Molecular Psychiatry (2012) 17, 1328–1339

www.nature.com/mp

ORIGINAL ARTICLE

Promoter polymorphisms in two overlapping *6p25* genes implicate mitochondrial proteins in cognitive deficit in schizophrenia

A Jablensky[1,2], D Angelicheva[3], GJ Donohoe[4], M Cruickshank[5,15], DN Azmanov[3], DW Morris[4], A McRae[6], CS Weickert[7], KW Carter[8], D Chandler[3], B Alexandrov[9], A Usheva[10], B Morar[3], PL Verbrugghe[3], A Filipovska[3], O Rackham[3], AR Bishop[9], KØ Rasmussen[9], M Dragovic[1], M Cooper[8], M Phillips[3], J Badcock[1], E Bramon-Bosch[11,12], OP Almeida[2,13], L Flicker[12,14], M Gill[4], A Corvin[4], S MacGregor[6] and L Kalaydjieva[3]

[1]*Centre for Clinical Research in Neuropsychiatry, The University of Western Australia, Perth, WA, Australia;* [2]*School of Psychiatry and Clinical Neurosciences, The University of Western Australia, Perth, WA, Australia;* [3]*Centre for Medical Research/Western Australian Institute for Medical Research, The University of Western Australia, Perth, WA, Australia;* [4]*Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine and Department of Psychiatry, Trinity College Dublin, Trinity College, Dublin, Ireland;* [5]*School of Biomedical, Biomolecular and Chemical Sciences, The University of Western Australia, Perth, WA, Australia;* [6]*Queensland Institute for Medical Research, Brisbane, QLD, Australia;* [7]*Schizophrenia Research Institute, School of Psychiatry, University of New South Wales and Neuroscience Research Australia, Randwick, NSW, Australia;* [8]*Telethon Institute for Child Health Research, Perth, WA, Australia;* [9]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA;* [10]*Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA;* [11]*Institute of Psychiatry, Kings College, London, UK;* [12]*NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust, London, UK;* [13]*Western Australian Centre for Health and Aging, Centre for Health Research, The University of Western Australia, Perth, WA, Australia and* [14]*School of Medicine and Pharmacology, The University of Western Australia, Perth, WA, Australia*

In a previous study, we detected a 6p25–p24 region linked to schizophrenia in families with high composite cognitive deficit (CD) scores, a quantitative trait integrating multiple cognitive measures. Association mapping of a 10 Mb interval identified a 260 kb region with a cluster of single-nucleotide polymorphisms (SNPs) significantly associated with CD scores and memory performance. The region contains two colocalising genes, *LYRM4* and *FARS2*, both encoding mitochondrial proteins. The two tagging SNPs with strongest evidence of association were located around the overlapping putative promoters, with rs2224391 predicted to alter a transcription factor binding site (TFBS). Sequencing the promoter region identified 22 SNPs, many predicted to affect TFBSs, in a tight linkage disequilibrium block. Luciferase reporter assays confirmed promoter activity in the predicted promoter region, and demonstrated marked downregulation of expression in the *LYRM4* direction under the haplotype comprising the minor alleles of promoter SNPs, which however is not driven by rs2224391. Experimental evidence from *LYRM4* expression in lymphoblasts, gel-shift assays and modelling of DNA breathing dynamics pointed to two adjacent promoter SNPs, rs7752203–rs4141761, as the functional variants affecting expression. Their C–G alleles were associated with higher transcriptional activity and preferential binding of nuclear proteins, whereas the G–A combination had opposite effects and was associated with poor memory and high CD scores. *LYRM4* is a eukaryote-specific component of the mitochondrial biogenesis of Fe–S clusters, essential cofactors in multiple processes, including oxidative phosphorylation. *LYRM4* downregulation may be one of the mechanisms involved in inefficient oxidative phosphorylation and oxidative stress, increasingly recognised as contributors to schizophrenia pathogenesis.
Molecular Psychiatry (2012) **17,** 1328–1339; doi:10.1038/mp.2011.129; published online 4 October 2011

Keywords: schizophrenia; cognitive deficit; gene expression; mitochondrial function

Correspondence: Professor L Kalaydjieva, Centre for Medical Research\Western Australian Institute for Medical Research, B-Block QEII Medical Centre, Perth, WA 6009, Australia.
E-mail: luba.kalaydjieva@uwa.edu.au
[15]Current address: 3 Early Life Epigenetics, Murdoch Childrens Research Institute, The Royal Children's Hospital, Melbourne, VIC, Australia.
Received 29 June 2011; revised 10 August 2011; accepted 1 September 2011; published online 4 October 2011

Introduction

Schizophrenia is a complex disorder with variable expression, ill-defined phenotype boundaries and poorly understood multifactorial aetiology, which involves a significant but likely heterogeneous genetic contribution.[1,2] Most diagnostic criteria remain symptom-based, relying largely on the interpretation of the patients' self-reported subjective experiences. As a result, the fundamental problem of 'connecting the phenotype with the genotype'[3] is likely to persist in the new era of powerful genetic technologies.

Hypotheses-guided endophenotype-based studies offer an approach to reducing the heterogeneity of schizophrenia as an alternative or a complement to symptom-based phenotypes. Consistent evidence suggests that measures of neurocognitive dysfunction provide the largest effect sizes among candidate endophenotypes.[4–6] Patterns of memory impairment, providing a strong signal against a background of generalised cognitive deficit (CD), have been replicated across studies and are present in a substantial proportion of schizophrenia patients.[7] As most of the neurocognitive tests tap into several component processes, composite endophenotypes, integrating multiple measures, are more likely to capture variation that is genetically influenced than single-feature endophenotypes.

The Western Australian Family Study of Schizophrenia (WAFSS) applied multi-domain endophenotyping to tease out subtypes based on objective measurement of cognitive dysfunction and to explore their genetic underpinnings. The cognitive measures employed were aggregated into a limited number of quantitative traits, using grade of membership analysis.[8–10] Grade of membership (GoM) is a version of latent structure analysis, which defines latent groups ('pure types') and allows individuals to resemble each pure types to a quantifiable degree. Two pure types represented >90% of schizophrenia patients, yielding distinct cognitive patterns: one of generalised cognitive deficit and one cognitively spared. The CD phenotype is relatively homogeneous and displays pervasive deficit across cognitive domains, with the most prominent dysfunctions involving verbal memory, sustained attention/working memory and general intelligence.[10] It is further characterised by early developmental delays, poor scholastic performance and social skills, and a clustering of soft neurological signs.[11]

Our genome-wide linkage analysis of schizophrenia, using ordered subsets with the composite CD scores as covariate, identified 6p25–p24 as the best region, with a logarithmic odds score of 3.3 contributed exclusively by the CD families.[10] The short arm of chromosome 6 is among the best replicated in linkage studies of schizophrenia, with two loci (6pter–p22.3 and 6p22.3–p21.1) pointing to the possible presence of more than one susceptibility gene.[12] Here we present the follow-up investigations of our 6p25–p24 linkage findings. Association mapping and functional analyses (luciferase reporter assays, site-directed mutagenesis, gene expression in lymphoblasts, gel-shift experiments and computer modelling) implicate two colocalising genes encoding mitochondrial proteins in CD in schizophrenia, with stronger evidence favouring *LYMR4* as the candidate. The data add to the accumulating evidence of mitochondrial dysfunction contributing to the pathogenesis of the disorder.[13]

Materials and methods

Subjects

The discovery sample comprised 583 subjects of European ancestry (>75% Anglo-Irish): 381 schizophrenia patients (mean age 33.9, range 17–60 years) recruited from consecutive admissions to psychiatric hospitals or community mental health centres, and 202 controls (mean age 39.6, range 17–76 years) randomly sampled from local telephone directories, or among Red Cross blood donors, and screened for psychopathology to exclude those with personal or family history of psychotic illness. Later in the study, the WAFSS sample was expanded by 126 cases (mean age 39.7, range 18–65 years) and 80 controls (mean age 39.1, range 18–59 years) to a total of 789 participants (507 cases, 282 controls) (demographics in Supplementary Table 1). Written informed consent was obtained from all subjects. The study was approved by the Human Research Ethics Committee of The University of Western Australia.

