Human Molecular Genetics, 2010, Vol. 19, No. 14 ***2841–2857***
doi:10.1093/hmg/ddq184
Advance Access published on May 10, 2010

SMARCA2 and other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate-specific evolution

Yann Loe-Mie[1,2,†], Aude-Marie Lepagnol-Bestel[1,†], Gilles Maussion[1], Adi Doron-Faigenboim[3], Sandrine Imbeaud[4], Hervé Delacroix[4], Lawrence Aggerbeck[4], Tal Pupko[3], Philip Gorwood[1], Michel Simonneau[1] and Jean-Marie Moalic[1,*]

[1]INSERM U675/U894, Centre Psychiatrie & Neurosciences, Université Paris-Descartes, 2 ter rue d'Alésia, 75014 Paris, France, [2]CNRS-UPR2589, Information Génomique & Structurale, 163 Avenue de Luminy, Marseille, France, [3]Department of Cell Research and Immunology, George S. Wise Faculty of Life Sciences, Tel-Aviv University, Tel Aviv 69978, Israel and [4]Centre de Génétique Moléculaire, FRE 3144, CNRS and Gif/Orsay DNA Microarray Platform (GODMAP), 91198 Gif sur Yvette, France

Received March 9, 2010; Revised and Accepted May 1, 2010

The *SMARCA2* gene, which encodes BRM in the SWI/SNF chromatin-remodeling complex, was recently identified as being associated with schizophrenia (SZ) in a genome-wide approach. Polymorphisms in *SMARCA2*, associated with the disease, produce changes in the expression of the gene and/or in the encoded amino acid sequence. We show here that an SWI/SNF-centered network including the *Smarca2* gene is modified by the down-regulation of *REST/NRSF* in a mouse neuronal cell line. REST/NRSF down-regulation also modifies the levels of *Smarce1*, *Smarcd3* and SWI/SNF interactors (*Hdac1*, *RcoR1* and *Mecp2*). *Smarca2* down-regulation generates an abnormal dendritic spine morphology that is an intermediate phenotype of SZ. We further found that 8 (*CSF2RA*, *HIST1H2BJ*, *NOTCH4*, *NRGN*, *SHOX*, *SMARCA2*, *TCF4* and *ZNF804A*) out of 10 genome-wide supported SZ-associated genes are part of an interacting network (including *SMARCA2*), 5 members of which encode transcription regulators. The expression of 3 (*TCF4*, *SMARCA2* and *CSF2RA*) of the 10 genome-wide supported SZ-associated genes is modified when the REST/NRSF-SWI/SNF chromatin-remodeling complex is experimentally manipulated in mouse cell lines and in transgenic mouse models. The REST/NRSF-SWI/SNF deregulation also results in the differential expression of genes that are clustered in chromosomes suggesting the induction of genome-wide epigenetic changes. Finally, we found that SMARCA2 interactors and the genome-wide supported SZ-associated genes are considerably enriched in genes displaying positive selection in primates and in the human lineage which suggests the occurrence of novel protein interactions in primates. Altogether, these data identify the SWI/SNF chromatin-remodeling complex as a key component of the genetic architecture of SZ.

INTRODUCTION

Recent studies support the hypothesis that the high heritability of schizophrenia (SZ) is linked to a combination of relatively common alleles, of small effect, and a few rare alleles with relatively large effects (1,2). Genome-wide association studies (GWAS) of SZ have identified 10 genes (*CSF2RA*, *HIST1H2BJ*, *NOTCH4*, *NRGN*, *SHOX*, *SMARCA2*, *TCF4*, *ZNF804A*, *PRSS16* and *PGBD1*) (3–10) but they confer

*To whom correspondence should be addressed. Email: jean-marie.moalic@inserm.fr
†Y.L.-M. and A.-M.L.-B. contributed equally to the paper.

relatively small increments in risk and explain only a small proportion of heritability (2).

Our working hypothesis attempts to explain this 'missing' heritability. We propose to identify modifier genes that are able to deregulate GWAS-supported SZ-associated genes in the absence of either the common or the rare variants in there GWAS-SZ genes. Second, these modifications in gene expression are sufficient to produce intermediate phenotypes such as modified dendrite or dendritic spine morphology. Third, it is coordinated changes in gene expression that generate these intermediate phenotypes.

A consequence of this hypothesis is that it should be possible to mimic the consequences, induced by a haplotype linked to the disease, on the levels of transcription and/or protein–protein interactions. In the case of the GWAS-supported *SMARCA2* gene, the polymorphisms associated with SZ induce changes in gene expression and in the nuclear localization of the protein (3). Thus, modification of the gene pathways controlling the expression of a given GWAS-supported SZ-associated gene should result in the production of this particular molecular phenotype.

This study is focused on *SMARCA2* on the basis of three specific arguments: (i) its association with SZ has been established both by the identification of structural variants (11) and by GWAS (3); (ii) SMARCA2, as part of the SWI/SNF complex, plays a key role in the epigenetic control of gene expression in SZ (12) and (iii) the components of the SWI/SNF chromatin-remodeling complex, including SMARCA2, have a strict stoichiometry (13,14).

The SWI/SNF complex is a mammalian chromatin-remodeling factor that contains either SMARCA2 (BRM) or SMARCA4 (BRG1) as the ATPase catalytic subunit (15). The SWI/SNF complex forms a super-complex with REST/NRSF and its interactors, SIN3A/B and RCOR1/2, both of which are associated with histone deacetylases HDAC1/2 (16). This super-complex binds to DNA via the REST/NRSF and the MECP2 components. The zinc finger domain of REST/NRSF recognizes a 21-bp conserved neuron-restrictive silencing element (RE1/NRSE) motif located within the control regions of target genes (17,18). MECP2 binds to DNA via a methyl-CpG-binding domain that is 85 amino acids long (19). Overexpression of *Dyrk1a*, which binds the SWI/SNF complex that interacts with REST/NRSF, induces a coordinated deregulation of multiple genes that are responsible for dendritic growth (20). In post-mitotic neurons, it was recently demonstrated that the expression of REST/NRSF led to microRNA-mediated de-repression of BAF53a, a subunit of SWI/SNF, which then led to defective dendritic morphogenesis (21). Importantly, the silencing of gene orthologs which encode SWI/SNF components and interactors induce similar changes in *Drosophila* dendrite phenotypes, indicating an evolutionarily conserved program of dendrite morphogenesis (22).

Here, we show that *REST/NRSF* down-regulation in a mouse neuronal cell line modifies a large gene repertoire that includes the Smarca2 gene. A second GWAS-supported SZ-associated gene, *TCF4*, is also deregulated. Down-regulation of *REST/NRSF* also alters the expression of genes encoding other SWI/SNF components (Smarce1 and Smarcd3) and SWI/SNF interactors (Hdac1, RcoR1 and Mecp2). In primary cultures of mouse cortical neurons, Smarca2 down-regulation generates an abnormal morphology for dendrite spines that is an intermediate phenotype of SZ.

We were able to generate an interacting network including 8 (*CSF2RA*, *HIST1H2BJ*, *NOTCH4*, *NRGN*, *SHOX*, *SMARCA2*, *TCF4* and *ZNF804A*) out of 10 genome-wide supported SZ genes. Five of the genes, including *SMARCA2*, encode transcription regulators. *DYRK1A* overexpression in a transgenic mouse model induces deregulation of TCF4 and *CSF2RA*. Altogether, these results show that the expression of 3 of the 10 genome-wide supported SZ-associated genes is altered when the REST/NRSF-SWI/SNF chromatin-remodeling complex is experimentally manipulated. As expected, these transcriptional modifications, which lead to an intermediate phenotype for dendrite spines, have epigenetic consequences as demonstrated by the chromosomal clustering of the differentially expressed genes.

As shown for *SMARCA2*, a risk allele results from a point mutation that induces a change in an amino acid located in a highly conserved region present in mammalian species (3). The amino acid associated with SZ is the ancestral residue found in the phylogenetic tree extending from the mouse to the chimpanzee. This change has an important functional consequence by altering the protein trafficking between the cytoplasm and the nucleus. This result illustrates the importance of single-amino-acid or protein domains, which may have a primate-accelerated evolution, as potential targets associated with the disease. Using several phylogenetic analyses, we found that SMARCA2 interactors and the GWAS-SZ genes are considerably enriched in genes displaying positive selection in the primate (10/20 and 3/9, respectively) clade and specifically in the human–chimpanzee clade (4/20 and 1/10, respectively). The identification of positive selected genes (PSGs) suggests the occurrence of novel protein interactions in primates.

On the basis of these data, we propose a multiple-step organization for the genetic architecture of SZ, which establishes REST/NRSF-SWI/SNF chromatin-remodeling complex proteins as novel potential therapeutic targets.

RESULTS

REST/NRSF down-regulation modifies Smarca2 gene expression in a mouse neuronal cell line

The *SMARCA2* gene displays copy number variations in SZ (11) and was recently found to be associated with the disease in a genome-wide approach (3). From these results and based upon a complex genetic architecture for SZ (2,23), we hypothesized that polygenic variations can converge to modify *SMARCA2* gene expression and lead to intermediate phenotypes specific for SZ.

To test this hypothesis, we explored the consequences of mouse *REST/NRSF* silencing on *Smarca2* gene expression. REST/NRSF is a molecular platform that links SWI/SNF proteins and their interactors (16,20,24). We used an siRNA strategy to repress the endogenous expression of Rest/Nrsf in a mouse neuroblastoma cell line (N18). As previously reported, this silencing approach induced neurite morphological changes as previously reported (25) (Fig. 1A). *Rest/Nrsf*

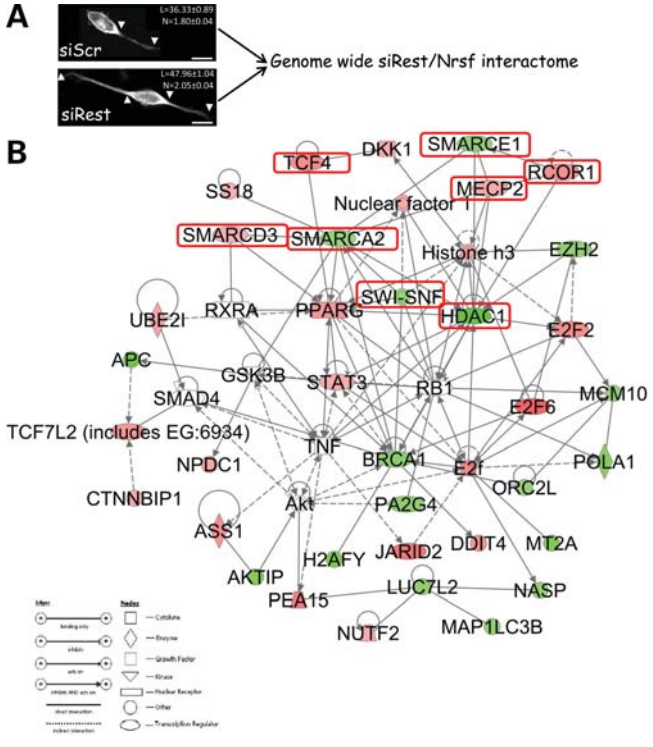


Figure 1. An SWI/SNF centered network deregulated by REST/NRSF silencing. (**A**) Rest/Nrsf silencing induced an increase in neuritic outgrowth as quantified in Lepagnol-Bestel *et al*. (25). Bar = 10 μm. (**B**) We compared the expression of genes between siRest/Nrsf (siRest) and control siScrambled (SiScr) transfected mouse N18 neuroblastoma cells (24 h of transfection). Six hundred and seventy-seven genes were differentially expressed (fold change >1.5 and <-1.5) in response to Rest/Nrsf silencing. Ingenuity functional pathway analysis identified a 44-gene regulatory network which incorporated 37 of the differentially expressed genes. This network is composed of genes which encode SWI/SNF molecules, their interactors, nuclear proteins and transcription factors, including SMARCA2 and TCF4. Six genes (*SMARCA2*, *SMARCE1*, *SMARCD3*, *HDAC1*, *RCOR1* and *MECP2*), highlighted in red boxes, are involved in chromatin-remodeling complexes. Up- and down-regulated genes are labeled in red and green, respectively.

silencing resulted in a decrease in the level of Rest/Nrsf protein in transfected cells (Supplementary Material, Fig. S1). Microarrays containing more than 40 000 mouse DNA targets were screened to identify genes with altered levels of expression in cells transfected with the Rest/Nrsf siRNA (siNrsf) when compared with the scramble (siScr) RNA duplexes. Independent microarray analyses performed on six siNrsf and six siScr cell populations that employed stringent criteria (transcripts up- or down-regulated with mean fold changes >1.5 or <-1.5, respectively, a nominal statistical significance of $P < 0.001$, mean log-expression over 6.0) were used to identify genes for which the levels of expression differed in siNrsf when compared with siScr transfected cells. Six hundred and seventy-seven genes were found to be differentially expressed (Supplementary Material, Table S1; GEO accession numbers no. GSE 14326).

