

Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder

Josephine Elia^{1,2,46}, Joseph T Glessner^{3,46}, Kai Wang³, Nagahide Takahashi⁴, Corina J Shtir⁵, Dexter Hadley³, Patrick M A Sleiman³, Haitao Zhang³, Cecilia E Kim³, Reid Robison⁶, Gholson J Lyon⁶, James H Flory³, Jonathan P Bradfield³, Marcin Imielinski³, Cuiping Hou³, Edward C Frackelton³, Rosetta M Chiavacci³, Takeshi Sakurai⁴, Cara Rabin⁷, Frank A Middleton⁸, Kelly A Thomas³, Maria Garris³, Frank Mentch³, Christine M Freitag⁹, Hans-Christoph Steinhausen^{10–12}, Alexandre A Todorov¹³, Andreas Reif¹⁴, Aribert Rothenberger¹⁵, Barbara Franke^{16,17}, Eric O Mick¹⁸, Herbert Roeyers¹⁹, Jan Buitelaar²⁰, Klaus-Peter Lesch¹⁴, Tobias Banaschewski²¹, Richard P Ebstein²², Fernando Mulas²³, Robert D Oades²⁴, Joseph Sergeant²⁵, Edmund Sonuga-Barke^{26–28}, Tobias J Renner²⁹, Marcel Romanos^{29,30}, Jasmin Romanos²⁹, Andreas Warnke²⁹, Susanne Walitza^{10,29}, Jobst Meyer³¹, Haukur Palmason³¹, Christiane Seitz³², Sandra K Loo³³, Susan L Smalley³³, Joseph Biederman¹⁸, Lindsey Kent³⁴, Philip Asherson¹⁰, Richard J L Anney³⁵, J William Gaynor³⁶, Philip Shaw⁷, Marcella Devoto^{37–40}, Peter S White^{41,42}, Struan F A Grant^{3,37,38}, Joseph D Buxbaum⁴, Judith L Rapoport⁷, Nigel M Williams⁴³, Stanley F Nelson⁵, Stephen V Faraone^{8,44} & Hakon Hakonarson^{3,36,37,45}

Attention deficit hyperactivity disorder (ADHD) is a common, heritable neuropsychiatric disorder of unknown etiology. We performed a whole-genome copy number variation (CNV) study on 1,013 cases with ADHD and 4,105 healthy children of European ancestry using 550,000 SNPs. We evaluated statistically significant findings in multiple independent cohorts, with a total of 2,493 cases with ADHD and 9,222 controls of European ancestry, using matched platforms. CNVs affecting metabotropic glutamate receptor genes were enriched across all cohorts ($P = 2.1 \times 10^{-9}$). We saw *GRM5* (encoding glutamate receptor, metabotropic 5) deletions in ten cases and one control ($P = 1.36 \times 10^{-6}$). We saw *GRM7* deletions in six cases, and we saw *GRM8* deletions in eight cases and no controls. *GRM1* was duplicated in eight cases. We experimentally validated the observed variants using quantitative RT-PCR. A gene network analysis showed that genes interacting with the genes in the *GRM* family are enriched for CNVs in ~10% of the cases ($P = 4.38 \times 10^{-10}$) after correction for occurrence in the controls. We identified rare recurrent CNVs affecting glutamatergic neurotransmission genes that were overrepresented in multiple ADHD cohorts.

ADHD is a common neuropsychiatric disorder with heritability estimates as high as 90% (refs. 1–3). Most neurodevelopmental disorders have been difficult to study by GWAS, apart from some progress

that has been made in autism^{4–6}. A GWAS was performed on 958 ADHD trios through the International Multicentre ADHD Genetics (IMAGE) study, but no locus reaching genome-wide significance was found^{7,8}. However, a PBAT (see URLs) analysis of the quantitative measures showed nominal significance of SNPs tagging *CDH13* (rs6565113) and *GFOD1* (rs552655)⁹. A SNP in linkage disequilibrium affecting *CDH13* has previously been reported in two independent ADHD GWAS¹⁰. Duplications of 16p13.11 have also been shown to be associated with ADHD¹¹. We previously reported CNV loci that we observed in the first 335 cases with ADHD that we recruited¹². Although in that previous study, no CNV loci met the criteria for significance ($P < 5 \times 10^{-4}$), one family in the study had a *GRM5* deletion that affected all three children with ADHD and was inherited from the mother, who also had ADHD.

Our discovery cohort included a total of 1,013 cases of European descent with ADHD recruited and genotyped at The Children's Hospital of Philadelphia (CHOP) and was comprised of 664 cases without available parental information and 349 cases from trios with complete information (Supplementary Tables 1 and 2). We established a minimum inclusion intelligence quotient threshold of 70 to exclude individuals with intellectual disability¹³. The control group included 4,105 healthy children of European ancestry that were 6–18 years old, 32% of which were female and 68% of which were male. We screened medical records and parental- or self-reported questionnaires for individuals with developmental delays or special

A full list of author affiliations appears at the end of the paper.

Received 24 June; accepted 28 October; published online 4 December 2011; doi:10.1038/ng.1013

Table 1 New CNVRs overrepresented in individuals with ADHD

CNVR (hg18)	CHOP cases ^a	CHOP controls ^b	Replication cases ^c	Replication controls ^d	Inherit	Combined <i>P</i>	OR (95% CI)	Type	Gene	Exon distance
chr11: 88,269,449–88,351,661	4	0	6	1	4:1:3 (62.5%)	1.36×10^{-6}	38.12 (5–298)	Del	<i>GRM5</i>	5,858
chr7: 126,525,124–126,536,202	3	0	5	0	0:1:0 (100%)	3.52×10^{-6}	Infinity	Del	<i>GRM8</i>	0
chr3: 7,183,953–7,197,236	4	0	2	0	0:2:0 (100%)	8.14×10^{-5}	Infinity	Del	<i>GRM7</i>	20,598
chr6: 146,657,076–146,694,047	5	2	3	0	2:0:0 (100%)	1.05×10^{-4}	15.24 (3–72)	Dup	<i>GRM1</i>	0
chr1: 72,317,292–72,328,395	4	0	1	0	0:3:0 (100%)	3.91×10^{-4}	Infinity	Dup	<i>NEGR1</i>	10,621
chr7: 153,495,598–153,564,827	5	0	3	2	1:2:0 (100%)	4.08×10^{-4}	15.24 (3–72)	Dup	<i>DPP6</i>	68,453
chr5: 65,027,976–65,046,520	4	0	2	1	2:0:2 (50%)	4.68×10^{-4}	22.85 (3–190)	Del	<i>SGTB/NLN</i>	0
chr1: 56,053,497–56,064,495	2	0	4	2	1:0:3 (25%)	1.54×10^{-3}	11.42 (2–57)	Del	<i>USP24^e</i>	80,234
chr19: 38,427,720–38,444,834	5	2	2	3	2:2:1 (80%)	4.95×10^{-3}	5.33 (2–17)	Del	<i>SLC7A10^e</i>	19,172
chr3: 1,844,168–1,859,889	4	0	3	6	2:4:0 (100%)	8.81×10^{-3}	4.44 (1–13)	Del	<i>CNTN4^e</i>	255,661
chr2: 81,419,297–81,446,082	2	0	2	3	1:0:1 (50%)	3.83×10^{-2}	5.07 (1–23)	Dup	<i>CTNNA2^e</i>	152,417
chr4: 113,772,340–113,788,584	2	0	2	3	0:0:0 (NA)	3.83×10^{-2}	5.07 (1–23)	Dup	<i>LARP7</i>	0