Diagnostic assessment was based on a modification of the Schedules for Clinical Assessment in Neuropsychiatry interviews,[14] scored using the OPCRIT algorithm.[15] Video-recorded interviews and clinical charts were independently reviewed by two senior clinicians who assigned consensus research International Classification of Diseases, 10th edition and Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition diagnoses of schizophrenia and spectrum disorders.[2] Patients and controls were administered a battery of tests assessing neurocognitive performance.[10] Data from the multiple neurocognitive domains were integrated into the CD and cognitively spared composite continuous traits[8–10] (Supplementary Methods and Supplementary Table 1). In addition, 485 cases were examined for soft neurological signs using the Neurological Evaluation Scale.[16]

An independent ethnically matched Irish replication sample (Trinity College, Dublin, Ireland) included 288 schizophrenia cases and 172 controls, where cognitive assessment similar to WAFSS allowed comparisons of individual test results and an estimation of CD scores (please see Phenotype definition section in Supplementary Methods).

The question whether the findings could be extended to normal cognitive function and age-related decline was addressed in a sample of 521 normal ageing men (age ⩾65 years, mean 76.6, range 71–87 years), recruited randomly from the electoral roll-all

living independently in the community and volunteering to participate in the Health In Men Study.[17,18] Memory performance had been assessed using the California Verbal Learning Test (CVLT-II).

SNP genotyping

A 10 Mb interval of the linked region was saturated with 1170 tagging single-nucleotide polymorphisms (SNPs) (MAF >0.1, $r^2>0.8$) from HapMap Rel21/phaseII-listed polymorphisms, supplemented with coding SNPs (dbSNP130). The SNPs captured variation in the longest transcripts and 10 kb upstream and downstream of 46 brain-expressed genes and mRNAs in the interval. We used the Illumina GoldenGate genotyping technology, with CEPH trio 1334 (Coriell Cell Repository, Camden, NY, USA) as internal controls. Data quality was assessed with the Illumina BeadStudio Genotyping Module (San Diego, CA, USA). Call rates were $\geqslant 98\%$ for all samples. Consistency (at $P>0.001$) with Hardy–Weinberg equilibrium in the control samples was analysed in PLINK 1.06.[19] Of the 1170 selected SNPs, two were monomorphic and one failed Hardy–Weinberg equilibrium testing, leaving 1167 variants in the statistical analyses (Supplementary Table 2). Replication samples were genotyped using Applied Biosystems Taqman® assays (Carlsbad, CA, USA).

Cell cultures

Lymphoblastoid cell lines (LCLs) from 249 WAFSS samples were used for gene expression and mitochondrial translation analysis. HEK293 and SH-SY5Y neuroblastoma cells served for luciferase reporter and gel-shift assays.

Statistical analysis of genetic association

The association was analysed with International Classification of Diseases, 10th edition/Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition-defined schizophrenia and with seven quantitative traits—composite CD scores and raw scores on individual traits contributing most significantly to the CD/cognitively spared classification: pre-morbid IQ (National Adult Reading Test); current IQ (Shipley Institute of Living Scale); verbal memory (Rey's Adult Verbal Learning Test immediate, RAVLT-IW, and delayed, RAVLT-DW, word recall); sustained attention (Continuous Performance Task identical pairs and degraded stimulus). Association with disease outcome was analysed in the schizophrenia–controls discovery sample using χ^2 goodness-of-fit test of allelic association (1-degree-of-freedom). Association with quantitative traits scores was analysed using linear regression on number of alleles as implemented in PLINK 1.06.[19] For each trait, phenotypes were randomly permuted as described below to calculate empirical *P*-values corrected for all markers tested. Permutation correction ensured that the *P*-values for non-normal traits (composite CD scores) were valid.

Permutation-based point-wise (uncorrected) *P*-values were calculated using adaptive permutation (--aperm in PLINK). For CD and RAVLT-DW scores, *P*-values corrected for multiple testing were obtained by calculating the highest test statistic across all SNPs from each of 10 000 permutations, with corrected empirical *P*-values obtained by comparing the observed test statistic with permutation results (--mperm in PLINK).

To test whether there was strong signal in the top 10 most significant SNPs, we employed a multiple test statistic (T)—the sum of the square of the *t*-statistics from the linear regression. Empirical significance was established by 10 000 permutations and comparing the observed T with that seen in each permutation replicate.

We also examined if the top 10 SNPs were more significantly clustered than expected by chance given the linkage disequilibrium (LD) pattern across the region. A clustering metric calculated on the empirical data was compared with the same metric applied to permutation replicates, where phenotype labels are repeatedly shuffled. Performing permutations in this way preserved the original LD structure. The *ad hoc* clustering metric was the variance of the physical distances (kb) between the top 10 SNPs, with end points explicitly included (even if the selected SNPs are not at the end points). This variance takes high values when there is strong clustering in the data and low values for no clustering (SNPs spread evenly over the region with similar inter-marker distances). The actual values of the association test statistic are not included, only the top 10 SNPs and their physical locations. The variance of the physical distances in the real data was compared with the variance from 10 000 permutation replicates and empirical *P*-values calculated. Although physical distances do not necessarily correlate highly with LD, as the observed LD structure is preserved in the permutations, any increased clustering due simply to LD is appropriately taken into account. This test is region-wide and no further correction for multiple testing is required.

Bioinformatics analysis

The genomic structure of the region harbouring the cluster of associated SNPs was examined in the UCSC Genome Browser. NCBI databases were used for checking polymorphisms in the region (dbSNP build 130) and expression data (GEOprofiles). LD was analysed in HaploView.[20] SNP-related changes in transcription factor binding sites (TFBS) were explored using TESS (http://www.cbil.upenn.edu/tess), CONSITE[21] and Genomatix (http://www.genomatix.de/en/produkte/genomatix-software-suite.html). The SNPExpress database[22] was searched for association between SNPs of interest and gene expression data. URLs are listed as Supplementary Information.

Genomic and functional analyses

Detailed experimental protocols are provided in Supplementary Methods.

Polymorphisms in the FARS2/LYRM4 *promoter region.* Since the top associated SNPs were located

within or close to the predicted overlapping *LYRM4/FARS2* promoters, we characterised comprehensively sequence variation in the region. A 1988 bp fragment (chr.6: 5 200 425–5 202 412, NT_034880.3) was sequenced in control samples from five parents–child trios and 20 singletons. A 240 bp subregion (chr.6: 5 200 663–5 200 902, NT_034880.3) was sequenced in 723 WAFSS samples.

