler R.Memory Scale-Revised. New York: Psychological Corporation; 1987.

42. Crawford JR, Deary IJ, Starr J, Whalley LJ. The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval. *Psychol.Med*. 2001:**31**: 451-458.

43. Nelson HE. National Adult Reading Test (NART) test manual (Part 1). Windsor, England: NFER-Nelson; 1982.

44. Nelson HE, Willison JR. National Adult Reading Reading Test (NART) test manual (Part 2). Windsor, England: NFER-Nelson; 1991.

45. Sinnwell JP, Schaid DJ. Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous. http://mayoresearch.mayo.edu/mayo/research/biostat/schaid.cfm: 2005

46. Lai CS, Gerrelli D, Monaco AP, Fisher SE, Copp AJ. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* 2003;**126**:2455-62.

Supplementary Table 1. Visual acuities in *RIMS1* and *PAX6* cohorts

	Visual acuity, right eye	Visual acuity, left eye
RIMS1 affected individuals		
II:2	6/60	6/60
II:6	3/36	3/60
II:8	6/9	6/9
III:1	6/36	6/36
III:2	6/9	6/9
III:5	3/60	6/60
IV:2	6/6	6/6
IV:3	6/18	3/60
PAX6 affected individuals*		
	6/36	6/24
	no perception of light	6/60
	n/a	n/a
	6/36	6/60
	6/60	6/60
	6/36	6/36
	no perception of light	no perception of light
	hand movements only	hand movements only
	3/60	3/60
	2/60	2/60
	1/60	1/60
	5/60	5/60
	3/60	3/60
	3/60	no perception of light

* Individuals from Thompson et al., 2004 (ref 7).

Supplementary Table 2: Association of individual tagging SNPs with cognitive measures. P values are shown both with and without sex factored into the analysis. Thick black lines depict sections the gene was divided into on the basis of pairwise D'.

***genotypes not significant predictors: there is no interaction (*sex by genotype*)**

	sex	rs2183066	rs4235866	rs4707954	rs1028387	rs10942989	rs1482567	rs1564609
g	0.03	0.23	0.51	0.36	0.11	0.08	0.55	0.27
g with sex	-	0.07	0.11	0.06	0.06	0.02	0.09	0.06
NART	0.55	0.17	0.5	0.75	0.61	0.83	0.02	0.08
NART with sex	-	0.3	0.63	0.89	0.78	0.92	0.06	0.16
Logical memory	0.13	0.96	0.4	0.79	0.87	0.84	0.35	0.47
Logical memory with sex	-	0.6	0.25	0.43	0.47	0.69	0.21	0.23
IQ 11	0.36	0.13	0.1	0.81	0.08	0.8	0.12	0.24
IQ 11 and sex	**-**	0.23	0.14	0.69	0.1	0.81	0.2	0.4
IQ 79	0.05	0.1	0.33	0.62	0.25	0.21	0.3	0.45
IQ 79 and sex	**-**	0.02	0.11	0.04	0.16	0.04	0.03	0.06
Raven's Matrices	0.0005	0.35	0.7	0.65	0.32	0.14	0.1	0.13
Raven's Matrices and sex	-	0.006*	0.005*	0.005*	0.007*	0.002*	0.0008*	0.003*
Verbal fluency	0.54	0.09	0.71	0.88	0.45	0.08	0.24	0.91
Verbal fluency and sex	**-**	0.13	0.75	0.88	0.64	0.14	0.2	0.92

