Objective: Attention deficit hyperactivity disorder (ADHD) is a common, highly heritable psychiatric disorder. Because of its multifactorial etiology, however, identifying the genes involved has been difficult. The authors followed up on recent findings suggesting that rare copy number variants (CNVs) may be important for ADHD etiology.

Method: The authors performed a genome-wide analysis of large, rare CNVs (<1% population frequency) in children with ADHD (N=896) and comparison subjects (N=2,455) from the IMAGE II Consortium.

Results: The authors observed 1,562 individually rare CNVs >100 kb in size, which segregated into 912 independent loci. Overall, the rate of rare CNVs >100 kb was 1.15 times higher in ADHD case subjects relative to comparison subjects, with duplications spanning known genes showing a 1.2-fold enrichment. In accordance with a previous study, rare CNVs >500 kb showed the greatest enrichment (1.28-fold). CNVs identified in ADHD case subjects were significantly enriched for loci implicated in autism and schizophrenia. Duplications spanning the CHRNA7 gene at chromosome 15q13.3 were associated with ADHD in single-locus analysis. This finding was consistently replicated in an additional 2,242 ADHD case subjects and 8,552 comparison subjects from four independent cohorts from the United Kingdom, the United States, and Canada. Presence of the duplication at 15q13.3 appeared to be associated with comorbid conduct disorder.

Conclusions: These findings support the enrichment of large, rare CNVs in ADHD and implicate duplications at 15q13.3 as a novel risk factor for ADHD. With a frequency of 0.6% in the populations investigated and a relatively large effect size (odds ratio=2.22, 95% confidence interval=1.5–3.6), this locus could be an important contributor to ADHD etiology.
Attention deficit hyperactivity disorder (ADHD) is one of the most common neuropsychiatric disorders in children (seen in 3%–5% of school-age children) (1) and adults (2%–4%) (2). Twin and adoption studies have shown high heritability of ADHD, with estimates averaging 76% (3). The etiology of ADHD is complex, with contributions from both genes and environmental factors. Until recently, research on the genetics of ADHD has focused mainly on common genetic variants in candidate gene studies. The effect sizes of most associated genetic variants identified have been small (4, 5). The few genome-wide association studies (GWAS) performed so far have been too underpowered to observe genome-wide significant associations (6–11), which fits well with the view of ADHD as a polygenic, multifactorial disorder, to which many common DNA variants (and environmental factors) of small effect contribute.

In contrast to this common disease-common variant hypothesis, rare genetic variants of moderate to large effect size have been found in a proportion of case subjects with other psychiatric disorders, such as schizophrenia and autism (12). So far, most such variants have been chromosomal aberrations and copy number variants (CNVs), deletions and duplications that encompass relatively large genomic segments spanning 1 kb to several megabases in size. Depending on their location, CNVs can influence gene expression through gene dosage effects or can directly influence protein function by excising or duplicating functional domains.

In ADHD, large, rare chromosomal aberrations have been reported to increase the risk for ADHD—for example, through the 22q11 deletion syndrome (40%–45% of patients with this syndrome have ADHD) (13)—as have rare chromosomal alterations, including a translocation involving SLC9A9 cosegregating with ADHD in an extended pedigree (14). The first published genome-wide analysis of CNVs in ADHD, which studied 335 ADHD child-parent trios and 2,026 healthy comparison subjects, failed to identify significant evidence for a higher rate of CNVs in patients (15). The study did, however, describe a number of the rare CNVs identified in ADHD patients, which spanned some intriguing candidate genes. A relatively large number of the CNVs occurred at loci that had previously been implicated in other disorders, especially autism, schizophrenia, and Tourette’s syndrome. A second CNV study analyzed 99 ADHD patients (16) and identified a duplication of the neuropeptide Y gene (NPY) cosegregating with disease in an extended pedigree. A recent genome-wide study of CNVs in ADHD (17), in which 366 children with ADHD and 1,047 comparison subjects were analyzed, found evidence for an overall increased burden of large, rare CNVs in the ADHD patients, which were also significantly enriched at loci that had previously been implicated in autism and schizophrenia. Moreover, locus-specific analysis revealed significant evidence that duplications at 16p13.11 were associated with ADHD, a finding that was replicated in an additional 825 patients and 35,243 comparison subjects. The most recent study (18) investigated 248 children with ADHD and their parents and observed de novo CNVs in 1.7% of the children and inherited CNVs in genes previously linked to ADHD or other neurodevelopmental disorders in 8%.