Postnatal LYRM4 *and* FARS2 *expression in the brain.* Postnatal developmental regulation of transcription was examined in our previously published microarray expression data[23] by specifically analysing *LYRM4* and *FARS2* levels. The expression analysis was conducted on total RNA from grey matter of the middle frontal gyrus (Brodman's area 46), obtained from 45 post-mortem brain samples grouped into neonates, infants, toddlers, school-age, teenagers, young adults and adults.[23]

Published LYRM4 *and* FARS2 *expression data.* RAW Affymetrix CEL files for microarray experiments were downloaded from the NCBI Gene Expression Omnibus database. Results from each study were normalised using the robust multi-array average algorithm available in the Affymetrix package of Bioconductor (Santa Clara, CA, USA). Mean log-normalised expression values were compared with a *t*-test to determine significant differences between sample groups for the probe sets representing *LYRM4* and *FARS2*.

Luciferase reporter assays. Promoter activity, the effect of DNA polymorphisms on expression, and the role of associated rs2224391 were examined in reporter assays. The polymerase chain reaction (PCR) amplified 1988bp fragment was cloned in a Firefly luciferase reporter plasmid in two orientations, following the direction of gene transcription—forward (*FARS2*) and reverse (*LYRM4*), each in two versions—homozygous for the major versus the minor SNP alleles haplotype. The individual effect of rs2224391 was analysed by site-directed mutagenesis targeting that SNP alone while preserving the remaining haplotype background, that is, introducing the minor rs2224391 allele into the major haplotype and *vice versa*. Plasmids were transiently co-transfected with a normalising *Renilla* luciferase reporter into neuroblastoma SH-SY5Y and HEK293 cells.

Gene expression in LCLs. LYRM4 (exons 1 and 2) mRNA levels were determined in 249 LCLs using quantitative PCR with the house-keeping gene *HPRT1* for normalisation. Association with promoter SNPs was analysed in SimHap.[24]

Gel-shift assays. Gel-shift experiments tested the sequence specificity and the effect of SNP alleles on the binding of nuclear proteins. A biotin-labelled oligonucleotide containing the SNPs of interest was incubated with nuclear extracts from SH-SY5Y cells. Formation of complexes with transcription machinery proteins was monitored on an electrophoretic gel. The specificity of binding was assessed in self- and cross-competition assays with non-labelled oligonucleotides, identical respectively to the test sequence or containing the alternative SNP alleles. Interference with DNA–protein interactions was investigated by the addition of antibodies targeting specific transcription factors, to determine the identity of the protein(s) driving the interaction.

Promoter DNA breathing dynamics. A recently developed computational approach[25,26] was applied independently to model the effect of SNPs on DNA breathing dynamics (the propensity of the sequence to form 'bubbles' necessary for the binding of the transcription machinery). Langevin molecular dynamic simulations, based on the Extended Peyrard–Bishop–Dauxois nonlinear model of DNA,[27] assessed bubble-formation probability, bubble lifetime and average strand separation as the mechanistic parameters characterising the transcriptional activity of the sequence. The simulations were run as described[25,28] on Linux clusters at the Los Alamos National Laboratory and Harvard Medical School.

Mitochondrial protein synthesis and steady-state mitochondria-encoded protein abundance. De novo mitochondrial protein synthesis was studied by incorporating ^{35}S-labelled methionine and cysteine in mitochondrial proteins in LCLs from subjects homozygous for the major or minor SNP haplotype, after inhibition of cytoplasmic protein translation with emetine. Steady-state levels of mitochondria-encoded proteins were analyzed by immunoblotting using mouse αCOX1 and αNDI antibodies and porin as a loading control.

Results

A 260 kb region associated with measures of cognitive dysfunction

Analysis of association with clinical schizophrenia identified SNPs scattered across multiple genes with marginal results (Supplementary Table 3), none of which withstood correction for multiple testing. In contrast, the use of quantitative cognitive measures identified a cluster of SNPs spanning 260 kb on 6p25.1, associated (at $P < 0.001$) with composite CD and memory (RAVLT-DW) scores (Figure 1, Table 1 and Supplementary Table 4). After correction for multiple SNPs, two variants remained significantly associated: rs17736905 with CD scores ($P = 0.029$) and rs2503812 with RAVLT-DW ($P = 0.026$). Analysis considering the summed test statistics of the top 10 SNPs generated a *P*-value of 0.0109 for CD scores and 0.0028 for RAVLT-DW (corrected for multiple markers), with study-wide *P*-values (corrected for eight phenotypes) 0.087 for CD and 0.0224 for RAVLT-DW. Testing if the clustering of the top 10 SNPs was more significant than expected, given the LD pattern,

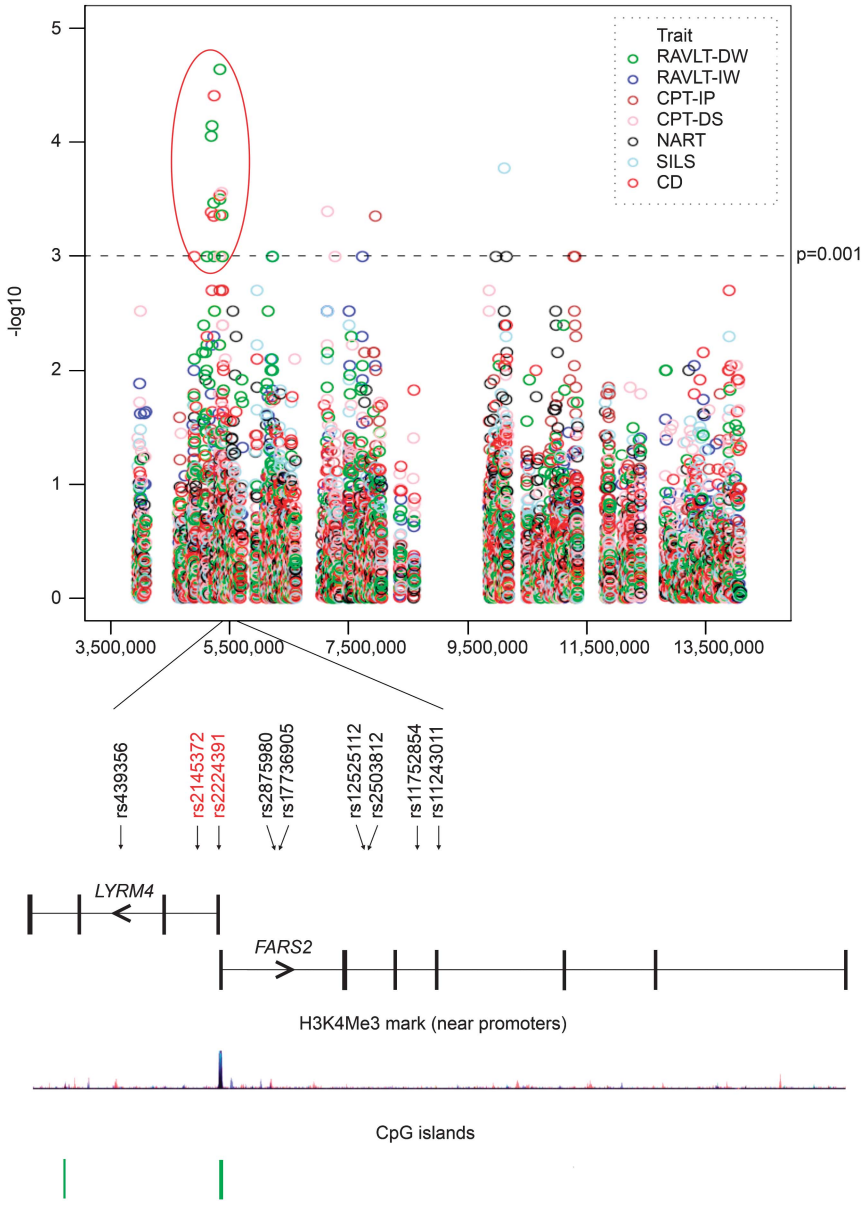

Figure 1 Association mapping of the 6p25–p24 linked region. Upper panel: Manhattan plot of the distribution of *P*-values for cognitive endophenotypes over a ~10 Mb region (chr.6: 3 957 928–14 090 460, NT_034880.3) covered by 1167 single-nucleotide polymorphisms (SNPs) in the discovery sample. The red ellipse outlines the 6p25.1 cluster of SNPs associated (at $P<0.001$) with composite cognitive deficit and verbal memory scores. Lower panel: Position of the cluster of associated SNPs (chr.6:5 117 780–5 376 339, NT_034880.3) relative to the *LYRM4* (−strand)–*FARS2* (+strand) genome structure. The two SNPs with strongest evidence of association (in red) are located around the overlapping promoters.