Loci significantly replicating are shown in bold. The 'inherit' column lists the inheritance pattern of each CNV from parents to cases in the format 'inherited from mother:inherited from father:de novo', with the percent of inheritance listed in parentheses. Note that information about the parents was not available for all subjects. Rare variants that were recurrent and observed to be enriched among cases with ADHD relative to controls and that were detected in multiple independent cohorts are reported. All GRM genes were directly affected by the CNVR. Regions listed represent the optimal overlap of cases and significance with respect to controls, as described in the Online Methods and **Supplementary Figure 10**. The closest gene is listed for each CNVR locus, as this is the gene most likely to be affected. For detailed counts from each contributing project, see **Supplementary Table 14**. ^a*n* = 1,013. ^b*n* = 4,105. ^c*n* = 2,493. ^d*n* = 9,222. ^eNo gene was directly affected, so the closest proximal gene is listed. Individual CNV boundaries are provided in **Supplementary Table 15**. OR, odds ratio; CI, confidence interval; del, deletion; dup, duplication. 'Replication' includes the combined IMAGE, PUWMA, IMAGE II, NIMH and Utah data.

educational needs. For the replication, we used 624 samples from the IMAGE cohort that met quality control criteria: the PUWMA consortium contributed 864 cases with ADHD and 1,258 parents, and the IMAGE II consortium contributed 787 cases with ADHD and 898 unrelated controls. We used an additional 128 cases from the US National Institute of Mental Health (NIMH) and 90 cases from the University of Utah for the replication. We genotyped the DNA samples on different platforms. To manage differences in CNV detection between the arrays, we used controls genotyped on platforms that matched the platforms used for the cases (**Supplementary Table 3**).

We used PennCNV (see URLs) to produce CNV calls for the cases and controls¹⁴. Ninety-three percent of the subjects had 8–45 CNV calls after quality control filtering (**Supplementary Fig. 1**). We called four different copy number states: 3,172 homozygous deletions (copy number of 0), 27,810 hemizygous deletions (copy number of 1), 14,806 hemizygous duplications (copy number of 3) and 581 homozygous duplications (copy number of 4). The raw data and the resulting CNV calls are shown in **Supplementary Figure 2**. CNV calls spanned between 3 and 598 SNPs, with an average of 14 SNPs per CNV call and an average CNV size of 62 kb.

Ninety-three percent of controls also had 8–45 CNV calls (**Supplementary Fig. 1**). Among these CNV calls, we identified 4,471 homozygous deletions, 49,726 hemizygous deletions, 27,032 hemizygous

duplications and 1,480 homozygous duplications. These CNV calls spanned between 3 and 708 SNPs, with an average of 12.8 SNPs per CNV call and an average CNV size of 53.6 kb. We also ran QuantiSNP¹⁵ to evaluate CNV calls that had a minimum of three SNPs on autosomes. We observed the same range of CNV calls (10–50 calls) and a 58% average direct overlap concordance with the PennCNV calls.

We did not detect any genome-wide significant associations in our genotype GWAS analysis (**Supplementary Note** and **Supplementary Tables 4–6**). However, we replicated SNPs in *GFOD1* in the families from the CHOP study using a transmission disequilibrium test (TDT) (the *P* values we found ranged from $P = 8 \times 10^{-4}$ to $P = 1 \times 10^{-2}$). The significance values we observed for previously reported ADHD SNPs^{10,16} are listed in **Supplementary Table 7**.

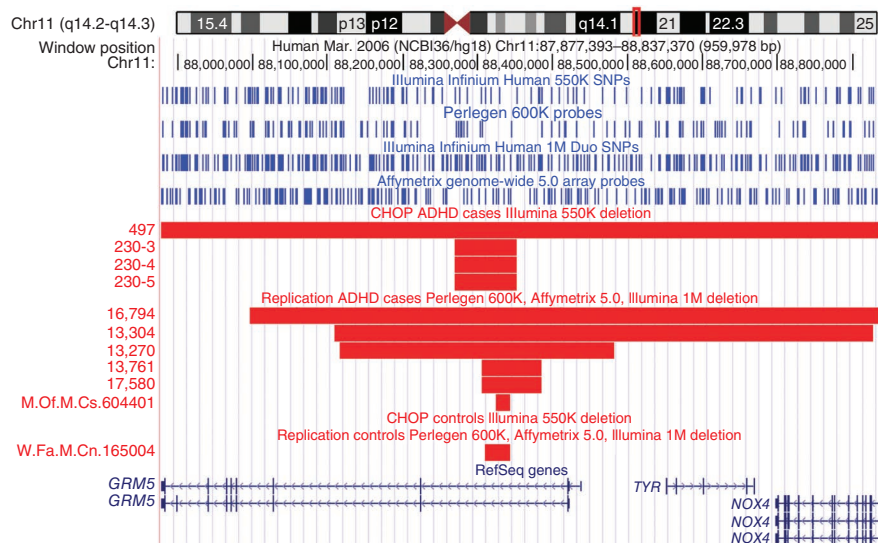
To identify CNVs associated with ADHD, we applied a segment-based approach for consecutive SNPs with more CNVs in cases than in controls^{4,14}. The genomic span for consecutive SNPs delineates the shared CNV regions (CNVRs). In the CHOP cohort, we identified ten CNVRs that were present in multiple cases but not in controls and two CNVRs that had a higher frequency in cases than in controls (**Table 1**). We observed the previously reported duplication of 16p13.11 (ref. 11) in three cases and no controls ($P = 0.013$). To ensure CNV reliability, we experimentally validated all CNVRs using quantitative RT-PCR (qRT-PCR) (**Supplementary Fig. 3**). To rule out