Forty-four genes constituted an interacting network (siRest/Nrsf network) centered on SWI/SNF, including the *SMARCA2* and *TCF4* genes (Fig. 1B). Importantly, this network is centered on a core of 27 genes with a common function (gene expression/transcription) and a *P*-value of 2.09×10^{-15} using Ingenuity Pathways Analysis. This core pathway includes *SMARCA2* and *TCF4* genes. These two genes are members of the 10-gene group associated with SZ by GWAS (1,3–10). A statistical comparison using a Fisher's exact test demonstrated enrichment in GWAS-SZ genes (*P*-value of 1.28×10^{-4}) with 2 genes out of 10 versus 44 out of 27 478 human gene entries in Ensembl (Ensembl release 56 in September 2009).

From the 677 genes differentially regulated by *Rest/Nrsf* silencing, we identified 30 that were part of the SZ database index (http://www.schizophreniaforum.org/res/sczgene/dbindex.asp), which includes 743 genes linked to SZ (26). Furthermore, 4 out of the 677 genes, namely *ACSL4/FACL4* (26), *DLGAP3/SAPAP3* (27,28), *ACCN2/ASIC1A* (29,30) and *CRIPT* (31), were also part of the Gene Ontology dendritic spine repertoire ($n = 200$). Importantly, these four genes encode proteins whose deregulation is sufficient to impact dendritic spine morphogenesis (27–30).

Next, we explored whether *REST/NRSF* silencing induces deregulation of *SMARCA2* in a human context. To this end, we used human (SH-5YSY) neuronal cell line and found that silencing of *REST/NRSF* induced a statistically significant decrease in *SMARCA2* gene expression (Supplementary Material, Fig. S2).

SWI/SNF components and interactors are under the transcriptional control of REST/NRSF

In the siRest/Nrsf interactome, six genes encode Rest/Nrsf interactors (Fig. 2A): *Smarca2* (FC = −1.50), *Smarce1* (FC = −1.57) and *Smarcd3* (FC = +1.51), which are members of the Swi/Snf complex, *Hdac1* (FC = −2.95), an N-terminal Rest/Nrsf repressor domain-interacting gene, *Rcor1* (FC = +2.16), a Rest/Nrsf co-repressor interacting with the C-terminal part of Rest/Nrsf, and *Mecp2* (FC = +1.68), a protein that may interact with both repressor domains of Rest/Nrsf. Quantitative polymerase chain reaction (PCR) measurements of the mRNAs encoding the SWI/SNF components and interactors Smarca2, Smarcd3, Smarce1, Hdac1, RcoR1 and Mecp2, isolated from N18 siNrsf and siScr treated cells, showed that these components are under the transcriptional control of REST/NRSF (Fig. 2B).

SWI/SNF components, including Smarca2 and SWI/SNF interactors, regulate dendrite growth and dendrite spine morphology

Using the same experimental paradigm that allowed us to identify the siRest/Nrsf interactome including SMARCA2, we previously demonstrated that the silencing of Rest/Nrsf modifies neurite outgrowth (25). This abnormal neurite outgrowth can be considered as part of an intermediate SZ phenotype (32).

We hypothesized that changes in the transcription of SWI/SNF components and interactors that are deregulated by *REST/NRSF* down-regulation may result in the production of intermediate phenotypes (such as changes in dendrite spine morphology) found in SZ (33,34). To test this hypothesis, we transfected mouse primary cortical neurons with sh-RNAs specific for the *Smarca2*, *Smarce1*, *Smarcd3*,

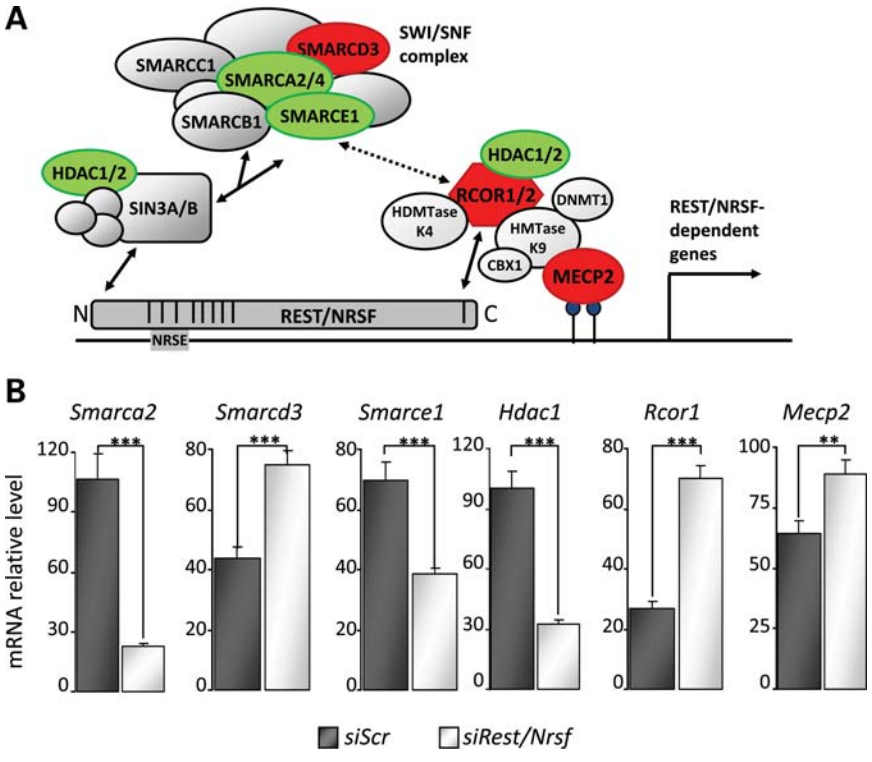


Figure 2. Genes encoding proteins of the SWI/SNF complex and its interactors are transcriptionally deregulated by REST/NRSF silencing. (**A**) Diagram of the REST/NRSF protein and its interacting partners including SWI/SNF complex proteins (adapted from Watanabe *et al*. (16). Genes encoding Hdac1, Smarca2, Smarcd3, Smarce1, Mecp2 and RcoR1 were deregulated (see Fig. 1). The vertical black lines in the REST/NRSF protein represent the zinc fingers which, in the N-terminus, mediate the binding to the 21 bp element (RE1/NRSE) in the control region of target genes. Colored proteins indicated that the corresponding genes were deregulated after siRest/Nrsf (with the same color code as in Fig. 1, i.e. red for up- and green for down-regulated genes). (**B**) RT–PCR quantification of transcript levels in N18 cells transfected with either scramble (siScr) or Rest/Nrsf (siRest) silencing RNA. Quantitative differences in the levels of transcription were noted for all six of the genes which encode proteins that are involved in chromatin-remodeling complexes and which are among the 677 differentially expressed gene repertoire and the Ingenuity regulatory network. $^{**}P < 0.01$; $^{***}P < 0.001$ (*t*-test). The nucleotide sequences of the primers used in these PCR reactions are given in Supplementary Material, Table S4.

Hdac1, *RcoR1* and *Mecp2* genes. Changes in the lengths of dendrites and dendrite spines have been reported in mouse models of SZ (32,35). Similar changes have also been documented in human post-mortem brain tissues from patients with autism and SZ (32,33,36,37). Interestingly, the genes that encode the SWI/SNF complex members are able to modulate neurite outgrowth in *Drosophila* (22). We previously reported that transfection of primary cultures of mouse neurons with sh-RNAs results in a change of ~50% in the expression of the targeted gene (38). We used a similar approach to decrease the expression of *Mecp2*, *Rcor1*, *Hdac1*, *Smarca2*, *Smarcd3* and *Smarce1*. We analyzed different neurite parameters, such as the lengths of the dendrites and the axons and the number of branches of dendrites (Fig. 3A and B). We found that decreases in the expression of five of these six genes induced significant changes in one of the above-mentioned neurite parameters. The length of dendrites (Fig. 3C1) is decreased in sh-Rcor1 cells when compared with control sh-NonSilencing transfected cells (38.55 ± 3.87 versus 55.11 ± 5.19 μm, $P < 0.01$), and the axonal length (Fig. 3C2) is increased in sh-Smarca2 and sh-Smarcd3 cells when compared with control sh-NonSilencing transfected cells (584.25 ± 49.97 and 605.46 ± 62.29 versus 449.28 ± 42.41 μm, $P < 0.01$). The dendritic ramification number (Fig. 3C3) is decreased in sh-Smarce1 and sh-Hdac1 cells when compared with control sh-NonSilencing transfected cells (1.13 ± 0.16 and 1.20 ± 0.16 versus 1.76 ± 0.18, $P < 0.01$). Furthermore, changes in the gene dosage of *Mecp2*, *Smarcd3*, *Smarca2* and *Rcor1* altered the parameters of dendrite spines, such as the length or the number of mushroom spines (Fig. 3D). There was a significant increase in the spine length (Fig. 3E1) in sh-Mecp2 and sh-Smarcd3 cells when compared with control sh-NonSilencing transfected cells (1.68 ± 0.91 and 1.62 ± 0.97 versus 1.47 ± 0.74 μm, $P < 0.0001$) and a significant increase in the number of total spines (Fig. 3E2) in sh-Mecp2, sh-Smarca2 and sh-Rcor1 cells when compared with control sh-NonSilencing transfected cells (83.84 ± 18.38; 99.00 ± 9.16 and 82.5 ± 16.67 versus 57.77 ± 16.54, $P < 0.001$). This effect is largely due to a significant increase in the number of mushroom spines.

A minimal interactome including eight GWAS-supported SZ genes

The siRest/Nrsf interactome includes *SMARCA2* and *TCF4*, two GWAS-supported SZ genes. These two genes can be linked via a single interactor, PPARG (Fig. 1B). This gene encodes a member of the peroxisome proliferator-activated

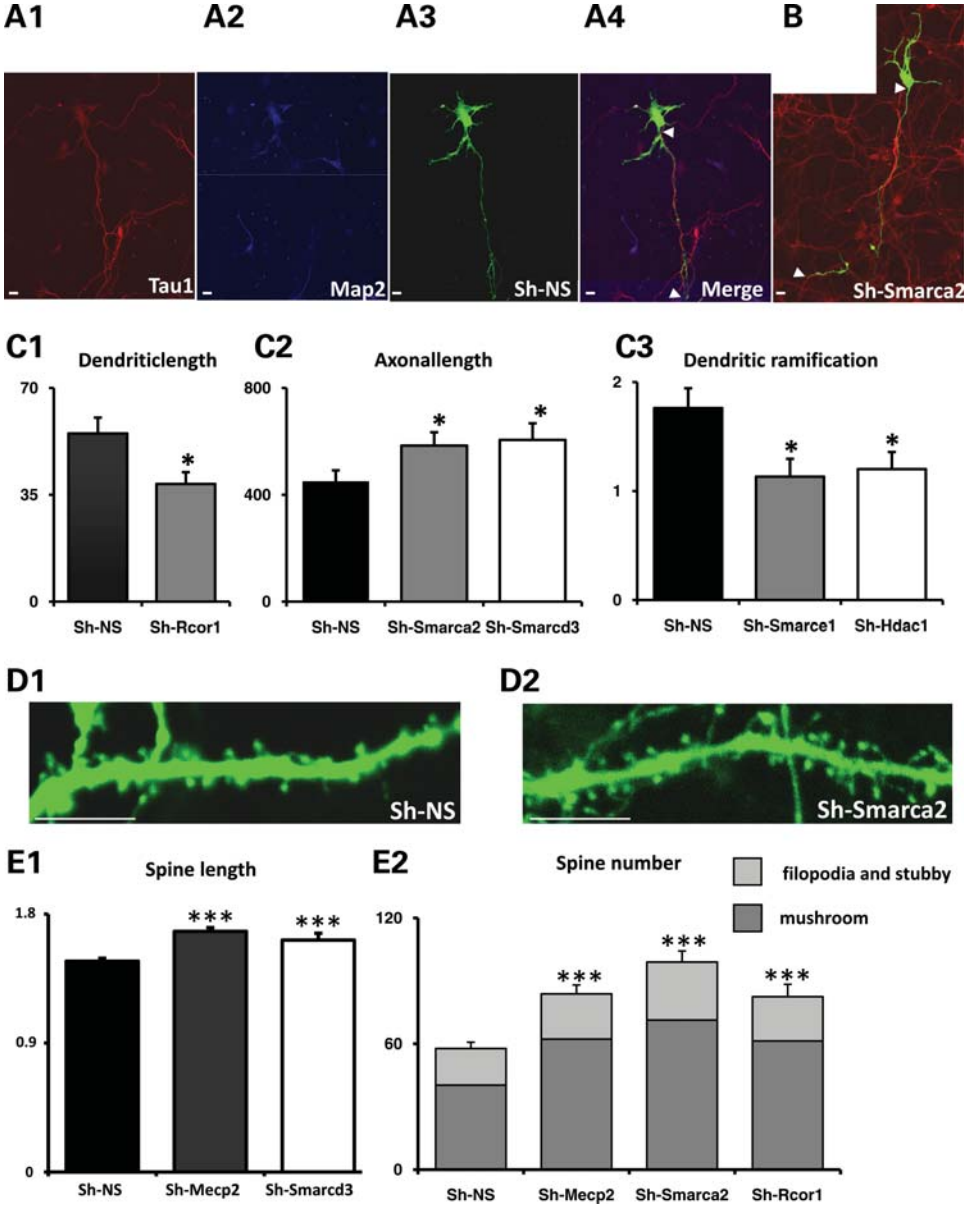


Figure 3. Effects of silencing of genes encoding SWI/SNF components on axon, dendrite and dendritic spine parameters of mouse cortical neurons. (**A**) Cortical neurons were transfected at day *in vitro* 2 (DIV2) with control GFP sh-NonSilencing (A3), fixed 48 h later and then labeled with Tau1 (A1) or Map2 (A2). (A4) The merged image. Bar = 10 μm. (**B**) Cortical neurons were transfected at DIV2 with GFP sh-Smarca2 (green), fixed 48 h later and then labeled with Tau1 (red). Bar = 10 μm. (**C**) Measurement of dendritic and axonal lengths and ramification numbers in neurons transfected with control sh-NonSilencing (Sh-NS) when compared with neurons transfected with sh-Rcor1, sh-Smarca2, sh-Smarcd3, sh-Smarce1 and sh-Hdac1. Note that sh-Rcor1 induces a significantly lower dendritic length (C1), sh-Smarca2 and sh-Smarcd3 induce a significantly higher axonal length (C2) and sh-Smarce1 and sh-Hdac1 induce a significantly lower dendritic ramification number (C3) when compared with the control sh-NonSilencing. $^{*}P < 0.01$. (**D**) Dendritic branches from a control sh-NonSilencing (Sh-NS) transfected cortical neuron (D1) and an sh-Smarca2 transfected cortical neuron (D2). Bar = 10 μm. (E1) Increase in dendritic spine length induced by sh-Mecp2 and sh-Smarcd3 compared with sh-NonSilencing (Sh-NS) transfected cortical neurons. (E2) Increase in the total number of dendritic spines in sh-Mecp2, sh-Smarca2 and sh-Rcor1 transfected cortical neurons when compared with control sh-NonSilencing transfected cortical neurons. Note that this increase is mainly due to a significant increase of mushroom spines with $^{***}P < 0.0001$.

