Molecular Psychiatry (2006) 11, 934–953

www.nature.com/mp

ORIGINAL ARTICLE

The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: association signals in *DRD4, DAT1* and 16 other genes

K Brookes[1], X Xu[1], W Chen[1], K Zhou[1], B Neale[1], N Lowe[2], R Aneey[2], B Franke[3], M Gill[2], R Ebstein[4], J Buitelaar[3], P Sham[1], D Campbell[1], J Knight[1], P Andreou[1], M Altink[3], R Arnold[5], F Boer[6], C Buschgens[3], L Butler[2], H Christiansen[7], L Feldman[8], K Fleischman[1], E Fliers[3], R Howe-Forbes[1], A Goldfarb[8], A Heise[9], I Gabriëls[10], I Korn-Lubetzki[11], R Marco[12], S Medad[8], R Minderaa[13], F Mulas[12], U Müller[14], A Mulligan[2], K Rabin[8], N Rommelse[15], V Sethna[1], J Sorohan[2], H Uebel[9], L Psychogiou[16], A Weeks[16], R Barrett[16], I Craig[1], T Banaschewski[9], E Sonuga-Barke[16], J Eisenberg[8], J Kuntsi[1], I Manor[8], P McGuffin[1], A Miranda[12], RD Oades[7], R Plomin[1], H Roeyers[10], A Rothenberger[9], J Sergeant[15], H-C Steinhausen[14], E Taylor[1], M Thompson[16], SV Faraone[17], P Asherson[1] and L Johansson[1]

[1]*MRC Social Genetic Developmental and Psychiatry Centre, Institute of Psychiatry, London, UK;* [2]*Department of Psychiatry, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland;* [3]*Department of Psychiatry, Radboud University, Nijmegen Medical Center, Nijmegen, The Netherlands;* [4]*S Herzog Memorial Hospital, Jerusalem, Israel;* [5]*Triversum, Alkmaar, Holland;* [6]*Amsterdam Medical Centre de Bascule, Amsterdam, Holland;* [7]*University Clinic for Child and Adolescent Psychiatry, Essen, Germany;* [8]*ADHD Clinic, Geha Mental Health Center, Petak-Tikvah, Israel;* [9]*Child and Adolescent Psychiatry, University of Göttingen, Göttingen, Germany;* [10]*Departments of Experimental Clinical Health Psychology, Ghent University, Ghent, Belgium;* [11]*Shaare Zedek Medical Center, Jerusalem, Israel;* [12]*Department of Developmental and Educational Psychology, University of Valencia, Valencia, Spain;* [13]*University Centre for Child and Adolescence Psychiatry, University Medical Centre, Groningen, The Netherlands;* [14]*Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland;* [15]*Vrije Universiteit, De Boelelaan, Amsterdam, Holland;* [16]*School of Psychology, University of Southampton, Highfield, Southampton, UK and* [17]*Departments of Psychiatry and Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA*

Attention deficit hyperactivity disorder (ADHD) is a common neurodevelopmental disorder, starting in early childhood and persisting into adulthood in the majority of cases. Family and twin studies have demonstrated the importance of genetic factors and candidate gene association studies have identified several loci that exert small but significant effects on ADHD. To provide further clarification of reported associations and identify novel associated genes, we examined 1038 single-nucleotide polymorphisms (SNPs) spanning 51 candidate genes involved in the regulation of neurotransmitter pathways, particularly dopamine, norepinephrine and serotonin pathways, in addition to circadian rhythm genes. Analysis used within family tests of association in a sample of 776 DSM-IV ADHD combined type cases ascertained for the International Multi-centre ADHD Gene project. We found nominal significance with one or more SNPs in 18 genes, including the two most replicated findings in the literature: DRD4 and DAT1. Gene-wide tests, adjusted for the number of SNPs analysed in each gene, identified associations with TPH2, ARRB2, SYP, DAT1, ADRB2, HES1, MAOA and PNMT. Further studies will be needed to confirm or refute the observed associations and their generalisability to other samples.
Molecular Psychiatry (2006) **11,** 934–953. doi:10.1038/sj.mp.4001869; published online 8 August 2006

Keywords: ADHD; association study; dopamine; noradrenaline; serotonin; candidate gene; linkage disequilibrium

Introduction

Attention deficit hyperactivity disorder (ADHD) is a common and highly heritable neurodevelopmental disorder[1] that affects 3–10% of children and 2–4% of adults,[2,3] depending on the measure used and the population sampled. The disorder is characterised by childhood onset of age-inappropriate and impairing hyperactivity, impulsivity and inattentiveness. ADHD is frequently accompanied by developmental problems, including dyslexia and specific and general learning difficulties and high levels of comorbidity with antisocial, mood, anxiety and substance use

Correspondence: Professor P Asherson, MRC Social Genetic Developmental Psychiatry, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK.
E-mail: p.asherson@iop.kcl.ac.uk
Received 6 March 2006; accepted 5 June 2006; published online 8 August 2006

disorders. ADHD children are at high risk of negative long-term outcomes including academic underachievement, unemployment, high accident rates and difficulty sustaining stable relationships.[4] Recognition and appropriate treatment of ADHD across the lifespan is therefore of considerable importance.[5,6]

The familial risk for ADHD is established with an estimated sibling risk ratio (λ_s = risk to siblings of ADHD probands/population risk) for broadly defined ADHD of around three- to four-fold. λ_s values for DSM-IV combined subtype are estimated to be slightly higher, between five- to six-fold.[7] The analysis from Todd *et al*.[7] suggests that DSM-IV combined type ADHD may be a genetically homogeneous subgroup, since this subtype fell within a single empirically derived latent class that showed high levels of subtype concordance in their data set of monozygotic and dizygotic twins. A review of 20 twin studies found the heritability of ADHD to be 76%.[8] Group heritabilities, derived from De Fries and Fulker (DF) regression analysis with various thresholds on ADHD symptom scales, suggests that genetic influences on ADHD are continuously distributed throughout the population. Both individual differences and DF approaches to bivariate twin analysis find a high level of shared genetic effects between ADHD and various comorbid disorders/traits including dyslexia,[9,10] conduct disorders and general cognitive ability.[11]

To date, there have been four published genome linkage scans of ADHD; three affected sibling pair studies and one of 16 multiplex pedigrees.[12–18] These studies have highlighted a number of potential linkage regions for further exploration, although there is as yet no clear consensus across the various data sets and no genes have been identified that account for linkage signals. Several of the linkage regions overlap in two or more of these studies, including regions of chromosomes 5p, 6q, 7p, 11q, 12q and 17p, suggesting that one or more loci of moderately large effect may exist.

Meta-analyses of candidate gene studies have successfully identified several genes of small effect associated with ADHD. The genes identified so far are involved in the regulation of catecholamine and other neurotransmitter pathways and follow on from *a priori* hypotheses of the role of dopamine pathways in particular. Both children and adults with ADHD show a rapid therapeutic response to methylphenidate and amphetamine, which is mediated by blockade of dopamine re-uptake with a subsequent increase in synaptic dopamine.

A role for the serotonin system is suggested since it is known to interact with dopamine in the control of impulsive behaviour,[19] and mediates stimulant response of hyperactivity in dopamine transporter knockout mice.[20] The therapeutic effects of atomoxetine, a specific noradrenergic reuptake inhibitor, directly implicate norepinephrine. Candidate genes suggested by these observations include genes that code for catecholamine neurotransmitter receptors, enzymes involved in the metabolism and degradation of the neurotransmitters and genes that regulate the vesicular release of neurotransmitters (synaptosomal proteins),[21] and influence neuronal membrane fluidity and responsiveness (fatty acid desaturases).[22]

Meta-analysis of genetic variants reported to be associated with ADHD in three or more studies found average odds ratios (OR) across studies to be in the order of 1.2–1.4. The genetic variants listed in this study include the 7-repeat allele of an exon 3 variable number tandem repeat (VNTR) in the dopamine D4 receptor gene (DRD4), a common allele of a microsatellite marker in the vicinity of the dopamine D5 receptor gene (DRD5), the 10-repeat allele of a VNTR marker in the 3′UTR region of the dopamine transporter gene (DAT1), the long repeat of a VNTR in the promotor region of the serotonin transporter gene (HTTLPR) and single-nucleotide polymorphisms (SNPs) located in the synaptosomal associated protein (SNAP-25) and the serotonin 5HT1B receptor gene (HTR1B). Assuming a simple additive model the combined effect size across the associated genetic variants is estimated to contribute around 3–4% to phenotypic variance leaving a substantial portion of the additive genetic influences on ADHD to be explained (Kuntsi *et al.*, under review).

Another group of candidate genes that have yet to be investigated are the circadian rhythm genes. The rationale for their inclusion in genetic studies of ADHD is that insomnia occurs in up to 54% of cases, pointing to a possible disturbance of circadian rhythms.[23] The circadian system comprises three key components: endogenous oscillators (clocks) generating a circadian rhythm, input pathways entraining the circadian rhythm to the astrophysical day, and output pathways distributing signals from the oscillator to the periphery. Animal models suggest a possible role of circadian rhythm genes in ADHD symptoms via regulation of acetylcholine release in the hippocampus, which is associated with attention, learning and memory.[24]

To further investigate genes previously associated with ADHD and extend the analysis to additional candidate genes, we completed an in depth analysis of 51 genes in a European multisite sample of ADHD combined type probands, siblings and parents. The study used the first set of 674 families collected for the International Multisite ADHD Gene project (the IMAGE project). We selected high priority candidate genes on the basis of prior evidence from pharmacological, neurobiological and genetic investigations.[25] By taking a 'biological systems' approach to candidate gene selection, including the most promising genes identified so far, we aimed to maximise our chances of detecting multiple associations with genes known or hypothesised to participate in the regulation of neurotransmitters pathways (dopamine, norepinephrine and serotonin), and circadian rhythms. We aimed to take advantage of contemporary high-density SNP maps to provide a comprehensive coverage of

each gene, by screening both tagging SNPs selected from HapMap,[26] and SNPs within known functional regions.

Subjects and methods

Subjects and sample collection

European Caucasian subjects were recruited from 12 specialist centres in eight countries: Belgium, Germany, Holland, Ireland, Israel, Spain, Switzerland and United Kingdom. Ethical approval for the study was obtained from National Institute of Health registered ethical review boards for each centre. Detailed information sheets were provided and informed consent obtained from the majority of children and from all of their parents. All ADHD probands and their siblings were aged 5–17 years at the time of entry into the study and access was required to one or both biological parents for DNA collection. Entry criteria for probands were a clinical diagnosis of DSM-IV combined subtype ADHD and having one or more full siblings available for ascertainment of clinical information and DNA collection. Exclusion criteria applying to both probands and siblings include autism, epilepsy, IQ <70, brain disorders and any genetic or medical disorder associated with externalising behaviours that might mimic ADHD.

The final data set used in this study consisted of 674 DSM-IV combined type probands with 808 siblings. One hundred and two of the siblings also had combined type ADHD making a total data set of 776 affected individuals. DNA was available for both parents in 614 families (90%) and from one parent in 49 families (7%). 93.5% of the ADHD cases were male subjects. The age range for both probands and siblings ranged from 5 to 17 years with the mean age of 11.2 (s.d. = 2.7) for probands and 11.2 (s.d. = 3.1) for siblings.

Since we evaluated all available siblings ($n = 808$), including a set of DSM-IV combined type ADHD cases ($n = 102$), we estimated the λ_s value for DSM-IV combined subtype to be around five using the population prevalence of 0.024 estimated from a recent survey in the United Kingdom.[27] Prior to entry into the study, all probands underwent clinical evaluations by a paediatrician or child psychiatrist and both existing and new patients were included in the study. Wherever possible, families withdrew stimulant medication for 1 week prior to research assessments to allow for more accurate ascertainment of the current level of ADHD symptoms and behaviours. Alternatively, we ensured as far as possible that ratings were based on medication-free periods. Probands were excluded from the study if the last medication free period was more than 2 years ago.