generated a *P*-value of 0.0010 for CD scores and 0.0039 for RAVLT-DW (study-wide *P*-values 0.008 for CD scores and 0.0312 for RAVLT-DW).

In the Irish sample, replication analysis of the four top SNPs (Table 1) supported association, with an identical direction of the effect, of memory performance (based on the Logical Memory delayed subtest from the Wechsler Memory Scale) with rs2145372 ($P=0.019$) and rs2224391 ($P=0.037$), and a trend for rs2875980 ($P=0.055$; correlated with rs17736905). Estimated CD scores were significantly associated with rs2145372 ($P=0.035$). No significant results were obtained for rs2503812 (Table 1). The two SNPs supported by the replication, rs2145372 (allele G) and rs2224391 (allele C), were analysed further in the newly recruited WAFSS participants and jointly in the entire WAFSS sample. Both showed association in the new sample, as well as in the entire data set, with highly significant *P*-values for both CD scores and memory performance (Table 1). Word list learning (as in RAVLT and CVLT-II) and logical memory (WMS-R) target similar or overlapping functions

Table 1 Genetic association of 6p25.1 SNPs and cognitive endophenotypes

SNP rs nos.	*Position in NT_034880.3*	*Gene position*	*Memory performance in cases*				*CD scores in entire sample*			
			P-value; β-coefficient (95% CI)				*P-value; β-coefficient (95% CI)*			
			Discovery WAFSS, N = 381	*Additional WAFSS, N = 126*	*Total WAFSS, N = 507*	*Irish sample, N = 288*	*Discovery WAFSS, N = 583*	*Additional WAFSS, N = 206*	*Total WAFSS, N = 789*	*Irish sample, N = 460*
439356	5 112 780	*LYRM4 IVS4*	0.00062; −0.82 (−1.29, −0.35)				0.00470; 0.06 (0.02, 0.1)			
2145372	5 185 069	*LYRM4 IVS1*	**7.94E-05**; −1.05 (−1.57, −0.53)	**0.014**; −0.77 (−1.79, 0.25)	**0.00023**; −0.94 (−1.47, −0.41)	**0.019**; −2.00 (−3.89 to −0.12)	**0.00041**; 0.09 (0.04, 0.13)	**0.002;** 0.13 (0.05, 0.21)	**3.02E-06;** 0.094 (0.05, 0.13)	**0.035;** 0.08 (0.00, 0.17)
2224391 (S11A)	5 200 935	*LYRM4 Ex1 and promoter region*	**6.98E-05**; −0.92 (−1.37, −0.47)	**0.007**; −1.27 (−2.19, −0.35)	**3.85E-05**; −0.99 (−1.48, −0.50)	**0.037**; −1.57 (−3.28, 0.15)	**0.00165;** 0.07 (0.03, 0.11)	**0.045;** 0.07 (−0.01, 0.15)	**0.00013;** 0.067 (0.03, 0.11)	0.135; 0.04 (−0.03, 0.12)
2875980	5 229 446	*FARS2 IVS1*	0.00036; −0.92 (−1.43, −0.42)			0.055; −1.53 (−3.42, 0.36)	0.00044; 0.08 (0.04, 0.13)			0.23; −0.03 (−0.11, 0.05)
17736905	5 231 510	*FARS2 IVS1*	0.00069; −1.01 (−1.58, −0.43)				**3.89E-05;** 0.11 (0.06, 0.16)			
12525112	5 329 498	*FARS2 IVS2*	0.00047; −0.95 (−1.47, −0.43)				0.00044; 0.08 (0.04, 0.13)			
2503812	5 330 603	*FARS2 IVS2*	**1.9E-05**; −0.98 (−1.44, −0.53)			0.24; −0.56 (−2.13, 1.01)	**0.00029'** 0.08 (0.04, 0.11)			0.102; 0.05 (−0.02, 0.12)
11752854	5 364 761	*FARS2 IVS3*	0.00036; −0.97 (−1.51, −0.43)				0.00058; 0.08 (0.04, 0.13)			
11243011 (N280S)	5 371 339	*FARS2 Ex 4*	0.00067; −0.92 (−1.45, −0.38)				0.00180; 0.08 (0.03, 0.12)			

Abbreviations: CD, cognitive deficit; CI, confidence interval; SNP, single-nucleotide polymorphism; WAFSS, Western Australian Family Study of Schizophrenia.
Bolded: SNPs analysed in Irish replication sample.

termed auditory memory, as demonstrated by factor analysis.[29,30]

In WAFSS cases, testing for association with soft neurological signs (prevalent in patients with CD) revealed nominally significant results for signs of impaired motor coordination: rs2145372, $P=0.033$ and rs2224391, $P=0.026$ (Table 2 and Supplementary Table 5). The other two components of the test battery showed no association.

In the sample of normal ageing men, verbal memory showed no association with either rs2145372 ($P=0.37$) or rs2224391 ($P=0.13$). This negative result could point to different mechanisms in age-related cognitive decline or to additional, age-related factors (for example, hypertension, stroke and diabetes, and so on),[18] masking the association in the Health In Men Study sample (mean age 76.6 years) compared to schizophrenia (mean age 33.9 years).

Two colocalising genes encoding mitochondrial proteins

The region contains two genes on the opposite DNA strands, *LYRM4* and *FARS2*, with overlapping predicted promoters (Figure 1). *FARS2* encodes the mitochondrial isoform of phenylalanine-tRNA synthetase, a member of an ancient family of enzymes, which catalyse the attachment of specific amino acids to their cognate tRNAs, thus 'establishing the rules of the genetic code'.[31] Mutations in other aminoacyl-tRNA synthetases (ARS) are known to result in disorders affecting the peripheral as well as central nervous system.[31,32] *LYRM4* encodes another mitochondrial protein (known in the biochemical literature as ISD11)—a eukaryote-specific component of the ISCU/NFS1/LYRM4 mitochondrial complex responsible for the biogenesis of Fe–S (iron–sulphur) clusters, essential cofactors in a variety of processes, including electron transfer during oxidative phosphorylation.[33–36] *LYRM4* is a stabiliser of the Fe–S cluster assembly platform, where it interacts directly with the complex-activating protein frataxin, which is mutated in Friedreich's ataxia.[33,34,36] Neither gene is an obvious candidate.

