Molecular Psychiatry (2010), 1–14

www.nature.com/mp

ORIGINAL ARTICLE

A common variant of the latrophilin 3 gene, *LPHN3,* confers susceptibility to ADHD and predicts effectiveness of stimulant medication

M Arcos-Burgos[1], M Jain[1], MT Acosta[1], S Shively[1,2], H Stanescu[1], D Wallis[1], S Domené[1], JI Vélez[1], JD Karkera[1], J Balog[1], K Berg[1], R Kleta[1], WA Gahl[1], E Roessler[1], R Long[1], J Lie[2], D Pineda[3], AC Londoño[3], JD Palacio[3], A Arbelaez[3], F Lopera[3], J Elia[4], H Hakonarson[4], S Johansson[5], PM Knappskog[5], J Haavik[5], M Ribases[6], B Cormand[7], M Bayes[8], M Casas[6,9], JA Ramos-Quiroga[6,9], A Hervas[10], BS Maher[11], SV Faraone[12], C Seitz[13], CM Freitag[13], H Palmason[14], J Meyer[14], M Romanos[15], S Walitza[15], U Hemminger[15], A Warnke[15], J Romanos[16], T Renner[15,16,17], C Jacob[16], K-P Lesch[15,16,17], J Swanson[18], A Vortmeyer[2], JE Bailey-Wilson[1], FX Castellanos[19], M Muenke[1]

[1]*National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA;* [2]*National Institute for Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA;* [3]*Neurosciences Group, University of Antioquia, Medellín, Colombia;* [4]*The Children's Hospital of Philadelphia, Philadelphia, PA, USA;* [5]*Department of Biomedicine, Section of Biochemistry and Molecular Biology, University of Bergen, Bergen, Norway;* [6]*Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Catalonia, Spain;* [7]*Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, CIBER-ER, IBUB, Barcelona, Catalonia, Spain;* [8]*Genes and Disease Program, Center for Genomic Regulation (CRG), CIBER-ESP, Centro Nacional de Genotipado (CeGen), Barcelona, Catalonia, Spain;* [9]*Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Catalonia, Spain;* [10]*Child and Adolescent Mental Health Unit, Hospital Mútua de Terrassa, Barcelona, Catalonia, Spain;* [11]*Departments of Psychiatry and Human Genetics, Virginia Institute of Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA;* [12]*Department of Psychiatry, SUNY Upstate Medical University, Syracuse, NY, USA;* [13]*Department of Child and Adolescent Psychiatry, Saarland University Hospital, Homburg, Saar, Germany;* [14]*Graduate School for Psychobiology, Division of Neuro-Behavioral Genetics, University of Trier, Trier, Germany;* [15]*Department of Child and Adolescent Psychiatry and Psychotherapy, University of Würzburg, Würzburg, Germany;* [16]*Department of Psychiatry and Psychotherapy, University of Würzburg, Würzburg, Germany;* [17]*Department of Molecular and Psychobiology, University of Würzburg, Würzburg, Germany;* [18]*UCI Child Development Center, University of California, Irvine, CA, USA and* [19]*New York University Child Study Center, New York, NY, USA*

Attention-Deficit/Hyperactivity Disorder (ADHD) has a very high heritability (0.8), suggesting that about 80% of phenotypic variance is due to genetic factors. We used the integration of statistical and functional approaches to discover a novel gene that contributes to ADHD. For our statistical approach, we started with a linkage study based on large multigenerational families in a population isolate, followed by fine mapping of targeted regions using a family-based design. Family- and population-based association studies in five samples from disparate regions of the world were used for replication. Brain imaging studies were performed to evaluate gene function. The linkage study discovered a genome region harbored in the Latrophilin 3 gene (*LPHN3).* In the world-wide samples (total *n* = 6360, with 2627 ADHD cases and 2531 controls) statistical association of *LPHN3* and ADHD was confirmed. Functional studies revealed that *LPHN3* variants are expressed in key brain regions related to attention and activity, affect metabolism in neural circuits implicated in ADHD, and are associated with response to stimulant medication. Linkage and replicated association of ADHD with a novel non-candidate gene (*LPHN3*) provide new insights into the genetics, neurobiology, and treatment of ADHD.
Molecular Psychiatry advance online publication, 16 February 2010; doi:10.1038/mp.2010.6

Keywords: ADHD; complex trait; gene; LPHN3; genetics; latrophilin

Correspondence: Dr M Muenke, National Human Genome Research Institute, National Institutes of Health, 35 Convent Drive, Bldg 35 room 1B203, Bethesda, MD, 20892-3717, USA.
E-mail: mamuenke@mail.nih.gov
Received 20 August 2009; revised 21 December 2009; accepted 23 December 2009

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is the most common behavioral disorder of childhood, with a high prevalence,[1] poor outcome,[2] substantial costs[3] and comorbidity with other behavioral and

emotional disorders.[4] Genetic factors are strongly implicated in the etiology of ADHD.[5] More than a decade ago, candidate gene association studies identified two genes with small effects,[6,7] and these associations have been replicated.[8] Many additional candidate genes have been evaluated, and some initial reports of association with ADHD have emerged, but these have not been replicated.[9] Several unbiased genome-wide studies have been conducted and reported, but none has identified additional genes that contribute to the genetic basis of ADHD. Following the suggestion of Risch and Merikangas,[10] the genome-wide association study (GWAS) approach has been used, but this did not produce a significant effect or even confirmed the replicated associations with the *DRD4* and *DAT* genes.[11]

The lack of success of GWAS to discover new genes related to ADHD is not unexpected. As for others disorders and traits with high heritability (such as height), the identified associated genes account for 5–10% of the estimated heritability, creating a mystery of 'dark' or unexplained heritability.[12] It has been assumed that many additional genes with small independent effects likely exist but have not been discovered. The integration of statistical and functional approaches has been recommended to uncover the 'dark matter' and begin to account for the heritability of common disorders.[13]

We identified large and multigenerational families from a genetic isolate, the Paisa population in Antioquia, Colombia.[4,14] The prevalence of ADHD is high and often comorbid with disruptive behaviors in the Paisa.[4] This makes the Paisa population well-suited for a linkage study to discover new genes that contribute to the etiology of ADHD. The initial study in our program of research evaluated ADHD symptoms in informative families ascertained in the Paisa and concluded that they would provide exceptionally good power to detect linkage even in the presence of heterogeneity.[14] A genome-wide linkage study[15] of 16 informative families with 433 individuals reported significant linkage of ADHD to a region on chromosome 4q13.2.[14–16] Fine mapping applied to nine linked families sharpened the linkage signal, disclosed new meiotic recombination events in ADHD-affected individuals, and in conjunction with the LOD-1 supported interval criterion narrowed the minimal critical region down to ~20 Mb (Table 1, Supplementary information).[15] Here, we report the next steps in this research program, which were designed to identify the gene responsible for the ADHD linkage signal and to elucidate its functional properties.

Materials and methods

Subjects

Detailed descriptions of clinical and demographic features of the Paisa sample for the linkage study have been published elsewhere.[4] We used 18 extended families with 433 individuals, and 137 nuclear

Table 1 Summary of the features of the different population samples where this study was performed

Sample	*Type of study*	*Total number of ascertained families*	*Number of individuals with ADHD*	*Number of unaffected individuals*	*Number of individuals with undefined phenotype*	*Total number of individuals*	*Clinical assessment and instruments used for ADHD phenotyping*	*References*
Original set of multigenerational and extended Paisa families[a]	Family-based	18	142	264	27	433	-Structured diagnostic interview-DICA-IV-P	4
Additional set of nuclear Paisa families for second stage analysis	Family-based	137	214	222	91	527	-Structured diagnostic interview-DICA-IV-P	4
US1	Family-based	316	630	647	88	1365	-Structured diagnostic interview - Vanderbilt, -SWAN, and -DICA-IV-P	17
US2	Family-based	322	356	644	0	1000	-Structured diagnostic interview - Vanderbilt, -SWAN, and -DICA-IV-P	18
Germany	Family-based	231	406	40	516	962	-Structured diagnostic interview I, II, WURS, CAARS, K-SADS-PL, CBCL	19
Norway	Case–control	0	450	585	0	1035	-Structured diagnostic interview. ASRS, WURS	20
Spain	Family-based	386	429	129	480	1038	-Structured diagnostic interview (K-SADS-PL)-Conners' Parent and Teacher Rating Scale: (CPRS-R: L, CTRS-R:L)	21
Total		1410	2627	2531	1202	6360		

[a]58 additional individuals were sampled for families showing linkage to 4q13.2 from the originally 375 reported Arcos-Burgos *et al.*[14]