Clinical procedures

Parental account of childhood symptoms (PACS). PACS is a semistructured, standardised, investigator-based interview developed as an instrument to provide an objective measure of child behaviour.[28] A trained interviewer administers PACS with parents, who are asked for detailed descriptions of the child's typical behaviour in a range of specified situations. Such situations are defined either by external events (e.g. watching television, reading a book or comic, playing alone, playing with friends, going to bed, traveling) or by behaviours shown (e.g. crying, worried talk, tempers, fighting with siblings). Interviewers then make their own ratings on the basis of a formal training and written definitions of the behaviours to be rated, on a four-point scale of severity and frequency in the previous week and previous year. Inter-rater reliability is high with product-moment correlations for pairs of interviewers ranging from 0.79 to 0.96. The Hyperactivity Subscale is made up of attention span (time spent on a single activity, rated separately for four different kinds of activity), restlessness (moving about during the same activities), fidgetiness (movements of parts of the body during the same activities), and activity level (rated for structured situations such as mealtimes and car journeys), with other subscales covering defiant, emotional and other comorbid disorders including autistic spectrum disorders.

Rating scales. Rating scales used to quantify ADHD symptoms included the Long Version of Conners' Parent Rating Scale (CPRS-R:L), Long Version of Conners' Teacher Rating Scale (CTRS-R:L),[29] parent version of the Strengths and Difficulties Questionnaires (SDQ) and teacher version of SDQ.[30] In order to exclude autism spectrum disorders that might confound the analysis of ADHD, both probands and siblings were screened using the Social Communication Questionnaire ($\geqslant 15$) in conjunction with the pro-social scale from the SDQ ($\leqslant 4$). Individuals falling above these thresholds were further evaluated using the autism spectrum disorder section of the PACS interview.

DSM-IV diagnoses. All raw data is centralised and stored on a secure database at the MRC Social Genetic Developmental Psychiatry research centre in London. A standardised algorithm is applied to PACS to derive each of the 18 DSM-IV ADHD items, providing operational definitions for each behavioural symptom. These are combined with items that scored 2 or 3 from the teacher rated Conners' ADHD subscale, to generate the total number of items from the DSM-IV symptom checklist. Situational pervasiveness was defined as some symptoms occurring within two or more different situations from the PACS interview, as well as the presence of one or more symptoms scoring 2 or more from the ADHD subscale of the teacher rated Conners'. Other ADHD symptom scales and measures of co-morbidity were not used in the diagnostic algorithm for the analyses in this paper.

DNA collection and genotype assays. Blood samples were sent immediately following collection by an International Courier service to Rutgers University Cell and DNA repository, New Jersey (RUCDR). These were either used to generate lymphocyte cell lines from which DNA was extracted, or DNA was extracted directly from a portion of the blood sample and lymphocytes cryopreserved for future recovery. In a few cases where individuals were not able to supply a blood sample, we used a mouth swab sampling technique and extracted the DNA at the SGDP laboratories in London. DNA stocks for the entire data set were collated in London where they were stored, organised and plated out for further analysis. Geneservice Ltd, Cambridge (UK) performed whole-genome amplification on all samples with $<100\,\mu$g stock DNA, using the REPLI-g kit (Quiagen Ltd, Crawley, UK). DNA samples were arrayed into 96-well plates at a concentration of $50\,$ng/μl and delivered to Illumina Inc. (San Diego, USA) under dry ice. In addition to the IMAGE sample, we included the panel of 30 CEPH trios of Northern European ancestry used in the HapMap project.[26]

All the SNP genotype assays were completed on a custom array using the Illumina Golden Gate Assay™ (Illumina Inc., San Diego, USA). In addition, we included the analysis of VNTR polymorphisms from exon three of the DRD4, the 3′-untranlsated region and intron 8 of the DAT1, and promotor region of the serotonin transporter (SERT) genes, which had previously been associated with ADHD.[8,28] Standard PCR protocols were used for all VNTR markers and amplified products visualised on 2% agarose under UV light, as previously described.[31–33]

SNP selection

We adopted a 'biological systems' approach by nominating 46 genes that were likely to exert an effect through regulation of dopamine, serotonin and norepinephrine neurotransmission, as well as six circadian rhythm genes. Selected genes fell into the following functional groups: dopamine receptors, serotonin receptors, norepinephrine receptors, neurotransmitter metabolic and catabolic enzymes, neuronal transporters, synaptic vesicle associated proteins, fatty acid desaturase enzymes and circadian rhythm genes (listed in Table 1). The selected genes included 23 that had previously been reported to contain polymorphic variants associated with ADHD in one or more studies, in addition to 28 genes that have yet to be investigated in ADHD or had not shown an association signal in previous studies.

For SNP selection we aimed at a comprehensive analysis of each gene using two main approaches. First, the direct association analysis of SNPs that fall in functional regions and have an increased chance of functional significance by virtue of their location. For this approach we targeted SNPs located within coding regions (synonymous and nonsynonymous), 5′ and 3′ untranslated regions, intron sequences within 300 bp from intron/exon boundaries, and one SNP per kb covering 5 kb upstream from the start of transcription or known 5′ regulatory regions including the promotor.

Second, we aimed to use indirect association to screen for association with common genetic variation across each gene. For this method we targeted a nonredundant set of tagging SNPs (tSNPs) that correlated highly with SNPs with minor allele frequency (MAF) $\geqslant 0.05$, using the CEPH panel from the HapMap database. We selected tSNPs using two methods. We used the 'CompleteLink routine' within CLUSTAG that implements a hierarchical clustering algorithm (http://hkumath.hku.hk/web/link/CLUSTAG/CLUSTAG.html.[34] Hierarchical clustering starts with a square matrix of pair-wise distances (as defined by $1\text{-}R^2$) between the SNPs to be clustered. The rationale is that the noncentrality parameter (NCP) for a tSNP is the product of the NCP at the causal SNP and the R^2. In order to ensure a high chance of detecting indirect association, we used a threshold for R^2 of $\geqslant 0.8$. The second method used for tSNP selection was the default algorithm in Haploview (http://www.broad.mit.edu/mpg/haploview) taken from Gabriel *et al.* (2002).[35] 95% confidence bounds on D-prime are generated and each comparison is called 'strong LD', 'inconclusive' or 'strong recombination'. A block is created if 95% of informative comparisons fall in the 'strong LD' category. The default setting for this algorithm sorts the list of all possible blocks, starting with the largest and adding blocks as long as they don't overlap with an already declared block. 39.6% (132 out of 437 tSNPs) of the CLUSTAG SNPs and 71.8% (241 out of 437 tSNPs) of the Gabriel method SNPs were shared. 7.0% (77 out of 1105, functional SNPs) were included in the tSNP criteria. To avoid redundancy of the marker information, where the two methods recommended selection of two different SNPs that fell within the same cluster defined by CLUSTAG, we preferentially selected the tSNP recommended by the CLUSTAG algorithm. If due to the constraints of the Illumina technology the tSNPs nominated by CLUSTAG could not be genotyped an alternative SNP located in the same cluster was selected.

SNP genotyping was completed using the Illumina high-throughput BeadArray™ technology (http://www.illumina.com). This set limits on the SNP selection, based on estimates from Illumina of the genotype success rate for each SNP. Data supplied by Illumina included a list of SNPs and their estimated genotype success rate for each gene, with an overall estimate of 'success rate' that takes into account both SNP validation status and designability for the array platform. In addition, markers less than 60 bp apart could not both be genotyped; when this occurred we selected the SNP with the highest estimation of overall success rate, followed by highest MAF. For tSNPs selected from HapMap, there were often additional SNPs that fell within a SNP cluster that could be selected in preference to one with a low

Table 1 The majority of genes to be included in this SNP screen have previously been investigated in relation to ADHD

Gene	*Replication status*	*Location (Chr)*	*Length (kb)*	*No. of SNPs*	*SNP density (SNPs/kb)*	*No. of SNPs (tagging)*	*No. of SNPs (functional)*
ADRA1A	*0*	*8p21*	117.3	41	2.86	17	24
ADRA1B	*0*	*5q33*	60.8	15	4.05	7	8
ADRA2A	+	*10q24*	8.6	9	0.96	1	8
ADRA2C	+	*4p16*	7.8	7	1.11	1	6
ADRB1	*0*	*10q24*	1.7	4	0.43	2	2
ADRB2	*0*	*5q32*	2	11	0.18	4	7
ADRB3	*0*	*8p12*	3.7	8	0.46	0	8
ADRBK1	*0*	*11q13*	20.2	7	2.89	4	3
ADRBK2	*0*	*22q11*	159	16	9.94	7	9
ARRB1	*0*	*11q13*	23.9	13	1.84	6	7
ARRB2	*0*	*17p13*	11	6	1.83	2	4
BDNF	+	*11p13*	71.8	20	3.59	8	16
CHRNA4	+	*20q13*	21.7	19	1.14	5	15
COMT	+	*22q11*	32.2	30	1.07	8	22
CSNK1E	*0*	*22q12*	31.8	17	1.87	10	8
DBH	+	*9q34*	28	33	0.85	18	15
DDC	+	*11p17*	90.7	41	2.21	19	22
DRD1	+	*5q31*	81.2	11	7.38	3	8
DRD1IP	+	*10q26*	16.5	6	2.75	4	3
DRD2	+	*11q23*	70.5	23	3.07	11	12
DRD3	−	*3q13*	55.2	28	1.97	9	19
DRD4	+	*11p15*	8.4	4	2.10	1	3
FADS1	−	*11q12*	17.2	7	2.46	2	7
FADS2	+	*11q12*	44.1	19	2.32	11	12
HES1	*0*	*3q28*	5.2	5	1.04	3	2
HES6	*0*	*2q37*	12.5	4	3.13	4	0
HTR1B	+	*6q13*	6.2	11	0.56	5	7
HTR1E	*0*	*6q14*	84	20	4.20	14	11
HTR2A	+	*13q14*	67.7	32	2.12	22	13
HTR3B	*0*	*11q23*	46.7	15	3.11	6	12
NET1	+	*16q12*	16.8	43	0.39	13	34
NFIL3	*0*	*9q22*	1.8	6	0.30	5	1
NURR1	*0*	*2q22*	15	6	2.50	5	4
PER1	*0*	*17p12*	4	13	0.31	8	5
PER2	*0*	*2q37*	48.7	19	2.56	8	13
PNMT	+	*17q21*	2.2	2	1.10	0	2
VMAT2	−	*10q25*	41.4	30	1.38	10	23
SLC6A1	*0*	*3p25*	51.5	35	1.47	16	22
DAT	+	*5p15*	57.6	32	1.80	12	27
SERT	+	*17q11*	42.8	14	3.06	7	9
SLC9A9	*0*	*3q21*	588.2	70	8.40	122	62
SNAP25	+	*20p11*	93.6	98	0.96	20	22
STX1A	−	*7q11*	14.9	38	0.39	4	8
SYT1	−	*12cen*	238.5	9	26.5	18	17
TPH1	−	*11p15*	24.8	14	1.77	5	9
TPH2	+	*12q15*	98.6	40	2.47	11	29
VAMP2	−	*17p12*	8.8	5	1.76	0	5
HTR2C	+	*Xp24*	331	23	14.39	0	23
MAOA	+	*Xp11*	95.6	16	5.98	0	16
MAOB	−	*Xp11*	120.8	10	12.08	0	10
SYP	+	*Xp11*	17.4	1	17.4	0	1

Replication status: '0' = not previously investigated; '+' = previous positive association reported; '−' = previous negative association reported. We investigated a total of 925 SNPs with minor allele frequency $\geqslant 0.02$, spanning a total region of 3.121 kb with an average SNP density of 1 SNP every 3.4 kb.
ADHD, attention deficit hyperactivity disorder; SNP, single-nucleotide polymorphism.

estimate of success rate. We set a lower limit for MAF of 0.05; however, we included a high proportion of nonvalidated SNPs and SNPs with unknown heterozygosity from 'functional' regions.

We grouped the selection criteria for SNPs as follows: Group 1 consisted of SNPs with known heterozygosity and previously genotyped on the Illumina platform; Group 2 consisted of SNPs with

known heterozygosity but not genotyped previously on the Illumina platform; Group 3 consisted of SNPs validated by proven 2hit-2-allele (every allele has been observed in at least two chromosomes; Group 4 consisted of other SNPs at various levels of validation. Illumina provided estimates of overall success on their platform: with a high designability rating Group 1 SNPs had a 90–95% success rate, Group 2 a 80–85% success rate, Group 3 a 30–50% success rate, and Group 4 an unknown success rate.