Table 2 Discovery, replication and combined significance of CNV regions

CNVR (hg18)	Discovery <i>P</i>	Replication <i>P</i>	Combined <i>P</i>	Permuted discovery <i>P</i>	Permuted replication <i>P</i>	Permuted combined <i>P</i>	Type	Gene
chr11: 88,269,449–88,351,661	1.53×10^{-3}	5.29×10^{-4}	1.36×10^{-6}	0.025	0.001	0.002	Del	<i>GRM5</i>
chr7: 126,441,593–126,621,501	7.74×10^{-3}	4.35×10^{-4}	3.52×10^{-6}	0.013	<0.001	<0.001	Del	<i>GRM8</i>
chr3: 7,183,953–7,197,236	1.53×10^{-3}	4.53×10^{-2}	8.14×10^{-5}	0.011	0.039	<0.001	Del	<i>GRM7</i>
chr6: 146,657,076–146,694,047	4.42×10^{-3}	9.63×10^{-3}	1.05×10^{-4}	0.006	<0.001	<0.001	Dup	<i>GRM1</i>
chr1: 72,317,292–72,328,395	1.53×10^{-3}	2.13×10^{-1}	3.91×10^{-4}	0.036	0.213	0.011	Dup	<i>NEGR1</i>
chr7: 153,495,598–153,564,827	1.53×10^{-3}	6.82×10^{-2}	4.08×10^{-4}	<0.001	0.058	<0.001	Dup	<i>DPP6</i>
chr5: 65,027,976–65,046,520	1.53×10^{-3}	1.17×10^{-1}	4.68×10^{-4}	0.003	0.108	0.001	Del	<i>SGTB/NLN</i>
chr1: 56,053,497–56,064,495	3.91×10^{-2}	2.12×10^{-2}	1.54×10^{-3}	0.035	0.024	<0.001	Del	<i>USP24</i>
chr19: 38,427,720–38,444,834	4.42×10^{-3}	2.89×10^{-1}	4.95×10^{-3}	0.002	0.262	0.007	Del	<i>SLC7A10</i>
chr3: 1,844,168–1,859,889	1.53×10^{-3}	4.12×10^{-1}	8.81×10^{-3}	0.008	0.416	0.015	Del	<i>CNTN4</i>
chr2: 81,419,297–81,446,082	3.91×10^{-2}	2.89×10^{-1}	3.83×10^{-2}	0.046	0.294	0.032	Dup	<i>CTNNA2</i>
chr4: 113,772,340–113,788,584	3.91×10^{-2}	2.89×10^{-1}	3.83×10^{-2}	0.033	0.288	0.042	Dup	<i>LARP7</i>

The top four most significant loci are shown in bold. Del, deletion; dup, duplication.

Figure 1 A deletion directly affecting *GRM5* that is exclusive to cases with ADHD and that was replicated in the IMAGE and PUWMA studies. Four hemizygous deletions in *GRM5* in cases with ADHD from the CHOP study that were replicated by two deletions and three larger deletions found in the IMAGE study and one deletion found in the PUWMA study. The SNP coverage of the Illumina 550K, Perlegen 600K, Illumina 1M and Affymetrix 5.0 arrays is shown by vertical blue lines. M.Of.M.Cs., Massachusetts General Hospital offspring male case; W.Fa.M.Cn., Washington University father male control.



false negatives in the controls, we performed qRT-PCR on 908 controls across the *GRM5* and *GRM7* loci, which confirmed that no CNVs were present (**Supplementary Fig. 4**). In addition, significantly associated loci were called by both QuantiSNP and PennCNV and were absent in controls.

Replication of the significant findings, including the ten case-specific CNVs from the discovery cohort, showed that three CNVs were exclusive to cases from the IMAGE, PUWMA, IMAGE II, NIMH and University of Utah studies, notably *GRM7* ($P = 8.14 \times 10^{-5}$), *GRM8* ($P = 3.52 \times 10^{-6}$) and *NEGR1* ($P = 3.91 \times 10^{-4}$) (**Table 1**). We observed a deletion in *GRM5* in ten cases with ADHD (10/3,506) and one control (1/13,327) ($P = 1.36 \times 10^{-6}$) (**Table 1**). We observed a duplication in *GRM1* in eight cases and two controls ($P = 1.05 \times 10^{-4}$). The odds ratios for *GRM5* and *GRM1* were 38.12 (95% confidence interval, 5–298) and 15.24 (95% confidence interval, 3–72), respectively (**Table 1**).

Thus, we identified four genes in multiple independent cohorts that belong to the metabotropic glutamate receptor gene family and are directly affected by CNVs (combined $P = 2.1 \times 10^{-9}$) (**Table 1** and **Table 2**). Several other CNV loci were enriched in cases with nominal significance (**Table 1** and **Table 2**). In **Figure 1**, we show representative CNV deletions at the *GRM5* locus (found in ten cases and one control) identified using the UCSC Genome Browser¹⁷. Experimental validation of the CNVs present in individuals from the IMAGE, PUWMA, IMAGE II, NIMH and University of Utah studies identified using qRT-PCR and raw B allele frequency (BAF) and log R ratio (LRR) plots is shown in **Supplementary Figures 5–7**. We also detected *GRM2* and *GRM6* deletions in single cases with ADHD from the CHOP and IMAGE II studies, respectively, that were not present in controls. Referencing CNV-associated loci to their genotype, a TDT analysis revealed the strongest support for association of *GRM7* to ADHD ($P = 8.35 \times 10^{-5}$) (**Supplementary Table 8**). We evaluated family-based CNV statistics of transmission disequilibrium and *de novo* events in a subset of 311 families from the CHOP study and 422 families from the IMAGE study (**Table 3** and **Supplementary Tables 9**

and **10**). We first verified trios by genotype inheritance. The Illumina CHOP data had a combined deletion and duplication *de novo* rate of 4.81%, with 16 deletion and 8 duplication CNVs, and the Perlegen IMAGE data had a *de novo* rate of 7.43%, with 2 deletion and 5 duplication CNVs. *GRM5* deletions were *de novo* in three cases (**Table 1**). We evaluated the association of homozygous deletion CNVs separately and found six loci with nominal significance, including a locus in *AKNAD1* (**Supplementary Table 11**).

We hypothesized that genes interacting with *GRM* genes would collectively have more CNVs enriched in cases compared to controls. We identified 228 genes within two degrees of relation to *GRM* genes in the merged human interactome using Cytoscape Software¹⁸. We evaluated these genes in the CHOP cohort for enrichment in cases with ADHD ($P < 0.05$). We detected 67 genes interacting with *GRM* genes (not including the *GRM* genes themselves) enriched for CNVs in cases compared to 16 such genes in controls, showing a threefold enrichment of the *GRM* network CNVs in individuals with ADHD ($P = 4.38 \times 10^{-10}$; **Fig. 2a**). We excluded large CNVs spanning multiple genes in the network to ensure that the network enrichment was not skewed. A GWAS analysis of *GRM* network genes did not result in a significant association (overall $P = 0.142$), confirming that the GWAS analysis did not capture association to CNVs. We then clustered the second degree *GRM* network to define interconnected modules of genes and scored the enrichment of gene ontology annotations (**Fig. 2b**).