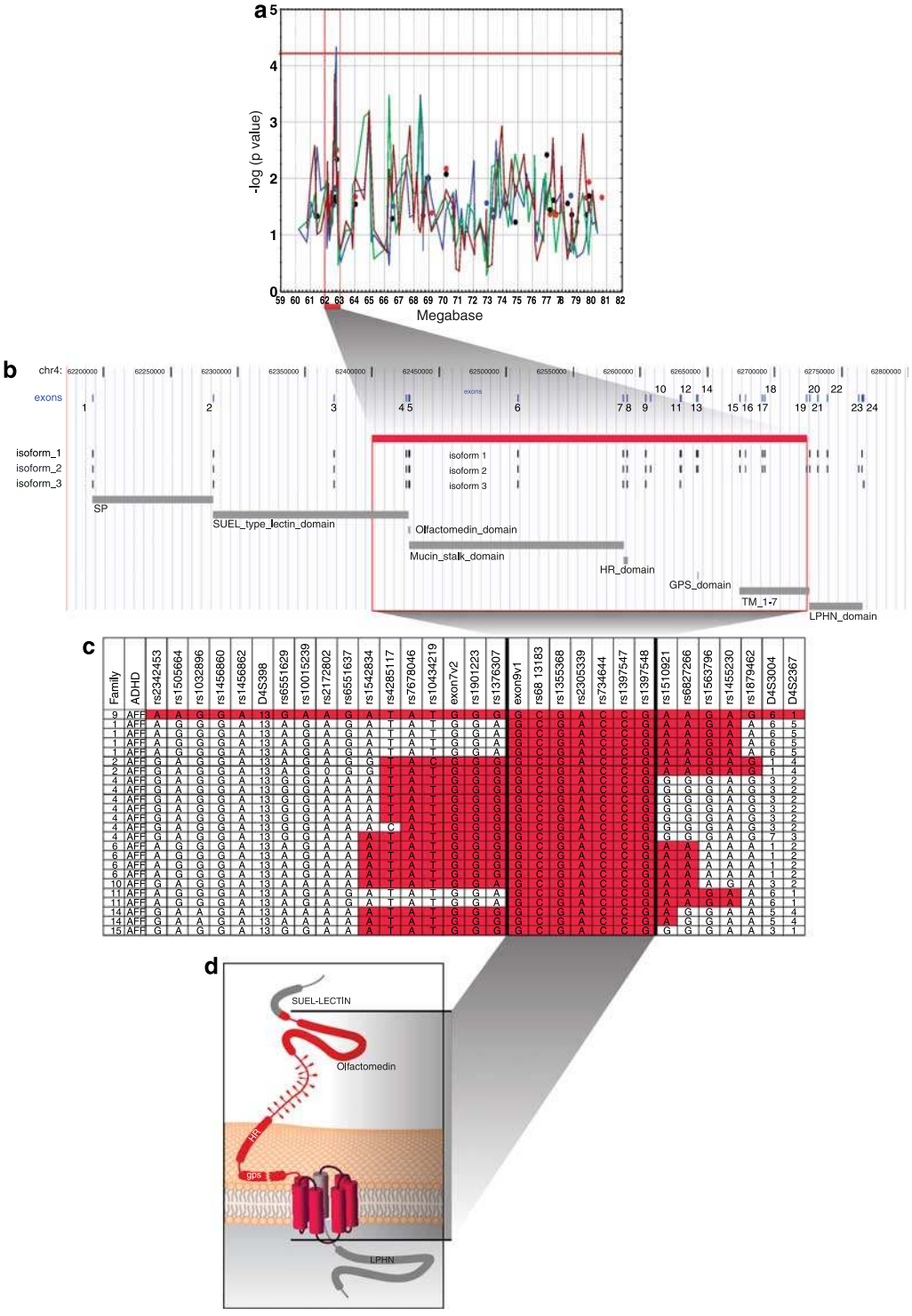
families with 527 individuals. The samples used for replication were ascertained from disparate regions of the world, namely from the United States (two independent samples, US1,[17] and US2[18]), Germany,[19] Norway[20] and Spain[21] (see Table 1). A total of 1410 families with 6360 individuals were evaluated. Approvals to conduct the study (Protocol no. 00-HG-0058) were obtained from the Institutional Review Boards of the National Human Genome Research Institute and those of the collaborating institutions.

Genotyping
DNA from the Colombian, US1, US2, German and Spanish samples was prepared from peripheral-blood specimens. DNA from the Norwegian sample was extracted either from whole blood or from saliva. The Illumina genotyping service was used for genotyping, except for the sample from Norway that was genotyped using Sequenom's MassARRAY iPLEX System (additional information in Supplementary information).

Screening of the minimal critical region in 4q13.2
Fine scale mapping with a resolution of ~68 kb (a third of the minimum distance able to provide full coverage in the Paisa population[22]) was conducted in the Paisa multigenerational sample of 18 families (433 individuals) and in the ADHD sample of 137 nuclear families (527 individuals) ascertained from the same genetic isolate. Areas of interest that are gene rich or carry potential candidates were covered at a higher density (Figure 1, Supplementary information). An empirical linkage disequilibrium map, built using control individuals, showed full coverage of the entire region and excluded uncovered gaps (Figure 2, Supplementary information).

Screening for sequence variations and SNP tagging
The finding that a shared ADHD susceptibility haplotype was embedded within the Latrophilin 3 (*LPHN3*) gene, and the absence of other known genes in the region, prompted further pursuit of the relationship between *LPHN3* variants and susceptibility to ADHD. Hence, the entire coding region of *LPHN3* was sequenced in eight individuals carrying two ($n=4$), one ($n=2$), or zero ($n=2$) copies of the susceptibility haplotype variants. In addition, 46 individuals from different families showing either nominal, suggestive or significant signals of linkage at the region were also sequenced using Surveyor (Transgenomics, Omaha, NE, USA). Furthermore, 139 unrelated patients diagnosed with ADHD from the United States and 52 commercially available normal controls were also studied. For the US and control samples, mutation detection was performed by PCR-based denaturing High Performance Liquid Chromatography analysis followed by direct bi-directional sequencing (variants identified are presented in Figure 3a, Supplementary information). Additional SNP selection and tagging employed a quality-based filter, a second population information-based filter (MAF, linkage disequilibrium blocks), and a third biological significance-based filter (coding, expression modulating, splicing modulating and assessing evolutionary conservation) (Figures 3b–e, Supplementary information). Databases used for the selection were SNPdb (NCBI) and UCSC genome browser. Quality filtering was based on the information available through the SNP track of the UCSC genome browser and by using AB SNPbrowser 3.0. Customized LD blocks were defined using AB SNPbrowser 3.0 and Haploview. Biological relevance was annotated on the UCSC genome browser with data available from UCSC (genome browser), NCBI (Entrez) and EBI (Ensembl). Conservation was based on the MAF 8X conservation track (courtesy of Elliott Margulies, NHGRI, NIH) and Exact Plus software (courtesy of Anthony Antonellis, NHGRI, NIH). The US1, the German and the Paisa sample were genotyped for the same set of SNPs ($n=120$) (Figure 3e and Supplementary Table 5, Supplementary information). The sample from Norway was genotyped using a two-stage design: 60 markers were first genotyped in 343 adult Norwegian ADHD patients and 336 controls (Supplementary Table 5, Supplementary information). All SNPs with a nominal P-value <0.15 (7 markers) were then included in a second round of genotyping using a new set of 107 cases and 249 controls. The US2 sample was genotyped for 77 SNP markers, and the Spanish sample was genotyped for 45 SNP markers (Supplementary Table 5, Supplementary information).

Genetic statistical analyses
Family-based association tests were applied by comparing allele frequencies for transmitted and untransmitted alleles using PLINK.[23] Parameters for excluding markers from analyses include a MAF less than or equal to 0.05, Hardy–Weinberg equilibrium deviations with P-values <0.0001, and a minimum genotype of 90%. Mendelian errors were checked using PEDSTATS.[24] Families with excessive Mendelian errors were excluded. SNPs with >2 Mendelian errors also were excluded. We tested unaffected siblings looking for overtransmission bias caused by evolutionary forces, such as selection *in utero* against deleterious mutations, meiotic drive and maternal–fetal incompatibility,[25] but found no significant overtransmission for any single marker or for any haplotype in the entire region. Linkage disequilibrium via cladistic analysis was performed using CLADH[26] with a Bonferroni-adjusted threshold global permutation P-value (10 000 permutations). With CLADH,[26] clades of haplotypes are tested for association with disease, exploiting the expected similarity of chromosomes with recent shared ancestry in the regions flanking the disease gene. We selected only one affected case from each pedigree. As controls, haplotypes of unrelated individuals without ADHD were chosen. To abrogate confounding from potentially undetected genotyping errors and the presence



of meiotic drift when the risk variants are the most common, we tested additional SNP markers harbored in the *LPHN3* gene, including those new variants disclosed by our sequencing analyses plus other potentially functional and phylogenetically conserved SNP variants (Figures 3a–e, in Supplementary information). For haplotype analyses both uncertain and infrequent variants (<1%) were excluded. To infer genotypes not assessed homogeneously in the samples, we performed *in silico* genotyping as implemented in IMPUTE.[27] The Hapmap CEPH sample was used as our comparison set of phased haplotypes. All genotypes inferred with >90% genotype probability were retained. Odds ratios in family-based samples were estimated by the number of transmitted target alleles/the number of untrans-