Of the 1536 SNPs included in the SNP array, there were 173 assays that completely failed across all samples. An additional 34 SNP assays failed only in the whole-genome amplified (WGA) DNA. Remaining genotypes from WGA DNA had similar genotype error, heterozygosity values and Hardy–Weinberg Equilibrium (HWE) distribution to nonamplified DNA, indicating a failure of the WGA method to amplify approximately 2% of loci, but no apparent allelic drop out or loss of quality compared to nonamplified DNA. GenCall quality control scores supplied by Illumina were higher than 0.25 in 99.78%, indicating a high average quality score with no samples dropping below the recommended quality control threshold of 0.2–0.25. As expected, SNP MAFs varied within the various validity groups used for selection of the remaining 1363 SNPs: MAF $\geqslant$5% were found in 88% of Group I SNPs, 86% of Group 2 SNPs, 21% of group 3 SNPs, and 20% of Group 4 SNPs. Using NCBI validation criteria (proven by cluster; proven by frequency; proven by submitter; proven by 2hit-2allele), 75.7% had MAF$\geqslant$2% and 70.2% MAF$\geqslant$5%, whereas for NCBI nonvalidated SNPs only 16.4 and 11.3% were polymorphic at the same MAF thresholds. The overall success rate was 925 (67.1%) SNPS with MAF$\geqslant$, and 860 (63.1%) with MAF$\geqslant$5%.

Pedigree error checks

Ten samples out of the total of 2937 samples ($<$0.04%) could not be genotyped even after multiple attempts and were removed from the data set. For the remaining samples (not including the 34 markers that failed on the WGA DNA) the average genotype drop out rate was 0.02%.

To investigate the random genotyping error rate, we included 15 duplicate DNA samples in the screening panel, which produced more than 40 000 genotype calls (1,363). All the duplicated genotypes were 100% consistent and the random genotyping error rate could therefore be estimated as lower than 0.005% (1/20 000). To investigate the systematic genotyping error rate on the Illumina platform, 30 CEPH trios screened in the HapMap project were also genotyped using our SNP panel. By comparison of 879 SNPs genotyped in both projects, there were 0.12% (92 out of 77 871) genotype inconsistencies. Since more than 95% of the HAPMAP genotypic data were not replicated on different platforms we estimate that half of the inconsistencies (0.06%) are due to systemic errors on our platform. The overall genotyping error (sum of the random and systematic error rates) would therefore be lower than 0.065% (0.06 + 0.005).

By running PEDCHECK for single point Mendelian inconsistencies on 987 polymorphic autosomal markers in the initial data set of 680 families, we identified 47 families with potential pedigree errors. Within these families the numbers of Mendelian errors range from 16 to 312 and could not be explained by genotyping errors. To identify the real problem of these pedigree errors which might be DNA swap, nonreported non paternity, nonreported half-sibs or any other unexpected errors, we performed the following steps and checked the reduction of Mendelian errors after each step: (1) Randomly assign any two subjects from a family as parents. (2) Drop one subject at a time from a family. (3) Insert erroneous subjects back to any empty place left by step 2. Using this procedure we identified 16 DNA swaps, eight nonpaternity and seven half-sib cases, which cleared up the Mendelian errors from 40 families. This left seven families with unknown pedigree errors that we removed from the analysis. The final data set for association analysis included 674 families that contained 156 sporadic Mendelian errors from 987 autosomal markers. The overall detection rate of Mendelian errors is therefore 0.02%, which is consistent with the estimated efficiency of SNP markers (13–75%, to detect such errors by identification of Mendelian errors and an estimated overall genotype error rate less than 0.065%). The large number of pedigree errors may be explained by the complicated procedures inherent within a large international cooperative process.

Hardy–Weinberg equilibrium (HWE)

We tested HWE in all common SNPs (MAF$\geqslant$0.05) using Pedstats. HWE tests were performed on groups consisting of IMAGE parents and offspring. Overall we identified 16 SNPs in the parent group, and 35 SNPs in the offspring group, with HWE $P<0.01$. Two of these SNPs were shared across the two groups and were removed from the analysis. One possible explanation for the increased rate of HW disequilibrium in the offspring is that association is confounded with the HWE test (Supplementary Table 1).

Association tests

The main analyses used transmission disequilibrium tests (TDT) implemented using the UNPHASED www.rfcgr.mrc.ac.uk/~fdudbrid/software/unphased) and WHAP (www.genome.wi.mit.edu/~shaun/whap) programs.

Single marker tests of association. For single marker analysis, we used a nominal approach, and two permutation tests to adjust for various multiple testing issues at the gene-wide level.[36] Although the processes of using more than one main statistic could create an additional multiple-testing problem, we consider that the increased level of information,

provides fuller knowledge of the association signal within each gene:

(1) *Nominal P-values*: Nominal *P*-values were obtained using different implementations of the TDT test in UNPHASED and WHAP. UNPHASED has the advantage of generating transmission ratios for each SNP and handles both autosomal and X-chromosome markers. Significance values calculated by UNPHASED was used to rank order the genes on the basis of the most significant SNP within each gene. WHAP was used as a secondary test to provide additional information about the association signal and confirmation that both methods worked properly; but does not handle X-chromosome markers. Significance values in WHAP are empirically derived by permutation and are therefore robust to asymptotic theory. We used 500 permutations of the data and 5,000 permutations for *P*-values <0.05.
(2) *Global-P*: For each gene, we derived the maximum SNP association from UNPHASED, adjusted for all SNPs within each gene. Global-*P*-values are determined in UNPHASED to compare the maximum observed SNP score in the observed data against each maximum of the SNP score in the permuted data set.
(3) *P-SUM values*: We derived the global evidence for association across each gene by summation of SNP *P*-values across each gene. *P*-SUM values were determined using WHAP to compare the sum of *P*-values for all SNPs within each gene, with the sum of the observed scores in the permuted data set. *P*-SUM significance levels may therefore reflect either single or multiple clusters of associated SNPs across a gene.

Multimarker tests of association. We decided *a priori* to restrict haplotype analyses to the investigation of genes that showed evidence for association with single SNP markers, in order to minimise type I error arising from performing multiple tests. We therefore restricted ourselves to genes that contained one or more SNPs achieving nominal significance ($P<0.05$). The overall aim of these analyses was to use these data to delineate as far as possible the putative associated regions.

For haplotype considerations, we adopted two main approaches. We applied the default block definition in HAPLOVIEW (http://www.broad.mit.edu/mpg/haploview) using the method of Gabriel *et al*.[35] In this method 95% confidence bounds on D prime are generated for each pair-wise comparison. An SNP block is formed if 95% of informative comparisons are in 'strong LD'. Although this method allows for multiple overlapping blocks, the default method is to identify the largest blocks and to avoid overlapping blocks. Haplotype tests of association were run using the UNPHASED program on blocks of SNP markers identified in HAPLOVIEW. To explore the data further, we were also interested to establish whether the analysis of consecutive SNPs that may not always be in strong LD with each other, could provide additional information. For this analysis we calculated 5-SNP haplotype associations using the sliding window method implemented in the WHAP program.

Results

The final data set consisted of 674 families with 776 affected individuals with DSM-IV combined subtype ADHD of which 598 (88.7%) of families had both parents, 73 (10.83%) had only one parent, and three (0.45%) had no parents. A summary of the final set of genes and SNPs used in this analysis is listed in Table 1.

Single marker analyses

A summary of the maximum significance values for each gene is provided in Figure 1 (detailed summary of *P*-values from UNPHASED and WHAP in online Supplementary Tables 2a, b and 3) and for the most significant genes with nominal *P*-values $\leqslant 0.05$ in Table 2. We identified six genes with nominal significant SNP scores $\leqslant 0.01$ and an additional 12 genes with nominal significant scores >0.01 and $\leqslant 0.05$. Using the Global-*P* gene-wide test of association, we identified three genes with Global-*P*-values $\leqslant 0.1$ (ADRB2, HES1, MAOA) and three additional genes with Global-*P*-values $\leqslant 0.05$ (TPH2, ARRB2,

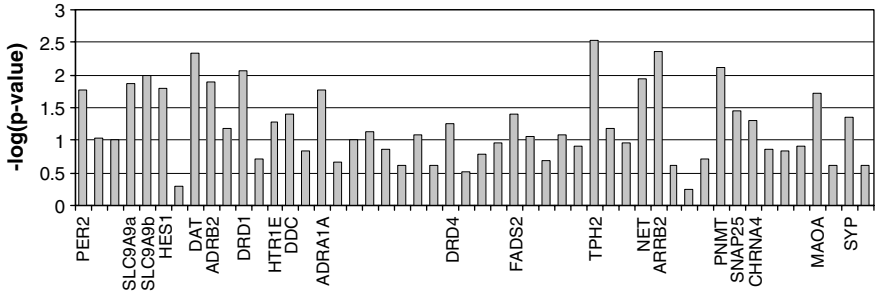

Figure 1 Maximum unadjusted *P*-values for each gene. Genes are listed in map order chromosome 1 through the X chromosome. Genes with one or more SNPs with nominal *P*-values <0.05 are labelled.

Table 2 Table of results for the 18 genes found to be suggestive of association in the SNP screen

Gene	*Nominal* P-*value*	*T*	*NT*	*OR*	*Global* P-*value*	*P_SUM Statistic*
TPH2	0.003	207	151	1.37	0.036	0.106
ARRB2	0.004	103	66	1.56	0.022	0.209
DAT1	0.005	349	278	1.26	0.119	0.014
PNMT	0.008	70	42	1.67	0.012	0.024
SLC9A9	0.01	74	46	1.61	0.485	0.114
NET	0.012	133	95	1.4	0.349	0.786
ADRB2	0.013	210	162	1.3	0.088	0.485
HES1	0.016	300	244	1.23	0.076	0.096
ADRA1A	0.017	283	229	1.24	0.443	0.387
PER2	0.017	31	15	2.07	0.124	0.419
MAOA	0.02	175	134	1.31	0.082	—
SNAP25	0.035	155	120	1.29	0.529	0.198
DDC	0.039	161	126	1.28	0.537	0.597
FADS2	0.039	284	237	1.2	0.389	0.727
SYP	0.045	180	114	1.25	0.034	—
CHRNA4	0.05	116	88	1.32	0.503	0.663
HTR1E	0.051	75	53	1.42	0.509	0.214
DRD4	0.055	34	20	1.7	0.199	0.321

Nominal *P*-value is the most significant SNP from UNPHASED analysis with the relevant number of transmitted (T) and nontransmitted (NT) alleles from heterozygote parents. Odds ratios (OR) and the significance values for the two gene-wide association tests are listed (Global-*P* and *P*-SUM). *P*-SUM could not be calculate for X-chromosome markers

Table 3 Haplotype analysis using 5-SNP sliding window method and analysed using UNPHASED

Gene	*Marker Window*	P-*value*	*T*	*NT*	*OR*	*Haplotype-specific* P-*value*
NET1	16-17-18-19-20	0.005	119	95	1.25	0.101
TPH2	36-37-38-39-40	0.007	206	151	1.36	0.004
PER2	3-4-5-6-7	0.016	188	160	1.18	0.133
ADRB2	4-5-6-7-8	0.024	137	98	1.40	0.011
HTR1E	9-10-11-12-13	0.031	15	8	1.88	0.144
MAOA	12-13-14-15-16	0.033	167	133	1.26	0.050
CHRNA4	11-12-13-14-15	0.046	14	3	7.12	0.008

Genes are ranked by nominal *P*-value ($P<0.05$) for the most significant haplotype block within each gene. SNP marker codes are available from Supplementary Table 2.

SYP). Using the P-SUM gene-wide test the most significant genes were DAT1 ($P<0.02$), PNMT ($P<0.03$), HES1 ($P<0.1$) and TPH2 ($P=0.1$).

We investigated a total of 925 SNPS with MAF $\geqslant$2% and would expect 46 SNPs at the 0.05 level and nine SNPs at the 0.01 level whereas we observed 51 SNPs and 12 SNPs at each significance value respectively. Although numerically greater than expectation, the overall distribution of *P*-values was not significantly increased from expectation. To quantify the probability that we had detected one or more true findings, we ranked the nominal *P*-values from the UNPHASED analysis and applied a false discovery rate of 50%. This identified 10 SNP markers that fell within three genes; TPH2, DAT1 and ARRB2.