The top two consistently associated SNPs, rs2224391 and rs2145372 ($D'=0.96$, $r^2=0.65$), are within or close to the shared promoters (Figure 1). The SNPExpress database reports association between rs2224391 and brain mRNA levels of both genes, with different directions of the effects: higher *FARS2* and lower *LYRM4*. Direct rs2224391 involvement in transcription regulation was suggested by all three bioinformatic programs (see Materials and methods) predicting loss of a GATA1/3 binding site due to the A>C transversion. At the same time, dbSNP listed additional polymorphisms in the predicted promoters, prompting us to characterise fully the diversity in our sample. Sequencing analysis of the predicted promoters identified 22 SNPs (Supplementary Figure 1) in a tight LD block ($D'\sim1.0$; $r^2>0.7$, broken between rs2224391 and rs4141761) with many predicted to affect TFBS.

The significantly associated SNPs around the putative overlapping promoters, the reported association of rs2224391 with mRNA levels and its predicted direct effect on TFBS, and the presence of other, correlated and possibly functional, polymorphisms in the promoter region focused our subsequent analyses on expression regulation.

LYRM4 *is developmentally regulated*

Expression microarray data in GEOProfiles showed rising *Lyrm4* ($P=0.001$), but not *Fars2* ($P=0.263$) brain mRNA levels from mouse embryonic day e11.5 (onset of neuronal differentiation) to day e13.5 (peak of neurogenesis) (GEOrecord GDS3442). Our analysis of prefrontal cortex *LYRM4* and *FARS2* mRNA levels in individuals aged 0–50 years demonstrated increasing expression parallel to postnatal brain maturation and intensifying neuronal activity and energy demands. We found a significant, high magnitude, increase for *LYRM4* mRNA ($r=0.75$, $P<0.00001$; ANOVA, $P=3.9E\text{-}5$) and less steep, but still significant

Table 2 Binary logistic regression analysis of SNPs rs2145372 and rs2224391 and the presence of neurological abnormalities in 485 WAFSS schizophrenia cases

SNP	*rs2145372*					*rs2224391*				
	B *(s.e.)*	*Wald*	P-*value*	*OR*	*95% CI*	B *(s.e.)*	*Wald*	P-*value*	*OR*	*95% CI*
Motor coordination[a]	0.50 (0.2)	4.55	0.033	1.65	1.04–2.62	0.45 (0.2)	4.98	0.026	1.57	1.06–2.33
Spontaneous movements[b]	0.12 (0.3)	0.12	0.729	1.12	0.58–2.20	0.30 (0.3)	1.21	0.271	1.36	0.79–2.33
Sequencing of motor acts[c]	0.18 (0.2)	0.51	0.476	1.20	0.73–1.95	−0.02 (0.2)	0.01	0.937	0.98	0.65–1.49

Abbreviations: CI, confidence interval; NES, Neurological Evaluation Scale; OR, odds ratio; SNP, single-nucleotide polymorphism; WAFSS, Western Australian family Study of Schizophrenia.
Affected subjects were examined using the NES.[16] The estimates above are based on the presence of one or more pathological sign in the respective category.
[a]Tandem walk, Romberg, diadochokinesis, finger–nose and finger–thumb tests.
[b]Adventitious overflow, mirror movements.
[c]Rhythm tapping, fist–ring and fist–edge–palm tests.

for *FARS2* ($r = 0.515$, $P < 0.001$; ANOVA, $P = 0.014$) (Supplementary Figure 2). The data suggest that both genes, particularly *LYRM4*, are upregulated in the first decade of human life and that increased synthesis of these gene products occurs until around adolescence, coincident with the typical onset of schizophrenia. Further GEOProfile searches revealed downregulation of *LYRM4* ($P = 0.009$), but not *FARS2* ($P = 0.205$) expression in the cerebellar cortex of schizophrenia patients compared with controls (GEOrecord GSE4036), and of *Lyrm4* ($P = 6.15e\text{-}05$), but not *Fars2* ($P = 0.058$) in the prefrontal cortex of mice with microdeletions in the locus syntenic to human 22q11.2 (GEOrecord GSE10784).

Lower LYRM4 expression under the minor promoter SNP haplotype

Reporter assays showed that luciferase expression in the test constructs was comparable to that under the strong SV40 promoter and 50- to 100-fold higher than that in the plasmid without promoter (Figure 2). This confirmed that promoters driving transcription in both orientations are present in the cloned fragments. The promoter SNPs had a pronounced effect in the *LYRM4* orientation, with lower expression under the minor haplotype (mean value 6.72 (s.d. 1.26) compared to 8.74 (s.d. 1.57) under the major haplotype, $P = 0.002$). A less significant effect in the opposite direction was observed in the *FARS2* orientation (mean value 10.83 (s.d. 1.94) under the major versus 12.82 (s.d. 2.25) under the minor haplotype, $P = 0.029$) (Figure 2). Site-directed mutagenesis, changing rs2224391 alleles while preserving the haplotype background, had a modest effect on expression in the *LYRM4* orientation ($P = 0.013$) and no effect in the *FARS2* orientation ($P = 0.24$) (Figure 2). Quantitative *LYRM4* mRNA analysis in 249 LCLs showed lack of effect of rs2224391 (mean level in minor allele homozygotes 0.89 (s.e. 0.08); mean levels in major allele carriers, hetero- and homozygotes, 0.95–0.97 (s.e. 0.02), $P = 0.65$). Taken together, these data pointed to *LYRM4* as the gene whose expression levels are markedly regulated by promoter SNPs, and suggested that rs2224391 alone cannot explain the haplotype effects, but may modulate them.

Two adjacent SNPs regulate LYRM4 *transcription*

In search of the functional variant(s), we next examined a tandem of adjacent SNPs, rs7752203 (C > G)–rs4141761 (A > G), 3 bp apart and predicted to be recognised jointly as TFBS. In the reporter assay, they were represented on the major haplotype as the C–G combination. Analysis of LCL mRNA levels showed significant *LYRM4* upregulation in the presence of the rs7752203–rs4141761–C–G haplotype (frequency 0.44): mean mRNA levels 0.99–1.04 in C–G carriers, hetero- and homozygotes, and 0.86 in non-carriers ($P = 0.00034$). *Post-hoc* analysis of the effect of rs7752203–rs4141761 on cognitive performance in the expanded WAFSS sample showed association of the G–A haplotype with lower RAVLT-DW (β-coefficient = −0.83; $P = 0.0025$) and higher CD scores (β-coefficient = 0.07; $P = 0.0004$).