GRMs are G-protein-coupled receptors involved in the modulation of excitatory synaptic transmission¹⁹. There are three receptor groups that are based on sequence homology, putative signal transduction mechanisms and pharmacologic properties²⁰. *GRM5* and *GRM1* are members of group 1 and are expressed in the basal

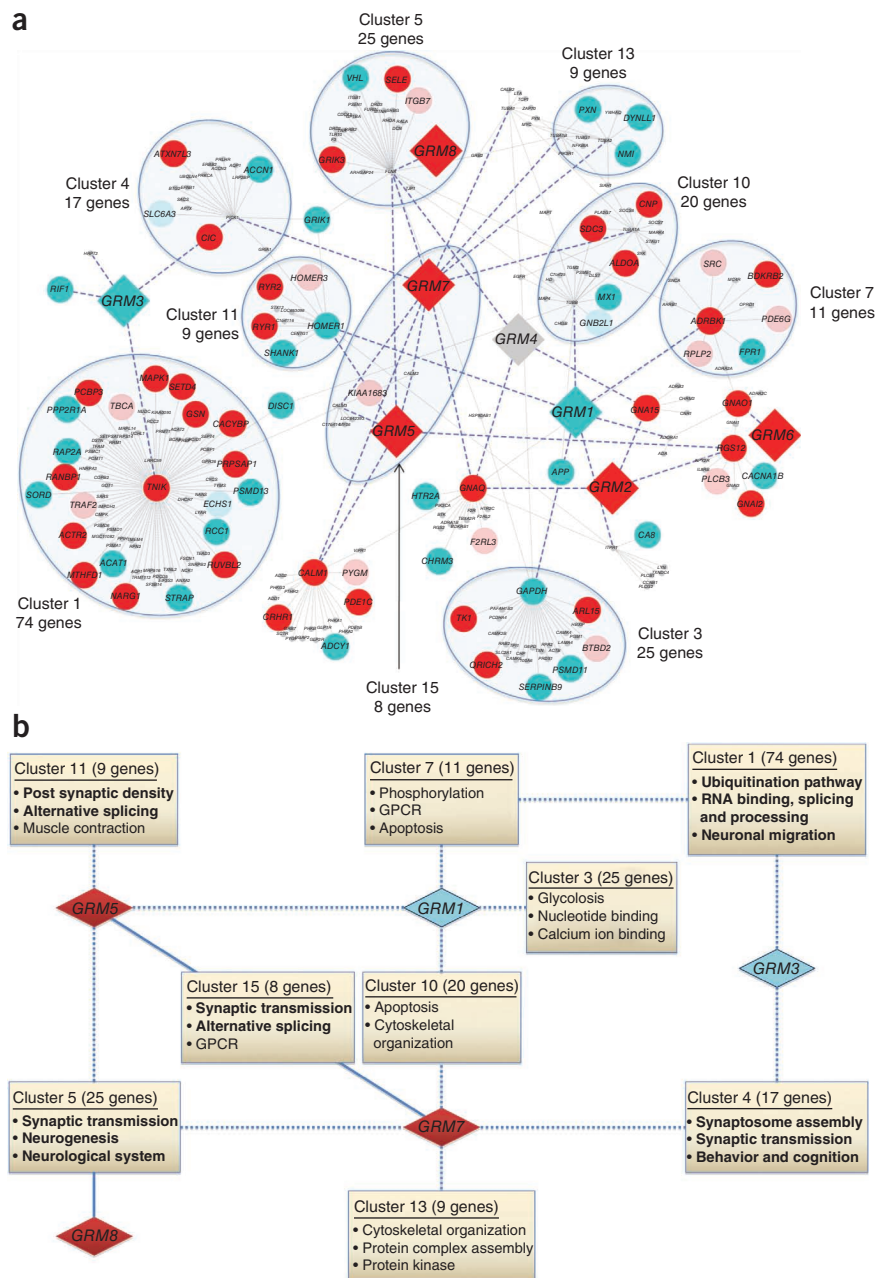
Table 3 Genome-wide analysis of *de novo* CNVs in the CHOP discovery cohort

CNV (hg18)	Inherited	<i>De novo</i>	Parent only	Type	Gene	Distance ^a
chr19: 15,992,679–15,997,923	15	6	15	Del	<i>LOC126536</i>	0
chr22: 38,384,374–38,403,731	4	4	13	Del	<i>CACNA11</i>	0
chr17: 71,112,486–71,120,734	12	3	16	Del	<i>KIAA1783</i>	0
chr12: 55,902,280–55,923,860	9	3	19	Del	<i>NDUFA4L2, NXPH4, SHMT2, STAC3</i>	0
chr19: 59,423,491–59,428,132	74	3	38	Dup	<i>LILRB3, LIR-3</i>	0
chr16: 87,694,595–87,778,383	32	2	21	Del	<i>AX748415, CDH15, LOC197322</i>	0
chr18: 65,358,832–65,367,619	33	2	21	Del	<i>DOK6</i>	0

^aThe respective CNVs directly affect the gene(s) at these loci. CNVs are visually validated as being either inherited or *de novo*. Confidence in the *de novo* calls was derived from multiple observations of the AA genotype in a specific parent and the BB genotype in child (or BB in the parent and AA in the child), suggesting a consistent parental origin. Del, deletion; dup, duplication.

Figure 2 *GRM* receptor gene interaction networks affected in ADHD. (a) *GRM* receptor genes are shown as large diamond-shaped nodes, and other genes within two degrees of interaction with *GRM* genes are shown as smaller circular nodes. Nodes are colored to represent the enrichment of the CNVs: dark red represents deletions enriched in cases, light red represents deletions enriched in controls, dark turquoise represents duplications enriched in cases, light turquoise represents duplications enriched in controls, and gray represents diploids that are devoid of CNVs. Thick blue dashed lines highlight edges that are connected to at least one *GRM* gene, and thin gray lines represent all other gene interactions. Highly connected modules enriched for significant functional annotations are highlighted by blue shaded ellipses. Details on the gene-based CNV observations are included in **Supplementary Table 16**, and the respective gene functional clusters are listed in **Supplementary Table 17**.

(b) A schematic overview showing the interaction of *GRM* receptors affected in ADHD with modules of genes enriched for functional significance. *GRM* receptor genes are shown as diamonds colored either turquoise or red to represent duplications and deletions, respectively, that were enriched in cases. Boxes highlight the functional modules defined by the network of interacting genes (a) that are significantly enriched for Gene Ontology annotations. The functional modules describe significant functional annotations and are labeled with the cluster name and the number of component genes in parenthesis. Functional annotations that may be particularly pertinent to the underlying pathophysiology of ADHD are shown in bold. The edges of the network connect *GRM* receptor genes to functional modules: solid lines indicate membership of the *GRM*-interacting gene in the functional module, and dotted lines indicate a first-degree relationship between *GRM* receptor genes and at least one component gene of a functional module.



ganglia and cerebellum²¹. These receptors activate phospholipase C and may have a role in addiction, anxiety and behavioral disorders²². *GRM7* and *GRM8* are members of group 3 and inhibit the cyclic AMP cascade. *GRM7* has been linked to anxiety²³ and is the most highly conserved *GRM* member across multiple species²⁴.