Haplotype analyses

The results of the haplotypes tests using the Gabriel method to define haplotype blocks and the sliding window approach are summarised in Table 3, Figure 2 and Supplementary Tables 4 and 5. Using the Gabriel method, we find nominal evidence for association with *ARDRB2* ($P<0.002$), *SNAP-25* ($P<0.004$) and *TPH2* ($P<0.01$). Using the 5-SNP sliding window, we find nominal evidence for association with *CHRNA4* ($P<0.05$), *HTR1A* ($P<0.03$), *PER2* ($P<0.04$), *NET1* ($P<0.005$), *TPH2* ($P<0.01$), *ADRB2* ($P<0.03$) and *MAOA* ($P<0.04$). The majority of haplotype window associations contained SNPs that were significant when considered on their own. Two additional associations were identified with haplotype windows in *NET1* and *CHRNA4* that

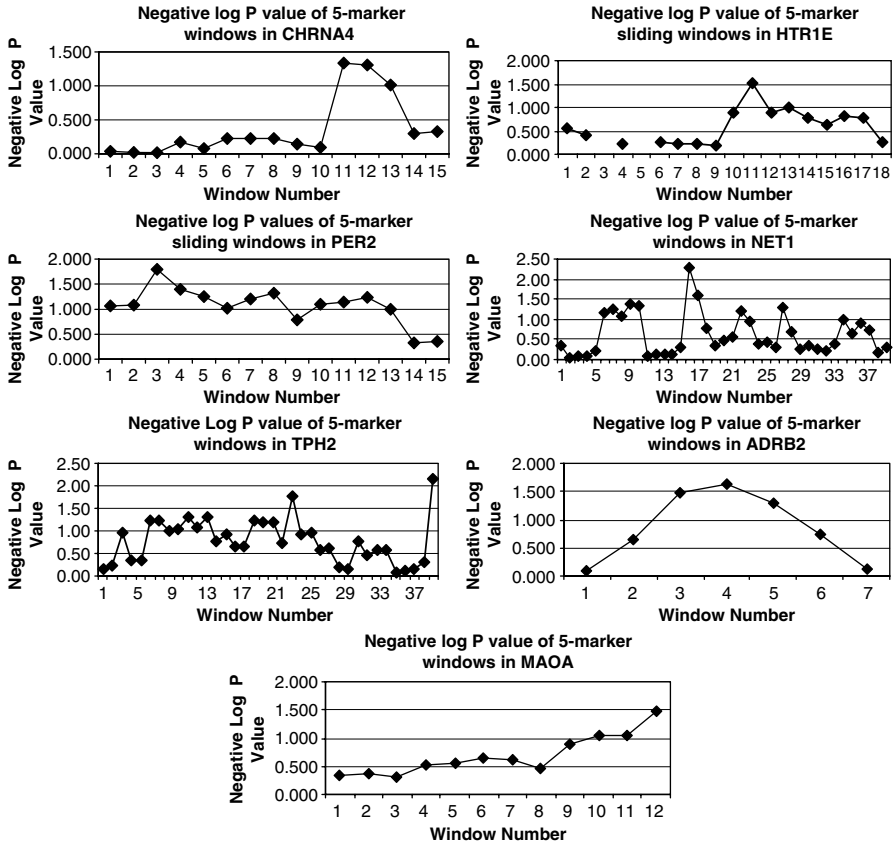

Figure 2 5-SNP sliding window haplotype analysis using UNPHASED (−log *P*-values), for seven genes with nominal haplotype *P*-values <0.05. Sliding Windows run 5′ to 3′ in SNP order. The window numbers are the same as the number of the first SNP within each 5-SNP set. SNP numbers for each gene can be found on online Supplementary Table 2.

contained no SNPs that had displayed evidence for association in the single marker tests.

Analysis of gene regions associated with ADHD

We identified 18 genes that show nominal evidence of association with ADHD from our SNP screen of 674 ADHD families (Table 2). These genes can be divided into three groups; those that have been replicated two or more times, those that have had single reports of association and those that are novel findings. For this purpose we assume that SNP replications should be allele or haplotype specific, since other scenarios are less likely to reflect true associations. These data and the match to our significant SNPs are summarised in Table 4.

Our SNP panel failed to assay some of the SNPs reported to be associated with ADHD in previous studies. For those SNPs that had been genotyped in the HapMap project, the LD maps generated from genotyping our custom SNP array on the HapMap Caucasian CEPH panel, were used to identify a SNP in our panel that acted as a reasonable proxy for indirect association. For this purpose, we integrated our CEPH panel genotypes with those from the HapMap PHASE I data release (Oct 2005) and built an LD map with 1947 additional SNPs within the genomic regions of our scan. In a few cases we were not able to clearly establish the identity of SNPs reported in the literature due to differences in nomenclature.

Multiple reports group

The first group contains genetic associations that have been reported two or more times in the ADHD literature. These genes include *DAT1*, *NET1*, *DRD4*, *MAOA* and *SNAP25*.

DAT1

A finding that has been widely reported is the association between ADHD and the 10-repeat allele of a VNTR in the 3′UTR region of *DAT1*. A nonsignificant overtransmission from heterozygote parents to their affected offspring was observed (allele-specific $P=0.21$), with an odds ratio (OR = 0.11) very close to the OR (1.12) estimated in a recent meta-analysis of available data.[8] More recently investigation of genetic variation in *DAT1* in English and Taiwanese samples suggested that the association is specific to a haplotype of the 10-repeat allele, with the 3-repeat allele of a second VNTR located within intron 8.[31] Analysis of these two markers in this sample revealed the same pattern, with only the 10-3 haplotype showing over transmission from heterozygote parents to their affected offspring (haplotype

Table 4 Table of literature findings for the 18 genes with one or more SNPs with nominal $P \leqslant 0.05$ from this study

Paper	*Gene*	*Polymorphism*	*Location*	*Finding in literature*	*SNP array*	*Finding from this study*
Sheehan *et al.*[53]	TPH2	rs1843809	Intron 5	T-allele	Y	Positive but with opposite allele
		rs1386493	Intron 5	C-allele	Y	Positive but with opposite allele
		rs1386497	Intron 8	A-allele	N	Clusters with rs1007023 ($r^2 = 1$): Positive
		rs2129572	Intron 4	NS	N	
		rs1023990	Intron 7	NS	N	Clusters with rs1386491 ($r^2 = 0.9$): Negative
		rs1487278	Intron 8	NS	N	
		rs1487275	Intron 8	NS	Y	Negative
		rs1487279	Intron 9	NS	N	
Walitza *et al.*[54]		rs4570625	Promoter (−703)	G-allele	Y	Negative
		rs11178997	Promoter (−473)	T-allele	N	
		rs4565946	Intron 2	NS	N	Clusters with rs4448731 ($r^2 = 0.9$): Negative
Brookes *et al.*[21]	SYP	rs2293945	3′UTR	Positive	N	
Brookes *et al.*[55]	FADS	rs498793	Intron 6	Positive	Y	Negative
Mill *et al.*[50,51]	SNAP25	Microsatellite	Intron 1	Positive	—	
		TG repeat	Promoter	Positive	—	
		A-2015T	Promoter	A-allele	N	
		G80609A	Intron 7	G-allele	N	
		rs3746544	3′UTR	T-T-haplotype	N	
		rs1051312	3′UTR	T-T-haplotype	N	
Brophy *et al.*[47]		rs3746544	3′UTR	NS	N	
		rs1051312	3′UTR	T-allele	N	
Barr *et al.*[48]		rs3746544	3′UTR	T-allele (T-C haplotype)	N	
		rs1051312	3′UTR	C-allele (T-C haplotype)	N	
Kustanovich *et al.*[49]		rs3746544	3′UTR	T-allele	N	
		rs1051312	3′UTR	T-C haplotype	N	
Kim *et al.* (unpublished)		rs3787283	Intron 6	Positive	Y	Negative
		rs2327269	outside gene	Positive	N	
Feng *et al.*[63]		rs6039806	Intron 2	C-allele	Y	Negative
		rs362549	Intron 4	A-allele	N	
		rs362987	Intron 4	A-allele	Y	Negative
		rs362998	Exon 5 aa 110	C-allele	N	
		rs1889189	5′ to gene	NS	Y	Negative
		rs362569	5′UTR	NS	N	Clusters with rs6039806 ($r^2 = 0.76$): Negative
		rs362549	Intron 3	NS	N	Clusters with rs362987 ($r^2 = 0.84$): Negative
		rs362986	Intron 4	NS	N	
		rs363006	Intron 6	NS	Y	Negative
		rs3746544	3′UTR	NS	N	
		rs1051312	3′UTR	NS	N	
Bobb *et al.*[38]		Microsatellite	5′UTR	NS	—	
Domschke *et al.*[44]	MAOA	30 bp repeat	Promoter	Positive (short-allele)	—	
		CA repeat	Intron 2	Positive	—	
		G941T	Exon 8	G-allele	—	
		A/G	Intron 12	NS	—	
Xu *et al.*[42]	NET	rs3785157	Intron 7	G-allele	N	Clusters with rs1861647 ($r^2 = 0.93$): Negative
		rs2242447	Intron 13	T-allele (trend)	Y	Negative
		rs998424	Intron 9	g-allele (trend)	N	

Table 4 Continued

Paper	*Gene*	*Polymorphism*	*Location*	*Finding in literature*	*SNP array*	*Finding from this study*
Bobb *et al.*[38]		rs998424	Intron 9	G-allele	N	
		rs3785157	Intron 7	T-allele	N	Clusters with rs1861647 ($r^2 = 0.93$): Negative
Barr *et al.*[64]		SNP	Exon 9	NS	N	
		rs998424	Intron 9	NS	N	
		rs2242447	Intron 13	NS	Y	Negative
McEvoy *et al.*[65]		C1148A	Intron 7	NS	—	
		G1389A	Intron 9	NS	—	
Comings *et al.*[66]		A1970G SNP		Positive	—	
De Luca[67]		G1389A	Intron 9	NS	—	
Kent *et al.*[68]		CHRNA4	Cfol SNP	Exon 5	Positive	—
Bobb *et al.*[38]		rs6090384	Intron 1 or 2	NS	Y	Negative
		rs2273505	Intron 1 or 2	NS	N	
		rs2273506	Exon 1 or 2	NS	Y	Negative
Comings *et al.*[69]		Dinucleotide	Intron 1	Positive	—	
Todd *et al.*[70]		SNP	5′ Intron 2	Positive (Inattent)	—	
Hawi *et al.*[39]	DAT	rs6347	Exon 9	$P = 0.22$	Y	Positive with WHAP: A-allele
		rs11564774	Exon 15	$P = 0.7$	N	
Brookes *et al.*[31]		rs40184	3′UTR	$P = 0.22$	Y	Positive G-allele
		rs1042098	3′UTR	$P = 0.121$	Y	Positive A-allele
		rs27072	3′UTR	$P = 0.28$	Y	Trend with WHAP G-allele
		VNTR × 2	3′UTR-intron 8	10-3 haplotype	—	Positive
Feng *et al.*[37]		rs27072	3′UTR	G-allele	Y	Trend with WHAP G-allele
		rs8179029	Intron 9	NS	Y	Negative
		rs3863145	3′UTR	NS	N	
Bobb *et al.*[38]		rs6347	Exon 9	A-allele (trend)	Y	Positive with WHAP A-allele
Rowe *et al.*[71]	DRD2	Taq1		Positive	—	
Comings *et al.*[72,73]		Taq1		Positive	—	
Todd and Lobos *et al.*[74]		Screened		NS	—	
Huang *et al.*[75]		Taq1		NS	—	
Muglia *et al.*[76]	DRD3	Ser/Gly	Exon 1	NS	—	
Payton *et al.*[56]		A/G	Exon 1	NS	—	
Barr *et al.*[77]		Ser/Gly	Exon 1	NS	—	
		MspI	Intron 5	NS	—	
Lowe *et al.*[40]	DRD4	rs1800955	Promoter	A-allele	N	
		rs747302	Promoter	C-allele	N	
		rs910455	Promoter	NS	N	
Payton *et al.*[56]		rs1800955	Promoter	NS	N	
Bellgrove *et al.*[41]		rs1800955	Promoter	A-allele	N	
Barr *et al.*[62]		rs1800955	Promoter	A-trend	N	
		rs747302	Promoter	C-trend	N	
Hawi *et al.*[57]	DDC	4 bp insertion	Exon 1	Trend	—	
		Microsatellite	3′	$P = 0.04$	—	
		Haplotype		$P = 0.025$	—	
		Microsatellite	5′	NS	—	
Comings *et al.*[78]	DBH	Taq1	Intron 5	NS	—	
Daly *et al.*[79]		Taq1	Intron 5	A2-allele	—	
Wigg *et al.*[80]		Taq1	Intron 5	A2-allele (trend)	—	
Roman *et al.*[81]		Taq1	Intron 5	A2-allele	—	
Inkster *et al.*[82]		Taq1	Intron 5	A2-allele (trend)	—	
Smith *et al.*[83]		Taq1	Intron 5	A1-allele	—	
Zhang *et al.*[84]		C-1021T	Promoter	Trend	—	
Zhang *et al.*[85]		C-1021T	Promoter	Positive	—	
Eisenberg *et al.*[86]	COMT	rs4680	Exon 2 aa 158	Positive (Val)	Y	Negative
Qian *et al.*[87]		rs4680	Exon 2 aa 158	Positive (Met)	Y	Negative