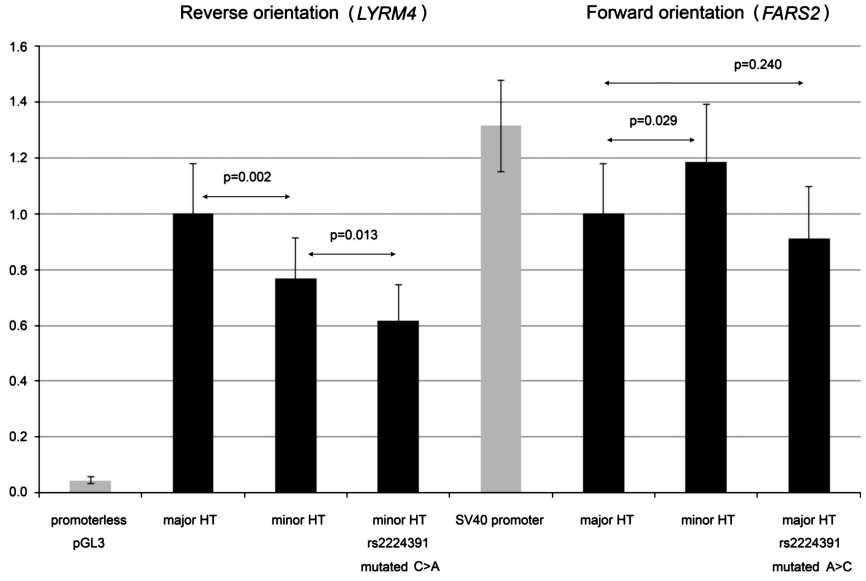

Figure 2 Effect of promoter polymorphisms on luciferase reporter expression in SH-SY5Y cells. Black bars show firefly luciferase expression in the test constructs, normalised against co-transfected *Renilla* luciferase. For each orientation, reverse (*LYRM4*) and forward (*FARS2*), expression under the major haplotype is set at 1, and that driven by the minor haplotype and rs2224391-mutated is presented as a proportion of this calibrator. Site-directed mutagenesis of rs2224391 was performed on the haplotype backgrounds leading to lower reporter expression. Control plasmids are shown in grey: negative (basic pGL3 plasmid lacking promoter) and positive (pGL3 with the strong SV40 promoter). Similar results were obtained in HEK293 cells (not shown).

Direct evidence of the involvement of these two SNPs in transcriptional regulation was obtained in gel-shift assays, where the biotin-labelled oligonucleotide probe containing the rs7752203–rs4141761–C–G sequence (and no other polymorphisms) identified two bands of presumably different protein composition, where specific binding was demonstrated by decreasing band intensity in the self-competition assay (Figure 3a). Preferential binding to the C–G sequence, especially of complex A, was supported by the cross-competition experiment where the G–A oligonucleotide was inefficient in displacing the C–G self-probe. The results suggest that the rs7752203–rs4141761 effect on transcription is direct and not a reflection of other correlated polymorphisms. Pretreatment with antibodies, targeting transcription factors predicted to recognise the oligonucleotide sequence, showed consistently diminished or absent band A after treatment with the antibody against nuclear factor-κB (NF-κB) subunit p50 (Figure 3b), suggesting that NF-κB may be the transcription factor involved in the preferential C–G binding.

Independently and blind to the experimental data, computer modelling revealed significant sequence effects on DNA breathing dynamics (Figure 3c). The simulations identified a fragment encompassing rs7752203–rs4141761, where the C–G sequence displayed high breathing probability, 9–10 bp bubble length and ~6 ps lifetime-parameters characterising sites with strong specific binding of the transcription machinery.[26]

Discussion

Our previous study of a modest number of schizophrenia families identified linkage to 6p25–p24 as a result of the dissection of the heterogeneous clinical phenotype by incorporating neurocognitive features in its characterisation.[10] In the present follow-up association analysis, clinical schizophrenia produced predictably unremarkable results, whereas the use of quantitative endophenotypes allowed us to refine a 10 Mb region with 46 positional candidates to a 260 kb interval containing only two genes. The strongest evidence of association, based on three different approaches to the statistical analysis, was obtained for the composite CD scores, in logical continuity with our linkage results, and with memory, a cognitive domain impaired early in the

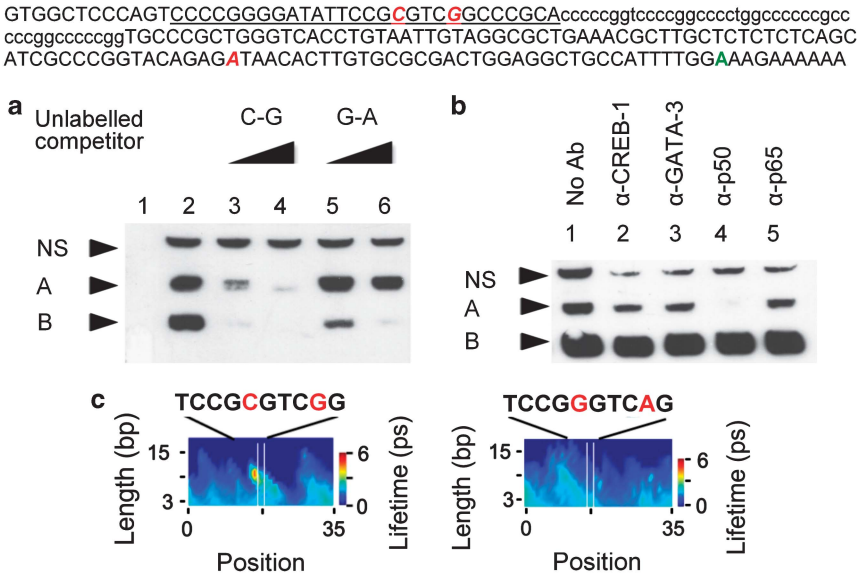

Figure 3 Gel-shift experiments and computer modelling of DNA breathing dynamics. At the top, sequence examined in the computer simulations, with the oligonucleotide probe used in gel-shift experiments underlined. Single-nucleotide polymorphisms (SNPs) rs7752203, rs4141761 and rs2224391 (left to right, italicised and in red) are shown as the major haplotype. The *LYRM4* transcription start site is bolded and in green. (**a**) Gel-shift assay with nuclear extracts prepared from neuroblastoma SH-SY5Y cells and rs7752203–rs4141761–C–G as probe. Unlabelled oligonucleotides used in the self- and cross-competition assays are shown on top, with triangles indicating increasing concentrations (20-/100-fold molar excess) of the competitor. Lanes: 1, no nuclear extract; 2, no competitor; 3 and 4, self-competitor; and 5 and 6, cross-competitor. Bands: NS, nonspecific; A and B, specific, with decreasing intensity in the self-competition experiment due to displacement of the labelled probe by the excess of unlabelled oligonucleotide. Preferential binding of the protein complex to the C–G sequence is suggested by the inability of the G–A oligonucleotide to compete out the labelled C–G probe. (**b**) Treatment of the nuclear extract with antibodies targeting specific transcription factors, before the addition of the oligonucleotide probe. The DNA–protein interaction is blocked by the antibody against nuclear factor-κB (NF-κB) subunit p50, pointing to NF-κB as the transcription factor driving complex A. (**c**) Computer models of DNA breathing dynamics. Strong breathing activity is predicted for a short sequence containing rs7752203–rs4141761–C–G (left), in contrast to the poor breathing potential of the G–A sequence (right). The white horizontal lines mark the SNP sites. Vertical axis, bubble length in bp; colour axis, bubble lifetime in picoseconds (predicted for DNA openings with amplitude >3.5 Å).

evolution of the disease and a major contributor to the overall CD.[4,5,37]

The identification of two colocalising genes, encoding mitochondrial proteins with no immediately obvious relationship to functional deficits in schizophrenia, made the choice of the better candidate a difficult one. Our data pointed to transcriptional regulation as the likely molecular mechanism and favoured *LYRM4* as the better candidate based on its developmental regulation, the marked effects of promoter polymorphisms on its expression and existing knowledge of its function. Contrary to bioinformatic predictions, our experimental data did not support rs2224391 and implicated rs7752203–rs4141761 as the regulatory polymorphisms. Experimental evidence was weak for *FARS2*: associated polymorphisms had a modest effect on expression and no effect on mitochondrial protein translation (Supplementary Figure 3). One should note however that our experimental system and the use of LCLs may not be an adequate model of living neurons; moreover, ARS mutations can lead to neurological disorders even when aminoacylation is unaffected.[31,32,38] Current understanding of proteins is far from complete, and *LYRM4/FARS2* co-regulation as a control mechanism for unknown functions remains a possibility.