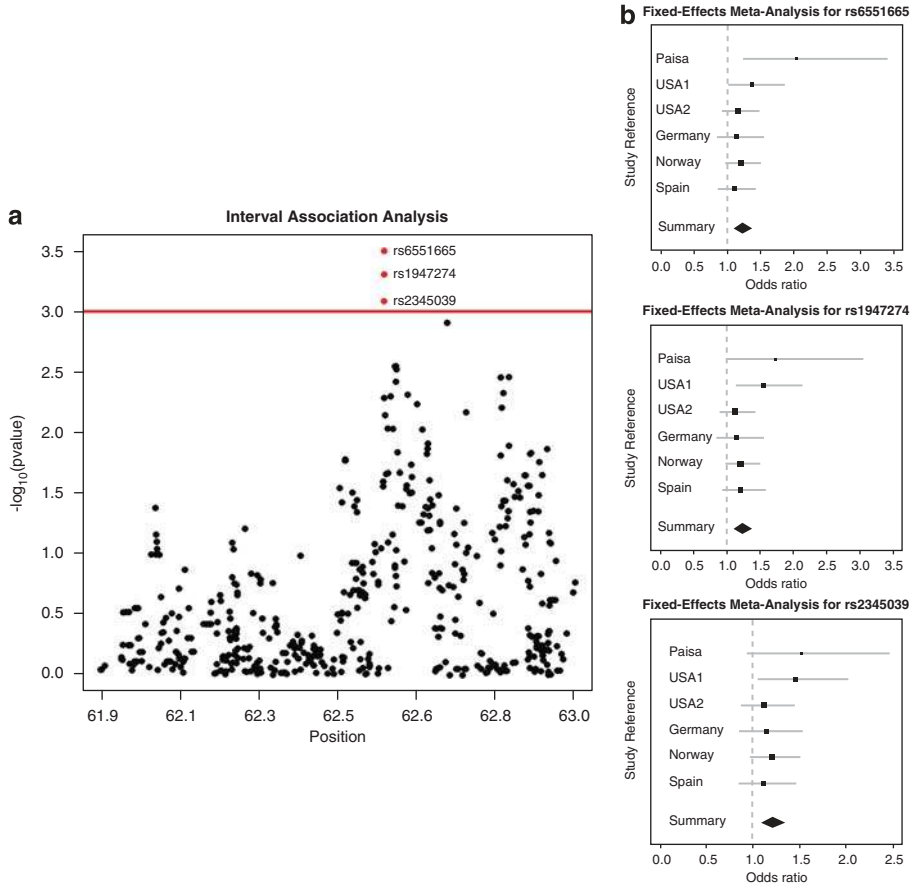


Figure 2 (**a**)Location at the *LPHN3* gene of the three markers, rs6551665, rs1947274 and rs2345039, reaching significance at the meta-analysis after 10 000 permutations. (**b**) Forest plot of odds ratios showing fixed effect results for rs6551665 (OR = 1.23, 95% CI 1.09–1.137, $P = 3.46 \times 10^{-4}$), rs1947274 (OR = 1.23, 95% CI 1.09–1.38, $P = 5.41 \times 10^{-4}$), and rs2345039 (OR = 1.21, 95% CI 1.08–1.35, $P = 8.97 \times 10^{-4}$), through populations.

Figure 1 (**a**) Results for association and linkage disequilibrium analyses on extended and nuclear ADHD families of the Paisa community (Antioquia, Colombia). On the *x* axis are base-pair coordinates (60–80 Mb) at 4q13.2; on the *y* axis, -log (p). Dots represent the distribution of the *Z*-statistic (single marker-based) for extended families: Black dots represent nuclear families; Red dots, the pooled families. The continuous lines of different colors represent the distribution of absolute values for -log (p) after testing for linkage disequilibrium using a permutation test on haplotype windows of different sizes. This level of significance was robust with respect to different haplotype windows (from 3 to 7 markers) and to correct for multiple comparisons. The Bonferroni-adjusted threshold *P*-value is represented as a red line at −log(p) = 4.25. A significant peak of association is located at the region between coordinates 62 and 63 Mb. (**b**) Annotated structure of *LPHN3* showing exons 1–24, the major protein structural domains, and the area of the susceptibility haplotype. The haplotype (in red) covers the olfactomedin domain, part of the mucin-like stalk, the hormone receptor domain, the G-Protein Coupled Receptor (GPCR) Proteolytic Site—domain and transmembrane helices 1–6 of the seven transmembrane spanning region. The isoforms represent the three splice site variants of *LPHN3*. (**c**) Haplotype variants transmitted in affected individuals from families with significant linkage of ADHD to 4q13.2. Comparison of these variants is carried out against F9 (row 1), the family providing the most to the linkage signal as published in Arcos-Burgos *et al*.[15] The shared region in orange exactly corresponds to the region mapped by family-based genetic association analyses. (**d**) Structure of LPHN3 with the protein domains that correlate with the ADHD susceptibility haplotype in red.

mitted alleles in heterozygous parents only.[28] Family-based odds ratios from the TDT tests in the Paisa, US1, US2, German, and Spanish samples were combined with the case–control odds ratio from the Norwegian sample using a meta-analytical approach.[29] The random-effects method[30] was used when significant heterogeneity was present and the fixed-effects method otherwise.[31]

Expression of LPHN3 in human brain

We studied the expression of identified genes in northern blots (Human Brain MTN Blots II and V) purchased from BD Biosciences (Palo Alto, CA, USA) with the gene probe synthesized based upon its cDNA sequence at Invitrogen (Carlsbad, CA, USA) and analyzed in brain tissues from amygdala, caudate nucleus, cerebellum, orbital frontal cortex, pontine

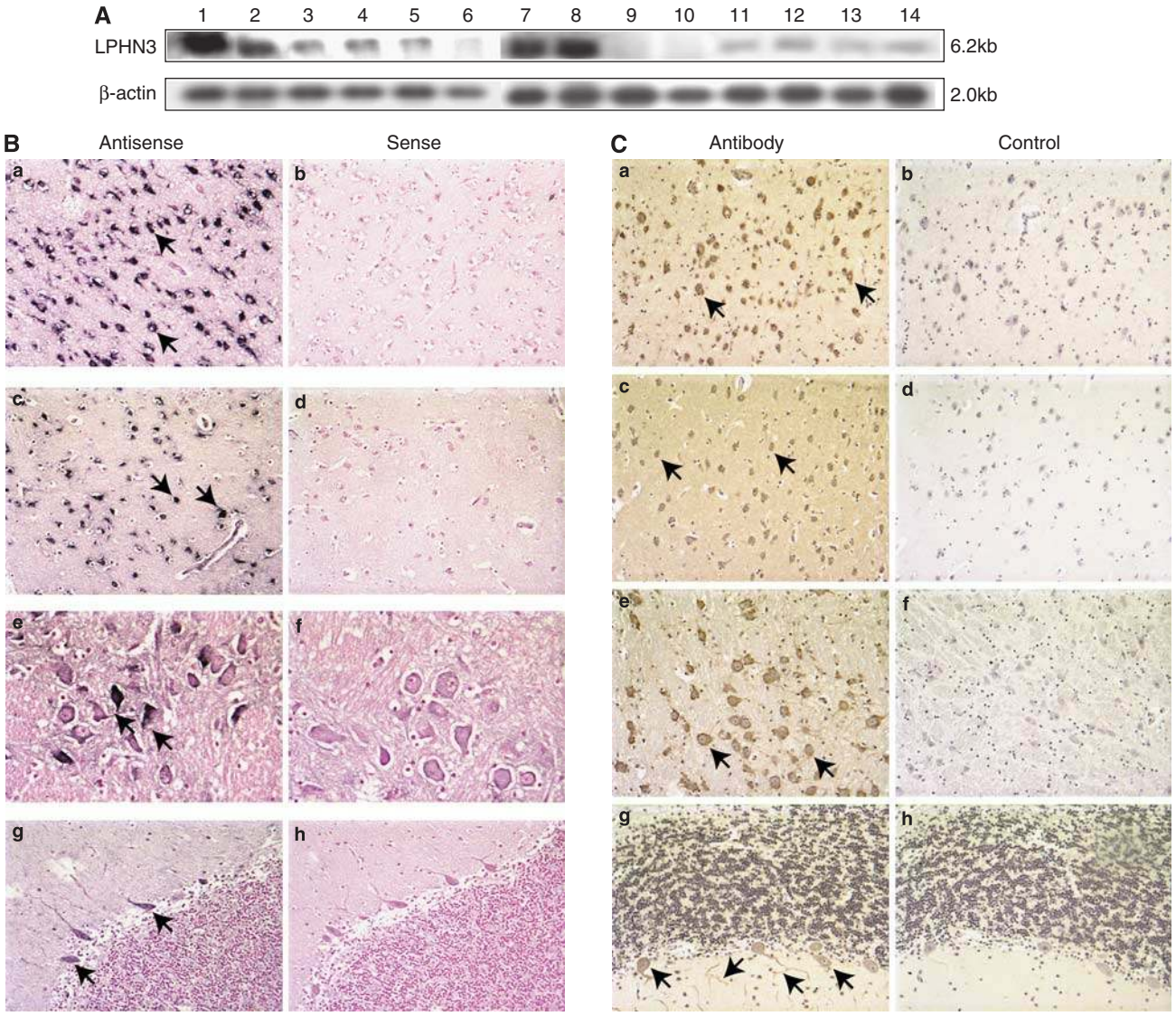


Figure 3 (**A**) Northern blots reveal expression of *LPHN3* mRNA in different regions of the human brain. Lane 1, Amygdala; Lane 2, Caudate Nucleus; Lane 3, Corpus Callosum; Lane 4, Hippocampus; Lane 5, Whole Brain Extract; Lane 6, Thalamus; Lane 7, Cerebellum; Lane 8, Cerebral Cortex; Lane 9, Medulla; Lane 10, Spinal Cord; Lane 11, Occipital Pole; Lane 12, Frontal Lobe; Lane 13, Temporal Lobe; and Lane 14, Putamen. (**B**) *In situ* hybridization of human brain tissue from a 2 year old reveals positive cytoplasmic signal with *LPHN3* antisense. *LPHN3* sense provided a negative control. Selected positive neurons are marked by arrows. (a, b), Amygdala; (c, d), Caudate Nucleus; (e, f), Pontine Nuclei; (g, h), Cerebellum (**c**) Immunohistochemistry of human brain tissue from a 2 year old reveals positive cytoplasmic immunoreactivity with *LPHN3* antibody. Primary *LPHN3* antibody was omitted for negative control. Selected positive neurons are marked by arrows. (a, b), Amygdala; (c, d), Caudate Nucleus; (e, f), Pontine Nucleus; (g, h), Cerebellum. Immunoreactive Purkinje cells are marked by arrows.

nuclei, cingulate gyrus, occipital cortex and thalamus. Probes were labeled with γ-^{32}P-ATP (GE Healthcare, North Arlington Heights, IL, USA) by T4 polynucleotide kinase (Invitrogen) and purified by Qiaquick Nucleotide Removal Kit (Qiagen, Valencia, CA, USA). Membranes were prehybridized with ExpressHyb (BD Biosciences) for 1 h at 42 °C and hybridized with labeled probes for 3 h at 42 °C. After hybridization, membranes were stringently washed and exposed on Kodak film at −70 °C for 10 days. Human brain tissue for *in situ* hybridization was obtained from the Brain and Tissue Bank for Developmental Disorders at the University of Maryland, Baltimore, MD, USA from male accident victims deceased at 2, 5, 8 and 30 years of age. Though reported as healthy normal individuals, there was no information regarding the mental health status.