In this study, we followed up on these findings by performing a genome-wide analysis of CNVs in 896 children with ADHD and 2,455 unrelated comparison subjects collected as part of the IMAGE II Consortium genome-wide association study (8), which represents the largest collection of ADHD cases studied to date.

Method

Participants

The 896 case subjects investigated in this study consisted of 1) samples collected by a subset of the International Multicenter ADHD Genetics (IMAGE) Project sites but not included in the IMAGE GWAS (7) and 2) samples collected at additional sites in the United Kingdom, Ireland, Germany, Switzerland, the Netherlands, and the United States and were assessed in a manner similar to that of the IMAGE samples. Case subjects were of European origin. At the collection sites, patients had been identified mainly through outpatient clinics, met DSM-IV criteria for ADHD, had been referred for assessment of
hyperactive, disruptive, or disorganized behavior, and had been clinically diagnosed as having ADHD (or hyperkinetic disorder, the most closely equivalent category in the ICD-10 nomenclature used by some of the clinics) using semistructured interviews with parents. All sites excluded subjects with an IQ below 70. The characteristics of the full case sample have been described in detail elsewhere (8); the characteristics of the sample included in the final analysis of the present study are summarized in Table S1 in the online data supplement that accompanies the online edition of this article. Twenty-three cases overlap with a recent CNV analysis by Williams et al. (17); 99 cases overlap with another CNV analysis using array comparative genomic hybridization technology (16). Results from these overlapping cases are clearly marked in Table S4 in the data supplement. For 600 cases, additional data on ADHD subtype and severity as well as presence of comorbid oppositional defiant disorder or conduct disorder were available. All case data were collected with the informed consent of parents and with the approval of each site’s institutional review board or ethical committee. The comparison samples (2,455 population subjects of European ancestry) were collected for a GWAS of schizophrenia and have been described elsewhere (19). Briefly, the comparison subjects were drawn from a nationally representative U.S. survey panel ascertained via random digit dialing. Subjects were screened for psychosis and bipolar disorder but not for ADHD. A blood sample was collected via phlebotomy services. Comparison subjects gave written consent for their biological materials to be used for medical research at the discretion of the National Institute of Mental Health.

Replication analysis was performed for the most frequent finding on chromosome 15q13. Replication samples included those from the recently published CNV study from Cardiff, U.K., excluding the 23 ADHD cases that overlap with the IMAGE II discovery sample (296 DSM-IV ADHD case subjects with IQ >70 and 1,047 comparison subjects genotyped on Illumina Human660W-Quad BeadChip [for the case subjects] or HumanHap550 BeadChip [for the comparison subjects]) (17); from the PUWMa (Pfizer-funded study from UCLA, Washington University, and Massachusetts General Hospital) sample of 692 DSM-IV ADHD case subjects and 1,101 comparison subjects genotyped on the Illumina 1M BeadChip as described elsewhere (9); from a Canadian study that included 247 DSM-IV ADHD case subjects and 2,357 comparison subjects genotyped on the Affymetrix 6.0 array (18); and from an unpublished sample that included 1,013 DSM-IV ADHD case subjects and 4,105 comparison children genotyped on the Illumina Infinium HumanHap550K BeadChip at Children’s Hospital of Philadelphia. In the latter study, ADHD case subjects of Northern European descent (ages 6–18) were recruited from pediatric and behavioral health clinics in the Philadelphia area. Exclusionary criteria included prematurity (<36 weeks), mental retardation, major medical or neurological disorders, pervasive developmental disorder, psychosis, and major mood disorders.
Statistical Analysis of Rare CNV Data

Quality control for samples, the procedures for CNV calling, and quality control for the CNV calls are described in the online data supplement. The genome-wide burden of rare CNVs was assessed according to either the number of rare CNVs per sample or the average rare CNV size per sample. Gene-centric burden analysis was performed by limiting the analysis to rare CNVs that overlapped with the list of genes (defined according to +/- 50 kb of the largest transcript) present in National Center for Biotechnology Information Build 36.1-hg18 (http://pngu.mgh.harvard.edu/~purcell/plink/res.shtml#glist). In accordance with other studies, the significance of the burden comparisons was assessed via permutation (10,000 permutations, one-sided test) using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink). Analyses were performed for all large, rare CNVs as well as by stratification according to CNV type (deletion or duplication) and size (>100 kb or >500 kb). Differences in the rates at which CNVs were called in males or females were assessed separately in case and comparison subjects using PLINK, with significance assessed via permutation (10,000 permutations, two-sided test).