LYRM4 (ISD11) is part of the assembly platform responsible for the biogenesis of Fe–S clusters, essential cofactors of a variety of proteins involved in electron transfer, enzymatic catalysis, DNA replication and repair, and iron homeostasis.[35,39,40] Energy metabolism is highly dependent on the availability of such proteins, as they include Krebs cycle and oxidative phosphorylation components. Defects in the different proteins of the Fe–S cluster assembly complex invariably affect energy metabolism, but show interesting differences in target tissue distributions and phenotypic effects. Mutations in the scaffold protein ISCU cause myopathy with exercise intolerance,[41,42] whereas frataxin deficiency leads to impaired mitochondrial energy metabolism affecting primarily the dorsal root ganglia, cerebellum and heart muscle.[43,44] Thus far, ISD11 has not been implicated in a human disease phenotype; however, its knockdown in yeast and human cells results in reduced levels and activity of Fe–S proteins,[33,34,45,46] and we note that lower levels of Fe–S proteins have been reported in post-mortem schizophrenia brains.[47–49] A commonality of the molecular effects of frataxin and ISD11 deficit is an intriguing hypothetical scenario, supported by the transcriptional co-repression of *LYRM4* in Friedreich's ataxia cells[45,46] and by our observation that SNPs, associated with CD scores and memory, were also associated with motor dyscoordination in our patients.

Mitochondrial dysfunction, which may be the end result of multiple converging processes, is central to senescence and neurodegeneration.[50–52] Our findings suggest that subtle chronic *LYRM4* downregulation could be one of the mechanisms behind impaired oxidative phosphorylation function and oxidative stress in schizophrenia, increasingly recognised as contributors to disease pathogenesis and specifically to impaired cognitive performance in affected subjects.[47,49,53–56] NF-κB, identified in our experiments as a transcription factor regulating *LYRM4* expression, is thought to be involved in neuronal growth, differentiation and plasticity, as well as in the response to oxidative stress—a function potentially relevant to *LYRM4* regulation.[57–59]

The evolution and findings of this study illustrate the long and convoluted road from genetic association to facing the complexity of molecular pathogenesis.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We thank patients, family members and volunteer controls for their participation. The study was supported by National Health and Medical Research Council of Australia Grant Nos. 37580400 and 37580900 to AJ and LK and Training Fellowship 634551 to DNA, with funding contribution from the North Metropolitan Health Area, Perth, Western Australia. Recruitment and genotyping of the Irish sample was supported by the Wellcome Trust and Science Foundation Ireland (SFI). Research performed at Los Alamos National Laboratory was carried out under the auspices of the National Nuclear Security Administration of the US Department of Energy under Contract No. DE-AC52-06NA25396. GD was supported by an NARSAD Young Investigator Award, and EB-B by an MRC (UK) Young Investigator Award. CSW is supported by the Schizophrenia Research Institute, utilising funds from the Macquarie Group Foundation and NSW Health. The SH-SY5Y cell line was a kind gift from Dr B Meloni (CNND, UWA) and the pcDNA3.1 (+) vector plasmid from Dr K Pfleger (WAIMR, UWA).

References

1 Jablensky A. Subtyping schizophrenia: implications for genetic research. *Mol Psychiatry* 2006; **11**: 815–836.
2 Jablensky A. The diagnostic concept of schizophrenia: its history, evolution, and future prospects. *Dialogues Clin Neurosci* 2010; **12**: 271–287.
3 Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat Genet* 2003; **33**: 228–237.
4 Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology* 1998; **12**: 426–445.
5 Gur RE, Calkins ME, Gur RC, Horan WP, Nuechterlein KH, Seidman LJ *et al.* The Consortium on the Genetics of Schizophrenia: neurocognitive endophenotypes. *Schizophr Bull* 2007; **33**: 49–68.
6 Reichenberg A. The assessment of neuropsychological functioning in schizophrenia. *Dialogues Clin Neurosci* 2010; **12**: 383–392.