A β-actin probe was used for normalization. The antisense probe for *LPHN3*, a 362-bp amplicon within

the final coding exon, was subcloned into Bluescript KS- (Stragene, San Diego, CA, USA) and synthesized using a T3 primer. Expression in fetal brain was more pronounced than in whole postnatal brain. Fibroblasts and testes also showed notable expression. Minimal expression was detected in thymus, lung, prostate, ovaries, heart, pancreas, liver and kidney. No expression could be detected in cDNAs from leukocytes, spleen, skeletal muscle, colon, small intestine and placenta (data not shown). The *in situ* hybridization experiments were performed at Histoserv (Germantown, MD, USA). In addition, a panel of normal human pooled cDNAs (BD Biosciences, Clontech, CA, USA) revealed robust expression of a single 3.7 Kb *LPHN3* amplicon (primers covering the entire open reading frame) in cerebellum, cerebral cortex, thalamus, amygdala, substantia nigra, hippocampus, spinal cord and retina (data not shown). For immunohistochemical analyses serial sections were taken from paraffin-embedded tissue blocks for immunohistochemical examinations. For antigen retrieval, sections were treated with 0.01 M sodium citrate buffer (pH 6.0) at 100 °C for 10 min. Sections were cooled at room temperature for 20 min and washed three times in PBS. Sections were then quenched for 20 min in a solution of 3 ml H_2O_2 and 180 ml methanol. After three washes in PBS, sections were incubated in 10% horse serum for 1 h. The LPHN3 primary antibody (1:500, Novus Biologicals, Littleton, CO, USA) was diluted in 2% horse serum, and the sections were incubated in a humidified chamber at 4 °C overnight. The sections were incubated with secondary antibody for 1 h followed by avidin–biotin complex incubation for 1 h and visualization with diaminobenzidine. The sections were counterstained with Mayer's hematoxylin for 10 min, dehydrated by graded ethanol washes of 95 and 100%, and rinsed in xylene before being mounted.

Proton magnetic resonance spectroscopy (^{1}H-MRS)

We explored potential differences between carriers and non-carriers of the susceptibility haplotype with proton magnetic resonance spectroscopy (^{1}H-MRS) to measure neuronal number or viability[32] as assessed by the ratio of *N*-acetylaspartate (NAA) to creatine (Cr). Brain regions that define the frontal–striatal–cerebellar circuit known to be dysfunctional in ADHD[33–36] were targeted. To obtain measures of metabolic brain activity using ^{1}H-MRS, T2-weighted high resolution anatomic images were obtained in the axial plane (TE = 103 ms, TR = 5910 ms, 3 mm slice thickness and 5:27 min:s imaging time). Axial images were oriented parallel to the orbitomeatal anatomical reference plane. The T2-weighted images were used to guide multi-voxel MR spectroscopy volume selection. The 2D chemical shift imaging (CSI) point-resolved spectroscopic sequence (PRESS) technique (TE = 30 ms, TR = 1500 ms, NEX = 3, resolution 10 mm × 10 mm × 10 mm, acquisition time = 6:05) was localized in the inferior vermis using the anatomical reference images. Three-dimensional CSI PRESS sequence (TE = 30 ms, TR = 1500 ms, NEX = 3, resolution 13.3 mm × 13.3 mm × 13.8 mm, acquisition time = 10:23) explored the center of the brain including the striatum, thalamus and the cingulate gyrus relative to anatomic images. Saturation bands around the 2D and 3D Volumes of Interest were used to prevent contamination of the spectra from subcutaneous fat signal. All MR data were obtained on a 1.5 T Symphony Master Class Siemens Clinical Imaging System using an 8 channel head array coil. The spectra were transferred offline to be processed automatically using the LCModel.[37] For this initial analysis, we focused on the ratio NAA to Cr within each spectral voxel.[38] Average data from voxels covering the left and right striatum (3–4 voxels), lateral (2 voxels) and medial (2 voxels) aspects of the thalamus, anterior (1 voxel), medial (1 voxel) and posterior (1 voxel) cingulate gyrus, and inferior vermis (2 voxels) were analyzed. Voxels containing cerebrospinal fluid were excluded from analyses. Criteria for acceptable reliability were those recommended by the LCModel provider. The imaged participants were neither sedated nor receiving medications for treatment of ADHD. Total scanning duration was 45 min. Absolute metabolite quantification was not attempted because of the requirement for markedly increased data acquisition time, which is particularly problematic for patients with ADHD. The general linear model was used to provide exploratory analyses to evaluate between-group differences with an uncorrected two-tailed α-level of 0.05.

Pharmacogenetic Study

An effectiveness evaluation of medication response was conducted in 240 children from the US1 sample[17] rated on and off stimulant medication on the Strengths and Weaknesses of ADHD-Symptoms and Normal-Behavior (SWAN) Scale.[39] By truncating the SWAN ratings, the standard ratings of symptom severity (just weaknesses) can be captured. The SWAN scale differs from other scales in its definition of items and in its scoring system, which give a closer approximation to a normal distribution for each item and of summary scores of overall ADHD and its two domains (Inattention and Hyperactivity–Impulsivity) in the general population than does any other ADHD scale that evaluates psychopathology.[39] In addition, the heritability of the SWAN scale has been shown.[39] Consistency scores were estimated for each interview, using a scale from 1 (unreliable) to 10 (fully reliable), according to the concordance between two questions. A full description of ascertainment and clinical diagnostic strategies of the US1 sample is presented elsewhere.[17]

We used statistical methods to identify clusters to define responders and non-responders using Latent Class Cluster Analysis.[40] Latent Class Cluster Analysis models containing 1 through 12 classes were fitted to the data using Latent GOLD 3.0.1 software (Statistical Innovations, Belmont, MA, USA). Latent GOLD uses both expectation/maximization (EM) and Newton–Raphson algorithms to find the maximum likelihood of each model after estimating model parameters.[40–42]

To avoid ending up with local solutions (a well-known problem in LCA), we used multiple sets of starting values as automatically implemented in Latent GOLD. Because we were dealing with sparse contingency tables, we estimated *P*-values associated with L^2 statistics by means of parametric bootstrap (500 replicates) rather than relying on asymptotic *P*-values. As covariates for the model, we used gender, ADHD medication usage and age. Age was included in our models as a continuous variable, as a categorical variable based on deciles (that is, 1–10, 10–20, etc.), and as a categorical variable using the age ranges we used previously (that is, children are between 4 and 11 years; adolescents 12–17; and adults are 18 years or older).[17] Our final models used the latter approach because it resulted in smaller bivariate residuals.

Initially, we did not consider the presence of interactions between variables and the basic assumption of local independence of the standard latent class model was supported. Next, we relaxed the local independence assumption by allowing for interactions between variables, as well as for direct effects of covariates on variables.[40,42,43] Latent GOLD calculates bivariate variable–variable and variable–covariate residuals that can be used to detect which pairs of observed variables are more strongly related. Therefore, bivariate residuals greater than 3.84 were included iteratively for each model to identify significant correlations between the associated variable–variable and variable–covariate pairs inside each class (for one d.f., bivariate residuals greater than 3.84 indicate statistical significance at the 0.05 level). As implemented in Latent GOLD, individuals are assigned posterior membership probabilities for belonging to each cluster based on their symptom profiles. Cases are then assigned to the cluster for which the posterior membership probability is highest. Based on this assignment, we compared cluster membership to our DSM-IV-based best-estimate clinical diagnoses of ADHD. Effects of covariates like dose, gender and age were controlled. We formed clusters based on two procedures. First, we used the traditional procedure based on ratings from just the OFF medication condition, which generated 7 clusters (subgroups). Then, we re-established clusters based on parent ratings from the ON medication condition, which identified 6 clusters. Then, by estimating conditional LCC probabilities of transition, we were able to define probabilities of effectiveness of stimulant treatment in ADHD. Significance of the transition between clusters was empirically obtained using 100 000 simulations assuming the multinomial distribution. Second, we used parent ratings from both the ON and OFF medication conditions to establish clusters. Based on these two approaches we designated each cluster of ADHD patients to being either responder or not responder to stimulant treatment.

The response trait was tested for association to the candidate SNPs using the TDT. Frequencies and proportions were estimated for categorical variables. Means and standard deviations were calculated for continuous variables. Categorical variables were compared using the χ^2-test. Continuous variables fitting both normality and homogeneity of variances were compared using the *t*-test for independent samples; otherwise they were tested using the non-parametric Mann–Whitney's *U*-test. Normality and variance homogeneity were tested with the Shapiro–Wilks and the Bartlett tests, respectively. Cohen's d-effect size of non-overlapped data was estimated for all variables using pooled variances.