receptor (PPAR) subfamily of nuclear receptors. PPARs form heterodimers with retinoid X receptors (RXRs) and these heterodimers regulate the transcription of several genes (39). We hypothesized that the 10 GWAS-supported SZ genes (gene list and references are in Table 1) form part of the same interactome.

We were able to generate a minimal interactome including 8 (*CSF2RA*, *HIST1H2BJ*, *NOTCH4*, *NRGN*, *SHOX*, *SMARCA2*, *TCF4* and *ZNF804A*) of the 10 GWAS-supported SZ genes linked via 6 interactors (Fig. 4, Table 1). Importantly, this network is centered on a core of five genes with a common function (cell death/apoptosis) and a *P*-value of 1.17×10^{-4} using Ingenuity Pathways Analysis. This core pathway includes CSF2RA, NRGN, SHOX, SMARCA2 and TCF4. Five GWAS-supported SZ genes (*HIST1H2BJ*, *SHOX*, *SMARCA2*, *TCF4* and *ZNF804A*) code for transcription

Table 1. The 10 genome-wide supported schizophrenia-associated genes and their chromosomal locations

Gene symbol	Gene name	Chromosome	References
CSF2RA	Colony-stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	Xp22.32/Yp11.3	4
SHOX	Short stature homeobox	Xp22.33/Yp11.3	4
ZNF804A	Zinc finger protein 804A	2q32.1	5,7,10
SMARCA2	SWI/SNF-related, matrix-associated, actin-dependent regulator or chromatin, subfamily a, member 2	9p22.3	3
HIST1H2BJ	Histone cluster 1, H2bj	MHC region/6p22.1	8,9
PRSS16	Protease, serine, 16 (thymus)	MHC region/6p21	8,9
PGBD1	PiggyBac transposable element derived 1	MHC region/6p22.1	9
NOTCH4	notch homolog 4 (*Drosophila*)	MHC region/6p21.3	6,9
NRGN	Neurogranin (protein kinase C substrate, RC3)	11q24	9
TCF4	Transcription factor 4	18q21.1	9

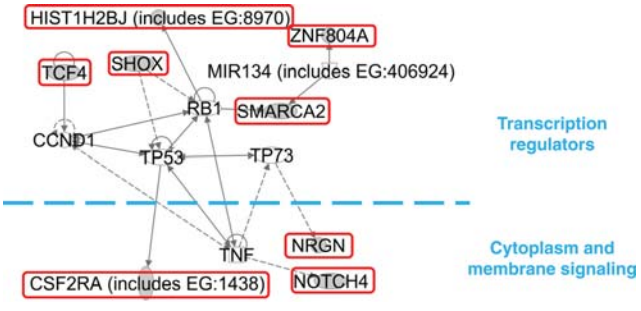


Figure 4. A minimal network including 8 out of 10 GWAS-SZ interactome genes. A minimal network of 14 genes, generated by Ingenuity functional pathway analysis, from the 10 GWAS-supported SZ genes (Table 1). The eight GWAS genes are red-boxed. Out of these eight GWAS genes, five genes encode transcriptional regulators and three genes code for cytoplasmic and membrane signaling proteins.

regulators and three encode proteins involved in cytoplasmic (NRGN, NOTCH4) and membrane (CSF2RA) signaling. Four of the interactors (TP53, TP73, RB1, CCND1) are nuclear proteins involved in chromatin dynamics and/or transcription regulation. In contrast, NOTCH4 is linked to this nuclear network via the tumor necrosis factor cytokine. The last interactor is a microRNA, MIR-134, that is specifically expressed in brain and is involved in dendritic spine regulation (40,41). Two GWAS-supported genes (*PRSS16*, *PGBD1*) were not present in this interactome. No interactors have been characterized for these two proteins either in the Ingenuity or the NCBI databases. *PRSS16* encodes a serine protease. Interestingly, in Alzheimer's disease, *PGBD1* was found to be a GWAS-supported gene (42). PiggyBac transposable element derived 1 (PGBD1) is part of the PGBD protein family that appears to be a novel protein expressed in primates that has no obvious relationship to other transposases or other known protein families (43–45).

From the siRest/Nrsf interactome (Fig.1B) and this minimal GWAS-supported SZ gene interactome (Fig. 4), it is possible to generate a minimal interactome that includes two GWAS-supported SZ genes, *TCF4* (FC = 2.26) and *SMARCA2* (FC = −1.50) that are deregulated by *Rest/Nrsf* silencing. Six other GWAS-supported genes are linked via six interactors including STAT3 and HDAC1, both of which are deregulated by *Rest/Nrsf* silencing (Fig. 5A).

On the basis of these results, we hypothesized that other manipulations of the Rest/Nrsf-SWI/SNF pathway would affect the GWAS-supported SZ gene interactome. We recently reported that DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin-remodeling complex to deregulate gene clusters involved in the production of neuronal dendrite abnormalities (20). Further analysis of that genome-wide microarray showed that both *TCF4* (FC = 1.14) and *CSF2RA* (FC = 1.21) are deregulated. A minimal interactome which includes the DYRK1A data is illustrated in Fig. 5B. This interactome includes the other GWAS-supported SZ genes along with nine interactors, including *NCOR2* which is deregulated by *DYRK1A* overexpression. Furthermore, genome-wide analysis demonstrated that this *in vivo DYRK1A* overexpression deregulated genes organized in statistically significant clusters when compared with a randomly generated repertoire ($P < 10^{-5}$) (Fig. 5C). Importantly, the deregulation of three out of eight GWAS-supported genes in these two genome-wide analyses is statistically significant (3/8 compared with 1065/51 520 for the summation of the two analyses, $P = 4.6 \times 10^{-4}$). Altogether, these results suggest that GWAS-supported SZ genes are part of a functional interactome. Modifications of the REST/NRSF-SWI/SNF-remodeling complex pathway can deregulate this interactome and induce epigenetic modifications as shown by the organization in clusters of the differentially expressed genes.

Primate-specific evolution of proteins involved in the SMARCA2 interactome

In their GWAS study, Koga *et al*. (3) discovered a *SMARCA2* risk allele, E1546, that induces a change of an amino acid (D1546E) which is located in a highly conserved region in mammalian species. At position 1546 of the protein, aspartic acid (D) is mutated to glutamic acid (E) for the alleles associated with SZ. Interestingly, this glutamic acid (E) is found commonly in the mammalian phylogenetic tree, from mouse to chimpanzee, which suggests that the risk allele exhibits the ancestral polymorphism (3). The study also identifies a key functional change resulting from this single amino acid change. The risk allele E1546 causes decreased nuclear localization (3). Altogether, these results illustrate that recent phylogenetic and evolutionary analyses are instrumental in identifying putative disease-associated changes in protein

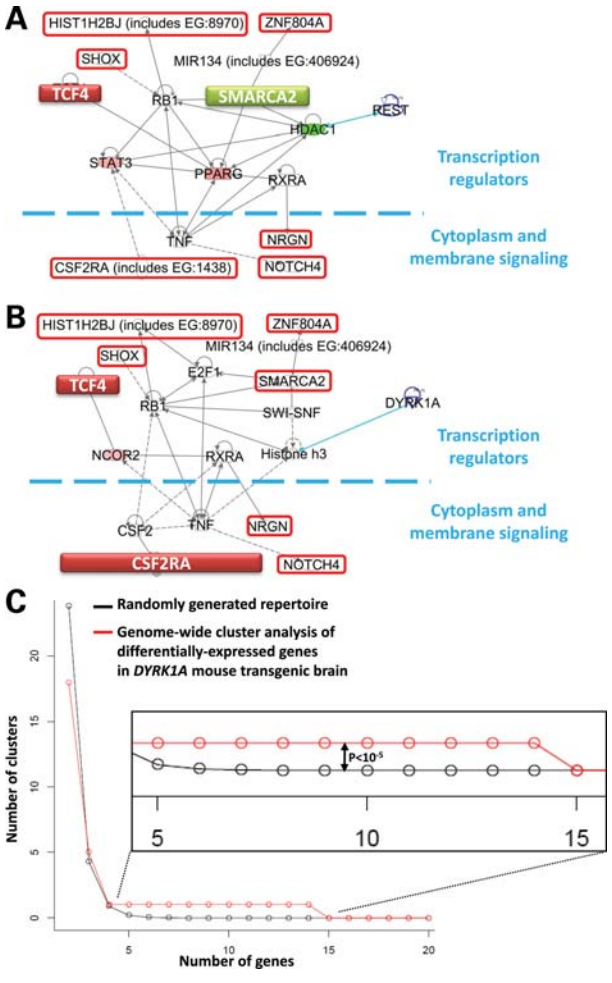


Figure 5. REST/NRSF-SWI/SNF pathway deregulation impacts GWAS-supported SZ gene networks with consequences on the chromosomal clustering of the differentially expressed genes. (**A**) Ingenuity minimal interactome based on GWAS-siRest/Nrsf differential genes and GWAS-SZ genes. A minimal network of 15 genes, generated by Ingenuity functional pathway analysis, using differentially expressed genes ($n = 677$) obtained by transfection of *siRest/Nrsf* in the murine neuroblastoma cell line and the 10 genome-wide supported schizophrenia-associated genes (Table 1). The GWAS genes are red-boxed and the two GWAS differentially expressed genes shown in Fig. 1 are in colored boxes (red for up- and green for down-regulated genes). (**B**) Ingenuity minimal interactome based on GWAS-*DYRK1A* transgenic mouse differential genes and GWAS-SZ genes. A minimal network of 17 genes, generated by Ingenuity functional pathway analysis, using differentially expressed genes ($n = 388$) in the 152F7 mouse model of DYRK1A-REST/NRSF-SWI/SNF deregulation (20) and the 10 genome-wide supported schizophrenia-associated genes (Table 1). The GWAS genes are red-boxed and the two GWAS differentially expressed genes shown in Fig. 1 are in colored boxes (red for up-regulated genes). (**C**) SWI/SNF deregulation and gene clustering in chromosomes. The number of clusters, including differentially expressed genes, as a function of the number of genes for differentially expressed genes in the 152F7 mouse model of DYRK1A-REST/NRSF-SWI/SNF deregulation (red line) when compared with a randomly generated repertoire (black line). We have previously shown that DYRK1A interacts with the SWI/SNF-remodeling complex (20).

sequences and, in particular, changes that result in modifications of their potential interactions and/or in other functional properties such as the sub-cellular locations of the proteins.

Using a genome-wide analysis of siRest/Nrsf differentially expressed genes, we identified a functional interactome centered on SMARCA2 (Fig. 1B). From this interactome and from available data on direct protein interactions (16), we defined a SMARCA2 interactome of 20 genes including REST/NRSF, C- and N-terminus REST/NRSF interactors and SWI–SNF complex components (Fig. 2A). This 20-gene repertoire includes *SMARCA2*, *SMARCD3*, *SMARCE1*, *HDAC1*, *RCOR1* and *MECP2* that are transcriptionally deregulated by *REST/NRSF* silencing (Fig. 2B).