7 Dickinson D, Harvey PD. Systemic hypotheses for generalized cognitive deficits in schizophrenia: a new take on an old problem. *Schizophr Bull* 2009; **35**: 403–414.
8 Woodbury MA, Clive J, Garson Jr A. Mathematical typology: a grade of membership technique for obtaining disease definition. *Comput Biomed Res* 1978; **11**: 277–298.
9 Manton KG, Korten A, Woodbury MA, Anker M, Jablensky A. Symptom profiles of psychiatric disorders based on graded disease classes: an illustration using data from the WHO International Pilot Study of Schizophrenia. *Psychol Med* 1994; **24**: 133–144.
10 Hallmayer JF, Kalaydjieva L, Badcock J, Dragovic M, Howell S, Michie PT *et al.* Genetic evidence for a distinct subtype of schizophrenia characterized by pervasive cognitive deficit. *Am J Hum Genet* 2005; **77**: 468–476.
11 Jablensky A. Challenging the genetic complexity of schizophrenia by use of intermediate phenotypes. In: Ritsner MS (ed). *The Handbook of Neuropsychiatric Biomarkers, Endophenotypes and Genes*. Springer Science: New York, NY, 2009, pp 41–56.
12 Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I *et al.* Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: schizophrenia. *Am J Hum Genet* 2003; **73**: 34–48.
13 Rosenfeld M, Brenner-Lavie H, Ari SG, Kavushansky A, Ben-Shachar D. Perturbation in mitochondrial network dynamics and in complex I dependent cellular respiration in schizophrenia. *Biol Psychiatry* 2011; **69**: 980–988.
14 Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R *et al.* SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry* 1990; **47**: 589–593.
15 Castle DJ, Jablensky A, McGrath JJ, Carr V, Morgan V, Waterreus A *et al.* The diagnostic interview for psychoses (DIP): development, reliability and applications. *Psychol Med* 2006; **36**: 69–80.
16 Buchanan RW, Heinrichs DW. The Neurological Evaluation Scale (NES): a structured instrument for the assessment of neurological signs in schizophrenia. *Psychiatry Res* 1989; **27**: 335–350.
17 Norman PE, Flicker L, Almeida OP, Hankey GJ, Hyde Z, Jamrozik K. Cohort profile: the Health In Men Study (HIMS). *Int J Epidemiol* 2009; **38**: 48–52.
18 Flicker L, Almeida O, Acres J, Le MT, Tuohy RJ, Jamrozik K *et al.* Predictors of impaired cognitive function in men over the age of 80 years: results from the Health in Men Study. *Age Ageing* 2005; **35**: 77–80.
19 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
20 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
21 Sandelin A, Wasserman WW, Lenhard B. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res* 2004; **32**: W249–W252.
22 Sanders MA, Verhaak RG, Geertsma-Kleinekoort WM, Abbas S, Horsman S, van der Spek PJ *et al.* SNPExpress: integrated visualization of genome-wide genotypes, copy numbers and gene expression levels. *BMC Genomics* 2008; **9**: 41.
23 Weickert CS, Elashoff M, Richards AB, Sinclair D, Bahn S, Paabo S *et al.* Transcriptome analysis of male–female differences in prefrontal cortical development. *Mol Psychiatry* 2009; **14**: 558–561.
24 Carter KW, McCaskie PA, Palmer LJ. SimHap GUI: an intuitive graphical user interface for genetic association analysis. *BMC Bioinform* 2008; **9**: 557.
25 Alexandrov BS, Gelev V, Monisova Y, Alexandrov LB, Bishop AR, Rasmussen KO *et al.* A nonlinear dynamic model of DNA with a sequence-dependent stacking term. *Nucleic Acids Res* 2009; **37**: 2405–2410.
26 Alexandrov BS, Gelev V, Yoo SW, Alexandrov LB, Fukuyo Y, Bishop AR *et al.* DNA dynamics play a role as a basal transcription factor in the positioning and regulation of gene transcription initiation. *Nucleic Acids Res* 2010; **38**: 1790–1795.
27 Peyrard M, Bishop AR. Statistical mechanics of a nonlinear model for DNA denaturation. *Phys Rev Lett* 1989; **62**: 2755–2758.
28 Alexandrov BS, Gelev V, Yoo SW, Bishop AR, Rasmussen KO, Usheva A. Toward a detailed description of the thermally induced dynamics of the core promoter. *PLoS Comput Biol* 2009; **5**: e1000313.
29 Tulsky DS, Price LR. The joint WAIS-III and WMS-III factor structure: development and cross-validation of a six-factor model of cognitive functioning. *Psychol Assess* 2003; **15**: 149–162.
30 Ivnik RJ, Smith GE, Petersen RC, Tangalos EG. Long-term stability and intercorrelations of cognitive abilities in older persons. *Psychol Assess* 1995; **7**: 155–161.
31 Park SG, Schimmel P, Kim S. Aminoacyl tRNA synthetases and their connections to disease. *Proc Natl Acad Sci USA* 2008; **105**: 11043–11049.
32 Antonellis A, Green ED. The role of aminoacyl-tRNA synthetases in genetic diseases. *Annu Rev Genomics Hum Genet* 2008; **9**: 87–107.
33 Adam AC, Bornhovd C, Prokisch H, Neupert W, Hell K. The Nfs1 interacting protein Isd11 has an essential role in Fe/S cluster biogenesis in mitochondria. *EMBO J* 2006; **25**: 174–183.
34 Wiedemann N, Urzica E, Guiard B, Muller H, Lohaus C, Meyer HE *et al.* Essential role of Isd11 in mitochondrial iron–sulfur cluster synthesis on Isu scaffold proteins. *EMBO J* 2006; **25**: 184–195.
35 Lill R. Function and biogenesis of iron-sulphur proteins. *Nature* 2009; **460**: 831–838.
36 Tsai C-L, Barondeau DP. Human frataxin is an allosteric switch that activates the Fe–S cluster biosynthesis complex. *Biochemistry* 2010; **49**: 9132–9139.
37 Skelley SL, Goldberg TE, Egan MF, Weinberger DR, Gold JM. Verbal and visual memory: characterizing the clinical and intermediate phenotype in schizophrenia. *Schizophr Res* 2008; **105**: 78–85.
38 Storkebaum E, Leitao-Goncalves R, Godenschwege T, Nangle L, Mejia M, Bosmans I *et al.* Dominant mutations in the tyrosyl-tRNA synthetase gene recapitulate in *Drosophila* features of human Charcot–Marie–Tooth neuropathy. *Proc Natl Acad Sci USA* 2009; **106**: 11782–11787.
39 Johnson DC, Dean DR, Smith AD, Johnson MK. Structure, function, and formation of biological iron–sulfur clusters. *Annu Rev Biochem* 2005; **74**: 247–281.
40 Lill R, Muhlenhoff U. Maturation of iron–sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annu Rev Biochem* 2008; **77**: 669–700.
41 Olsson A, Lind L, Thornell LE, Holmberg M. Myopathy with lactic acidosis is linked to chromosome 12q23.3–24.11 and caused by an intron mutation in the ISCU gene resulting in a splicing defect. *Hum Mol Genet* 2008; **17**: 1666–1672.
42 Mochel F, Knight MA, Tong WH, Hernandez D, Ayyad K, Taivassalo T *et al.* Splice mutation in the iron–sulfur cluster scaffold protein ISCU causes myopathy with exercise intolerance. *Am J Hum Genet* 2008; **82**: 652–660.
43 Pandolfo M, Pastore A. The pathogenesis of Friedreich ataxia and the structure and function of frataxin. *J Neurol* 2009; **256**(Suppl 1): 9–17.
44 Schmucker S, Puccio H. Understanding the molecular mechanisms of Friedreich's ataxia to develop therapeutic approaches. *Hum Mol Genet* 2010; **19**: R103–R110.
45 Shan Y, Napoli E, Cortopassi G. Mitochondrial frataxin interacts with ISD11 of the NFS1/ISCU complex and multiple mitochondrial chaperones. *Hum Mol Genet* 2007; **16**: 929–941.
46 Shi Y, Ghosh MC, Tong WH, Rouault TA. Human ISD11 is essential for both iron–sulfur cluster assembly and maintenance of normal cellular iron homeostasis. *Hum Mol Genet* 2009; **18**: 3014–3025.
47 Prabakaran S, Swatton JE, Ryan MM, Huffaker SJ, Huang JT, Griffin JL *et al.* Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. *Mol Psychiatry* 2004; **9**: 684–697.
48 Martins-de-Souza D, Gattaz WF, Schmitt A, Rewerts C, Marangoni S, Novello JC *et al.* Alterations in oligodendrocyte proteins, calcium homeostasis and new potential markers in schizophrenia anterior temporal lobe are revealed by shotgun proteome analysis. *J Neural Transm* 2009; **116**: 275–289.

49 Martins-de-Souza D, Harris LW, Guest PC, Bahn S. The role of energy metabolism dysfunction and oxidative stress in schizophrenia revealed by proteomics. *Antioxid Redox Signal* 2011; **15**: 2067–2079.
50 Lin MT, Beal MF. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 2006; **443**: 787–795.
51 Pickrell AM, Moraes CT. What role does mitochondrial stress play in neurodegenerative diseases? *Methods Mol Biol* 2010; **648**: 63–78.
52 Navarro A, Boveris A. Brain mitochondrial dysfunction in aging, neurodegeneration and Parkinson's disease. *Front Aging Neurosci* 2010; **2**: 34.
53 Hazlett EA, Buchsbaum MS, Jeu LA, Nenadic I, Fleischman MB, Shihabuddin L *et al.* Hypofrontality in unmedicated schizophrenia patients studied with PET during performance of a serial verbal learning task. *Schizophr Res* 2000; **43**: 33–46.
54 Volz HR, Riehemann S, Maurer I, Smesny S, Sommer M, Rzanny R *et al.* Reduced phosphodiesters and high-energy phosphates in the frontal lobe of schizophrenic patients: a (31)P chemical shift spectroscopic-imaging study. *Biol Psychiatry* 2000; **47**: 954–961.
55 Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophr Res* 2001; **49**: 1–52.
56 Bitanihirwe BK, Woo TU. Oxidative stress in schizophrenia: an integrated approach. *Neurosci Biobehav Rev* 2011; **35**: 878–893.
57 Pizzi M, Spano P. Distinct roles of diverse nuclear factor-kappaB complexes in neuropathological mechanisms. *Eur J Pharmacol* 2006; **545**: 22–28.
58 Massa PT, Aleyasin H, Park DS, Mao X, Barger SW. NFkappaB in neurons? The uncertainty principle in neurobiology. *J Neurochem* 2006; **97**: 607–618.
59 Morgan MJ, Liu ZG. Crosstalk of reactive oxygen species and NF-kappaB signaling. *Cell Res* 2011; **21**: 103–115.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)