Table 4 Continued

Paper	*Gene*	*Polymorphism*	*Location*	*Finding in literature*	*SNP array*	*Finding from this study*
Bobb *et al.*[38]		rs4680	Exon 2 aa 158	NS	Y	Negative
Thapar *et al.*[88]		rs4680	Exon 2 aa 158	Positive (Val)	Y	Negative
Curran *et al.*[89]	SERT	rs1050565	exon 12 aa 443	$P = 0.0045$	N	
		rs1487871	Intron 3	$P = 0.033$	N	
		rs2020930	Not in gene	$P = 0.035$	Y	
		rs2020937	UTR	$P = 0.092$	N	
		rs2020942	Intron 1	$P = 0.079$	Y	Negative
		rs140701	Intron 7	$P = 0.013$	Y	Negative
		T/G SNP	3′UTR	$P = 0.13$	—	
Kent *et al.*[90]		T/G SNP	3′UTR	T-allele	—	
Li *et al.*[91]	HTR2A	rs6313	Exon 1 aa 34	C-allele	Y	Negative
Bobb *et al.*[38]		rs6313	Exon 1 aa 34	NS	Y	Negative
		rs6311	Locus	NS	Y	Negative
		rs6314	Exon 3 aa 452	NS	Y	Negative
Hawi *et al.*[92]	HTR1B	rs6296		G-allele	Y	Negative
Quist *et al.*[93]		rs6296		G-allele	Y	Negative
Bobb *et al.*[38]		rs6296	Exon 1 aa 287	NS	Y	Negative
		rs6298	Exon 1 aa 43	Trend	Y	Negative
Li *et al.*[94]		rs6296		$P = 0.052$	Y	Negative
Bobb *et al.*[38]	HTR2C	rs6318		NS	Y	Negative
Li *et al.*[95]		C-759T		C-allele	—	
		G-697C		G-allele	—	

The table tabulates the author, reference, gene, rs number is known and gene location. SNP array Y = SNPs included in the Illumina array. Findings in the literature list nominal significant associations with risk allele identified whether possible, NS = nonsignificant, positive = significant but associated allele unknown. Findings from this study are listed in the last column.

specific $P < 0.06$, OR = 1.19); consistent with the 10-3 haplotype association in the previous study.

Two SNPs have been highlighted in earlier studies; rs27072[34] and rs6347.[38,39] We detected a trend with over transmission of the same allele (A-allele) in rs6347 and no association with rs27072. We did however detect additional evidence for association in the 3′UTR, intron 10 and intron 13. LD between the associated SNPs was generally weak with the strongest relationship between rs1042098 and rs3776513 ($r^2 = 0.378$). One of the associated SNPs (rs40184) shows some LD with rs27072 ($r^2 = 0.274$).

No positive findings of association have been reported in the published literature with SNPs in the 5′ flanking region. We observed four SNPs in this region with $P < 0.05$. Three of the associated SNPs (rs2652511, rs10070282 and rs2550946) show very strong LD with an average R^2 around 0.97, while rs11564750 shows weak LD with this 3-SNP cluster (R^2 approximately 0.14).

DRD4

Numerous studies have investigated the association between ADHD and the 7-repeat allele of a highly variable VNTR in exon 3 of *DRD4*. We also detected overtransmission of this allele from heterozygote parents to their affected offspring (allele specific $P < 0.09$) with an OR (1.18) that was very close to the average OR (1.17) estimated in a recent meta-analysis of available data; consistent with a small but significant genetic association.

Several additional markers in DRD4 have been reported to be associated with ADHD including the A-allele of rs180955 in two independent studies[41,41] and weak evidence for association of the C-allele of rs747302; both located in the promotor region. We were unable to genotype these two markers, but we did observe nominal association with rs9195457. We have no information on the LD structure between this marker and the two previously reported SNPs.

NET1

Two previous studies reported association with two SNPs (rs998424, rs3785157) in *NET1* in addition to a third SNP (rs2242447) in one of the studies.[38,42] We genotyped rs2242447 and a proxy marker (rs1861647) for rs3785157 ($r^2 = 0.84$ in CEPH panel). Neither of these two markers showed evidence for association in our sample. The SNPs found to be nominally associated in this study lie within the first (rs3785143) and fifth (rs11568324) introns and should be treated as novel findings.

MAOA

Jiang *et al.*[43] using a Chinese sample reported evidence for association with a dinucleotide repeat located within intron 2, which was followed by a

reported trend for association with this marker in an independent Caucasian sample.[44] A 30 bp VNTR with 2–5 repeats in the promotor region that has been shown to regulate *MAOA* transcription levels and a SNP (G941T) in exon 8 have also been studied. Manor *et al*.[45] found the 4-repeat allele of the VNTR to be associated ($P = 0.036$) with preference for maternal transmission. However, two studies[44,46] identified association with the short alleles of the VNTR and one with the G-allele of the G941T.[44]

In our study we found nominal significance with five SNPs in addition to two SNPs with $P < 0.1$. The region containing these SNPs spans approximately 31 kb from intron 5 to the 3′UTR with an average r^2 of 0.8, suggesting strong LD. This region incorporates G941T reported in one of the earlier studies and may therefore reflect the same association signal; however, we were unable to determine the rs reference number for that SNP.

SNAP-25

Four studies have implicated the 3′UTR region with the association of two SNPs.[47–50] Interpretation of these data has been difficult with various combinations of alleles found to be significant across the three studies: T-allele of rs1051312,[48] T-C haplotype of rs1051312 and rs3746544[49,50] and the T-T haplotype of the two SNPs.[50] We were unable to genotype these two markers for technical reasons. We did however investigate two alternative SNPs in the 3′UTR and did not detect an association signal.

Other SNAP25 markers reported to be associated with ADHD include a SNP in the 5′ flanking region (A-2015T),[50] a microsatellite marker in intron 1[51] and the G-allele of a SNP in intron 7 (G-80609A).[50] More recently SNP associations were reported with several SNPs in introns 2 and 4 and exon 5. We included two of the four associated SNPs reported in this study, neither of which gave an association signal. We did however detect one nominally associated cluster consisting of two SNPs in the 5′UTR (rs363020, rs362567, $r^2 = 0.95$). We were unable to find the A-2015T SNP on NCBI.

Single report group

The second group contains SNP associations with ADHD reported once in the literature. This includes *CHRNA4*, *TPH2*, *FADS2* and *SYP*.[52–55]

CHRNA4

Association between ADHD and *CHRNA4* has previously been reported with two SNPs located in exon 2 and the 5′ end of intron 2.[52] However, we found no evidence for association with these two markers (rs6090384, rs2273506) or from analysis of haplotypes in the region. The SNP associated in our sample is in the 5′ flanking region and is a novel finding.

TPH2

Sheenan *et al*.[53] reported three single marker associations in *TPH2*: the T-allele of rs1843809, the C-allele of rs1386493 and the A-allele of rs1386497. We found association to the first two of these markers; however the risk alleles were not the same in the two studies. A third marker (rs1007023) that was in perfect LD ($R^2 = 1$) with rs1386497 in HapMap was associated in our sample ($P = 0.004$). Across these four SNPs the average R^2 across the region is greater than 0.8. Analysis of haplotypes, while nominally significant, provided no additional evidence for the association. In addition, we investigated one of two SNPs in the 5′-flanking region, reported by Walitza *et al*.,[54] but found no evidence of association.

SYP and FADS2

Association with two SNPs in *SYP* (rs2293945) and *FADS2* (rs498793) were reported in an independent sample of English combined subtype probands;[21,55] Only rs498793 in *FADS2* could be genotyped on the Illumina platform and this did not replicate the earlier finding. The SNP in *FADS2* showing nominal association in this study (rs174611) is located within intron 7 and shows weak LD with rs498793 ($r^2 = 0.1$ in CEPH panel of HapMap). The associated SNP in *SYP* found in this study (rs5906754) is located in the 5′ flanking region, whereas rs2293945 is located in the 3′UTR.

DDC

Association to markers of *DDC* was reported by Hawi *et al*.[57] who observed increased transmission of a 4 bp insertion/deletion in exon 1 and the 213 bp allele of a microsatellite marker (D7S2422) in the 3′-end of the gene.[44] The SNP associated in this study (rs11575454) is located within intron 8; we have no information on LD between these various markers.

Novel finding group

The third group included *HES1*, *PER2*, *SLC9A9*, *PNMT*, *ADRA1A*, *HTR1E*, *ADRB2* and *ARRB2*. None of these findings approach study wide or genome wide levels of significance, so these should all be considered as speculative findings of potential significance until further replication studies have been completed. Of these, *ARRB2* was the strongest novel finding, with a nominal $P < 0.005$ and Global-$P < 0.02$ with an SNP rs7208257 located in intron 5.

P-SUM was significant for *PNMT* though neither SNP was particularly common (MAF $\leqslant 0.05$). The R^2 correlation between these two markers was only 0.001, suggesting that association with these two low frequency SNPs are independent of each other; despite the two markers being located less than 1 kb apart at the 5′-end of the gene. Haplotype analysis of these two markers gives a nominal *P*-value of 0.0016. Conditional testing of the contribution of each of these markers to the haplotype association using WHAP suggested that both loci contribute to the haplotype association.

SLC9A9 displayed nominally associated SNPs in several regions across the gene; intron 14 ($P < 0.04$),

intron 12 ($P<0.02$), intron 6 ($P<0.02$), intron 5 ($P<0.03$), intron 2 ($P<0.05$, and the 5 flanking regions ($P<0.02$). SLC9A9 was the largest gene region we investigated, spanning a total distance of 578 kb and including the analysis of 167 SNPs. Although none of these SNPs were significant when adjusted for the total number of SNPs investigated in this gene, we identified a total of 16 SNPs with $P<0.05$ and 29 with $P<0.1$, compared to the expected number of 8 and 17 SNPs, respectively.

Discussion

We have completed an association screen of 51 candidate genes in a sample of 776 DSM-IV combined type ADHD subjects and their parents. We aimed to provide high coverage of each gene by selecting SNPs within or close to known functional regions, in addition to selecting a set of tSNPs that characterised common variation across each gene. Of 1536 SNPs included in the SNP array, 925 SNPs with low estimated error rate and MAF $\geqslant 0.02$ were included in the final analysis. In addition, to the analysis of SNPs, we investigated association with four VNTR markers that had shown evidence for association from meta-analytic studies in DRD4, DAT1 and SERT.

We highlighted the most promising associations by ranking genes on the basis of nominal *P*-values for the most significant SNPs within each gene; this identified 18 genes containing one of more SNPs with a nominal $P\leqslant 0.05$. We adjusted for testing multiple SNPs across each gene using two permutation tests; the global *P*-value statistic that identified the most significant SNP adjusting for the number of SNPs in each gene, and the *P*-SUM statistic that summarised the overall evidence for association by summating *P*-values across each gene. The permutation tests provided the best evidence for association with TPH2 (Global-$P=0.051$, *P*-SUM <0.1), ARRB2 (Global-$P<0.02$), SYP (Global-$P<0.04$), DAT1 (Global-$P=0.11$, *P*-SUM <0.02) and PNMT (Global-$P=0.02$, *P*-SUM <0.03). In addition to more marginal evidence for HES1 (*P*-SUM <0.1, Global-$P=0.08$), ADRB2 (Global-$P=0.09$), HES1 (Global-$P=0.08$) and MAOA (Global-$P=0.07$).

Ten out of the 18 top ranked genes had previously been reported to show some evidence for association in previous studies; however, not always with the same markers or the same alleles. Of particular interest are the VNTR markers in DAT1 and DRD4 that are both prominent findings in the current association literature for ADHD. For both of these genes we found evidence for over transmission of the previously identified risk alleles, consistent with small but significant effects. The OR estimated for these alleles were very close to those established from meta-analysis, and fits well with expectation from this relatively large multisite study.