Results

Gene discovery

The next stage of the positional cloning method of our research program was based on the analysis of fine mapping. Partitioned and joint family-based association,[44] marker,[45] and haplotype-based[26] analyses revealed, after correction by multiple comparisons, a significant area of association with ADHD, delimited by the SNP markers rs1901223 and rs1355368 ($P = 3.1 \times 10^{-3}$, marker based; $P = 2.7 \times 10^{-5}$, haplotype based) (Figure 1a) (Supplementary Tables 2, 3, and 4). The region of association is located on chromosome 4 at 62.4–62.7 Mb (UCSC coordinates) within the *LPHN3* gene encompassing exons 4 through 19 (Figure 1b). Empirical analysis of meiotic events in ADHD-affected subjects from linked families pointed to the same *LPHN3* gene region (Figure 1c). The ADHD region of association involves sequences coding for the olfactomedin, hormone receptor, the G-protein-coupled receptor proteolytic site, and transmembrane domains of *LPHN3* (Figure 1d), and contains the bulk of the variability conferred by *LPHN3* splice isoforms (Figure 1b).

Figure 2a shows three markers that passed a test of heterogeneity and were significant after adjusting for multiple tests: rs6551665 (OR = 1.23, 95% CI 1.09–1.37, $P = 3.46 \times 10^{-4}$) and rs1947274 (OR = 1.23, 95% CI 1.09–1.38, $P = 5.41 \times 10^{-4}$) and rs2345039 (OR = 1.21, 95% CI 1.08–1.35, $P = 8.97 \times 10^{-4}$) (Table 2, Figure 2b). Marker rs6551665 was genotyped in all of the samples. In the Hapmap CEPH sample, the Paisa, and the US1 sample, these three markers belong to a common LD block. Despite our conservative approach (the correction of statistical significance due to multiple testing for this gene), *P*-values reported here have strongly overcome estimated threshold values.

Screening for sequence variations

We sequenced all *LPHN3* exons and intron–exon boundaries and identified two SNPs of potential significance for the expression profile of LPHN3. Both SNPs are in the haplotype block associated with ADHD and are located at splice sites. One SNP, rs1397548 (G > A; major allele in Caucasians: G), is located at the last base of the last codon of exon 15. Even though this base change itself does not change the translation (synonymous change, that is, p.P937P), it may affect proper splicing. This splice site variant decreases the splice site score from 93 to 80, suggesting a decrease of

Table 2 Meta-analysis results for three significant markers after correction for multiple comparisons

	Single nucleotide polymorphism ([T:UT] / OR / 95%CI / weight)		
Population	**rs6551665 [G:A]**	**rs1947274 [C:A]**	**rs2345039 [C:G]**
PAISA	[45:22]/2.0454/1.2283–3.4054/14.7761	[33:19]/1.7368/0.9876–3.0538/12.0576	[41:27]/1.5185/0.9342–2.4680/16.2794
US1	[98:63]/1.5555/1.1335–2.1345/38.3478	[89:61]/1.4590/1.0533–2.0207/36.1933	[99:72]/1.375/1.0149–1.8625/41.6842
US2	[151:134]/1.1268/0.8930–1.4219/70.9964	[127:113]/1.1238/0.8722–1.4480/59.7958	[148:127]/1.1653/0.9193–1.4770/68.3490
GERMANY	[96:84]/1.1428/0.8527–1.5316 /44.8	[89:78]/1.1410/0.8419–1.5463 /41.5688	[88:77]/1.1428/0.8417–1.5516 /41.0666
NORWAY[a]	[271/397:245/433]/1.2064/0.9681–1.5032 /79.3648	[281/397:253/431]/1.2057/0.9697–1.4991/80.9692	[311/391:283/431]/1.2113/0.9805–1.4963 /86.0086
SPAIN	[127:115]/1.1043/0.8581–1.4212 /60.3512	[120:99]/1.2121/0.9289–1.5816 /54.2465	[109:98]/1.1122/0.8466–1.4610 /51.6038
Fixed-effects meta-analysis (OR/95%CI/*P*-value)	**1.2258/1.0964–1.3705/3.46E-4**	**1.2275/1.0929–1.3786/5.41E-4**	**1.2094/1.0810–1.3530/8.97E-4**

[a]Case–control data.

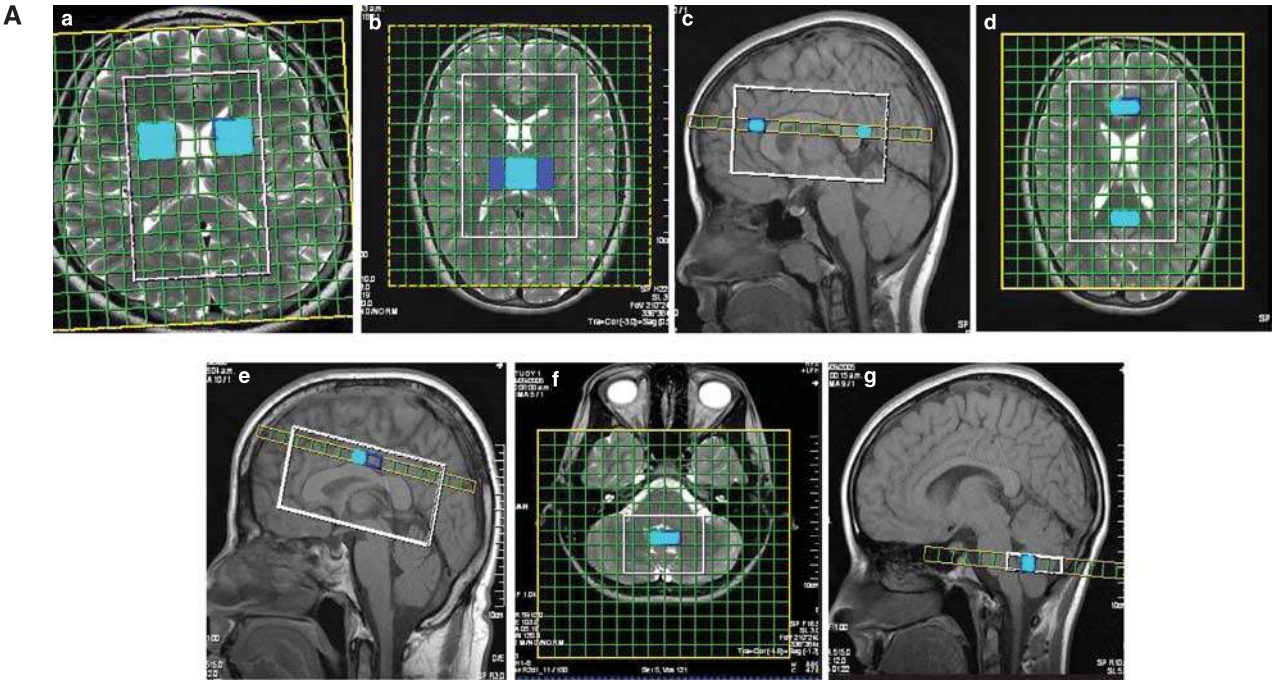
the recognition of this altered splice site leading to a skipping of exon 15.[46] The deletion of exon 15, containing 210 bases, is an in-frame deletion and would generate a shortened protein with potentially altered functions. The other SNP, rs2305339 (A > G; major allele in all populations studied so far: A), is located at the −4 position of intron 10. Sequence variations at this position could also cause aberrant splicing, but are less well-defined. Regular splicing and some aberrant splicing may result, with generation of a regular full length product and potentially an mRNA with a deletion of exon 11. A deletion of exon 11, containing 184 bases, is out of frame and would create a premature stop codon at amino acid position 643. Thus, this variant could generate a certain degree of haploinsufficiency, but importantly it does not predict a complete loss of function for this protein. While preliminary, corroborating the functional effect of these variants will require additional studies.

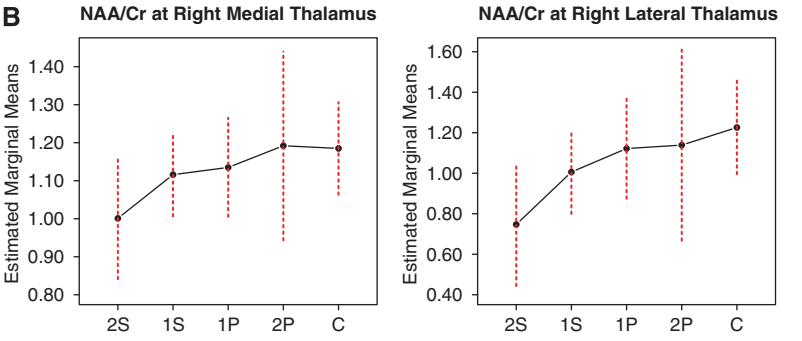
Expression of LPHN3

Northern blot analysis showed significant expression of *LPHN3* mRNA in amygdala, caudate nucleus, cerebellum and cerebral cortex (Figure 3A, lanes 1, 2, 7, 8). Lower expression was detected in corpus callosum, hippocampus, whole brain extract, occipital pole, frontal lobe, temporal lobe, and putamen (Figure 3A, lanes 3, 4, 5, 11, 12, 13, 14). No expression was detected in thalamus, medulla and spinal cord (Figure 3A, lanes 6, 9, 10). *LPHN3* expression examined by *in situ* hybridization of formalin-fixed tissues of brain regions from humans of different ages showed a strong cytoplasmic signal in neurons of the amygdala, caudate nucleus, pontine nucleus and in Purkinje cells of the cerebellum at all ages tested, that is, at 2, 5, 8 and 30 years; tissues of a 2 year old are shown in Figure 3B. Involvement of the prefrontal cortex, cerebellum, amygdala[47] and temporal lobes has been implicated in ADHD.[48] Weak cytoplasmic signals for *LPHN3* were observed in a subset of cingulate gyrus neurons in the 2- and 5 year olds, but not in the 8- and 30 year olds, and in indusium griseum neurons in the 2 year old (data not shown). Areas of the brain that revealed *LPHN3* expression by *in situ* hybridization were also consistently immunoreactive using anti-LPHN3 antibody (Figure 3C).