To perform locus-specific tests of association, we first defined test regions according to the genomic boundaries for each CNV identified in the entire sample. Where multiple CNVs identified in different samples overlapped, they were merged to create a single locus that encompassed all overlapping CNVs. PLINK was then used to determine the number of CNVs present within each test region in case and comparison subjects. Locus-specific tests of association were made using PLINK, again with the significance being assessed via permutation (10,000 permutations, one-sided test).

To assess whether the CNVs identified in our ADHD cohort were significantly enriched for loci previously implicated in schizophrenia or autism, we first defined the genomic coordinates for a list of single genes and genomic regions containing contiguous sets of genes that had previously been reported to harbor CNVs associated with a greater risk of autism (20) or schizophrenia (21–25). We then counted the number of CNVs larger than 100 kb in the case and comparison subjects that occurred within, or completely or partially overlapped, each locus. We also tested the overall significance of case-control comparisons for the total burden of CNVs at these loci using logistic regression analysis. To allow for the possibility that any significant overlap was caused by differences in the size of CNVs in the case and comparison subjects, we included CNV size as an independent variable.

Validation of a Rare CNV on 15q13.3

A total of 41 subjects were included in the validation study using quantitative real-time polymerase chain reaction analysis, including eight ADHD 15q13.3 duplication carriers (ADHD carriers) and their family members, eight randomly selected ADHD subjects without duplications (ADHD noncarriers), six comparison subjects with the duplication (control carriers), and four random comparison subjects (non-ADHD). The characteristics of each subject are summarized in Table S2 in the online data supplement, and procedures are explained in Table S3.

Replication of Duplications at 15q13.3

The eight duplications identified in the ADHD patient sample at 15q13.3 spanned a consensus region of approximately 420 kb that was clearly defined by two segmental duplications (chr15:29,811,982–30,232,981, National Center for Biotechnology Information Build 36.1-hg18). From each replication cohort, we selected all duplications in the case and comparison subjects that spanned the 15q13.3 consensus region by at least 60%. Tests of association were then performed for each replication sample by Fisher’s exact test and for meta-analysis by logistic regression with sample site included as an independent variable. The Breslow-Day test was used to test for heterogeneity between replication samples.

Results
A total of 732 ADHD cases passed quality control; 84% of subjects were male, and the mean age was 10.4 years. Most had combined-type ADHD (81%), and the remainder had primarily inattentive (14%) or primarily hyperactive-impulsive ADHD (5%). In the subsample for which information on comorbid disorders was available, 18% had comorbid conduct disorder and 47% had comorbid oppositional defiant disorder. A detailed breakdown by site has been described (8) and is summarized in Table S1 in the online data supplement. After exclusion of common (minor allele frequency >0.01) CNVs, all association analyses were based on 1,562 rare CNVs larger than 100 kb (460 in case subjects and 1,102 in comparison subjects; see Table S4 in the online data supplement). There was no significant difference in the rate at which CNVs >100 kb were called in males compared with females in either case or comparison subjects (data not presented).

We observed a significant excess of rare CNVs >100 kb in our ADHD case subjects relative to comparison subjects, with a rate 1.15 times higher and a proportion of subjects carrying at least one rare CNV >100 kb 1.13 times higher (Table 1). There was no significant evidence to suggest that the rare CNVs identified in the ADHD cases were on average longer than those identified in the comparison subjects, a finding in line with that of a previous study using a different data set (17).

Limiting our analysis to the largest CNVs suggested that the higher rate in ADHD cases was strongest for rare CNVs >500 kb (1.28 times higher in the ADHD cases, p=0.032; Table 1). The rate of rare CNVs >500 kb observed in ADHD cases was 12.2%, which is in accordance with the rate of 12.5% reported in the previous study using analogous methodology and a different data set (17). While there was a difference in the gender distribution between the ADHD cases and the population-based comparison sample (which was 48% male), there was no significant difference in the rate at which CNVs >100 kb were called in males compared with females in either case or comparison subjects (data not presented). When we restricted the analysis to CNVs >100 kb that spanned genes, we found a significantly greater burden of such CNVs in case subjects, which was strongest for duplications (Table 1). There was no evidence of a greater burden of non-gene-centric CNVs >100 kb (minimum p=0.11; data not presented). As over 90% of CNVs >500 kb spanned at least one gene, the gene-centric burden analysis was limited to CNVs >100 kb.