DAT1 was one of the genes for which we had gene-wide evidence for association using the *P*-SUM statistic. For this gene, we found evidence for SNP associations in three regions of the gene. First, we detected an association signal (analysed independently of the two VNTR markers) for SNPs at the 3′-end of the gene. Second, we observed novel evidence for association with SNPs in the 5′ flanking region. Finally, we observed marginal evidence for association with rs6347, a nonsynonymous SNP in exon 9 that showed weak association in one previous study.[96] For both DAT1 and DRD4 we conclude that small but significant genetic effects have been demonstrated when these data are considered alongside the existent literature.

Of the other genes, TPH2 showed a significant Global-*P*-value and associated haplotype. The risk haplotype identified in this study was, however, the protective haplotype in the previous study from Sheenan *et al*.[53] further studies will be needed to clarify the relationship of genetic variants in TPH2 with ADHD. For SYP, FADS2 and DDC, we identified nominally associated SNPs that did not appear to be in LD with nominally associated SNPs reported in previous studies. For NET1 two previous studies reported weak evidence for association with SNPs in introns 7 and 9; however, the associated alleles were different in the two reports and we did not find evidence for association in this region. We did however find evidence for association from single markers and haplotype analysis with SNPs within the region spanning introns one to five. MAOA is an interesting gene involved in the catabolism of catecholamine neurotransmitters with evidence of extensive LD across the gene locus; we obtained evidence for association consistent with one previous report. For SNAP-25 we were not able to replicate the associations reported to be present at the 3′ end of the gene, but did detect an association signal in the 5′ UTR, consistent with findings reported by Mill *et al*.[50] Finally, we detected several novel associations with the strongest evidence coming from PNMT and SLC9A9, in addition to evidence for association with two genes in the circadian rhythm system.

With the exception of DRD4 and DAT1, we cannot draw firm conclusions. Further investigations will therefore be required to clarify whether genetic variations in the genes investigated in this study are associated with ADHD. The predominant reason for the difficulty in identifying association risk alleles is the very small effect sizes observed for even the most replicated association findings with ADHD. This is demonstrated by the data from DRD4 and DAT1, where we can only conclude that a believable level of evidence has been obtained due to the numerous preceding studies and conclusions from meta-analyses of world data, involving several thousand diagnosed cases. In the case of DAT1 a recent meta-analysis found no overall net effect of the 10-repeat allele in Caucasian populations despite a large number of positive reports; evidence of heterogeneity and haplotype associations in the region are however consistent with the association (Li *et al*., 2006) and more work is required to identify functional

variants in the region.[96] The small OR identified for polymorphisms within these two genes are similar to those observed for both nominally significant and nonsignificant findings in this study, indicating the need for future studies to focus on several large collaborative data sets to detect, and subsequently replicate key findings.

In this study we faced several difficulties. Although we were able to screen many of the associated SNPs from previous studies, in some cases SNPs were predicted to have a low chance of genotyping successes on the Illumina platform, and were therefore left out of the SNP array. This was the case for SNPs in SNAP-25 (rs3787283; rs6039806 and rs362987, TPH2 (rs1487275; rs4570625), NET1 (rs2242447) and CHRNA4 (rs6090384 and rs2273506). For seven of these SNPs we were able to perform indirect assays, using a proxy SNP in strong LD ($R^2 > 0.8$) with the reported SNP. We were able to achieve this since we included analysis of our SNP array on the CEPH Caucasian panel used in the HapMap project; enabling us to link by LD relationship many non-HapMap SNPs included in our study, with SNPs on the HapMap database (in addition to SNPs already on the HapMap database).

Another problem was that several published SNP findings did not include 'rs' number nomenclature. We were able to identify the majority of these SNPs using PCR primer sequences, to ensure correct identification of the relevant database SNP. There were a few cases where we could not determine the rs number or their precise location, and these may not currently be on public databases. To ensure ease of replication and minimise the potential for errors by misidentification of SNPs, we recommend that future investigators use the rs number nomenclature and stick to database SNPs whenever possible. Novel SNPs verified by individual studies, should be submitted for inclusion into the public databases.

Inescapable from the genetic analysis of complex genetic disorders in studies of this type are problems of multiple testing. Rather than test everything we can possibly test, we have made some considerations in the name of keeping the number of tests limited. From a single-marker standpoint, at a bare minimum, we would only perform a single analysis for each SNP. We expanded on that minimum, by utilising two single marker analytic techniques, implemented in UNPHASED and WHAP. These single marker approaches differ subtly but are largely correlated. That correlation alleviates the multiple testing problem that arises from the use of multiple techniques, as the number of novel tests conducted under a secondary or tertiary analysis are only as great as the amount of signal independently assessed by that technique. As we observe, a correlation of 0.70 between our *P*-values for the single marker analysis in WHAP and UNPHASED, in effect, we have only carried out 1.3 times as many tests with these two approaches, rather than double. In our opinion, this is an acceptable burden for the added information about the signal in the data set, and confirmation that both methods are working properly.

The second consideration for multiple testing is how to treat the haplotype analysis. Haplotype analysis differs from single marker analysis in that the number of possible approaches is much greater (e.g. sliding window analysis vs haplotype block vs tSNPs; all haplotypes vs most significant; all haplotypes vs some percentage cutoff). Additionally, the interpretation of a haplotype association signal in the absence of single marker association is by no means clear. By conditioning on nominal significance as our criteria for conducting haplotype analysis, we limited our testing burden greatly. Obviously, we may have missed some associations by not exhaustively looking at our data, but we feel more confident about the signals we do detect in our sample.

The third consideration is how to provide some adjustment for the number of SNPs tested. We approached this by empirically adjusting for the number of tests across each gene. Although we could have provided adjustment at the study wide level or the genome wide level, none of the SNP or gene-wide associations would have passed such stringent criteria. Adjustment for the analysis of 51 genes in our gene-wide significance tests would require a significance level of 0.00099 after Bonferroni correction, whereas our best gene-wide statistic was 0.012 for PNMT. For risk alleles of very small effect, achieving genome-wide significance levels may not be a realistic target within the limitations of feasible sample sizes, even where very large data sets are collated through large multisite collaborations. Whether such stringent test criteria are required remains a controversial issue, with concerns about 'throwing the baby out with the bathwater' suggesting that in the context of replication, necessary significance thresholds may be dramatically lower than genome-wide significance levels.[36]

In the future, association studies of ADHD will draw on two main resources. First, the results of linkage studies have already identified several regions with a high chance of containing one or more susceptibility loci for ADHD. The results of our linkage scan during 2006 when combined with linkage data from other sites will further delineate the regions worth pursuing. Replication of several loci in some of the existing linkage data (5p, 6q, 7p, 11q, 12q and 17p) is encouraging; since linkage is a low powered method these data suggest that several loci of more moderate effect size may exist. However, the ability of association to detect such loci will depend critically on the whether linkage signals are the result of common risk variants or multiple rare variants. Although this cannot be determined in advance of high-density association studies, previous data from other complex behavioural disorders suggest that association findings within linkage regions do not generally account for the entire linkage signal, suggesting the existence of more than one risk variant within linkage regions.

The ability to perform high-density genome-wide screens for common risk variants using contemporary genome wide SNP arrays is expected to greatly speed the process of novel gene identification for complex disorders.[58] However, the ability to detect meaningful 'signals' above the background 'noise' generated by investigating hundreds of thousands of genetic markers is likely to present a major hindrance to progress. Successful identification of novel gene associations will need to utilise large initial studies to reliably detect putative association, and large replication samples. Statistical information will subsequently need to be supplemented by functional analysis and interpretation that seek to make biological sense of association findings.[36]

Despite the availability of high-density SNP information from public databases and use of the highly efficient Illumina platform for customised SNP analysis, we still found difficulties in selecting a sufficiently representative sets of informative SNPs. There were two main constraints. First, we found that many SNPs within or close to 'functional' gene regions did not appear on the HapMap database, and could not therefore be knowingly detected by selection of tSNPs. Out of 1105 SNPs selected for their location within functional regions, only 381 (34.8%) had been genotyped in the HapMap CEPH panel; so no LD information was available for these markers at the start of the investigation. We have now genotyped all markers included in this study in the CEPH panel and these data will be presented elsewhere. Many of the 'functional' SNPs had no NCBI validation status and within this group we found a very low rate of polymorphism, around 84%. This rate is much lower than the 30–50% estimated in the recent past. This suggests that as the number of validated SNPs increases, the proportion of residual (nonvalidated) SNPs that are polymorphic is decreasing.

Another limiting factor was the estimated chance of a successful assay on the Illumina platform. Although in general sufficient SNPs could be selected for a high-density screen of each gene, there were instances when SNPs that we wished to select had a low chance of successful genotyping. Across the genes that we selected for investigation in this study, we identified 7902 potential database SNPs. Of these 982 (12.4%) had a design rating of zero, meaning that it was not possible to design a working genotype assay; and 381 (4.8%) had a design rating of 0.5, meaning an approximate 50% chance of designing a working assay. Analysis was also incomplete for X-linked genes, due to the lack of information on polymorphic markers in HapMap and nonimplementation of chromosome X analyses in tagging programs such as CLUSTAG and HAPLOVIEW.

A potential limitation of this study was site heterogeneity since we collected ADHD probands and their families from 12 sites across eight European countries. There are however several factors that minimise our concerns. Many of the European sites involved in this study have more than a decade long track record of collaborative research as members of the European Network for Hyperactivity Disorders (Eunethydis) consortium, which has yearly meetings devoted to collaborative studies of ADHD, and many of the issues related to collaborative clinical research, such as agreement on diagnostic methods, have been resolved. In this study, we have been able to adopt the same diagnostic procedures across all sites, and investigators performing clinical diagnostic procedures have received a common training package. The PACS interview used as the main diagnostic instrument has the advantage for multisite studies that cultural differences in perception of behavioural thresholds for each symptoms item are minimised, since investigators make objective ratings of observed behaviours. The same algorithm is then applied to these behavioural observations, establishing operational criteria for calling each symptom.

We have further sort to reduce aetiological heterogeneity by targeting DSM-IV combined subtype ADHD and excluding individuals with possible autism, providing a sample with considerable clinical homogeneity. We chose to identify cases with a combination of home and school based measures (PACS for home and other situations and Conners for classroom). This provides a refined phenotype of children whose problems are pervasive across situations. Situation specific hyperactivity (children whose hyperactivity presents only at home or only at school) can be included in ADHD if only one measure is used, but this goes against the spirit of both ICD-10 and DSM-IV. It is not at all clear that problems seen only at school, or only at home, have the same biological basis. Indeed, some evidence from twin studies indicates that to some extent there may be different genetic influences on behaviour in the two settings.[59] Ho *et al*.[60] surveyed a school population with parent and teacher questionnaires, and selected those who showed hyperactive behaviour at home only, at school only, or pervasively across both situations. The home specific group showed less evidence of cognitive problems and more evidence of family conflict and adversity, whereas the school specific group showed more evidence of academic learning difficulties. A study of subgroups in a clinically referred population in the UK found that the behaviour problems in the school-situational group tended to appear later, only after school entry.[54] It was only in the group identified both by parents and by teachers that the typical neurodevelopmental associations of ADHD emerged. This pervasive group was identified in both the Ho *et al*.[60] and Taylor *et al*.[59] studies, using the same measures (teacher Conners and PACS) as in our investigation.

The potential for population heterogeneity was minimised by restricting ascertainment to individuals of white European origin, comparable to a multisite collection of European-Americans. We also restricted our analyses to allelic tests of transmission from heterozygote parents (TDT), limiting the impact of population differences in MAFs. Although use of the

TDT provides excellent protection from population stratification effects, there is some evidence that analysis of case–control data may be associated with stronger genetic effects. This may be due at least in part to the increased severity of ADHD probands in samples where collecting DNA from parents is not a requirement for inclusion into a study.[61] We will include case–control investigations in some future studies using this sample.

The possibility that risk alleles may differ in prevalence is handled in our statistical procedures that assume homogeneity. The power to detect associations is therefore determined by the average effect size in the combined sample. As discussed above, it is reassuring that this large multisite study provided estimates of OR for the known VNTR alleles in DRD4 and DAT1 that were very close to those estimated from a recent meta-analysis of multisite data. Interpretation of these data is limited in one way; if we detect an OR of 1.5, we do not know whether this is a large effect in a subset of samples, or a modest effect across the entire sample. Although this can be formally tested, preliminary inspection suggested no site heterogeneity for the SNP associations, and since heterogeneity analysis does not add power for detection of gene associations, we have not presented these data here. Our genotyping strategy minimised the potential for site differences due to systematic genotype error, by using the same DNA extraction procedures, collecting all DNA at one site, partial randomisation of samples across DNA plates, and use of the high accuracy Illumina platform.