Evidence for glutamatergic involvement in ADHD is emerging from diverse fields. Although association studies investigating *GRM* genes and transporters have reported mixed results^{25–28}, a GWAS examining the methylphenidate response in children with ADHD found an association with a SNP in *GRM7* (rs3792452)²⁹. *GRIN2A* was found to be associated with ADHD in a linkage study²⁵, and *GRIN2B* was also found to be associated with ADHD using a TDT³⁰. Magnetic resonance spectroscopy showed an increased glutamatergic tone in the frontal and striatal brain of subjects with ADHD^{31–33} that normalized with stimulants and atomoxetine³⁴. Mice in the ADHD *Slc6a3* knockout model remain responsive to methylphenidate and lack the dopamine transporter³⁵, and the hyperactivity that is increased by *N*-methyl-D-aspartic acid (NMDA) receptor blockers is suppressed by drugs that increase

glutamatergic transmission³⁶. Increased *Slc6a3* and *Drd4* expression in the midbrain were reported in rats that had an increase of glutamate transporter in the striatum³⁷, suggesting that decreases in dopamine alter glutamate signaling. *GRIN2A* disruption increased *N*-methyl D-aspartate and serotonin metabolism in the frontal cortex and striatum of mice and increased locomotor activity that had been reversed by dopamine or serotonin receptor antagonists³⁸. Dysregulated expression of glutamatergic pathway genes has been observed in spontaneously hypertensive rat models^{39–42} and in ADHD rat models with exposure to polychlorinated biphenyls⁴¹. Increased concentrations of glutamate were also reported in the neurometabolism of ADHD brains, which is consistent with altered glutamate transmission in ADHD³³.

Apart from genes in the *GRM* family, we detected association at eight other loci, four of which directly affect genes (**Table 1**). *DPP6* has been associated with amyotrophic lateral sclerosis in previous GWAS^{43,44}, and CNVs have been associated with autism⁴⁵. *DPP6* and *CTNNA2*

were also implicated in a previous GWAS¹⁰. *NLN* is responsible for the metabolic inactivation of neural peptides, including neuropeptide Y, which has been implicated in ADHD^{46,47}. *SLC7A10* modulates glutamatergic synapse transmission. *LARP7* is important for small nuclear ribonucleoprotein integrity. *NEGR1* encodes the neural cell adhesion and growth molecule, which is associated with obesity⁴⁸.

ADHD is phenotypically complex. In addition to ADHD, one of the three siblings with a deletion in *GRM5* that we studied also had symptoms of social avoidance, one sibling had coexisting obsessive compulsive symptoms, and one sibling was free of comorbidity. Assessment of the mother of these three siblings using the adult ADHD self-report scale⁴⁹ indicated a likelihood of ADHD. In subjects with *GRM7* CNVs, one had comorbid anxiety and another had coexisting oppositional defiant disorder. In subjects with *GRM1* CNVs, one had comorbid minor depression that was considered secondary to the ADHD symptoms, two had oppositional defiant disorder and a third had obsessive compulsive symptoms. Subjects with *GRM* CNVs showed a truncated normal distribution of intelligence quotient and all had intelligence quotients above 75 (Supplementary Fig. 8). In the IMAGE cohort, two subjects with *GRM* CNVs had a comorbidity of conduct disorder and four had oppositional defiant disorder.

Evaluation of the genes interacting with *GRM* genes for CNV frequency in cases and controls allowed for the inclusion of marginally significant loci, given the prior knowledge of robust association of the *GRM* receptor gene family. Whereas most reported disease-associated CNVs are rare (<1%), they have strong correlation to disease. For example, the *GRM* CNVs associated with ADHD confer large effect sizes (see the odds ratios in Table 1). Based on loci that are significant individually, 3.66% of the cases with ADHD have the newly discovered CNVs, and this number increases to 9.94% when genes interacting with *GRM* genes are included. Major hubs of the network include *TNIK*⁵⁰, *GNAQ*⁵¹ and *CALM1* (ref. 52) (Fig. 2a). The network gene *GRIK1* was previously associated with the hyperactive and impulsive symptoms of ADHD⁹.

Taken together, our CNV analysis shows that the *GRM* gene family and genes interacting with it are enriched for CNVs in individuals with ADHD. Several of these genes are crucial in the process of synaptic transmission, in neurogenesis and in neuronal processes that may be defective in ADHD. The observed *GRM* gene modules regulate RNA binding, processing and alternative splicing, which are processes known to influence brain-specific synaptic activity^{53,54}. Abnormal brain connectivity has also been observed in developmental brain disorders with cognitive dysfunction, including ADHD^{55,56}.

In conclusion, using two-stage genome-wide association for high-resolution CNV detection, we identified 12 loci showing enrichment of CNVs in cases with ADHD compared to controls and successfully replicated four genes using independent datasets of cases with ADHD and healthy controls genotyped on matched case-control platforms. The four replicating genes belong to the metabotropic glutamate receptor gene family. Extended studies identified over 200 genes interacting with glutamate receptors that were collectively affected by CNVs, suggesting that up to 10% of individuals with ADHD may be enriched for *GRM* network variants. This *GRM*-interacting gene network defines a set of functional modules that, when affected by CNVs, may contribute to the pathogenesis of ADHD. Therefore, enrichment of CNVs in genes within this molecular system that are associated with ADHD has suggested new susceptibility mechanisms and is likely to spur assessment of additional variations, including single-base changes, and expression and functional assays to evaluate the biological effects of these CNVs. Future work will determine whether clinical studies using selective *GRM* agonists as a potential treatment for ADHD are warranted in individuals with ADHD and variants in *GRM* genes.

URLs. PBAT, <http://www.biostat.harvard.edu/~clange/default.htm>; PennCNV, <http://www.openbioinformatics.org/penncnv/>; ParseCNV, <http://parsecnv.sourceforge.net/>; transformation into LRR and BAF values using PennCNV, http://www.openbioinformatics.org/penncnv/penncnv_tutorial_affy_gw6.html.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturegenetics/>.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all the children with ADHD and their families who participated in this study and all the control subjects who donated blood samples to The Children's Hospital of Philadelphia (CHOP) for genetic research purposes. We thank the technical staff at the Center for Applied Genomics, CHOP for generating the genotypes used in this study and the medical assistants, and nursing and medical staff who recruited the subjects. We thank the Center for Biomedical Informatics for bioinformatics support. We also thank S. Kristinsson, L.A. Hermansson and A. Krisbjörnsson for their software design and contribution to the study.