The 1,562 CNVs included in this study segregated into 912 independent loci (see Table S5 in the online data supplement). Genome-wide locus-specific analysis identified one region (chr15:28,231,568–30,571,466) that was nominally associated with ADHD (p=0.012), although this finding did not survive correction for genome-wide testing (p=0.79). Nevertheless, post hoc analysis of this locus revealed that the association was primarily contributed to by eight duplications in 732 ADHD cases, compared with six in the 2,010 comparison subjects (p=0.016 uncorrected), all of which spanned a consensus region of approximately 420 kb (chr15:29,811,982–30,232,981), defined by two segmental duplications (Figure 1). Validation of the CNV using a different genotyping method confirmed the presence of the variant in the ADHD cases and showed that it was inherited in all families for which both parents were available for testing (see Figure S1 in the online data supplement). While post hoc analyses failed to reveal any significant evidence that overall CNV carrierhip was associated with ADHD subtype, ADHD symptom dimension, or presence of oppositional defiant disorder, we did observe a nominally significant association (p=0.03 uncorrected) between conduct disorder and carriers of 15q13.3 duplications.

We attempted to replicate the observation of an excess of duplications at 15q13.3 by studying an additional 2,242 ADHD case subjects and 8,552 comparison subjects in four independent samples of European Caucasian descent from the United States, the United Kingdom, and Canada (Table 2). Duplications spanning chr15:29,811,982–30,232,981 were found in case and comparison subjects from all samples (Figure 1), and they were indeed enriched in the ADHD patients across the replication samples (p=0.00275; Table 2). Combined analysis of all samples investigated (2,966 cases, 10,556 comparison subjects) produced highly significant evidence that duplications at 15q13.3 are associated with ADHD (p=0.000178; odds ratio=2.22, 95% confidence interval=1.46–3.38).
It was previously reported that CNVs identified in ADHD cases are enriched at loci that harbor CNVs associated with schizophrenia and autism (17). In the present study, we observed that 18 of 460 (3.9%) CNVs >100 kb identified in the case subjects overlapped with one of the 32 loci previously implicated in autism (20), compared with only 20 of 1,102 (1.8%) of the CNVs identified in comparison subjects (p=0.009; Table 3), representing a rate 2.16 times greater in case subjects relative to comparison subjects. We also observed that ADHD case subjects had a 1.49-fold excess of CNVs located at the eight loci previously implicated in schizophrenia relative to comparison subjects (5.4% compared with 3.6%, p=0.03; Table 3). The lists of regions previously implicated in autism and schizophrenia are not independent of each other (they have five regions in common).

Discussion

Until recently, the common disease-common variant hypothesis has been used to explain the occurrence of most cases of psychiatric disorders. Recent studies showing a higher frequency of rare CNVs in psychiatric patients have challenged this view. Such rare CNVs have also been described in ADHD patients (15, 16, 18) and have been

Segmental duplications are labeled using the nomenclature defined by Szafranski et al. (26). FISH=fluorescent in situ hybridization.

found to be enriched in this population (17). Our results in this study, a CNV analysis in the largest clinical sample hitherto investigated, support the findings of the previous study (17) reporting an increased burden of rare CNVs in ADHD patients. Whereas the latter study investigated only CNVs larger than 500 kb (showing enrichment in both deletions and duplications in the patients), we show here that an increased burden is also observed when CNVs down to 100 kb in size are considered.

While most CNVs occurred only in single patients in our study, there was some overlap with the findings from the earlier CNV studies in ADHD (15–18), and some CNVs have been linked to ADHD in other ways (see Table S6 in the online data supplement). These CNVs mark genes that might be of particular relevance to ADHD and would make good candidates for further study.

As also noted in earlier CNV studies in ADHD (15–17), we found significant evidence that CNV regions in ADHD patients overlapped with loci implicated by CNVs in autism and schizophrenia. Although schizophrenia and ADHD do not typically co-occur, ADHD and autism cooccur in patients more often than would be expected by chance, and they share heritability (27). The overlap in CNV loci among disorders suggests pleiotropy of genes predisposing to psychiatric disorders (28–31). Additional factors seem to be necessary to explain the specificity of a clinical phenotype. On the other hand, pleiotropy might also imply that the clinical classification tools for psychiatric disorders do not match the biological underpinnings of such disorders (28).