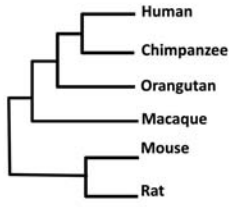
We hypothesized that SMARCA2 and its interactors display changes in amino acid residues or in protein motifs linked to human-specific evolution. We used complementary approaches to analyze the coding sequences of the components of the SMARCA2 interactome using sequences available in a six-species phylogeny (human, chimpanzee, macaque, orangutan, rat and mouse) (Fig. 6A1 and A2, Supplementary Material, Table S2). We first analyzed average Ka/Ks ratios in rodents and in primates. Significant differences in the average Ka/Ks ratios were observed for the SMARCA2 interactome genes (primate: 0.105 ± 0.022; rodents: 0.043 ± 0.014; $P = 0.026$; $n = 20$). These results indicate that the average rate of protein evolution for these genes was faster in primates than in rodents (Supplementary Material, Fig. S3).

We next aimed to characterize the selection type and intensities acting on the 20 genes comprising the SMARCA2 interactome. Specifically, we tested whether the fasted evolutionary rate observed above can be explained by positive selection. We first compared a model in which positive selection is allowed in all taxa against a model in which positive selection is not allowed (site-specific models M8 versus M8a (46)). Only MECP2 and DNMT1 showed a significant pattern of positive selection based on this comparison (Fig. 6A1). We suspected that additional genes may evolve under a positive selection type of selection in specific lineages, and because the M8–M8a comparison assumes the same selection for all branches, we refined this test by applying branch-site-specific models (47–51). Specifically, we compared two models, one that allows for positive selection only in the primate clade versus a model that does not allow for positive selection. As a control, a model that allows only positive selection in the rodent clade was compared against a model that does not allow for positive selection. Ten out of 20 SMARCA2 interactome genes support positive selection in the primate clade according to this branch-site model analysis. In the control test, no genes showed positive selection specific to rodents. These results show that changes in selection pattern has occurred somewhere during the evolution of primates, at least for some of the genes. We next wanted to better determine where in the primate clade this change has occurred. To this end, we compared a model that specifically allows for positive selection only in the chimpanzee–human clade against a model that does not allow for positive selection. We found that *SMARCA4*, *SMARCE1*, *SMARCC1* and *SMARCD3* support positive selection specific to the human–chimpanzee clade (Fig. 6A1). The significantly higher average rate of protein evolution observed above for the primate genes can be either an adaptive evolution of these genes in primates or it may result from functional relaxation in these genes during primate evolution (52). Our analyses argue against the latter hypothesis for the vast majority of the 20 genes comprising the SMARCA2 interactome.

A1 SMARCA2 Interactome

Gene name	Site Model	Branch-Site Models		
	M8-M8a	Rodent	Primate	Human Chimp
MECP2	7.36E-22	1	4.34E-37	1
SMARCA2	0.996	0.996	3.86E-20	1
DNMT1	4.07E-03	1	9.35E-18	1
SMARCA4	1	1	1.38E-12	0
SMARCE1	0.826	1	3.52E-09	0
EHMT2	0.969	1	2.20E-07	0.953
SMARCC1	0.245	0.986	2.40E-04	2.39E-09
HDAC2	0.999	1	8.26E-03	1
SIN3A	1	0.993	2.40E-02	1
RCOR2	0.785	1	3.73E-02	0.613
SMARCD3	1	1	0.053	5.36E-03
REST	0.077	1	0.257	0.921
SMARCB1	1	1	0.469	1
HDAC1	0.352	1	1	1
RCOR1	1	1	1	1
RCOR3	1	0.843	1	1
AOF2	1	1	1	1
CBX1	1	1	1	1
SMARCD2	1	1	1	1
SMARCD1	1	1	1	1

A2 Phylogeny used in SMARCA2 interactome analysis



B GWAS-SZ Interactome

Gene name	Site Model	Branch-Site Models		
	M8-M8a	Rodent	Primate	Human Chimp
CSF2RA	1.98E-03	1	2.83E-02	2.61E-07
HIST1H2BJ	1	NC	1	1
NOTCH4	1	0.755	1	1
NRGN	1	1	1	1
PRSS16	0.174	0.275	1	1
SHOX	1	NC	1	1
SMARCA2	0.996	0.996	3.86E-20	1
ZNF804A	6.91E-03	0.208	1	1
TCF4	8.24E-22	1.91E-04	4.34E-120	1
PGBD1	0.463	NC	NC	0.932
	Present only in Human, Chimpanzee, Macaque and Orangutan			

Figure 6. Likelihood ratio tests (LRTs) used to detect positive selection in the genes of the SMARCA2 interactome and the GWAS-SZ interactome. (**A**) LRTs used to analyze the 20 genes of the SMARCA2 interactome with the *P*-values shown in A1. Panel A2 shows the phylogenic tree used in SMARCA2 interactome analysis. (**B**) *P*-values of the different LRTs used to analyze the 10 genes of the GWAS-SZ interactome. The tree used for the GWAS-SZ interactome analysis differs from that in A2 as follows: CSF2RA (excluding rat and macaque), HIST1H2BJ (including x_tropicalis excluding rat, mouse and macaque), NOTCH4 (excluding macaque), SHOX (including x_tropicalis, gorilla and zebrafish excluding rat, mouse, chimpanzee orangutan and macaque) and PGBD1 (excluding rat and mouse). Positive selected genes (PSGs) are indicated in grey cells.

A recent genome-wide scan for PSGs in mammals reported that 45 out of 14 425 genes analyzed displayed positive selection in primates with 10 out of 14 558 displaying positive selection only in the human lineage (53). Our results indicate that PSGs are considerably enriched in the SMARCA2 interactome (with Fisher's exact test values of $P = 4.85 \times 10^{-20}$ and 2.55×10^{-9} for primate and human PSGs, respectively).

Furthermore, we analyzed whether specific domains of proteins encoded by the SMARCA2 interactome genes display positive selection. We used a site-specific Ka/Ks analysis based on the branch-site codon model that lets the Ka/Ks ratio vary both among sites and among lineages. Using this analysis, we found amino acid domains presenting positive selection for SMARCA2 (Fig. 7A). Koga *et al*. (3) reported a change in the amino acid at position 1546 (D1546E) as being associated with SZ. The risk allele exhibits the ancestral polymorphism (E instead of D). As expected, this codon is detected as undergoing positive selection in our Ka/Ks analysis (circle in Fig. 7A). Other specific positively selected sites were also identified in the inner bromo-domain site of SMARCA2 (Fig. 7A). Similarly, MECP2 displays positively selected codons along the whole protein as well as inside the methyl-CpG-binding sites and in the repression domain (Fig. 7B).

We also focused our analysis on the *REST/NRSF* gene as a modifier gene for SMARCA2. The branch-site codon model analysis shows that the Ka/Ks ratio of the *REST/NRSF* gene was higher in primates (0.398) than in rodents (0.278) (Supplementary Material, Fig. S3A). The Ka/Ks values in primates under this model are shown in black in Fig. 7C. As a control, we also show the Ka/Ks values in a model allowing for positive selection only in the rodent clade (the Ka/Ks values for the rodent clade under this model are shown in grey in Fig. 7C). Notably, differences in selection forces on the *REST/NRSF* gene were also found within primates. The average Ka/Ks ratio was found to be approximately twice as large in the lineage leading to human compared with that in the lineage leading to the chimpanzee (human Ka/Ks = 0.703; chimpanzee Ka/Ks = 0.353; Supplementary Material, Fig. S3B). However, the current available data are insufficient to determine whether this primate-specific alleviated evolutionary rate is the result of relaxation of selection or positive selection. To support the later explanation, a McDonald–Kreitman test (54,55) was used to test for human-specific pattern of positive selection in REST/NRSF. The test examines whether the non-synonymous-to-synonymous ratio for divergence between the human and chimpanzee species (Da/Ds) is larger than the ratio for polymorphisms within the human species (Pa/Ps). The fraction, α, of fixed amino acid substitutions driven by positive selection is then estimated as $1-(\text{Pa/Ps})/(\text{Da/Ds})$ (56). The McDonald–Kreitman test gave a positive value of α (0.25) for REST/NRSF as expected for a Darwinian positive selection. Having shown that selection forces acting on REST/NRSF have changed after the rodent–primate divergence, we next analyzed the specific sites undergoing positive selection. Sites undergoing positive

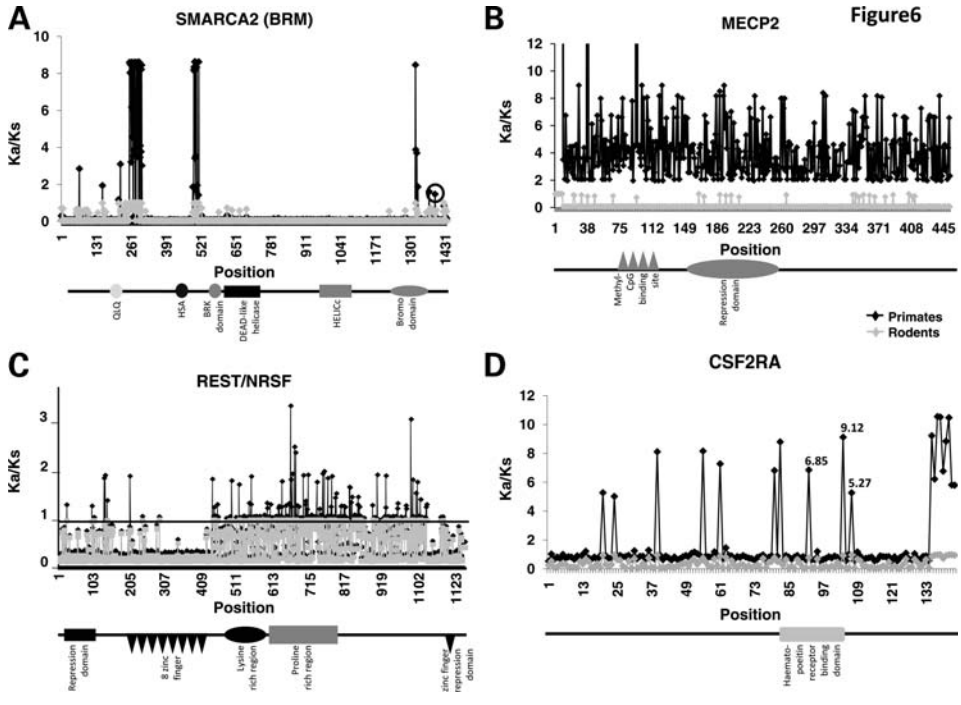


Figure 7. Identification of the protein domains displaying a primate-accelerated evolution using a branch-site model test for four members of the REST/NRSF-SWI/SNF-remodeling complex. Rodent Ka/Ks values are represented by squares in grey and primate Ka/Ks values by dots in black. Peaks above the horizontal line (shown on graph C) indicate Ka/Ks values greater than 1. A schematic representation of the proteins, with their known domains according to Homologene, is shown under the graphs. (**A**) Site-specific Ka/Ks analysis along the *SMARCA2* primate and rodent coding regions. The D amino acid at position 1546 is shown by a black circle. This amino acid is mutated to E (D1546E) in patients with Schizophrenia (3). This variant is detected by the SNP rs2296212 in their GWAS. Note that because of gaps in sequence alignments, in our sequence, the D amino acid is at position 1394 rather than at position 1546. Ka/Ks analysis identifies this site, and others, as undergoing positive selection. (**B**) Site-specific Ka/Ks analysis along the *MECP2* primate and rodent coding regions. (**C**) Site-specific Ka/Ks analysis along the *REST/NRSF* primate and rodent coding regions. (**D**) Site-specific Ka/Ks analysis along the *CSF2RA* primate and rodent coding regions. The Ka/Ks values are shown for the three amino acids located inside the hematopoietin receptor binding domain. This domain starts at amino acid 134 and ends at amino acid 165 in the CSF2RA human protein. Because of gaps in the sequence alignments, this domain is from alignment position 83 to position 107 on the graph shown.

selection are identified using the model that allows for positive selection only in the primate clade. The region from amino acids 443 to 1021 had many sites with Ka/Ks ratios larger than 1. This region includes putative PEST-like sequences. These sequences have been found in proteins such as the transcription factors E1A, c-myc, p53, c-fos and v-myb, which are destined for rapid degradation and which have intracellular half-lives of less than 2 h (57). These motifs are conserved among mouse, rat and human REST orthologous proteins, and the PEST–FIND score (P–F) was calculated according to Bromberg *et al*. (58). The motifs are located between amino acids 690–729 (P–F: 18.3), 668–727 (P–F: 16.4) and 852–869 (P–F: 14.2). This feature strongly suggests that REST, as a transcription factor, may have a rapid turnover and/or its activity may be post-translationally regulated by proteolysis. Interestingly, the major Ka/Ks peak found in the human REST protein (amino acids ∼550 to ∼750) is observed in the region where PEST sequences are found in rodents. Furthermore, we found a low level of alignment between amino acids 595 and 815 among primates and several other species (Supplementary Material, Table S3).

As internal control of our evolutionary studies, we also analyzed BRCA1 which is known to interact with the SWI/SNF complex (59) and which displays bona fide positive selection (60). Using the M8–M8a site model test, we found that positive selection is acting on this gene ($P = 2 \times 10^{-4}$). No positive selection was detected in the rodent clade using the branch-site model test. Positive selection was detected, however, in the primate clade ($P = 0.026$). Sites showing the pattern of positive selection along the *BRCA1* coding regions, based on site-specific Ka/Ks analysis, are shown in Supplementary Material, Fig. S4.