We thank the IMAGE, IMAGE II and PUWMa consortium investigators and the NIMH for making the genotype data available. Funding support for International Multi-Center ADHD Genetics (IMAGE) and IMAGE II Projects was provided by US National Institutes of Health (NIH) grant R01MH62873 to S.V.F., and the genotyping of samples was provided through the Genetic Association Information Network (GAIN). The dataset used for the IMAGE analyses described in this manuscript were obtained from the database of Genotype and Phenotype (dbGaP), which is found at <http://www.ncbi.nlm.nih.gov/gap>, through dbGaP accession numbers phs000016.v2.p2 (ADHD IMAGE), phs000020.v2.p1 (depression) and phs000019.v1.p1 (psoriasis). Samples and associated phenotype data for the GAIN Major Depression: Stage 1 Genome-wide Association In Population Based Samples Study (principal investigator, P.F. Sullivan, University of North Carolina) were provided by D.I. Boomsma and E. de Geus, Vrije Universiteit Amsterdam (principal investigators, Netherlands Twin Register); B.W. Penninx, Vrije Universiteit Medical Center Amsterdam; F. Zitman, Leiden University Medical Center, Leiden; and W. Nolen, University Medical Center Groningen (principal investigators and site principal investigators, Netherlands Study of Depression and Anxiety). Samples and associated phenotype data for the Collaborative Association Study of Psoriasis were provided by J.T. Elder (University of Michigan, Ann Arbor, Michigan), G.G. Krueger (University of Utah, Salt Lake City, Utah), A. Bowcock (Washington University, St. Louis, Missouri) and G.R. Abecasis (University of Michigan, Ann Arbor, Michigan). Samples and associated phenotype data for the International Multi-Center ADHD Genetics Project were provided by the following investigators: S.V.F. (principal investigator), R.J.L.A., P.A., J.S., R.P.E., B.F., M. Gill, A. Miranda, F. Mulas, R.D.O., H.R., A. Rothenberger, T.B., J. Buitelaar, E. Sonuga-Barke and H.-C.S. (site principal investigators), M. Daly, C. Lange, N. Laird, J. Su and B. Neale (statistical analysis team). Samples and associated phenotype data were accessed through an authorized data access request by H.H. and S.F.A.G. Data collection for the PUWMa sample was supported by NIH grants to S.V.F., J. Biederman, S. Smalley, R. Todd and A.A.T. GWAS genotyping of the PUWMa samples was completed by Genizon and was provided through a grant for genotyping services to S.V.F. The PUWMa consortium represents a Pfizer-funded collaboration among the University of California Los Angeles, Washington University and Massachusetts General Hospital. We thank M.C. O'Donovan, M. Gill, M.J. Owen, P.A. Holmans, A. Thapar, B.M. Neale and A. Miranda for contributing DNA samples and phenotypes to the study and for editing the manuscript. We thank G. DePalma, T. Töpner, A. Guiney and H. Zhang for providing samples for the qRT-PCR CNV validation.

The study was supported by an Institutional Development Award to the Center for Applied Genomics from CHOP (H.H.), which funded all of the discovery genome-wide genotyping for this study. This work was additionally supported in part by NIH grant K23MH066275 (J.E.), University of Pennsylvania National Center for Research Resources Clinical and Translational Science Awards grant UL1-RR-024134 (J.E.), by a Research Development Award from the Cotswold Foundation (H.H. and S.F.A.G.) and by US Department of Health and Human Services grant IRC2MH089924 (H.H.). N.T., T.S. and J.D.B. received support from the Seaver Foundation and a Conte Center for Neuroscience of Mental Disorders grant from the NIMH (P50MH066392).

Sample and phenotype data collection of parts of the IMAGE II cohort was supported by the Deutsche Forschungsgemeinschaft (KFO 125, SFB 581 and SFB

TRR 58 to K.-P.L. and A. Reif, ME 1923/5-1 and ME 1923/5-3 to C.M.F. and J.M.) and the Bundesministerium für Bildung und Forschung (BMBF 01GV0605 to K.-P.L.).

AUTHOR CONTRIBUTIONS

H.H. and J.E. designed the CHOP study and supervised the data analyses and interpretation. S.V.F., M.G., P.A. and J. Buitelaar designed the IMAGE and IMAGE II studies. S.V.F. designed the PUWMA study and coordinated the analyses for the IMAGE, IMAGE II and PUWMA studies. J.T.G. and K.W. conducted the statistical analyses. C.E.K. and E.C.F. directed the stage 1 genotyping. J.D.B. coordinated the validation analyses. N.T. performed the qRT-PCR validation of the CNVs. J.T.G. and H.H. drafted the manuscript. J.E. collected the CHOP samples. C.R., P.S. and J.L.R. collected the NIMH samples. C.M.F., H.-C.S., A.A.T., A. Reif, A. Rothenberger, B.F., E.O.M., H.R., J. Buitelaar, K.-P.L., L.K., T.B., R.P.E., F.M., R.D.O., J.S., E.S.-B., T.J.R., M.R., J.R., A.W., S.W., J.M., H.P., C.S., S.K.L., S.L.S., J. Biederman, L.K., P.A. and R.J.L.A. collected data for the IMAGE, IMAGE II and PUWMA projects. J. Biederman, E.O.M., S.V.F., S.K.L., S.L.S. and A.A.T. collected samples for the PUWMA study. F.A.M. genotyped the IMAGE II data. H.H. directed and D.H. and J.T.G. performed the gene interaction network and functional enrichment analyses. All authors contributed to the manuscript preparation. S.F.A.G. accessed the public domain data, assisted with the interpretation of the data and edited the manuscript. All other authors contributed samples and/or were involved with data mining and processing.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturegenetics/>.