In this study, we were also able to perform a regional analysis testing each locus carrying a CNV for association with ADHD. Despite earlier identification of ADHD case subjects carrying duplications at 16p13 (17), in this sample there was no evidence for association between duplications at this locus and ADHD (two CNVs were found in case subjects, six in comparison subjects). However, we did identify significant associations of duplications at 15q13.3 with ADHD. Notably, we replicated this observation in a total of 2,242 independent ADHD cases and 8,552 comparison subjects from four different sites, including the study by Stergiakouli et al. (32). Duplications were identified at 15q13.3 in all studies and using all different platforms for CNV detection used, with odds ratios ranging from 0.91 to 3.99. Specifically, our data implicate duplications spanning a region of approximately 420 kb (chr15:29,811,982–30,232,981), which is flanked by two segmental duplications. However, as with all CNV analyses of single-nucleotide polymorphism (SNP) array data, our study had
limited resolution to establish the nature of potentially complex rearrangements at this locus; therefore, we cannot exclude the possibility that some of the duplications identified at 15q13 are of a more complex nature. The presence of the 15q13.3 duplication also seemed to modulate the ADHD phenotype, as carriers had a higher lifetime rate of comorbid conduct disorder.

Rare CNVs in this locus (deletions and duplications) have previously been implicated in several psychiatric disorders (e.g., autism, schizophrenia, intellectual disability), as well as nonpsychiatric conditions, such as epilepsy (33), albeit with reduced penetrance. The duplicated region contains a plausible candidate gene for ADHD, **CHRNA7** (Mendelian Inheritance in Man code *118511), which encodes the a7 subunit of the neuronal nicotinic acetylcholine receptor, a homo-oligomeric ion channel involved in calcium signaling in the brain. The a7 nicotinic acetylcholine receptor participates in an ADHD-relevant pathway by mediating dopamine release (34). Dopamine dysregulation is strongly implicated in ADHD; in fact, a7 receptor agonists show modest efficacy for the treatment of ADHD (35). Two candidate gene studies of microsatellite markers and a SNP in and near this gene in ADHD have been negative (36, 37). However, a recent study implicates the receptor in the response to stress and shows that maternal genotype has a strong effect on offspring phenotype (38). This might suggest that this gene is a particularly interesting candidate for parent-of-origin and gene-environment interaction studies in ADHD.

Do our findings imply that ADHD behaves as a mono-genic disease in the patients carrying CNVs? This study does not provide evidence that any of the rare CNVs identified in ADHD behave as highly penetrant variants; as overlap is routinely observed in findings between patients and comparison subjects. From this, we can conclude that these CNVs are neither necessary nor sufficient to cause ADHD. This is consistent with other studies of rare CNVs segregating in extended pedigrees, which did not report perfect cosegregation of risk variants with ADHD (16, 39) or autism (40). Therefore, while we accept that particularly for de novo variants we cannot exclude a high penetrance, we expect that most rare CNVs implicated in this and other studies are moderate risk factors for ADHD that interact with other DNA risk variants or environmental factors to cause the disorder.

Our study has both strengths and weaknesses. Clearly, the study’s large sample size is an important strength, as is the availability of several replication cohorts. The potential weakness of using two different genotyping platforms for case and comparison subjects has been addressed by concentrating on SNPs represented on both arrays, by strict quality control, and by analysis of only large, rare CNVs, which, in accordance with previous findings (17, 21), can be reliably called. Unfortunately, we could not assess the inheritance of most of our rare CNVs. Finally, our sample is not well suited for studying additional phenotypic variation, given the limited phenotypic range caused by the fact that most case subjects suffered from the most severe, combined form of ADHD. Even larger studies with more phenotypic variability might be necessary to investigate the effects of CNVs on the ADHD subtypes and correlates.

In conclusion, our study provides further evidence for a role of large, rare CNVs in ADHD. The replicated association between ADHD and duplications on chromosome 15q13.3, increasing ADHD risk with an odds ratio of 2.22, is one of the strongest risk factors for ADHD identified thus far and, with a frequency >0.6% in the population, could be an important contributor to ADHD etiology.
The authors thank the patients and the family members who provided data for this project and the many research coworkers who helped collect and manage the data. They also thank Benjamin Neale for data retrieval.
References


34. Seipel AT, Yakel JL: The frequency-dependence of the nicotine-induced inhibition of dopamine is controlled by the alpha7 nicotinic receptor. J Neurochem 2010; 114:1659–1666