Primate-specific evolution of proteins encoded by GWAS-supported SZ genes

Manipulation of the REST/NRSF-SWI/SNF pathway (Fig. 5) indicated that three GWAS-supported SZ genes can be deregulated, suggesting that these GWAS-supported SZ genes can be part of an interactome (GWAS-SZ interactome).

On the basis of these results, we investigated whether the proteins encoded by GWAS-supported SZ genes, other than *SMARCA2*, present primate-specific evolution. Interestingly, *PGBD1* encodes a protein of the PiggyBac family. PiggyBac proteins are transposases related to the transposase of the canonical PiggyBac transposon. This gene product is primate-specific with expression occurring in the chimpanzee, macaque, orangutan and in humans, with a specific expression in the brain (45,61). University of California, Santa Cruz (UCSC) phylogenetic conservation profiles suggest that

PGBD1 can be a chimeric primate gene obtained by capture of the transposase gene from a mobile element, as for SETMAR (45,61) (Supplementary Material, Fig. S5).

For the 10 GWAS-SZ genes, we analyzed the encoded proteins using the four models previously described in Figure 6A1. Three out of the 10 genes (*CSF2RA*, *ZNF804A* and *TCF4*) support positive selection according to the M8–M8a test. When positive selection is tested specifically in the primate clade, *CSF2RA*, *SMARCA2* and *TCF4* showed significant primate-specific pattern of positive selection. Branch-site model analysis indicated that only CSF2RA displays positive selection in the human–chimpanzee lineage (Fig. 6B). As for the SMARCA2 interactome genes, the GWAS-SZ interactome genes display a considerable enrichment in PSGs (with Fisher's exact test values of $P = 2.86 \times 10^{-6}$ and 7.53×10^{-3} for primate and human PSGs, respectively).

We used the site-specific Ka/Ks analysis based on the branch-site codon model described for SMARCA2, MECP2 and REST/NRSF analysis (Fig. 7A–C) and also found amino acid domains presenting positive selection for CSF2RA (Fig. 7D). Importantly, positive selection was additionally found for three amino acids located inside the hematopoietin receptor binding domain (62) of CSF2RA which is a component of the granulocyte/macrophage colony-stimulating factor receptor complex (63).

DISCUSSION

There is accumulating evidence for the concept that developmental neuropsychiatric diseases as diverse as autism and SZ combine distinct, individually rare, genetic variants (11,64) as well as common haplotypes (3–10). We focused this study on *SMARCA2* because this gene has been identified as being associated with SZ both on the basis of GWAS (3) and CNV studies (11).

As most common variants identified so far in GWAS, such as SMARCA2, confer relatively small increments in risk and explain only a small proportion of familial clustering, the question arises as to the explanation for the remaining 'missing' heritability (2). Here, we examined possible modifier genes of *SMARCA2* as potential sources for this missing heritability.

Our working hypothesis, on the basis of our previous studies (20), was that SMARCA2, as part of the SWI/SNF-remodeling complex, can be impacted by REST/NRSF and by DYRK1A. We have shown that DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin-remodeling complex to deregulate gene clusters involved in the production of an abnormal phenotype for dendrites (20). Furthermore, expression of REST in post-mitotic neurons led to a subunit switch in SWI/SNF-remodeling complexes which induces changes in dendrite outgrowth (21).

Here, we found that silencing of *REST/NRSF* in a neuronal cell line is able to deregulate both *SMARCA2* and *TCF4*, two GWAS-supported genes, and to induce a change in neurite length. Similarly, *in vivo*, *DYRK1A* overexpression deregulates GWAS-supported *TCF4* and *CSF2RA*. This deregulation is associated with abnormalities of neurite outgrowth (20). Importantly, we found that silencing of *REST/NRSF* deregulates five genes encoding SWI/SNF components and interactors. The products of the genes are involved in direct protein–protein interactions (Fig. 2A). This intriguing result indicates that a subset of genes coding for the SWI/SNF complex and its interactors are part of a transcriptional operon.

Altogether the results suggest that GWAS-supported genes such *SMARCA2* can be transcriptionally deregulated by modifier genes and that their deregulation may induce a specific phenotype of neurite outgrowth.

Using Ingenuity interactome analyses, we were able to show that 8 out of the 10 GWAS-supported genes are part of an interacting network. This small group of genes is linked either by protein–protein interactions or by their involvement in the same signaling pathway. This network organization has not been identified previously and it may have key functional consequences. One possible consequence of these interactions could be that the variation in expression of any one of these GWAS-supported genes induces a change in a unique neurobiological pathway which leads to abnormal outgrowth of neurites. Recent studies, such as the one concerning cellular clock mechanisms in a human cell model (65), have shown the feasibility of a genome-wide search for modifiers. This approach, which uses a simple readout based on a reporter gene, is directly transposable for the characterization of genes able to modify the level of expression of the *SMARCA2* gene.

Abnormalities in the length of neurites and in the spines of dendrites as defining an intermediate phenotype was established in SZ, in part, from the analyses of post-mortem brain tissues which showed abnormalities in the number of spines of dendrites that were statistically significant (34,36,66). This phenotypic trait needs to be better characterized in terms of its molecular mechanism. Both cell-autonomous and non-cell-autonomous factors, such as interactions with axonal afferents of other neurons, may induce an abnormal density of dendrite spines. Interestingly, such axonal afferent interactions have been found to be abnormal in recent studies of auditory signal integration in SZ (66). The SMARCA2 ortholog, BRM, induces a similar abnormal phenotype of dendrites in *Drosophila*. BRM silencing results in an abnormal dendrite routing that depends on neurite–neurite interactions. This suggests that, at least in *Drosophila*, the SMARCA2 dendrite phenotype is partly non-cell-autonomous (22). In mouse models of SZ, such as the DISC1 knockin, axons, dendrites and dendrite spines all appear to be affected (Lepagnol-Bestel *et al.*, submitted). We report here that, in mouse cortical neurons, a decrease in the Smarca2 transcript level induces changes in both the length of axons and the number of dendritic spines. Taken together, these results suggest that it is important to fully characterize an intermediate phenotype that can be validated as a bona fide readout for SZ. Genome-wide studies may be designed to identify additional modifiers of such bona fide SZ-linked intermediate phenotypes. In this respect, induced pluripotent stem cells from patients with SZ may be instrumental for fully identifying such intermediate phenotypes *in vitro* (23).

In Fig. 8, we propose a general scheme that illustrates how an intermediate phenotype, such as abnormalities in dendrites and their spines, can be generated by impacting on the SWI/SNF complex (green arrow (1)). The SWI/SNF complex and its inter-

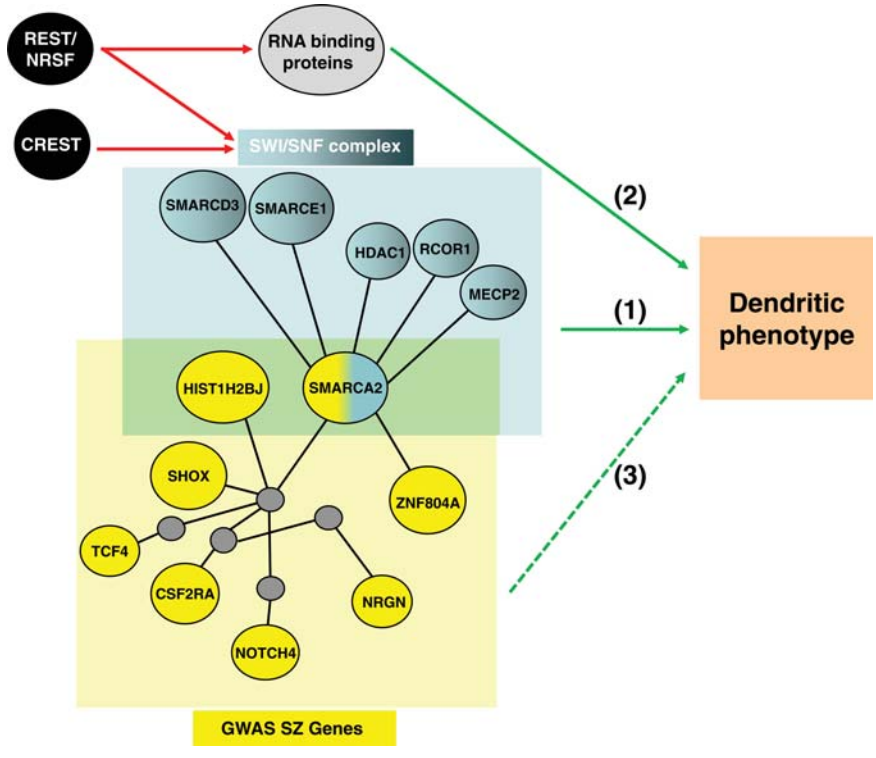


Figure 8. Functional network linking an intermediate phenotype (dendrite and dendrite spine abnormalities) with the SWI/SNF complex and GWAS SZ genes. Protein–protein interactions were established from EntrezGene. Other interactions were validated from our microarray data. Protein–protein interactions are shown by black lines, transactivation or transrepression by red arrows. SWI/SNF complex proteins are colored blue and the GWAS-SZ genes are yellow with two common elements (SMARCA2 and HIST1H2BJ). Intermediate interactors of GWAS-SZ genes defined in Fig. 4 are grey. Changes in expression of either SWI/SNF complex proteins (green arrow (1)), RNA-binding proteins (green arrow (2)) or GWAS-SZ genes (dashed green arrow (3)) may contribute to the same intermediate phenotype affecting dendrites and dendrite spines (green arrows). The hypothesis that all GWAS-supported SZ genes, including SMARCA2, leads to a common intermediate phenotype remains to be demonstrated, as indicated by the dashed green arrow (3).

actors, which include two GWAS-supported SZ genes, *SMARCA2* and *HIST1H2BJ*, can be deregulated not only by REST/NRSF, but also by other transcriptional regulators such as CREST (67,68). Thus, as modifiers of SMARCA2 and HIST1H2BJ, *REST/NRSF* and *CREST* could be involved in the production of the intermediate phenotype that may characterize SZ. Any one of the GWAS-supported SZ genes could also generate the intermediate phenotype (dashed green arrow (3)). Among the pathways deregulated by *REST/NRSF* silencing which are expected to contribute to the intermediate phenotype, the RNA-binding protein compartment is expected to be an important element (green arrow (2)). This compartment is involved in RNA trafficking from the nucleus to both the cytoplasm and to the spines of dendrites (69). We found that this compartment was enriched in our genome-wide study and that some of the proteins that are encoded by these genes have domains with primate-accelerated evolution.

We identified that Rest/Nrsf repression caused expression changes in 677 genes. Interestingly, 30 of which are in the SZ gene database and 4 are in the gene ontology (GO) dendritic spine repertoire. As *Rest/Nrsf* repression caused down-regulation of *Smarca2* and *Mecp2*, respectively, it would be important to define how many genes out of the 30 in the SZ gene database and of the 4 GO dendritic spine repertoire are linked to *Smarca2* and *Mecp2* deregulation. Further studies using silencing approaches of *Rest/Nrsf*, *Smarca2* and *Mecp2* in a similar experimental model will be required to address this important issue.

For SMARCA2, the polymorphism associated with SZ changes a codon that is specific for humans to a codon that is found phylogenetically conserved from mouse to chimpanzee (3). The resulting amino acid change alters the trafficking of SMARCA between the cytoplasm and the nucleus. This human-specific polymorphism, which is found in SMARCA2, raises the question as to positive selection in other proteins involved in SZ-linked pathways. An evolutionary hypothesis related to SZ has already been proposed by several groups (70–73). Indeed, the products of a few genes associated with SZ, *DISC1* (74), *GABRB2* (75) and *KCNH2* (76), have been found recently to display positive selection. In this study, we found that out of 20 SMARCA2 interactome genes, 10 display positive selection in primates and 4 specifically in the human–chimpanzee clade. Similarly, out of the 10 GWAS-SZ interactome genes, 3 display a positive selection in primates and 1 in the human–chimpanzee clade. These results suggest that genes associated with SZ (GWAS-SZ interactome) and genes encoding interactors of SMARCA2 (the SMARCA2 interactome) have a peculiar evolutionary pressure in the primate clade. Possible novel protein interactions resulting from these PSGs need further studies to unravel potential targets linked to SZ.

What are the evolutionary implications of such a high proportion of positively selected genes among the SZ-associated gene GWAS repertoire? This question may be inserted into a more general evolutionary framework integrating the environmental pressures which have operated during human evolution on genes influencing the susceptibility to some common diseases (77) and the complexities of neurodevelopmental disorders (23). A potential link between complex diseases and human adaptation to environmental changes recently emerged from an evolutionary model where genetic variants are proposed to have the potential to modify phenotype variability via epigenetic mechanisms (78).

Intermediate phenotypes such as dendritic spine abnormalities in SZ can result from changes targeting both conserved and positively selected genes. On the basis of studies with *Drosophila* that indicate that SWI/SNF components are involved in dendrite morphogenesis (22), it is clear that fundamental mechanisms for dendritic growth are conserved, phylogenetically, throughout evolution. However, our identification of PSGs and positively selected domains in proteins potentially involved in the morphogenesis of dendrite spines suggests that novel protein–protein interactions may occur in humans.