Published online at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Derks, E.M. *et al.* Genetic and environmental influences on the relation between attention problems and attention deficit hyperactivity disorder. *Behav. Genet.* **38**, 11–23 (2008).
- Wood, A.C., Rijdsdijk, F., Saudino, K.J., Asherson, P. & Kuntsi, J. High heritability for a composite index of children's activity level measures. *Behav. Genet.* **38**, 266–276 (2008).
- Haberstick, B.C. *et al.* Genetic and environmental contributions to retrospectively reported DSM-IV childhood attention deficit hyperactivity disorder. *Psychol. Med.* **38**, 1057–1066 (2008).
- Glessner, J.T. *et al.* Autism genome-wide copy number variation reveals ubiquitous and neuronal genes. *Nature* **459**, 569–573 (2009).
- Wang, K. *et al.* Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459**, 528–533 (2009).
- Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368–372 (2010).
- Franke, B., Neale, B.M. & Faraone, S.V. Genome-wide association studies in ADHD. *Hum. Genet.* **126**, 13–50 (2009).
- Neale, B.M. *et al.* Genome-wide association scan of attention deficit hyperactivity disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B**, 1337–1344 (2008).
- Lasky-Su, J. *et al.* Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B**, 1345–1354 (2008).
- Lesch, K.P. *et al.* Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. *J. Neural. Transm.* **115**, 1573–1585 (2008).
- Williams, N.M. *et al.* Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet* **376**, 1401–1408 (2010) Epub 2010 Sep 29.
- Elia, J. *et al.* Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol. Psychiatry* **15**, 637–646 (2010).
- Elia, J., Gai, X., Hakonarson, H. & White, P.S. Structural variations in attention-deficit hyperactivity disorder. *Lancet* **377**, 377–378 (2011).
- Wang, K. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
- Colella, S. *et al.* QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res.* **35**, 2013–2025 (2007).
- Zhou, K. *et al.* Meta-analysis of genome-wide linkage scans of attention deficit hyperactivity disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B**, 1392–1398 (2008).
- Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
- Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
- Taniura, H., Sanada, N., Kuramoto, N. & Yoneda, Y. A metabotropic glutamate receptor family gene in *Dictyostelium discoideum*. *J. Biol. Chem.* **281**, 12336–12343 (2006).
- Conn, P.J. & Pin, J. Pharmacology and functions of metabotropic glutamate receptors. *Annu. Rev. Pharmacol. Toxicol.* **37**, 205–237 (1997).
- Berthele, A. *et al.* Expression of metabotropic glutamate receptor subtype mRNA (mGluR1–8) in human cerebellum. *Neuroreport* **10**, 3861–3867 (1999).
- Koob, G.F., Sanna, P.P. & Bloom, F.E. Neuroscience of addiction. *Neuron* **21**, 467–476 (1998).
- Cryan, J.F. Antidepressant and anxiolytic-like effects in mice lacking the group III metabotropic glutamate receptor mGluR7. *Eur. J. Neurosci.* **17**, 2409–2417 (2003).
- Makoff, A., Pillinga, C., Harrington, K. & Emson, P. Human metabotropic glutamate receptor type 7: Molecular cloning and mRNA distribution in the CNS. *Brain Res. Mol. Brain Res.* **40**, 165–170 (1996).
- Turic, D. *et al.* Follow-up of genetic linkage findings on chromosome 16p13: evidence of association of N-methyl-D aspartate glutamate receptor 2A gene polymorphism with ADHD. *Mol. Psychiatry* **9**, 169–173 (2004).
- Mick, E. & Faraone, S.V. Genetics of attention deficit hyperactivity disorder. *Child Adolesc. Psychiatr. Clin. N. Am.* **17**, 261–284 (2008).
- Turic, D. *et al.* A family based study implicates solute carrier family 1-member 3 (SLC1A3) gene in attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **57**, 1461–1466 (2005).
- Elia, J. *et al.* Candidate gene analysis in an on-going genome-wide association study of attention-deficit hyperactivity disorder: suggestive association signals in ADRA1A. *Psychiatr. Genet.* **19**, 134–141 (2009).
- Mick, E., Neale, B., Middleton, F.A., McGough, J.J. & Faraone, S.V. Genome-wide association study of response to methylphenidate in 187 children with attention-deficit/hyperactivity disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B**, 1412–1418 (2008).
- Dorval, K.M. *et al.* Association of the glutamate receptor subunit gene GRIN2B with attention-deficit/hyperactivity disorder. *Genes Brain Behav.* **6**, 444–452 (2007).
- Jin, Z., Zang, Y.F., Zeng, Y.W., Zhang, L. & Wang, Y.F. Striatal neuronal loss or dysfunction and choline rise in children with attention-deficit hyperactivity disorder: a 1H-magnetic resonance spectroscopy study. *Neurosci. Lett.* **315**, 45–48 (2001).
- MacMaster, F.P., Carrey, N., Sparkes, S. & Kusumakar, V. Proton spectroscopy in medication-free pediatric attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **53**, 184–187 (2003).
- Courovoise, H., Hooper, S.R., Fine, C., Kwock, L. & Castillo, M. Neurometabolic functioning and neuropsychological correlates in children with ADHD-H: preliminary findings. *J. Neuropsychiatry Clin. Neurosci.* **16**, 63–69 (2004).
- Carrey, N. *et al.* Metabolite changes resulting from treatment in children with ADHD: a 1H-MRS study. *Clin. Neuropharmacol.* **26**, 218–221 (2003).
- Gainetdinov, R.R. *et al.* Role of serotonin in the paradoxical calming effect of psychostimulants on hyperactivity. *Science* **283**, 397–401 (1999).
- Gainetdinov, R.R., Mohn, A.R., Bohn, L.M. & Caron, M.G. Glutamatergic modulation of hyperactivity in mice lacking the dopamine transporter. *Proc. Natl. Acad. Sci. USA* **98**, 11047–11054 (2001).
- Masuo, Y., Ishido, M., Morita, M. & Oka, S. Effects of neonatal 6-hydroxydopamine lesion on the gene expression profile in young adult rats. *Neurosci. Lett.* **335**, 124–128 (2002).
- Miyamoto, K. *et al.* Involvement of enhanced sensitivity of N-methyl-D-aspartate receptors in vulnerability of developing cortical neurons to methylmercury neurotoxicity. *Brain Res.* **901**, 252–258 (2001).
- Russell, V., Allie, S. & Wiggins, T. Increased noradrenergic activity in prefrontal cortex slices of an animal model for attention-deficit hyperactivity disorder—the spontaneously hypertensive rat. *Behav. Brain Res.* **117**, 69–74 (2000).
- Russell, V.A. Dopamine hypofunction possibly results from a defect in glutamate-stimulated release of dopamine in the nucleus accumbens shell of a rat model for attention deficit hyperactivity disorder—the spontaneously hypertensive rat. *Neurosci. Biobehav. Rev.* **27**, 671–682 (2003).
- DasBanerjee, T. *et al.* A comparison of molecular alterations in environmental and genetic rat models of ADHD: a pilot study. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B**, 1554–1563 (2008).
- Sagvolden, T. *et al.* The spontaneously hypertensive rat model of ADHD—the importance of selecting the appropriate reference strain. *Neuropharmacology* **57**, 619–626 (2009).
- Del Bo, R. *et al.* DPP6 gene variability confers increased risk of developing sporadic amyotrophic lateral sclerosis in Italian patients. *J. Neurol. Neurosurg. Psychiatry* **79**, 1085 (2008).
- Cronin, S., Tomik, B., Bradley, D.G., Slowik, A. & Hardiman, O. Screening for replication of genome-wide SNP associations in sporadic ALS. *Eur. J. Hum. Genet.* **17**, 213–218 (2009).
- Marshall, C.R. *et al.* Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 477–488 (2008).
- Lesch, K.P. *et al.* Genome-wide copy number variation analysis in ADHD: association with neuropeptide Y gene dosage in an extended pedigree. *Mol. Psychiatry* **16**, 491–503 (2011).
- Oades, R.D., Daniels, R. & Rascher, W. Plasma neuropeptide Y levels, monoamine metabolism, electrolyte excretion, and drinking behavior in children with attention-deficit hyperactivity-disorder (ADHD). *Psychiatry Res.* **80**, 177–186 (1998).