Concluding remarks

We have conducted a large multi site association study of 674 DSM-IV ADHD probands to screen 925 SNPs in 52 candidate genes. Our findings are consistent with the two most replicated findings in ADHD research, with VNTR markers in DRD4 and DAT1. We identified 16 additional genes with one or more nominally associated SNPs, including eight genes that passed gene-wide tests that controlled for the number of SNPs analysed in each gene. Since none of the findings were significant if adjusted at the study wide or genome-wide level, we are unable to identify any specific genetic variant that are more likely to be true risk factors for ADHD. Since overall we found more significant SNPs that expected, further investigations of these SNP markers are required in large replication samples, to enable positive identification of risk alleles from these findings. Future association studies of ADHD are likely to need several samples of around 1000 probands or more to identify the most prominent genetic risk factors for this important common disorder.

Acknowledgments

The IMAGE project is a multisite, international effort supported by NIH Grant R01MH62873 to SV Faraone. Site Principal Investigators are Philip Asherson, Tobias Banaschewski, Jan Buitelaar, Richard P Ebstein, Stephen V Faraone, Michael Gill, Ana Miranda, Robert D Oades, Herbert Roeyers, Aribert Rothenberger, Joseph Sergeant, Edmund Sonuga-Barke, and Hans-Christoph Steinhausen. Senior co-investigators are Margaret Thompson, Pak Sham, Peter McGuffin, Robert Plomin, Ian Craig and Eric Taylor. Chief Investigators at each site are Rafaela Marco, Nanda Rommelse, Fernando Mulas, Wai Chen, Henrik Uebel, Hanna Christiansen, U Mueller, Cathelijne Buschgens, Barbara Franke, Lamprini Psychogiou. Other investigators are Marieke Altink, Ellen Fliers, Ruud Minderaa and Alysa Doyle. We thank all the families who kindly participated in this research.

References

1 Asherson P. IMAGE consortium. Attention-deficit hyperactivity disorder in the post-genomic era. *Eur Child Adolesc Psychiatry* 2004; **13**(Suppl 1): 150–170.
2 Burd L, Klug MG, Coumbe MJ, Kerbeshian J. Children adolescents with attention deficit-hyperactivity disorder: 1. Prevalence cost of care. *J Child Neurol* 2003; **18**: 555–561.
3 Kessler RC, Adler L, Ames M, Barkley RA, Birnbaum H, Greenberg P *et al.* The prevalence effects of adult attention deficit/hyperactivity disorder on work performance in a nationally representative sample of workers. *J Occup Environ Med* 2005; **47**: 565–572.
4 Biederman J, Faraone SV. Attention-deficit hyperactivity disorder. *Lancet* 2005; **366**: 237–248.
5 Taylor E, Dopfner M, Sergeant J, Asherson P, Banaschewski T, Buitelaar J *et al.* European clinical guidelines for hyperkinetic disorder – first upgrade. *Eur Child Adolesc Psychiatry* 2004; **13**(Suppl 1): 17–30.
6 Asherson P. Clinical assessment and treatment of attention deficit hyperactivity disorder in adults. *Expert Rev Neurotherapeutics* 2005; **5**: 525–539.
7 Todd RD, Lobos EA, Sun LW, Neuman RJ. Mutational analysis of the nicotinic acetylcholine receptor alpha 4 subunit gene in attention deficit/hyperactivity disorder: evidence for association of an intronic polymorphism with attention problems. 2003; **8**: 103–108.
8 Faraone SV, Perlis RH, Doyle AE, Smoller JH, Goralnick JJ, Holmergen MA *et al.* Molecular genetics of attention-deficit/hyperactivity disorder. *Biol Psychiatry* 2005; **57**: 1313–1323.
9 Willcutt EG, Pennington BF, DeFries JC. Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am J Med Genet* 2000; **96**: 293–301.
10 Gilger JW, Pennington BF, DeFries JC. A twin study of the etiology of comorbidity: attention-deficit hyperactivity disorder and dyslexia. *J Am Acad Child Adolesc Psychiatry* 1992; **31**: 343–348.
11 Kuntsi J, Eley TC, Taylor A, Hughes C, Asherson P, Caspi A *et al.* Co-occurrence of ADHD and low IQ has genetic origins. *Am J Med Genet B Neuropsychiatr Genet* 2004; **124**: 41–47.
12 Arcos-Burgos M, Castellanos X, Pineda D, Lopera F, Palcio JD, Palacio LJ *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.
13 Bakker SC, van der Meulen EM, Buitelaar JK, Sandkuijl LA, Pauls DL, Monsuur AJ *et al.* A whole-genome scan in 164 Dutch sib pairs with attention-deficit/hyperactivity disorder: suggestive evidence for linkage on chromosomes 7p and 15q. *Am J Hum Genet* 2003; **72**: 1251–1260.
14 Fisher SE, Francks C, McCracken JT, McGough JJ, Marlow AJ, MacPhie IL *et al.* A genomewide scan for loci involved in attention-deficit/hyperactivity disorder. *Am J Hum Genet* 2002; **70**: 1183–1196.

15 Smalley SL, Kustanovich V, Minassian SL, Stone JL, Ogdie MN, McGough JJ *et al.* Genetic linkage of attention-deficit/hyperactivity disorder on chromosome 16p13, in a region implicated in autism. *Am J Hum Genet* 2002; **71**: 959–963.
16 Ogdie MN, Macphie IL, Minassian SL, Yang M, Fisher SE, Francks C *et al.* A genomewide scan for attention-deficit/hyperactivity disorder in an extended sample: suggestive linkage on 17p11. *Am J Hum Genet* 2003; **72**: 1268–1279.
17 Ogdie MN, Fisher SE, Yang M, Ishii J, Francks C, Loo SK *et al.* Attention deficit hyperactivity disorder: fine mapping supports linkage to 5p13, 6q12, 16p13, and 17p11. *Am J Hum Genet* 2004; **75**: 661–668.
18 Hebebrand J, Dempfle A, Saar K, Thiele H, Herpetz-Dahlmann B, Linder M *et al.* A genome-wide scan for attention-deficit/hyperactivity disorder in 155 German sib-pairs. *Mol Psychiatry* 2005; **11**: 196–205.
19 Winstanley CA, Theobald DE, Dalley JW, Robbins TW. Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology* 2005; **30**: 669–682.
20 Gainetdinov RR, Wetsel WC, Jones SR, Levin ED, Jaber M, Caron MG. Role of serotonin in the paradoxical calming effect of psychostimulants on hyperactivity. *Science* 1999; **283**: 397–401.
21 Brookes KJ, Knight J, Xu X, Asherson P. DNA pooling analysis of ADHD and genes regulating vesicle release of neurotransmitters. *Am J Med Genet B Neuropsychiatr Genet* 2005; **139**: 33–37.
22 Hallahan B, Garland MR. Essential fatty acids and mental health. *Br J Psychiatry* 2005; **186**: 275–277.
23 Tjon Pian Gi CV, Broeren JP, Starreveld JS, Versteegh FG. Melatonin for treatment of sleeping disorders in children with attention deficit/hyperactivity disorder: a preliminary open label study. *Eur J Pediatr* 2003; **162**: 554–555.
24 Sei H *et al.* Increase of hippocampal acetylcholine release at the onset of dark phase is suppressed in a mutant mice model of evening-type individuals. *Neuroscience* 2003; **117**: 785–789.
25 Castellanos FX, Tannock R. Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nat Rev Neurosci* 2002; **3**: 617–628.
26 Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature* 2005; **437**: 1299–1320.
27 Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders. *J Am Acad Child Adolesc Psychiatry* 2003; **42**: 1203–1211.
28 Taylor E, Schachar R, Thorley G, Wieselberg M. Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients.. *Br J Psychiatry* 1986; **149**: 760–767.
29 Conners CK. Conners' Rating Scales-Revised: Technical Manual. 2003 Sixth Printing MHS.
30 Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry* 1997; **38**: 581–586.
31 Brookes K, Mill J, Guindalini C, Curran S, Xu X, Knight J *et al.* A common haplotype of the dopamine transporter gene associated with attention-deficit/hyperactivity disorder and interacting with maternal use of alcohol. *Arch Gen Psychiatry* 2006; **63**: 74–81.
32 Brookes KJ, Xu X, Chen CK, Huang YS, Wu YY, Asherson P. No evidence for the association of DRD4 with ADHD in a Taiwanese population within-family study. *BMC Med Genet* 2005; **6**: 31.
33 Xu X, Mill J, Chen CK, Brookes K, Taylor E, Asherson P. Family-based association study of serotonin transporter gene polymorphisms in attention deficit hyperactivity disorder: no evidence for association in UK and Taiwanese samples. *Am J Med Genet B Neuropsychiatr Genet* 2005; **139**: 11–13.
34 Ao SI, Yip K, Cheung D, Fong PY, Melhado I, Sham PC. CLUSTAG: hierarchical clustering and graph methods for selecting tag SNPs. *Bioinformatics* 2005; **21**: 1735–1736.
35 Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B *et al.* The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
36 Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 2004; **75**: 353–362.
37 Feng Y, Wigg KG, Makkar R, Ickowicz A, Pathare T, Tannock R *et al.* Sequence variation in the 3′-untranslated region of the dopamine transporter gene and attention-deficit hyperactivity disorder (ADHD). *Am J Med Genet B Neuropsychiatr Genet* 2005; **139**: 1–6.
38 Bobb AJ, Addington AM, Sidransky E, Gornick EM, lerch JP, Greenstein JK *et al.* Support for association between ADHD and two candidate genes: NET1 and DRD1. *Am J Med Genet B Neuropsychiatr Genet* 2005; **134**: 67–72.
39 Hawi Z, Lowe N, Kirley A, Gruenhage F, Nothen M, Greenwood T *et al.* Linkage disequilibrium mapping at DAT1, DRD5 and DBH narrows the search for ADHD susceptibility alleles at these loci. *Mol Psychiatry* 2003; **8**: 299–308.
40 Lowe N, Kirley A, Mullins C, Fitzgerald M, Gill M, Hawi Z. Multiple marker analysis at the promoter region of the DRD4 gene and ADHD: evidence of linkage and association with the SNP -616. *Am J Med Genet B Neuropsychiatr Genet* 2004; **131**: 33–37.
41 Bellgrove MA, Hawi Z, Lowe N, Kirley A, Robertson IH, Gill M. DRD4 gene variants and sustained attention in attention deficit hyperactivity disorder (ADHD): effects of associated alleles at the VNTR and −521 SNP. *Am J Med Genet B Neuropsychiatr Genet* 2005; **136**: 81–86.
42 Xu X, Knight J, Brookes K, Mill J, Sham P, Craig I *et al.* DNA pooling analysis of 21 norepinephrine transporter gene SNPs with attention deficit hyperactivity disorder: no evidence for association. *Am J Med Genet B Neuropsychiatr Genet* 2005; **134**: 115–118.
43 Jiang S, Xin R, Lin S, Qian Y, Tang G, Wang D *et al.* Linkage studies between attention-deficit hyperactivity disorder and the monoamine oxidase genes. *Am J Med Genet* 2001; **105**: 783–788.
44 Domschke K, Sheehan K, Lowe N, Kirley A, Mullins C, O'Sullivan R *et al.* Association analysis of the monoamine oxidase A and B genes with attention deficit hyperactivity disorder (ADHD) in an Irish sample: preferential transmission of the MAO-A 941G allele to affected children. *Am J Med Genet B Neuropsychiatr Genet* 2005; **134**: 110–114.
45 Manor I, Tyano S, Mel E, Eisenberg J, Bachner-Melman R, Kotler M *et al.* Family-based and association studies of monoamine oxidase A and attention deficit hyperactivity disorder (ADHD): preferential transmission of the long promoter-region repeat and its association with impaired performance on a continuous performance test (TOVA). *Mol Psychiatry* 2002; **7**: 626–632.
46 Lawson DC, Turic D, Langely K, Pay KM, Govan CF Norton N *et al.* Association analysis of monoamine oxidase A and attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 2003; **116**: 84–89.
47 Brophy K, Hawi Z, Kirley A, Fitzgerald M, Gill M. Synaptosomal-associated protein 25 (SNAP-25) and attention deficit hyperactivity disorder (ADHD): evidence of linkage and association in the Irish population. *Mol Psychiatry* 2002; **7**: 913–917.
48 Barr CL, Feng Y, Wigg K, Bloom S, Roberts W, Malone M *et al.* Identification of DNA variants in the SNAP-25 gene and linkage study of these polymorphisms and attention-deficit hyperactivity disorder. *Mol Psychiatry* 2000; **5**: 405–409.
49 Kustanovich V, Merriman B, McGough J, McCracken JT, Smalley SL, Nelson SF. Biased paternal transmission of SNAP-25 risk alleles in attention-deficit hyperactivity disorder. *Mol Psychiatry* 2003; **8**: 309–315.
50 Mill J, Richards S, Knight S, Curran S, Taylor E, Asherson P. Haplotype analysis of SNAP-25 suggests a role in the aetiology of ADHD. *Mol Psychiatry* 2004; **9**: 801–810.
51 Mill J, Curran S, Kent L, Gould A, Huckett L, Richards S *et al.* Association study of a SNAP-25 microsatellite and attention deficit hyperactivity disorder. *Am J Med Genet* 2002; **114**: 269–271.
52 Todd RD, Lobos EA, Sun LW, Neuman RJ. Mutational analysis of the nicotinic acetylcholine receptor alpha 4 subunit gene in attention deficit/hyperactivity disorder: evidence for association of an intronic polymorphism with attention problems. *Mol Psychiatry* 2003; **8**: 103–108.
53 Sheehan K, Lowe N, Kirley A, Mullins C, Fitzgerald M, Gill M *et al.* Tryptophan hydroxylase 2 (TPH2) gene variants associated with ADHD. *Mol Psychiatry* 2005; **10**: 944–949.
54 Walitza S, Renner TJ, Dempfle A, Konrad K, Wewetzer Ch, 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.