In this study, we have shown that *Rest/Nrsf* silencing induced down-regulation of *Smarca2* and up-regulation of both *Mecp2* and *Tcf4* genes. Risk alleles corresponding to one linkage disequilibrium block in the *SMARCA2* gene were associated with low *SMARCA2* expression levels in the post-mortem prefrontal cortex (3). Neurodevelopment diseases with mental retardation are known to be linked to MECP2 loss of function and to down-regulation of TCF4. Mutations in the X-linked methyl-CpG-binding protein 2 (MECP2), encoding a transcriptional repressor, cause Rett syndrome, a progressive neuro-developmental disorder and one of the most common causes of mental retardation in females (79). *TCF4* haploinsufficiency causes severe mental retardation with breathing abnormalities (Pitt–Hopkins syndrome) (80). Interestingly, we can expect that the up-regulation of *Mecp2* reported here could induce phenotypic changes as reported in Collins *et al*. (81). Mild overexpression of *MeCP2* in mice induces an initial cognitive phenotype with enhanced motor and contextual learning and enhanced synaptic plasticity in the hippocampus. Similarly, overexpression of Tcf4 induces a modified cognitive phenotype in transgenic mice with profound deficits in contextual and cued fear conditioning and sensorimotor gating (82). Altogether, these results suggest that down-regulation of *Smarca2* and up-regulation of both *Mecp2* and *Tcf4* genes can induce contrast cognitive phenotypes distinct of *Mecp2* knockout and *Tcf4* haploinsufficiency.

Genome-wide association studies have identified and published common variants whose allele frequencies are statistically correlated with various illnesses and traits. However, the vast majority of such variants have no established biological relevance to disease or clinical utility for prognosis or treatment (83). Interestingly, *SMARCA2* is one of the rare examples of genes identified both in GWAS and structural variation studies and for which biological relevance to disease was demonstrated (3). Here, we present additional evidence of such biological relevance of *SMARCA2* deregulation with the quantification of dendritic spine phenotype changes in mouse cortical neurons.

Altogether, our results show the importance of modifiers of GWAS-supported genes in SZ and demonstrate that genome-wide approaches in mouse models may be instrumental for identifying novel therapeutic targets for SZ.

MATERIALS AND METHODS

Cell line cultures and silencing

Mouse N18 and human SH-5YSY neuroblastoma cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum in 100 mm dishes. N18 cells were transfected close to confluence with 13 μg/μl of scrambled siRNA (siScr) or Nrsf siRNA (siNrsf), using Lipofectamine (Invitrogen), as described by the manufacturer. RNA was extracted 24 h after transfection. siRNA sense and antisense sequences (Supplementary Material, Table S4) were previously described (25) and were identical to the target sequences designed for the mouse Nrsf gene by Kim *et al*. (84).

Total RNA preparation

Mouse N18 and human SH-5YSY cells were homogenized in Trizol reagent (Gibco-BRL, Life Technologies), treated with DNase I (Ambion) and processed according to the manufacturer's protocol.

Pangenomic microarrays

Agilent Whole Mouse Genome oligomicroarrays (GEO accession no. GPL2872, Agilent Technologies, Palo Alto, CA, USA) were used. They contain 60-mer DNA probes synthesized *in situ* in a 44k format. Of 44 290 spots, 2756 are controls. The remaining 41 534 spots represent 33 661 unique transcripts which correspond to 20 202 unique human genes. Six independent (NRSF silencing overexpression, six individual samples) measurements were carried out for each group of biological conditions using exchanged dye-labeled RNA targets (i.e. Cy3 and Cy5 dye-swapping experiments). The integrity of the total RNA was determined using an Agilent Technologies 2100 Bioanalyzer and the RNA 6000 Lab-Chip® kit3. The total RNA concentration was determined using a NanoDrop ND-1000 spectrophotometer. For microarray experiments, target preparation, hybridization and washing were done according to the manufacturer's instructions. Briefly, 1 μg aliquots of total RNA were labeled using the Agilent Linear Amplification/Labeling kit (Agilent Technologies). After checking the labeling efficiency and the product integrity, 750ng Cy3- and 750ng Cy5-labeled targets were mixed and incubated on an Agilent microarray slide for 17 h at 65°C, in a rotating oven, using an Agilent *in situ* hybridization kit. The slides were washed and then any traces of water were removed by centrifugation at 300 g for 1 min. Slides were scanned using a GenePix 4000B Microarray Scanner (Molecular Devices, Sunnyvale, CA, USA) at 5 μm resolution. The photo multiplier tube levels of the two channels at 635 nm (for Cy5) and 532 nm (for Cy3) were

balanced to limit the number of saturated pixels (<0.005%) for generating gray-scale 16-bit TIFF image files. The resulting images were processed using the GenePix Pro (v6.0.1.26) image analysis software. This included defining the spots, measuring the intensities, flagging the spots having inadequate quality control parameters and evaluating local background. The pre-processing of the data and its quality assessment were done using the MAnGO tools suit, an R script that allows integrated analysis of two-color microarrays (85). The background level was calculated using morphological operators (a short closing followed by a large opening) and subtracted. Raw data were normalized using the print-tip lowess method (86,87), i.e. local regression normalization within an artificial print-tip block.

To adequately identify the differentially expressed genes, including those with low variations in expression, both an analysis of variance (ANOVA) using the Genesis Client system (88) and a *P*-value ranking strategy using *z*-statistics in the ArrayStat 1.0 software (Imaging Research, Inc.) were applied for statistical analysis. In order to adjust the *P*-values derived from each test, a false discovery rate (FDR) procedure was used (89). Statistical tests were computed and a gene's expression was declared to be significantly different between the conditions being tested when the adjusted *P*-value was <1%. A filtering procedure excluded data points as being unreliable if they corresponded to probe sets associated with low signal intensities (Amean < 6.0), or if they displayed a limited magnitude of change ($-0.4 < M < 0.4$).

For large-scale multiple testing situations, such as with microarray experiments, the control of the FDR, the proportion of false-positives among all significant tests, is crucial. In our report, the false-positive rate, due to the number of statistical tests performed, was controlled by multiple testing corrections, a *P*-value adjustment converting *P*-value error rates on individual genes into error rates on the entire list of genes. The adjustment method used on our data set is the Benjamini and Hochberg (90) step-up procedure, which was implemented to adjust the raw *P*-values. The correction of *P*-values allowed the determination of the gene sets that are to be selected with an FDR of 0.01. Gene sets which did not meet the criterion were not selected.

Regulatory network exploration

Networks of genes were defined according to Calvano *et al*. (91) using Ingenuity pathways analysis (http://www.ingenuity.com). Precisely, selected data sets were uploaded as tab-delimited txt files containing the Agilent Probe ID for 'focus genes'; each probe identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. The program queries the database for interactions between focus genes and all other gene objects stored in the knowledge base and generates a set of networks. The expression levels are not used to weight the genes; the program computes a score for each network according to the fit of the network to the set of focus genes. The score indicates the likelihood of the focus genes being found in a given network due to random chance. Biological networks of interacting genes and other functional groups were built on this assumption. A score of >2 indicates that there is a <1 in 100 chances that the focus genes were assembled randomly into a network due to a random chance.

Gene ontology terms enrichment analysis

Genes were categorized into various ontologies using the Unigene cluster IDs (Build #163, June 7) via the eGOn (explore Gene Ontology), a tool for mapping microarray data onto the Gene Ontology structure (92). Precisely, the gene identifiers were submitted to the server, and eGOn performed a search over the publicly available gene databases to retrieve GO terms that have been annotated to these genes. The information comes from all the three top-level branches in the ontology: molecular function, biological process and cellular component.

A Master-Target test was performed based on the assumption that under the null hypothesis the genes on the lists act independently within each list. If a particular ontology term was found enriched for modulated genes, it was concluded that the ontology term is likely to be involved in the process. A two-sided jack-knife Fisher's exact probability was used to calculate the *P*-value. Adjusted *P*-values were calculated using the step-up procedure of Benjamini and Hochberg (90); only GO categories that had adjusted *P*-values of <0.05 were reported.

Real-time quantitative RT–PCR analysis

Reverse transcription (RT) was carried out at 37°C for 50 min in a 20 μl RT mixture containing 1 μg of total RNA, 200 U of M-MLV reverse transcriptase (Invitrogen), 0.5 mM of each dNTP and 6.25 μM random hexamers (Invitrogen). The cDNAs generated were amplified by real-time PCR, using probes (obtained from PROLIGO Genset) labeled at their 5′ ends with a fluorogenic reporter dye and at their 3′ ends with a quencher dye (TAMRA). Sequences of primers and probes are available on request. PCR assays had a final reaction volume of 20 μl, and contained 2 U of *Taq* polymerase (Applied Biosystems), 10 μM primers and fluorogenic probe. PCR was carried out over 40 cycles of 95°C for 15 s, 60°C for 1 min, 50°C for 1 min. We used the Opticon2 sequence detection system (MJ Research/Biorad), with the Opticon Monitor software for data analysis. For each group, the cDNAs synthesized from total RNA were serially diluted to cover the 0.19–50 ng range for specific mRNAs and 0.98–250 pg range for 18S rRNAs. These serial dilutions were used to construct standard curves for 18S and for each gene of interest and to calculate the amounts of RNA for 18S and the genes of interest corresponding to the PCR products generated from individual cDNAs for each experimental group. Each Q-PCR signal was normalized with respect to 18S, as previously described (25).

sh-RNA vectors

To silence mouse *Mecp2, Rcor1, Hdac1, Smarca2/Brm, Smarcd3/Baf60c* and *Smarce1/Baf57*, we used sh-RNA p-GIPZ vectors from the OPENBIOSYSTEMS clone library, and control Non-Silencing sh-RNA as well (scrambled):

Gene name	Clone/IMAGE ID	Item number
Mecp2	V2LMM_10472	RMM4431-99337103
RcoR1	V2LMM_204999	RMM4431-98696647
Hdac1	V2LMM_7452	RMM4431-99435591
Smarca2 (Brm)	V2LMM_78811	RMM4431-98975859
Smarcd3 (Baf60c)	V2LMM_10434	RMM4431-99211408
Smarce1 (Baf57)	V2LMM_54434	RMM4431-99435803

Primary cell culture, transfection and neurite analysis

E17.5 mouse cortical neurons were enzymatically dissociated (0.25% trypsin, DNase), mechanically triturated with a flamed Pasteur pipette and plated on 35 mm dishes (8×10^5 cells per dish) coated with poly-DL-ornithine (Sigma), in DMEM (Invitrogen) supplemented with 10% fetal bovine serum. Four hours after plating, DMEM was replaced with Neurobasal® medium (Invitrogen) supplemented with 2 mM glutamine and 2% B27 (Invitrogen). For the analyses of dendrites and axons, primary neuronal cultures were transfected after 2 days in culture and analyzed 48 h later. To analyze the spines of dendrites, neurons were transfected after 7 days in culture and analyzed on day 21. The spines of dendrites were identified as filopodia, stubby or mushroom based on the classification of Bourne and Harris (93). The dendrite spine length consists in the measurement of the neck of the mushroom spines present on 75 μm dendrite fragment samples, cells were transfected with the sh-RNA described above using Lipofectamine and Plus-Reagent (Invitrogen), as described by the manufacturer. The analysis of neurites and spines was carried out with the ImageJ software (Wayne Rasband, NIH).

Immunochemistry

Cells were fixed by incubation at room temperature for 20 min, in 4% paraformaldehyde in phosphate-buffered saline (PBS). They were permeabilized for 10 min in 0.3% Triton in PBS and blocked for 30 min with 3% bovine serum albumin and 0.1% Triton in PBS. The cells were then incubated overnight at 4°C with primary antibodies (chicken polyclonal anti-MAP2, 1/10 000, Abcam and mouse monoclonal anti-Tau1, 1/500, Chemicon) in the blocking solution. Cells were washed three times with PBS and incubated for 1 h at room temperature with the secondary antibodies (Alexa-555 anti-mouse IgG or Alexa-647 anti-chicken IgG, 1/1000, Invitrogen) in the blocking solution. Finally, the cells were washed three times in PBS and mounted with Prolong Gold Mounting Medium (Invitrogen). Immunofluorescence was observed with a TCS_4D confocal imaging system (Leica Instruments), using standard filter sets.

Statistical analysis

All gene repertoires were evaluated with the Shapiro-Wilk test to estimate the normality of the different groups. According to the normality test result, Ka/Ks mean ratios from paired samples were compared either by a paired *t*-test or by a Wilcoxon test, whereas the Ka/Ks mean ratios from independent groups were compared either by an unpaired *t*-test or by a Mann–Whitney test. The averages of the relative levels of mRNA expression were compared by ANOVA and *t*-tests. All statistical tests were performed with SPSS software.