	rs9360524	rs1482574	rs1482574	rs11756248	rs2496517	rs2697433	rs4256398	rs9442769
g	0.93	0.93	0.93	0.26	0.07	0.18	0.95	0.32
g with sex	0.39	0.31	0.31	0.14	0.07	0.04	0.21	0.09
NART	0.99	0.33	0.33	0.08	0.09	0.2	0.12	0.24
NART with sex	0.96	0.46	0.46	0.17	0.17	0.27	0.19	0.39
Logical memory	0.32	0.94	0.94	0.07	0.02	0.15	0.67	0.22
Logical memory with sex	0.44	0.54	0.54	0.13	0.03	0.15	0.63	0.19
IQ 11	0.83	0.63	0.63	0.99	0.96	0.8	0.92	0.81
IQ 11 and sex	0.93	0.64	0.64	0.88	0.97	0.92	0.94	0.9
IQ 79	0.58	0.79	0.79	0.04	0.06	0.27	0.5	0.04
IQ 79 and sex	0.4	0.27	0.27	0.03	0.02	0.02	0.06	0.1
Raven's Matrices	0.61	0.75	0.75	0.17	0.06	0.03	0.61	0.17
Raven's Matrices and sex	0.03*	0.01*	0.01*	0.008*	0.006*	0.0006*	0.008*	0.004*
Verbal fluency	0.54	0.98	0.98	0.21	0.67	0.02	0.57	0.17
Verbal fluency and sex	0.55	0.85	0.85	0.33	0.6	0.05	0.73	0.18

	rs9342944	rs1416546	rs1361311	rs2815715	rs9351921
g	0.63	0.47	0.6	0.5	0.56
g with sex	0.21	0.1	0.19	0.14	0.12
NART	0.98	0.53	0.89	0.28	0.53
NART with sex	0.89	0.66	0.96	0.37	0.63
Logical memory	0.19	0.24	0.52	0.41	0.54
Logical memory with sex	0.17	0.14	0.45	0.25	0.39
IQ 11	0.16	0.33	0.39	0.7	0.43
IQ 11 and sex	0.19	0.37	0.56	0.59	0.6
IQ 79	0.53	0.89	0.61	0.44	0.16
IQ 79 and sex	0.28	0.21	0.05	0.1	0.02
Raven's Matrices	0.94	0.84	0.7	0.91	0.21
Raven's Matrices and sex	0.01*	0.002*	0.007*	0.01*	0.001*
Verbal fluency	0.13	0.56	0.28	0.07	0.88
Verbal fluency and sex	0.4	0.78	0.36	0.09	0.96

Supplementary Table 3. Haplotype association statistics.

Scores shown are the P-values for the global score statistic calculated for the overall haplotype effect for each gene section (labelled by dbSNP rs number) for each cognitive measure. P values are uncorrected for multiple testing.

	rs10498879-rs2040055	rs9342903-rs1564609	rs7743295-rs1482574	rs12213714-rs2496531	rs1015946-rs2807530	rs2746200-rs10943011	rs2815736-rs9446692
g	0.37 (0.38)	0.45 (0.45)	0.29 (0.31)	0.27 (0.27)	0.91 (0.91)	0.30 (0.34)	0.70 (0.74)
NART	0.34 (0.37)	0.45 (0.47)	0.15 (0.12)	0.04 (0.03)	0.70 (0.72)	0.47 (0.51)	0.28 (0.29)
Logical memory	0.46 (0.48)	0.92 (0.89)	0.91 (0.91)	0.31 (0.31)	0.68 (0.67)	0.21 (0.21)	0.30 (0.35)
Verbal Fluency							
Total	0.22 (0.23)	0.20 (0.20)	0.78 (0.77)	0.86 (0.85)	0.86 (0.86)	0.22 (0.19)	0.02 (0.05)
IQ 11	0.98 (0.98)	0.26 (0.28)	0.36 (0.39)	0.21 (0.21)	0.98 (0.97)	0.94 (0.94)	0.17 (0.23)
IQ 79	0.16 (0.16)	0.30 (0.30)	0.16 (0.14)	0.09 (0.09)	0.52 (0.53)	0.14 (0.15)	0.85 (0.82)
Raven's Matrices	0.51 (0.53)	0.38 (0.37)	0.07 (0.06)	0.36 (0.34)	0.77 (0.78)	0.38 (0.36)	0.57 (0.61)

(scores in parenthesis are global simulation P values)