Brain chemistry

For ^{1}H-MRS we chose target regions (striatum, cingulate gyrus, and cerebellar vermis) based on previous evidence of anatomic abnormalities and added a region that has been difficult to quantify volumetrically (the medial and lateral thalamus[48]). The voxel placement is shown in Figure 4A. MRS was performed on a total of 33 individuals selected to differ on the haplotype defined by variants of SNPs rs6551665, rs1947274 and rs2345039, with 15 individuals (ADHD-affected = 13, ADHD-unaffected = 2) having the risk haplotype, 10 individuals (ADHD-affected = 1, ADHD-unaffected = 9) with the protective haplotype, and eight individuals (all unaffected) with





C

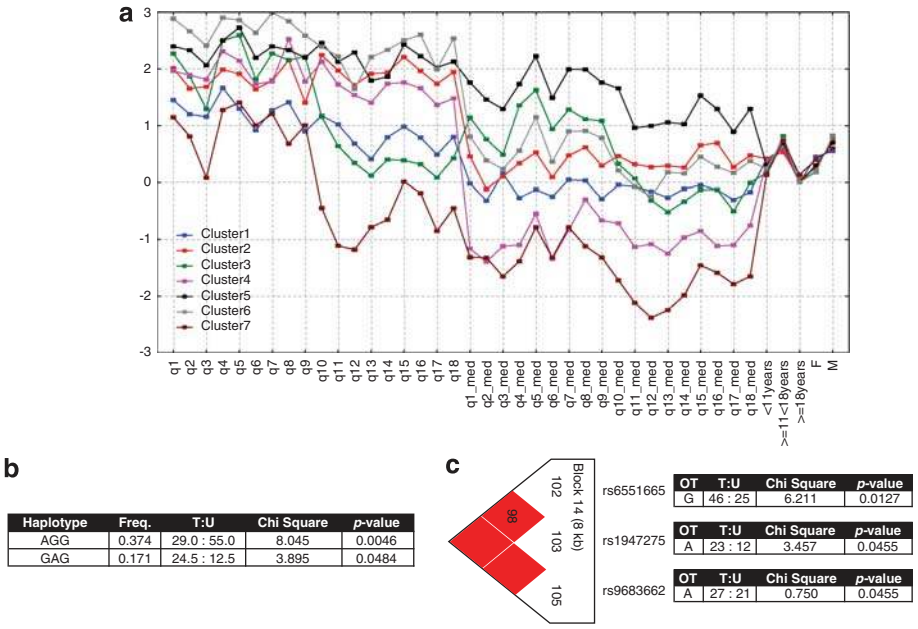
GLM term	Estimate	Std. Error	Pr(>|t|)
Intercept	1.4850	0.2846	0.0000
Sex	-0.1548	0.3795	0.6869
log(Age)	-0.0718	0.0852	0.4080
Haplotype (Yes: 1; No: 0)	6.4989	2.4889	0.0153
ADHD Status (Yes: 1, No: 0)	0.0556	0.0895	0.5408
Sex*log(Age)	-0.0001	0.1121	0.9992
Sex*Haplotype	-6.6810	2.5141	0.0138
log(Age)*Haplotype	-1.8001	0.6638	0.0122
Sex*log(Age)*Haplotype	1.8547	0.6726	0.0110

Figure 4 (**A**) ^{1}H-MRS brain imaging. Turquoise squares show voxel selection at different brain structures for which ^{1}H-MRS ratios were acquired. (a) right and left striatum (3–4 voxels); (b), right and left lateral (2 voxels, blue) and medial (2 voxels) thalamus; (c and d), right and left anterior (1 voxel) and posterior (1 voxel) cingulate gyrus; e, right and left mid-cingulate gyrus (1 voxel); (f and g), posterior–inferior vermis (2 voxels). (**B**). Plots of marginal means for levels of NAA/Cr at the right medial and lateral thalamus (*y* axis), showing an inverse relationship between the NAA/Cr ratio and the number of copies of the haplotype carrying variants of susceptibility of the markers rs6551665, rs1947274 and rs2345039. On the *x* axis are points for 2S (two copies of the susceptibility haplotype), 1S (one copy of the susceptibility haplotype), 1P (one copy of the protective haplotype), 2P (two copies of the protective haplotype), and controls (no copies of either haplotype). (**c**) Adjacent to the plot is the GLM incorporating the effects of sex, haplotype carrier status, affection status and age.

haplotype variants differing from the susceptibility and protective haplotype variants. Across all individuals carrying the haplotype with susceptibility variants, NAA/Cr was significantly decreased in the left lateral thalamus ($P<0.01$), left medial thalamus ($P<0.05$), and the right striatum ($P<0.05$), and significantly increased in the inferior–posterior cerebellar vermis[49] ($P<0.05$). To control for ADHD status, we examined the effect of the number of copies of haplotypes containing susceptibility and protective variants on the NAA/Cr ratio. The NAA/Cr ratio increased monotonically in the right medial and lateral thalamus in relation to the number of copies of the protective haplotype (Figure 4B). Carriers of two copies of the susceptibility haplotype had the lowest levels of NAA/Cr; the ratio increased in a dosage-dependent fashion, reaching the highest levels in individuals with two copies of the protective haplotype and in controls, that is, those individuals with variants different from the susceptibility and protective haplotypes ($P<0.05$).

Response to stimulant medication

ADHD individuals belonging to clusters 2, 4, 6 and 7 in Figure 5a were defined as responders to stimulant treatment. No significant difference was found between responders and non-responders for demographic or clinical variables such as sex, age,



Haplotype	Freq.	T:U	Chi Square	*p*-value
AGG	0.374	29.0 : 55.0	8.045	0.0046
GAG	0.171	24.5 : 12.5	3.895	0.0484

OT	T:U	Chi Square	*p*-value
G	46 : 25	6.211	0.0127

OT	T:U	Chi Square	*p*-value
A	23 : 12	3.457	0.0455

OT	T:U	Chi Square	*p*-value
A	27 : 21	0.750	0.0455

Figure 5 (**a**) Profile plot of clusters derived from latent class cluster analyses applied to symptoms obtained by the SWAN Scale for ADHD Symptoms screened while being OFF and ON medication. Attention deficit (Inattention) items abbreviated for caption OFF medication (q1-q9): q1-careless, inattentive, q2-sustains attention poorly, q3-appears to not listen, q4-poor follow through, q5-disorganized, q6-avoids/dislikes sustained mental effort, q7-loses needed objects, q8-easily distracted, and q9-often forgetful. Hyperactivity/impulsivity items (q10–q18): q10-fidgets or squirms, q11-cannot stay seated, q12-restless, q13-loud, noisy, q14-always 'on the go', q15-talks excessively, q16-blurts out, q178-impatient and q18-intrusive. Similar sequence is followed for items while being ON medication, for example, (q1_med–q18_med). Age categories and sex are also presented. Clusters 2, 4, 6 and 7 were defined as those where the stimulant treatment was effective. (**b**) Marker and haplotype wise association analyses of variants associated to ADHD susceptibility to stimulant medication response. (**c**) LD block containing the SNP rs6551665 and two additional surrogate markers, rs1947275 and rs9683662: OT: over transmitted, T: Transmitted, U: not-transmitted.

ADHD subtype and type of stimulant medication used for the treatment of ADHD symptoms (Supplementary Table 6, Supplementary information). Significant association, both at the marker ($P<0.05$) and haplotype ($P<0.01$) level, was found for response to stimulant medication and SNP marker rs6551665 (which was also associated with ADHD) (Figure 5b and c). In order to control for a possible confounder effect of the LPHN3-associated SNPs to ADHD in any pharmacogenetic protocol, we designed an experiment looking for association between genotype and baseline SWAN scores to dissect any effect correlated with specific symptom dimensions of ADHD. Thus far, a quantity Δ quantifying the change in the SWAN scale was defined as follows:

For the question i and the individual j, we calculated:

$$\Delta_{i,j} = |Q_{i,j,ON} - Q_{i,j,OFF}|$$

$$i = 1,2,\ldots,18,\ j = 1,2,\ldots,240,$$

which compares SWAN scale scores for all of the questions in each individual when they were ON vs OFF medication. Given $\Delta_{i,j}$ we performed a Likelihood Ratio Test (LRT) between two models: Model 1: $\Delta \sim Sex + Age + Genotype$, where Genotype refers to G/G, G/A and A/A of rs6551665, being A/A the reference genotype for comparisons, and Model 2: $\Delta \sim Sex + Age$. The LRT discloses a significant improvement of the model while including the genotype variable (data not presented). To determine the significance and direction of the Genotype effect, an ANOVA for Model 1 was conducted. Significant positive genotype effects were identified for questions 2, 3 and 9 of the Inattention component (Q2: $F_{(2,160)}=4.96$, $P=0.008093$; Q3: $F_{(2,160)}=4.32$, $P=0.01487$; Q9: $F_{(2,160)}=4.80$, $P=0.009424$) as well as for question number 18 of the H/I component (Q18: $F_{(2,160)}=3.53$, $P=0.03167$). We concluded that genotypes with either one or two copies of the G allele had a positive effect on Δ, and the strongest response to medication.