Sequence acquisition and analysis

Sequences for the human (*Homo sapiens*) genes were obtained from NCBI molecular databases (October 30, 2006). Rodent [rat (*Rattus norvegicus*) and mouse (*Mus musculus*)] orthologous genes were determined from RefSeq database (October 2006) using the reciprocal best BLAST hits (BLASTP, 2.2.13 version (94)). Primate [macaque (*Macaca mulatta*), orangutan (*Pongo pygmaeus*) and chimpanzee (*Pan troglodytes*)] orthologous genes were determined from UCSC data available on April 2008. For orthologs not found with this procedure, a reciprocal BLAST search was performed from ENSEMBL genes (ENSEMBL release 41, October 2006). All the orthologous genes were further validated using HomoloGene database (HomoloGene release 62) and synteny characteristics. Orthologous gene coding sequences were aligned in frame using the MUSCLE V3.6 program (95).

Ka/ks analysis

Orthologous sequences from human, chimpanzee, macaque, orangutan, rat and mouse were analyzed to infer the Ka/Ks ratios. The codeml program from the PAML package was used (96). The Ka/Ks ratios in the rodent clade and in the primate clade were computed using the branch-codon model (47,48). Specifically, given a multiple sequence alignment and a phylogeny, the following parameters were estimated using the maximum-likelihood paradigm: the tree branch lengths, the transition/transversion bias, the codon frequencies and three different Ka/Ks parameters (denoted by ω according to Yang and Nielsen (48)), one for the rodent clade, one for the primate clade and one for the branch connecting these two clades. The comparison of Ka/Ks ratios between different species in the same clade (i.e. the primates) was achieved similarly by assuming an independent ω ratio for each branch in the clade.

The above test determines whether there are selection differences between two clades. It does not explicitly test for positive selection. Thus, we compared the M8 model (allowing for positive selection at some sites) versus the M8a model (in which positive selection is not allowed) using the likelihood ratio test (38).

In order to test for adaptive evolution in specific lineages (e.g. within the primate) for the REST/NRSF gene and its direct interactors, we used the branch-site codon model (49). This model allows testing for positive selection in a specific clade, accounting for the variability of selection forces among sites. Specifically, two distributions over the ω parameter were assumed: one for the forward branches (the primate clade) and one for the backward branches (the remaining branches, which include the rodent clade). The ω distribution allows for positive selection in the forward branches only. Computations were repeated when the rodent clade was considered as forward branches, thus allowing the inference of positive selection in the rodent clade. Finally, a likelihood ratio test was performed comparing a model which

allows for positive selection in the forward branches against a model that does not. As this test is applied to all putative interactors, a correction for multiple testing was used: the false discovery approach suggested by Benjamini and Hochberg (90). This test was repeated to test for positive selection specific for the human–chimpanzee clade. These branch-site-specific tests also provide site-specific Ka/Ks estimates.

McDonald–Kreitman analysis

We used the Ka/Ks ratios for single-nucleotide polymorphisms (SNPs) and divergences to estimate the proportion of amino acid substitutions driven by positive selection according to Smith and Eyre-Walker (55). We estimated the fraction of fixed amino acid substitution as $\alpha = 1 - (\mathrm{Pa}/\mathrm{Ps})/(\mathrm{Da}/\mathrm{Ds})$, where Pa is the number of non-synonymous human polymorphisms, Ps the number of synonymous human polymorphisms, Da the number of non-synonymous substitutions and Ds the number of synonymous substitutions. Human polymorphisms (P) were obtained from NCBI dbSNP. Only HapMap validated SNPs displaying a frequency >0.05 (according to Ramensky *et al*. (56)) were included in the calculation. Divergences (D) were taken from our human–chimpanzee orthologous sequence comparisons.

Cluster estimation

For the two microarray experiments corresponding to GEO accession numbers GSE 14326 and GEO accession numbers GSE 14021, GSE 14030, super-series GSE 14105 and GSE 14072, we selected differentially expressed genes as indicated. The positions of transcription start sites were downloaded from UCSC on mm9 assembly. If genes had multiple Transcription start site (TSS), we took the mean of these positions. Genes were grouped together if their start sites were located in a 1 Mb window. If two clusters shared at least one gene, the two clusters were merged. To estimate statistical significance, we generated random gene sets. These gene sets were constructed from the genes expressed in the relevant microarray (based on the threshold for expression: Amean ≥ 6 for siRest/Nrsf, Amean ≥ 5.5 for 152F7 trangenic mice). We selected, as expressed, as many genes as there were differentially expressed genes in the experiment. We generated 10 000 random gene sets and compared their characteristics with the observed differentially expressed gene sets. For each criterion (the number of clustered genes, the number and length of clusters), we inferred statistical significance by comparison with the fraction of random genes sets which satisfied the criteria.

PEST motive analysis

PEST motives were analyzed using PESTfind Analyis webtool (EMBnet).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

A.D.-F. is a post-doctorate fellow of the Edmond J. Safra Bioinformatics Program at Tel-Aviv University. We thank Dr Nicole Faucon-Biguet for providing human SH-5YSY line and Dr Richard Violette for comments on the manuscript.

Conflict of Interest statement. None declared.

FUNDING

This work was partly supported by INSERM, the CNRS, the ANR06-neuro FRAXAmRNP, the *Université Paris Sud*, Association Française du Syndrome de RETT, Foundation Orange and the Jérôme Lejeune Foundation (to M.S.). A.-M.L.-B. and G.M. were supported by ANR and Orange Foundation, respectively. T.P. was supported by the Israel Science Foundation (878/09).

REFERENCES

1. Owen, M.J., Williams, H.J. and O'Donovan, M.C. (2009) Schizophrenia genetics: advancing on two fronts. *Curr. Opin. Genet. Dev.*, **19**, 266–270.
2. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A. *et al.* (2009) Finding the missing heritability of complex diseases. *Nature*, **461**, 747–753.
3. Koga, M., Ishiguro, H., Yazaki, S., Horiuchi, Y., Arai, M., Niizato, K., Iritani, S., Itokawa, M., Inada, T., Iwata, N. *et al.* (2009) Involvement of SMARCA2/BRM in the SWI/SNF chromatin-remodeling complex in schizophrenia. *Hum. Mol. Genet.*, **18**, 2483–2494.
4. Lencz, T., Morgan, T.V., Athanasiou, M., Dain, B., Reed, C.R., Kane, J.M., Kucherlapati, R. and Malhotra, A.K. (2007) Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol. Psychiatry*, **12**, 572–580.
5. O'Donovan, M.C., Craddock, N., Norton, N., Williams, H., Peirce, T., Moskvina, V., Nikolov, I., Hamshere, M., Carroll, L., Georgieva, L. *et al.* (2008) Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat. Genet.*, **40**, 1053–1055.
6. Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F. and Sklar, P. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, **460**, 748–752.
7. Riley, B., Thiselton, D., Maher, B.S., Bigdeli, T., Wormley, B., McMichael, G.O., Fanous, A.H., Vladimirov, V., O'Neill, F.A., Walsh, D. and Kendler, K.S. (2010) Replication of association between schizophrenia and ZNF804A in the Irish Case–Control Study of Schizophrenia sample. *Mol. Psychiatry*, **15**, 29–37.
8. Shi, J., Levinson, D.F., Duan, J., Sanders, A.R., Zheng, Y., Pe'er, I., Dudbridge, F., Holmans, P.A., Whittemore, A.S., Mowry, B.J. *et al.* (2009) Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*, **460**, 753–757.
9. Stefansson, H., Ophoff, R.A., Steinberg, S., Andreassen, O.A., Cichon, S., Rujescu, D., Werge, T., Pietiläinen, O.P., Mors, O., Mortensen, P.B. *et al.* (2009) Common variants conferring risk of schizophrenia. *Nature*, **460**, 744–747.
10. Steinberg, S., Mors, O., Børglum, A.D., Gustafsson, O., Werge, T., Mortensen, P.B., Andreassen, O.A., Sigurdsson, E., Thorgeirsson, T.E., Böttcher, Y. *et al.* (2010) Expanding the range of ZNF804A variants conferring risk of psychosis. *Mol. Psychiatry*, in press.
11. Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A. *et al.* (2008) Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*, **320**, 539–543.
12. Tsankova, N., Renthal, W., Kumar, A. and Nestler, E.J. (2007) Epigenetic regulation in psychiatric disorders. *Nat. Rev. Neurosci.*, **8**, 355–367.

13. Chen, J. and Archer, T.K. (2005) Regulating SWI/SNF subunit levels via protein–protein interactions and proteasomal degradation: BAF155 and BAF170 limit expression of BAF57. *Mol. Cell. Biol.*, **25**, 9016–9027.
14. Harikrishnan, K.N., Chow, M.Z., Baker, E.K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J.M., Jones, P.L., Sif, S. and El-Osta, A. (2005) Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat. Genet.*, **37**, 254–264.
15. Roberts, C.W. and Orkin, S.H. (2004) The SWI/SNF complex—chromatin and cancer. *Nat. Rev. Cancer*, **4**, 133–142.
16. Watanabe, H., Mizutani, T., Haraguchi, T., Yamamichi, N., Minoguchi, S., Yamamichi-Nishina, M., Mori, N., Kameda, T., Sugiyama, T. and Iba, H. (2006) SWI/SNF complex is essential for NRSF-mediated suppression of neuronal genes in human nonsmall cell lung carcinoma cell lines. *Oncogene*, **25**, 470–479.
17. Bruce, A.W., Donaldson, I.J., Wood, I.C., Yerbury, S.A., Sadowski, M.I., Chapman, M., Göttgens, B. and Buckley, N.J. (2004) Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. *Proc. Natl Acad. Sci. USA*, **101**, 10458–10463.
18. Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of *in vivo* protein–DNA interactions. *Science*, **316**, 1497–1502.
19. Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905–914.
20. Lepagnol-Bestel, A.M., Zvara, A., Maussion, G., Quignon, F., Ngimbous, B., Ramoz, N., Imbeaud, S., Loe-Mie, Y., Benihoud, K., Agier, N. *et al.* (2009) DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. *Hum. Mol. Genet.*, **18**, 1405–1414.
21. Yoo, A.S., Staahl, B.T., Chen, L. and Crabtree, G.R. (2009) MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature*, **460**, 642–646.
22. Parrish, J.Z., Kim, M.D., Jan, L.Y. and Jan, Y.N. (2006) Genome-wide analyses identify transcription factors required for proper morphogenesis of *Drosophila* sensory neuron dendrites. *Genes Dev.*, **20**, 820–835.
23. Sebat, J., Levy, D.L. and McCarthy, S.E. (2009) Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet.*, **25**, 528–535.
24. Reisman, D., Glaros, S. and Thompson, E.A. (2009) The SWI/SNF complex and cancer. *Oncogene*, **28**, 1653–1668.
25. Lepagnol-Bestel, A.M., Maussion, G., Ramoz, N., Moalic, J.M., Gorwood, P. and Simonneau, M. (2007) Nrsf silencing induces molecular and subcellular changes linked to neuronal plasticity. *Neuroreport*, **18**, 441–446.
26. Allen, N.C., Bagade, S., McQueen, M.B., Ioannidis, J.P., Kavvoura, F.K., Khoury, M.J., Tanzi, R.E. and Bertram, L. (2008) Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat. Genet.*, **40**, 827–834.
27. Meloni, I., Parri, V., De Filippis, R., Ariani, F., Artuso, R., Bruttini, M., Katzaki, E., Longo, I., Mari, F., Bellan, C., Dotti, C.G. and Renieri, A. (2009) The *XLMR* gene ACSL4 plays a role in dendritic spine architecture. *Neuroscience*, **159**, 657–669.
28. Kindler, S., Rehbein, M., Classen, B., Richter, D. and Bockers, T.M. (2004) Distinct spatiotemporal expression of SAPAP transcripts in the developing rat brain: a novel dendritically localized mRNA. *Brain Res. Mol. Brain Res.*, **126**, 14–21.
29. Schutt, J., Falley, K., Richter, D., Kreienkamp, H.J. and Kindler, S. (2009) Fragile X mental retardation protein regulates the levels of scaffold proteins and glutamate receptors in postsynaptic densities. *J. Biol. Chem.*, **284**, 25479–25487.
30. Zha, X.M., Wemmie, J.A., Green, S.H. and Welsh, M.J. (2006) Acid-sensing ion channel 1a is a postsynaptic proton receptor that affects the density of dendritic spines. *Proc. Natl Acad. Sci. USA*, **103**, 16556–16561.
31. Passafaro, M., Sala, C., Niethammer, M. and Sheng, M. (1999) Microtubule binding by CRIPT and its potential role in the synaptic clustering of PSD-95. *Nat. Neurosci.*, **2**, 1063–1069.
32. Bellon, A. (2007) New genes associated with schizophrenia in neurite formation: a review of cell culture experiments. *Mol. Psychiatry*, **12**, 620–629.
33. Hill, J.J., Hashimoto, T. and Lewis, D.A. (2006) Molecular mechanisms contributing to dendritic spine alterations in the prefrontal cortex of subjects with schizophrenia. *Mol. Psychiatry*, **11**, 557–566.
34. Kolluri, N., Sun, Z., Sampson, A.R. and Lewis, D.A. (2005) Lamina-specific reductions in dendritic spine density in the prefrontal cortex of subjects with schizophrenia. *Am. J. Psychiatry*, **162**, 1200–1202.
35. Kvajo, M., McKellar, H., Arguello, P.A., Drew, L.J., Moore, H., MacDermott, A.B., Karayiorgou, M. and Gogos, J.A. (2008) A mutation in mouse Disc1 that models a schizophrenia risk allele leads to specific alterations in neuronal architecture and cognition. *Proc. Natl Acad. Sci. USA*, **105**, 7076–7081.
36. Lewis, D.A. and Gonzalez-Burgos, G. (2008) Neuroplasticity of neocortical circuits in schizophrenia. *Neuropsychopharmacology*, **33**, 141–165.
37. Raymond, G.V., Bauman, M.L. and Kemper, T.L. (1996) Hippocampus in autism: a Golgi analysis. *Acta Neuropathol.*, **91**, 117–119.
38. Maussion, G., Carayol, J., Lepagnol-Bestel, A.M., Tores, F., Loe-Mie, Y., Milbreta, U., Rousseau, F., Fontaine, K., Renaud, J., Moalic, J.M. *et al.* (2008) Convergent evidence identifying MAP/microtubule affinity-regulating kinase 1 (MARK1) as a susceptibility gene for autism. *Hum. Mol. Genet.*, **17**, 2541–2551.
39. Tan, N.S., Michalik, L., Desvergne, B. and Wahli, W. (2005) Multiple expression control mechanisms of peroxisome proliferator-activated receptors and their target genes. *J. Steroid Biochem. Mol. Biol.*, **93**, 99–105.
40. Schratt, G.M., Tuebing, F., Nigh, E.A., Kane, C.G., Sabatini, M.E., Kiebler, M. and Greenberg, M.E. (2006) A brain-specific microRNA regulates dendritic spine development. *Nature*, **439**, 283–289.
41. Fiore, R., Khudayberdiev, S., Christensen, M., Siegel, G., Flavell, S.W., Kim, T.K., Greenberg, M.E. and Schratt, G. (2009) Mef2-mediated transcription of the miR379–410 cluster regulates activity-dependent dendritogenesis by fine-tuning Pumilio2 protein levels. *EMBO J.*, **28**, 697–710.
42. Bertram, L. and Tanzi, R.E. (2009) Genome-wide association studies in Alzheimer's disease. *Hum. Mol. Genet.*, **18**, R137–R145.
43. Sarkar, A., Sim, C., Hong, Y.S., Hogan, J.R., Fraser, M.J., Robertson, H.M. and Collins, F.H. (2003) Molecular evolutionary analysis of the widespread PiggyBac transposon family and related 'domesticated' sequences. *Mol. Genet. Genomics*, **270**, 173–180.
44. Feschotte, C. and Pritham, E.J. (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.*, **41**, 331–368.
45. Cordaux, R. and Batzer, M.A. (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.*, **10**, 691–703.
46. Swanson, W.J., Nielsen, R. and Yang, Q. (2003) Pervasive adaptive evolution in mammalian fertilization proteins. *Mol. Biol. Evol.*, **20**, 18–20.
47. Yang, Z. (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol. Biol. Evol.*, **15**, 568–573.
48. Yang, Z. and Nielsen, R. (1998) Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J. Mol. Evol.*, **46**, 409–418.
49. Yang, Z. and Nielsen, R. (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.*, **19**, 908–917.
50. Yang, Z., Wong, W.S. and Nielsen, R. (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.*, **22**, 1107–1118.
51. Zhang, J., Nielsen, R. and Yang, Z. (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.*, **22**, 2472–2479.
52. Dorus, S., Vallender, E.J., Evans, P.D., Anderson, J.R., Gilbert, S.L., Mahowald, M., Wyckoff, G.J., Malcom, C.M. and Lahn, B.T. (2004) Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. *Cell*, **119**, 1027–1040.
53. Kosiol, C., Vinar, T., da Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., Nielsen, R. and Siepel, A. (2008) Patterns of positive selection in six mammalian genomes. *PLoS Genet.*, **4**, e1000144.
54. McDonald, J.H. and Kreitman, M. (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.
55. Smith, N.G. and Eyre-Walker, A. (2002) Adaptive protein evolution in *Drosophila*. *Nature*, **415**, 1022–1024.
56. Ramensky, V.E., Nurtdinov, R.N., Neverov, A.D., Mironov, A.A. and Gelfand, M.S. (2008) Positive selection in alternatively spliced exons of human genes. *Am. J. Hum. Genet.*, **83**, 94–98.