55 Brookes K, Chen W, Xu X, Taylor E, Asherson P. Association of fatty acid desaturase genes with Attention Deficit Hyperactivity disorder. *Biol Psychiatry* 2006, in press.
56 Payton A, Holmes J, Barrett JH, Hever T, Fitzpatrick H, Trumper AL *et al.* Examining for association between candidate gene polymorphisms in the dopamine pathway and attention-deficit hyperactivity disorder: a family-based study. *Am J Med Genet* 2001; **105**: 464–470.
57 Hawi Z, Foley D, Kirley A, McCarron M, Fitzgerald M, Gill M. Dopa decarboxylase gene polymorphisms and attention deficit hyperactivity disorder (ADHD): no evidence for association in the Irish population. *Mol Psychiatry* 2001; **6**: 420–424.
58 Kruglyak L. Power tools for human genetics. *Nat Genet* 2005; **37**: 1299–1300.
59 Taylor E, Everitt B, Thorley G, Schacher R, Rutter M, Wieselberg M. Conduct disorder and hyperactivity: II. A cluster analytic approach to the identification of a behavioural syndrome. *Br J Psychol* 1986; **149**: 768–777.
60 Ho TP, Luk ES, Leung PW, Taylor E, Lieh-Mak F, Bcaon-Shone J. Situational versus pervasive hyperactivity in a community sample. *Psychol Med* 1996; **26**: 309–321.
61 West A, Langley K, Hamshere ML, Kent L, Craddock N, Owen MJ *et al.* Evidence to suggest biased phenotypes in children with attention deficit. *Mol Psychiatry* 2002; **7**: 962–966.
62 Barr CL, Feng Y, Wigg KG, Schachar R, Tannock R, Roberts W *et al.* 5′-untranslated region of the dopamine D4 receptor gene and attention-deficit hyperactivity disorder. *Am J Med Genet* 2001; **105**: 84–90.
63 Feng Y, Crosbie J, Wigg K, Pathare T, Ickowicz A, Schachar R *et al.* The SNAP25 gene as a susceptibility gene contributing to attention-deficit hyperactivity disorder. *MolPsychiatry* 2005; **10**: 998–1005.
64 Barr CL, Kroft J, Feng Y, Wigg K, Roberts W, Malone M *et al.* The norepinephrine transporter gene attention-deficit hyperactivity disorder. *Am J Med Genet* 2002; **114**: 255–259.
65 McEvoy B, Hawi Z, Fitzgerald M, Gill M. No evidence of linkage or association between the norepinephrine transporter (NET) gene polymorphisms and ADHD in the Irish population. *Am J Med Genet* 2002; **114**: 665–666.
66 Comings DE, Gade-Andavolu R, Gonzalez N, Wu S, Muhleman D, Blake H *et al.* Multivariate analysis of associations of 42 genes in ADHD, ODD and conduct disorder. *Clin Genet* 2000; **58**: 31–40.
67 De Luca V, Muglia P, Jain U, Kennedy JL. No evidence of linkage or association between the norepinephrine transporter (NET) gene MnlI polymorphism and adult ADHD. *Am J Med Genet B Neuropsychiatr Genet* 2004; **124**: 38–40.
68 Kent L, Middle F, Hawi Z, Fitzgerald M, Gill M, Feehan C *et al.* Nicotinic acetylcholine receptor alpha4 subunit gene polymorphism and attention deficit hyperactivity disorder. *Psychiatr Genet* 2001; **11**: 37–40.
69 Comings DE, Gade-Andavolu R, Gonzalez N, Wu S, Muhleman D, Blake H *et al.* Comparison of the role of dopamine, serotonin, and noradrenaline genes in ADHD, ODD and conduct disorder: multivariate regression analysis of 20 genes. *Clin Genet* 2000; **57**: 178–196.
70 Todd RD, Lobos EA, Sun LW, Neuman RJ. Mutational analysis of the nicotinic acetylcholine receptor alpha 4 subunit gene in attention deficit/hyperactivity disorder: evidence for association of an intronic polymorphism with attention problems. *Mol Psychiatry* 2003; **8**: 103–108.
71 Rowe DC, Van den Oord EJ, Stever C, Giedinghagen LN, Gard JM, Cleveland HH *et al.* The DRD2 TaqI polymorphism and symptoms of attention deficit hyperactivity disorder. *Mol Psychiatry* 1999; **4**: 580–586.
72 Comings DE, Comings BG, Muhleman D, Dietz G, Shahbahrami B, Tast D *et al.* dopamine D2 receptor locus as a modifying gene in neuropsychiatric disorders. *JAMA* 1991; **266**: 1793–1800.
73 Comings DE. Clinical and molecular genetics of ADHD and Tourette syndrome. Two related polygenic disorders. *Ann NY Acad Sci* 2001; **931**: 50–83.
74 Todd RD, Lobos EA. Mutation screening of the dopamine D2 receptor gene in attention-deficit hyperactivity disorder subtypes: preliminary report of a research strategy. *Am J Med Genet* 2002; **114**: 34–41.
75 Huang YS, Lin SK, Wu YY, Chao CC, Chen CK. A family-based association study of attention-deficit hyperactivity disorder and dopamine D2 receptor TaqI A alleles. *Chang Gung Med J* 2003; **26**: 897–903.
76 Muglia P, Jain U, Kennedy JL. A transmission disequilibrium test of the Ser9/Gly dopamine D3 receptor gene polymorphism in adult attention-deficit hyperactivity disorder. *Behav Brain Res* 2002; **130**: 91–95.
77 Barr CL, Wigg KG, Bloom S, Schachar R, Tannock R, Roberts W *et al.* Further evidence from haplotype analysis for linkage of the dopamine D4 receptor gene and attention-deficit hyperactivity disorder. *Am J Med Genet* 2000; **96**: 262–267.
78 Comings DE, Gade-Andavolu R, Gonzalez N, Blake H, Wu S, MacMurray JP. Additive effect of three noradrenergic genes (ADRA2a, ADRA2C, DBH) on attention-deficit hyperactivity disorder and learning disabilities in Tourette syndrome subjects. *Clin Genet* 1999; **55**: 160–172.
79 Daly G, Hawi Z, Fitzgerald M, Gill M. Mapping susceptibility loci in attention deficit hyperactivity disorder: preferential transmission of parental alleles at DAT1, DBH and DRD5 to affected children. *MolPsychiatry* 1999; **4**: 192–196.
80 Wigg K, Zai G, Schachar R, Tannock R, Roberts W, Malone M *et al.* Attention deficit hyperactivity disorder and the gene for dopamine Beta-hydroxylase. *Am J Psychiatry* 2002; **159**: 1046–1048.
81 Roman T, Schmitz M, Polanczyk GV, Eizirik M, Rohde LA, Hutz MH. Further evidence for the association between attention-deficit/hyperactivity disorder and the dopamine-beta-hydroxylase gene. *Am J Med Genet* 2002; **114**: 154–158.
82 Inkster B, Muglia P, Jain U, Kennedy JL. Linkage disequilibrium analysis of the dopamine beta-hydroxylase gene in persistent attention deficit hyperactivity disorder. *Psychiatr Genet* 2004; **14**: 117–120.
83 Smith KM, Daly M, Fischer M, Yiannoutsos CT, Bauer L, Barkley R *et al.* Association of the dopamine beta hydroxylase gene with attention deficit hyperactivity disorder: genetic analysis of the Milwaukee longitudinal study. *Am J Med Genet B Neuropsychiatr Genet* 2003; **119**: 77–85.
84 Zhang HB, Wang YF, Li J, Wang B, Yang L. Association of dopamine beta-hydroxylase polymorphism with attention deficit hyperactivity disorder in children. *Beijing Da Xue Xue Bao* 2004; **36**: 290–293.
85 Zhang HB, Wang YF, Li J, Wang B, Yang L. Association between dopamine beta hydroxylase gene and attention deficit hyperactivity disorder complicated with disruptive behavior disorder. *Zhonghua ErKeZa Zhi* 2005; **43**: 26–30.
86 Eisenberg J, Mei-Tal G, Steinberg A, Tartakovsky E, Zohar A, Gritsenko I, Nemanov L, Ebstein RP. Haplotype relative risk study of catechol-O-methyltransferase (COMT) and attention deficit hyperactivity disorder (ADHD): association of the high-enzyme activity Val allele with ADHD impulsive-hyperactive phenotype. *Am J Med Genet* 1999; **88**: 497–502.
87 Qian Q, Wang Y, Zhou R, Li J, Wang B, Glatt S, Faraone SV. Family-based and case-control association studies of catechol-*O*-methyltransferase in attention deficit hyperactivity disorder suggest genetic sexual dimorphism. *Am J Med Genet B Neuropsychiatr Genet* 2003; **118**: 103–109.
88 Thapar A, Langley K, Fowler T, Rice F, Turic D, Whittinger N *et al.* Catechol O-methyltransferase gene variant and birth weight predict early-onset antisocial behavior in children with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 2005; **62**: 1275–1278.
89 Curran S, Purcell S, Craig I, Asherson P, Sham P. The serotonin transporter gene as a QTL for ADHD. *Am J Med Genet B Neuropsychiatr Genet* 2005; **134**: 42–47.
90 Kent L, Doerry U, Hardy E, Parmar R, Gingell K, Hawi Z *et al.* Evidence that variation at the serotonin transporter gene influences susceptibility to attention deficit hyperactivity disorder (ADHD): analysis and pooled analysis. *MolPsychiatry* 2002; **7**: 908–912.
91 Li J, Wang Y, Qian Q, Wang B, Zhou R. Association of 5-HT(2A) receptor polymorphism and attention deficit hyperactivity disorder in children. *Zhonghua Yi Xue Za Zhi* 2002; **82**: 1173–1176.

92 Hawi Z, Dring M, Kirley A, Foley D, Kent L, Craddock N *et al*. Serotonergic system and attention deficit hyperactivity disorder (ADHD): a potential susceptibility locus at the 5-HT(1B) receptor gene in 273 nuclear families from a multi-centre sample. *MolPsychiatry* 2002; **7**: 718–725.
93 Quist JF, Barr CL, Schachar R, Roberts W, Malone M, Tannock R *et al*. The serotonin 5-HT1B receptor gene and attention deficit hyperactivity disorder. *MolPsychiatry* 2003; **8**: 98–102.
94 Li J, Wang Y, Zhou R, Zhang H, Yang L, Wang B *et al*. Serotonin 5-HT1B receptor gene and attention deficit hyperactivity disorder in Chinese Han subjects. *Am J Med Genet B Neuropsychiatr Genet* 2005; **132**: 59–63.
95 Li J, Wang YF, Zhou RL, Zhang HB, Wang B, Yang L. Association between serotonin 2C gene polymorphisms and attention deficit hyperactivity disorder comorbid or not comorbid with learning disorder. *Beijing Da Xue Xue Bao* 2004; **36**: 366–369.
96 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.

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