A similar approach was used to quantify the Genotype effect on both the Inattention and Hyperactivity/Impulsivity (H/I) dimensions of the SWAN scale. A Likelihood Ratio Test (LRT) between two models: Model 1: $\Delta_{Inattention} \sim Sex + Age + Genotype$

Model 2: $\Delta_{Inattention} \sim Sex + Age$, where $\Delta_{Inattention} = |S_{Inattention,ON} - S_{Inattention,OFF}|$, with $S_{Inattention,ON}$ and $S_{Inattention,OFF}$, are, respectively, the sum of the SWAN scores for the Inattention questions (questions 1–9) ON and OFF medication. Analogously, we performed an LRT test to compare models with the same structure as

above but using $\Delta_{H/I} = |S_{H/I,ON} - S_{H/I,OFF}|$. Here, $S_{H/I,ON}$ and $S_{H/I,OFF}$ are, respectively, the sum of the SWAN scores for the H/I questions (questions 10–18). Significant positive genotype effects were identified for the Inattention dimension ($F_{(2,159)} = 3.71$, $P = 0.02646$) but not for the H/I dimension ($F_{(2,159)} = 2.41$, $P = 0.09726$).

Discussion

The research program that was initiated in 2002[14] and produced intermediate results soon after[15] was successful in locating and characterizing a novel gene that contributes to the genetic etiology of ADHD. The unbiased linkage scan identified an ADHD susceptibility region on chromosome 4.[15,16] Here, we provide the results of fine mapping that narrowed the region to the coding sequence of *LPHN3* (exons 4 through 19) that contains important functional domains and most of the splice isoform variability of the gene.

Iterative and expanded analyses of multiple generations of ADHD families from a genetic isolate in Colombia provided a consistent and reliable presence of association and linkage at the marker and haplotype level of LPHN3 variants. Association with ADHD was replicated in US, German and Spanish families and in a case–control analysis of a Norwegian sample.

The effect was larger in the Paisa sample than in the samples used for replication. This may reflect the unique characteristic of the Paisa isolate or the method of ascertainment of the sample, which was from families with a high prevalence of ADHD (about 30%) and a subtype highly comorbid with conduct disorder and substance use.[4,16] The replicated association suggests that regulation of *LPHN3* expression may be involved in the pathogenesis of ADHD. *LPHN3* is a member of the *LPHN* subfamily of G-protein-coupled receptors (GPCRs). *LPHNs* have seven transmembrane regions as well as long N-terminal extracellular sequences containing a 19-amino acid signal peptide (GPCR proteolytic site, GPS domain), and a serine/threonine-rich glycosylation region[50] (Figure 1d). *LPHN1* and *LPHN2* serve as receptors for α-latrotoxin, a component of the venom of the black widow spider (*latrodectus mactans*); α-latrotoxin interacts with neuronal GPCRs to stimulate exocytosis of GABA-containing presynaptic vesicles.[51] *LPHN3* is the most brain-specific *LPHN*[50,52]; in fact, other GPCRs, such as DRD4 and DRD5, have been associated directly with ADHD.[53]

The association of *LPHN3* with increased risk *(OR ~ = 1.2)* points to new neuro-molecular mechanisms related to ADHD. This increased risk appears consistent with the role of *LPHN3* in neuronal transmission and maintenance of neuron viability. In fact, the dosage of the *LPHN3* susceptibility haplotype varied inversely with the ratio of NAA/Cr, a measure of the neuronal number thought to be abnormal in ADHD. Moreover, mice lacking the closely related *LPHN1*, although viable and fertile, attend poorly to their offspring, resulting in increased neonatal mortality.[54]

The interpretation of the 1.2 OR figure must be placed in the context of its potential impact for the clinical and epidemiological practice when dealing with ADHD in the general population. We think that the Population Attributable Risk (PAR), as a measure of epidemiological impact, can provide a figure at a glance about the consequences (prevalence and outcome) of an association between an exposure factor (*LPHN3* common variant conferring susceptibility) and a disease (ADHD) at the population level.[55–57] Specifically, the PAR defines the proportion of ADHD cases that could be treated if it were possible to control for the effects of the *LPHN3* common variant conferring susceptibility to ADHD. The PAR is a function of the relative risk and the probability of exposure (Pe) given that a person has the disease. Because family-based samples provide OR instead of relative risk, and for a highly prevalent disorder, such as ADHD, the OR is not a good estimator of relative risk, we calculated PAR% for marker rs65511665 with the case–control-based sample from Norway, as proposed by Hildebrandt *et al*.[58] The 95% confidence intervals were estimated as $\hat{\theta} \pm z_{1-\alpha/2}\sqrt{\hat{\sigma}^2_{\hat{\theta},B}}$ where $\hat{\theta}$ is the *PAR%* point estimation, $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution, and $\hat{\sigma}^2_{\hat{\theta},B}$ is the simulation-based variance of $\hat{\theta}$ given that B random samples are drawn. To approach this variance, a total of $B = 1\,00\,000$ samples were generated from a multinomial distribution as proposed elsewhere.[55,56] Thus far, the *PAR%* for the marker rs6551665 in the Norway sample is 8.99 (95% CI = 3.90–14.12) meaning that by controlling the effect of the *LPHN3* common variant conferring susceptibility to ADHD, a reduction of ~9% in the ADHD incidence in the Norway population would be observed. Though this figure is encouraging, we anticipate that additional population-based epidemiological studies would be needed, as well as clinical and pharmacogenetic trials, to confirm the impact of this gene in the general population.

The spatial and temporal expression of *LPHN3* also supports its role in the pathogenesis of ADHD. *LPHN3* is expressed in regions of the brain implicated in ADHD (that is, the amygdala, caudate nucleus, pontine nucleus and cerebellar Purkinje cells[48]) and at earlier ages in human post-mortem brains[59] (that is, at a time in brain development when ADHD is considered to emerge[48]).

It is very intriguing that the same variant (SNP marker rs6551665) associated with susceptibility to ADHD is also associated with response to stimulant medication. This opens a window for the evaluation of molecular substrates of ADHD and development of new drugs targeting new genes and brain pathways involved in ADHD.

Conflict of interest

None of the authors at the National Institutes of Health has any conflict of interest. Authors from other

institutions contributed samples for replication studies only.

Acknowledgments

We are grateful to the families who participated in this research. This research was supported by the Division of Intramural Research, NHGRI, NIH and in part by COLCIENCIAS, Grants 1115-04-12010, 11150418083, and by the Deutsche Forschungsgemeinschaft (KFO 125/1-1, SFB 581). This study used the high-performance computational capabilities of the SGI Origin 2000 system at the Center for Information Technology, National Institutes of Health, Bethesda, MD and the services of the NHGRI Genomics Core under the supervision of Marypat S Jones and Chandra Settara. Darryl Leja provided graphical assistance to the figures. N Steigerwald provided technical assistance in the DNA German sample processing. Mahim Jain, Sharon B Shively and Sabina Domené are doctoral students in the NIH Graduate Partnerships Program, at Oxford University, George Washington University and the University of Buenos Aires, respectively. Some of this work is to be presented to the above programs in partial fulfillment of the requirements for the PhD degree. Authors from Spain are grateful to Anna Bielsa, Xavier Gastaminza, Rosa Bosch, Monica Fernandez Anguiano and Sílvia Rodríguez-Ben for their participation in the clinical assessment and to Cristina Sánchez-Mora for laboratory assistance. MR is a recipient of a Miguel de Servet contract from the 'Ministerio de Sanidad y Política Social', Spain. Financial support was received from 'Instituto de Salud Carlos III-FIS' (PI041267, PI042010, PI040524, PI080519) and 'Agència de Gestió d'Ajuts Universitaris i de Recerca-AGAUR' (2009GR971). SNP genotyping of the Spanish samples was performed at the Barcelona node of the 'Centro Nacional de Genotipado' (CEGEN; www.cegen.org).