57. Rechsteiner, M. and Rogers, S.W. (1996) PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.*, **21**, 267–271.
58. Bromberg, K.D., Ma'ayan, A., Neves, S.R. and Iyengar, R. (2008) Design logic of a cannabinoid receptor signaling network that triggers neurite outgrowth. *Science*, **320**, 903–909.
59. Hill, D.A., de la Serna, I.L., Veal, T.M. and Imbalzano, A.N. (2004) BRCA1 interacts with dominant negative SWI/SNF enzymes without affecting homologous recombination or radiation-induced gene activation of p21 or Mdm2. *J. Cell Biochem.*, **91**, 987–998.
60. Huttley, G.A., Easteal, S., Southey, M.C., Tesoriero, A., Giles, G.G., McCredie, M.R., Hopper, J.L. and Venter, D.J. (2000) Adaptive evolution of the tumour suppressor BRCA1 in humans and chimpanzees. Australian Breast Cancer Family Study. *Nat. Genet.*, **25**, 410–413.
61. Cordaux, R., Udit, S., Batzer, M.A. and Feschotte, C. (2006) Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc. Natl Acad. Sci. USA*, **103**, 8101–8106.
62. LaPorte, S.L., Juo, Z.S., Vaclavikova, J., Colf, L.A., Qi, X., Heller, N.M., Keegan, A.D. and Garcia, K.C. (2008) Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system. *Cell*, **132**, 259–272.
63. Hansen, G., Hercus, T.R., McClure, B.J., Stomski, F.C., Dottore, M., Powell, J., Ramshaw, H., Woodcock, J.M., Xu, Y., Guthridge, M. *et al.* (2008) The structure of the GM-CSF receptor complex reveals a distinct mode of cytokine receptor activation. *Cell*, **134**, 496–507.
64. Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E. *et al.* (2008) Large recurrent microdeletions associated with schizophrenia. *Nature*, **455**, 232–236.
65. Zhang, E.E., Liu, A.C., Hirota, T., Miraglia, L.J., Welch, G., Pongsawakul, P.Y., Liu, X., Atwood, A., Huss, J.W. 3rd, Janes, J. *et al.* (2009) A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell*, **139**, 199–210.
66. Sweet, R.A., Henteleff, R.A., Zhang, W., Sampson, A.R. and Lewis, D.A. (2009) Reduced dendritic spine density in auditory cortex of subjects with schizophrenia. *Neuropsychopharmacology*, **34**, 374–389.
67. Aizawa, H., Hu, S.C., Bobb, K., Balakrishnan, K., Ince, G., Gurevich, I., Cowan, M. and Ghosh, A. (2004) Dendrite development regulated by CREST, a calcium-regulated transcriptional activator. *Science*, **303**, 197–202.
68. Wu, J.I., Lessard, J., Olave, I.A., Qiu, Z., Ghosh, A., Graef, I.A. and Crabtree, G.R. (2007) Regulation of dendritic development by neuron-specific chromatin-remodeling complexes. *Neuron*, **56**, 94–108.
69. Sutton, M.A. and Schuman, E.M. (2006) Dendritic protein synthesis, synaptic plasticity, and memory. *Cell*, **127**, 49–58.
70. Crespi, B., Summers, K. and Dorus, S. (2007) Adaptive evolution of genes underlying schizophrenia. *Proc. Biol. Sci.*, **274**, 2801–2810.
71. Crow, T.J. (2006) March 27, 1827 and what happened later—the impact of psychiatry on evolutionary theory. *Prog. Neuropsychopharmacol. Biol. Psychiatry*, **30**, 785–796.
72. Doi, N., Hoshi, Y., Itokawa, M., Usui, C., Yoshikawa, T. and Tachikawa, H. (2009) Persistence criteria for susceptibility genes for schizophrenia: a discussion from an evolutionary viewpoint. *PLoS One*, **4**, e7799.
73. Moalic, J.M., Le Strat, Y., Lepagnol-Bestel, A.M., Ramoz, N., Loe-Mie, Y., Maussion, G., Gorwood, P. and Simonneau, M. (2010) Primate-accelerated evolutionary genes: novel routes to drug discovery in psychiatric disorders. *Curr. Med. Chem.*, **17**, 1300–1316.
74. Bord, L., Wheeler, J., Paek, M., Saleh, M., Lyons-Warren, A., Ross, C.A., Sawamura, N. and Sawa, A. (2006) Primate disrupted-in-schizophrenia-1 (DISC1): high divergence of a gene for major mental illnesses in recent evolutionary history. *Neurosci. Res.*, **56**, 286–293.
75. Lo, W.S., Xu, Z., Yu, Z., Pun, F.W., Ng, S.K., Chen, J., Tong, K.L., Zhao, C., Xu, X., Tsang, S.Y. *et al.* (2007) Positive selection within the Schizophrenia-associated GABA(A) receptor beta2 gene. *PLoS One*, **2**, e462.
76. Huffaker, S.J., Chen, J., Nicodemus, K.K., Sambataro, F., Yang, F., Mattay, V., Lipska, B.K., Hyde, T.M., Song, J., Rujescu, D. *et al.* (2009) A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nat. Med.*, **15**, 509–518.
77. Di Rienzo, A. and Hudson, R.R. (2005) An evolutionary framework for common diseases: the ancestral-susceptibility model. *Trends Genet.*, **21**, 596–601.
78. Feinberg, A.P. and Irizarry, R.A. (2010) Evolution in health and medicine Sackler colloquium: stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl Acad. Sci. USA*, **107**, 1757–1764.
79. Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. and Zoghbi, H.Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.*, **23**, 185–188.
80. Brockschmidt, A., Todt, U., Ryu, S., Hoischen, A., Landwehr, C., Birnbaum, S., Frenck, W., Radlwimmer, B., Lichter, P., Engels, H. *et al.* (2007) Severe mental retardation with breathing abnormalities (Pitt–Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum. Mol. Genet.*, **16**, 1488–1494.
81. Collins, A.L., Levenson, J.M., Vilaythong, A.P., Richman, R., Armstrong, D.L., Noebels, J.L., David Sweatt, J. and Zoghbi, H.Y. (2004) Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.*, **13**, 2679–2689.
82. Brzozka, M.M., Radyushkin, K., Wichert, S.P., Ehrenreich, H. and Rossner, M.J. (2010) Cognitive and sensorimotor gating impairments in transgenic mice overexpressing the schizophrenia susceptibility gene Tcf4 in the brain. *Biol. Psychiatry*, in press.
83. McClellan, J. and King, M.C. (2010) Genetic heterogeneity in human disease. *Cell*, **141**, 210–217.
84. Kim, C.S., Hwang, C.K., Choi, H.S., Song, K.Y., Law, P.Y., Wei, L.N. and Loh, H.H. (2004) Neuron-restrictive silencer factor (NRSF) functions as a repressor in neuronal cells to regulate the mu opioid receptor gene. *J. Biol. Chem.*, **279**, 46464–46473.
85. Marisa, L., Ichanté, J.L., Reymond, N., Aggerbeck, L., Delacroix, H. and Mucchielli-Giorgi, M.H. (2007) MAnGO: an interactive R-based tool for two-colour microarray analysis. *Bioinformatics*, **23**, 2339–2341.
86. Smyth, G.K. (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article 3.
87. Smyth, G.K. and Speed, T. (2003) Normalization of cDNA microarray data. *Methods*, **31**, 265–273.
88. Sturn, A., Quackenbush, J. and Trajanoski, Z. (2002) Genesis: cluster analysis of microarray data. *Bioinformatics*, **18**, 207–208.
89. Klipper-Aurbach, Y., Wasserman, M., Braunspiegel-Weintrob, N., Borstein, D., Peleg, S., Assa, S., Karp, M., Benjamini, Y., Hochberg, Y. and Laron, Z. (1995) Mathematical formulae for the prediction of the residual beta cell function during the first 2 years of disease in children and adolescents with insulin-dependent diabetes mellitus. *Med. Hypotheses*, **45**, 486–490.
90. Benjamini, Y.a.H.Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. Ser. B*, **57**, 289–300.
91. Calvano, S.E., Xiao, W., Richards, D.R., Felciano, R.M., Baker, H.V., Cho, R.J., Chen, R.O., Brownstein, B.H., Cobb, J.P., Tschoeke, S.K. *et al.* (2005) A network-based analysis of systemic inflammation in humans. *Nature*, **437**, 1032–1037.
92. Beisvag, V., Jünge, F.K., Bergum, H., Jølsum, L., Lydersen, S., Günther, C.C., Ramampiaro, H., Langaas, M., Sandvik, A.K. and Laegreid, A. (2006) GeneTools—application for functional annotation and statistical hypothesis testing. *BMC Bioinformatics*, **7**, 470.
93. Bourne, J.N. and Harris, K.M. (2008) Balancing structure and function at hippocampal dendritic spines. *Annu. Rev. Neurosci.*, **31**, 47–67.
94. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
95. Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
96. Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.*, **13**, 555–556.