References

1 Biederman J, Faraone SV. Attention-deficit hyperactivity disorder. *Lancet* 2005; **366**: 237–248.
2 Faraone SV, Biederman J, Mennin D, Gershon J, Tsuang MT. A prospective four-year follow-up study of children at risk for ADHD: psychiatric, neuropsychological, and psychosocial outcome. *J Am Acad Child Adolesc Psychiatry* 1996; **35**: 1449–1459.
3 Pelham WE, Foster EM, Robb JA. The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. *J Pediatr Psychol* 2007; **32**: 711–727.
4 Palacio JD, Castellanos FX, Pineda DA, Lopera F, Arcos-Burgos M, Quiroz YT *et al.* Attention-deficit/hyperactivity disorder and comorbidities in 18 Paisa Colombian multigenerational families. *J Am Acad Child Adolesc Psychiatry* 2004; **43**: 1506–1515.
5 Biederman J, Petty CR, Wilens TE, Fraire MG, Purcell CA, Mick E *et al.* Familial risk analyses of attention deficit hyperactivity disorder and substance use disorders. *Am J Psychiatry* 2008; **165**: 107–115.
6 Cook Jr EH, Stein MA, Krasowski MD, Cox NJ, Olkon DM, Kieffer JE *et al.* Association of attention deficit disorder and the dopamine transporter gene. *Am J Hum Genet* 1995; **56**: 993–998.
7 LaHoste GJ, Swanson JM, Wigal SB, Glabe C, King N, Kennedy JL *et al.* Dopamine D4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. *Mol Psychiatry* 1996; **1**: 121–124.
8 Thapar A, Langley K, Owen MJ, O'Donovan MC. Advances in genetic findings on attention deficit hyperactivity disorder. *Psychol Med* 2007; **37**: 1681–1692.
9 Li D, Sham PC, Owen MJ, He L. Meta-analysis shows significant association between dopamine system genes and attention deficit hyperactivity disorder (ADHD). *Hum Mol Genet* 2006; **15**: 2276–2284.
10 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516–1517.
11 Neale BM, Lasky-Su J, Anney R, Franke B, Zhou K, Maller JB *et al.* Genome-wide association scan of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**: 1337–1344.
12 Maher BS, Riley BP, Kendler KS. Psychiatric genetics gets a boost. *Nat Genet* 2008; **40**: 1042–1044.
13 McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. *Hum Mol Genet* 2008; **17**: R156–R165.
14 Arcos-Burgos M, Castellanos FX, Lopera F, Pineda D, Palacio JD, Garcia M *et al.* Attention-deficit/hyperactivity disorder (ADHD): feasibility of linkage analysis in a genetic isolate using extended and multigenerational pedigrees. *Clin Genet* 2002; **61**: 335–343.
15 Arcos-Burgos M, Castellanos FX, Pineda D, Lopera F, Palacio JD, Palacio LG *et al.* Attention-deficit/hyperactivity disorder in a population isolate: linkage to loci at 4q13.2, 5q33.3, 11q22, and 17p11. *Am J Hum Genet* 2004; **75**: 998–1014.
16 Jain M, Palacio LG, Castellanos FX, Palacio JD, Pineda D, Restrepo MI *et al.* Attention-deficit/hyperactivity disorder and comorbid disruptive behavior disorders: evidence of pleiotropy and new susceptibility loci. *Biol Psychiatry* 2007; **61**: 1329–1339.
17 Acosta MT, Castellanos FX, Bolton KL, Balog JZ, Eagen P, Nee L *et al.* Latent class subtyping of attention-deficit/hyperactivity disorder and comorbid conditions1. *J Am Acad Child Adolesc Psychiatry* 2008; **47**: 797–807.
18 Elia J, Ambrosini P, Berrettini W. ADHD Characteristics: I. Concurrent Co-morbidity Patterns in Children & Adolescents. *Child Adolesc Psychiatry Ment Health* 2008; **2**: 15.
19 Walitza S, Renner TJ, Dempfle A, Konrad K, Wewetzer C, Halbach A *et al.* Transmission disequilibrium of polymorphic variants in the tryptophan hydroxylase-2 gene in attention-deficit/hyperactivity disorder. *Mol Psychiatry* 2005; **10**: 1126–1132.
20 Johansson S, Halleland H, Halmoy A, Jacobsen KK, Landaas ET, Dramsdahl M *et al.* Genetic analyses of dopamine related genes in adult ADHD patients suggest an association with the DRD5-microsatellite repeat, but not with DRD4 or SLC6A3 VNTRs. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**: 1470–1475.
21 Ribases M, Ramos-Quiroga JA, Hervas A, Bosch R, Bielsa A, Gastaminza X *et al.* Exploration of 19 serotoninergic candidate genes in adults and children with attention-deficit/hyperactivity disorder identifies association for 5HT2A, DDC and MAOB. *Mol Psychiatry* 2009; **14**: 71–85.
22 Service S, DeYoung J, Karayiorgou M, Roos JL, Pretorious H, Bedoya G *et al.* Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet* 2006; **38**: 556–560.
23 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
24 Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics graphics quality assessment for gene mapping data. *Bioinformatics* 2005; **21**: 3445–3447.
25 Zollner S, Wen X, Hanchard NA, Herbert MA, Ober C, Pritchard JK. Evidence for extensive transmission distortion in the human genome. *Am J Hum Genet* 2004; **74**: 62–72.
26 Durrant C, Zondervan KT, Cardon LR, Hunt S, Deloukas P, Morris AP. Linkage disequilibrium mapping via cladistic analysis of single-nucleotide polymorphism haplotypes. *Am J Hum Genet* 2004; **75**: 35–43.

27 Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906–913.
28 Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.
29 Kazeem GR, Farrall M. Integrating case-control and TDT studies. *Ann Hum Genet* 2005; **69**: 329–335.
30 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–188.
31 Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; **22**: 719–748.
32 Carrey N, MacMaster FP, Fogel J, Sparkes S, Waschbusch D, Sullivan S *et al.* Metabolite changes resulting from treatment in children with ADHD: a 1H-MRS study. *Clin Neuropharmacol* 2003; **26**: 218–221.
33 Moore CM, Biederman J, Wozniak J, Mick E, Aleardi M, Wardrop M *et al.* Differences in brain chemistry in children and adolescents with attention deficit hyperactivity disorder with and without comorbid bipolar disorder: a proton magnetic resonance spectroscopy study. *Am J Psychiatry* 2006; **163**: 316–318.
34 Sun L, Jin Z, Zang YF, Zeng YW, Liu G, Li Y *et al.* Differences between attention-deficit disorder with and without hyperactivity: a 1H-magnetic resonance spectroscopy study. *Brain Dev* 2005; **27**: 340–344.
35 Yeo RA, Hill D, Campbell R, Vigil J, Brooks WM. Developmental instability and working memory ability in children: a magnetic resonance spectroscopy investigation. *Dev Neuropsychol* 2000; **17**: 143–159.
36 Jin Z, Zang YF, Zeng YW, Zhang L, Wang YF. Striatal neuronal loss or dysfunction and choline rise in children with attention-deficit hyperactivity disorder: a 1H-magnetic resonance spectroscopy study. *Neurosci Lett* 2001; **315**: 45–48.
37 Provencher SW. Automatic quantitation of localized *in vivo* 1H spectra with LCModel. *NMR Biomed* 2001; **14**: 260–264.
38 Castillo M, Smith JK, Kwock L. Correlation of myo-inositol levels and grading of cerebral astrocytomas. *AJNR Am J Neuroradiol* 2000; **21**: 1645–1649.
39 Swanson JM, Elliott GR, Greenhill LL, Wigal T, Arnold LE, Vitiello B *et al.* Effects of stimulant medication on growth rates across (3 years in the MTA follow-up. *J Am Acad Child Adolesc Psychiatry* 2007; **46**: 1015–1027.
40 Vermunt JK, Magidson J. LatentGOLD. [3 01]2003. Statistical Innovations, Inc. 2-10-2003.
41 Magidson J, Vermunt JK. Latent class analysis. In: *Handbook of quantitative methodology for the social sciences*, Kaplan D (ed). Sage Publications: Newbury Park, CA, USA, 2003.
42 Vermunt JK, Magidson J. Latent class cluster analysis. In: *Applied latent class analysis*, Hagenaars JA & McCutcheon AL (eds) Cambridge University Press, Cambridge, UK, 2002; 89–106.
43 Hagenaars JA. Latent structure models with direct effects between indicators: local dependence models. *Sociol Meth Res* 1988; **16**: 379–405.
44 Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006; **38**: 209–213.
45 Martin ER, Bass MP, Gilbert JR, Pericak-Vance MA, Hauser ER. Genotype-based association test for general pedigrees: the genotype-PDT. *Genet Epidemiol* 2003; **25**: 203–213.
46 Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics functional implications in gene expression. *Nucleic Acids Res* 1987; **15**: 7155–7174.
47 Plessen KJ, Bansal R, Zhu H, Whiteman R, Amat J, Quackenbush GA *et al.* Hippocampus and amygdala morphology in attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 2006; **63**: 795–807.
48 Krain AL, Castellanos FX. Brain development and ADHD. *Clin Psychol Rev* 2006; **26**: 433–444.
49 Anderson CM, Polcari AM, Lowen SB, Renshaw PF, Teicher MH. Effects of methylphenidate on functional magnetic resonance relaxometry of the cerebellar vermis in children with ADHD. *Am J Psychiatry* 2002; **159**: 1322–1328.
50 Sugita S, Ichtchenko K, Khvotchev M, Sudhof TC. alpha-Latrotoxin receptor CIRL/latrophilin 1 (CL1) defines an unusual family of ubiquitous G-protein-linked receptors. G-protein coupling not required for triggering exocytosis. *J Biol Chem* 1998; **273**: 32715–32724.
51 Linets'ka MV, Storchak LH, Himmelreich NH. Effect of synaptosomal cytosolic [3H]GABA pool depletion on secretory ability of alpha-latrotoxin]. *Ukr Biokhim Zh* 2002; **74**: 65–72.
52 Ichtchenko K, Khvotchev M, Kiyatkin N, Simpson L, Sugita S, Sudhof TC. alpha-latrotoxin action probed with recombinant toxin: receptors recruit alpha-latrotoxin but do not transduce an exocytotic signal. *EMBO J* 1998; **17**: 6188–6199.
53 Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA *et al.* Molecular genetics of attention-deficit/hyperactivity disorder. *Biol Psychiatry* 2005; **57**: 1313–1323.
54 Tobaben S, Sudhof TC, Stahl B. Genetic analysis of alpha-latrotoxin receptors reveals functional interdependence of CIRL/latrophilin 1 and neurexin 1 alpha. *J Biol Chem* 2002; **277**: 6359–6365.
55 Benichou J. Methods of adjustment for estimating the attributable risk in case-control studies: a review. *Stat Med* 1991; **10**: 1753–1773.
56 Benichou J, Gail MH. Methods of inference for estimates of absolute risk derived from population-based case-control studies. *Biometrics* 1995; **51**: 182–194.
57 Spiegelman D, Hertzmark E, Wand HC. Point and interval estimates of partial population attributable risks in cohort studies: examples and software. *Cancer Causes Control* 2007; **18**: 571–579.
58 Hildebrandt M, Bender R, Gehrmann U, Blettner M. Calculating confidence intervals for impact numbers. *BMC Med Res Methodol* 2006; **6**: 32.
59 Ichtchenko K, Bittner MA, Krasnoperov V, Little AR, Chepurny O, Holz RW *et al.* A novel ubiquitously expressed alpha-latrotoxin receptor is a member of the CIRL family of G-protein-coupled receptors. *J Biol Chem* 1999; **274**: 5491–5498.

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