48. Renström, F. *et al.* Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden. *Hum. Mol. Genet.* **18**, 1489–1496 (2009).
49. Kessler, R.C. *et al.* Patterns and predictors of attention-deficit/hyperactivity disorder persistence into adulthood: results from the national comorbidity survey replication. *Biol. Psychiatry* **57**, 1442–1451 (2005).
50. Potkin, S.G. *et al.* FBIRN. A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. *Schizophr. Bull.* **35**, 96–108 (2009).
51. Wang, X., Bao, X., Pal, R., Agbas, A. & Michaelis, E.K. Transcriptomic responses in mouse brain exposed to chronic excess of the neurotransmitter glutamate. *BMC Genomics* **11**, 360 (2010).
52. de Lanerolle, N.C., Eid, T. & Lee, T.S. Genomic expression in the epileptogenic hippocampus and psychiatric co-morbidities. *Curr. Psychiatry Rev.* **6**, 135–144 (2010).
53. Ule, J. *et al.* An RNA map predicting Nova-dependent splicing regulation. *Nature* **444**, 580–586 (2006).
54. Ule, J. *et al.* Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* **37**, 844–852 (2005).
55. Murias, M., Swanson, J.M. & Srinivasan, R. Functional connectivity of frontal cortex in healthy and ADHD children reflected in EEG coherence. *Cereb. Cortex* **17**, 1788–1799 (2007).
56. Wang, L. *et al.* Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Hum. Brain Mapp.* **30**, 638–649 (2009).

¹Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. ²Department of Child and Adolescent Psychiatry, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ³Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ⁴Laboratory of Molecular Neuropsychiatry, Department of Psychiatry, Mount Sinai School of Medicine, New York, New York, USA. ⁵Department of Human Genetics and Psychiatry, University of California Los Angeles, Los Angeles, California, USA. ⁶Department of Psychiatry, University of Utah, Salt Lake City, Utah, USA. ⁷Child Psychiatry Branch, National Institute of Mental Health, Bethesda, Maryland, USA. ⁸Department of Neuroscience and Physiology, State University of New York Upstate Medical University, Syracuse, New York, USA. ⁹Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Goethe University, Frankfurt am Main, Germany. ¹⁰Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland. ¹¹Aalborg Psychiatric Hospital, Aarhus University Hospital, Aarhus, Denmark. ¹²Institute of Psychology, University of Basel, Basel, Switzerland. ¹³Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA. ¹⁴Attention Deficit Hyperactivity Disorder Clinical Research Network, Unit of Molecular Psychiatry, Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany. ¹⁵Child and Adolescent Psychiatry, University of Göttingen, Göttingen, Germany. ¹⁶Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. ¹⁷Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Centre for Neuroscience, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. ¹⁸Department of Psychiatry, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, USA. ¹⁹Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium. ²⁰Radboud University Nijmegen Medical Centre, Donders Institute for Brain Cognition and Behaviour, Department of Cognitive Neuroscience, Nijmegen, The Netherlands. ²¹Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany. ²²Department of Psychology, National University of Singapore, Queenstown, Singapore. ²³Department of Neuropaediatrics, La Fe University Hospital, Faculty of Medicine, Valencia, Spain. ²⁴Clinic for Child and Adolescent Psychiatry, University of Duisburg-Essen, Essen, Germany. ²⁵Vrije Universiteit, De Boelelaan, Amsterdam, The Netherlands. ²⁶School of Psychology, Institute for Disorder on Impulse and Attention, University of Southampton, Highfield, Southampton, UK. ²⁷Department of Experimental Clinical and Health Psychology, Ghent University, Dunantlaan, Ghent, Belgium. ²⁸Child Study Center, New York University, New York, New York, USA. ²⁹Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany. ³⁰Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Munich, Munich, Germany. ³¹Institute of Psychobiology, Department of Neurobehavioral Genetics, University of Trier, Trier, Germany. ³²Department of Child and Adolescent Psychiatry, Saarland University, Homburg, Germany. ³³Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, California, USA. ³⁴Bute Medical School, St. Andrews, Scotland, UK. ³⁵Department of Psychiatry, Trinity Centre for Health Sciences, St. James's Hospital, Dublin, Ireland. ³⁶Division of Cardiothoracic Surgery, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ³⁷Division of Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ³⁸Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. ³⁹Dipartimento di Medicina Sperimentale, University La Sapienza, Rome, Italy. ⁴⁰Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. ⁴¹Center for Biomedical Informatics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ⁴²Division of Oncology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ⁴³Department of Psychological Medicine, School of Medicine, Cardiff University, Cardiff, Wales, UK. ⁴⁴Department of Psychiatry, State University of New York Upstate Medical University, Syracuse, New York, USA. ⁴⁵Division of Pulmonary Medicine, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. ⁴⁶These authors contributed equally to this work. Correspondence should be addressed to H.H. (hakonarson@email.chop.edu).

ONLINE METHODS

Illumina Infinium assay for CNV discovery. We performed high-throughput genome-wide SNP genotyping using the InfiniumII HumanHap550 BeadChip (Illumina) at the Center for Applied Genomics at the CHOP. The genotype data content together with the intensity data from the genotyping array provided high confidence for the CNV calls. A simultaneous analysis of intensity data and genotype data in the same experimental setting established a highly accurate definition for normal diploid states and any deviation thereof. To call CNVs, we used PennCNV and QuantiSNP. PennCNV combines multiple sources of information, including LRR and BAF, SNP spacing, trained Hidden Markov Models and the population frequency of the B allele to generate CNV calls. The replication case and control cohorts performed genome-wide SNP genotyping using the Perlegen 600K, Illumina 1M and Affymetrix 5.0 arrays. Raw X and Y values were \log_{10} normalized and clustered to establish the BAF and LRR values using the PennCNV-Affy protocol (Online Methods and **Supplementary Table 12**).

CNV quality control. We calculated quality control measures on our HumanHap550 GWAS data based on statistical distributions to exclude DNA samples of poor quality and false-positive CNVs. The first quality control threshold used was the percentage of attempted SNPs that were successfully genotyped. Only samples with call rate >98% were included. The genome-wide intensity signal should have as little noise as possible. Only samples within the s.d. of the normalized intensity ($LRR < 0.35$) were included. All samples were required to be of European descent based on principle components analysis (**Supplementary Fig. 9**). Furthermore, case and control matching was insured by calculating a genomic inflation factor (which was 1.024) between groups. Wave artifacts roughly correlating with GC content resulting from hybridization bias of low full-length DNA quantity are known to interfere with the accurate inference of copy number variations⁵⁷. Only samples where |GC base pair wave factor (GCWF)| < 0.05 were accepted. If the count of CNV calls made by PennCNV exceeds 70 (**Supplementary Fig. 1**), the DNA quality is usually poor. Therefore, only samples with CNV call count <70 were included. One sample from any duplicate samples (such as those from monozygotic twins) was excluded. **Supplementary Table 13** lists the number of samples excluded using each quality control measure.

Statistical analysis of CNVs. The CNV frequency between cases and controls was evaluated at each SNP using a Fisher's exact test. We only considered loci that were nominally significant between cases and controls ($P < 0.05$) when cases in the CHOP discovery cohort had the same variation, which was replicated in the IMAGE, PUWMA or IMAGE II studies, or loci that were not observed in any of the control subjects and were validated using an independent method. We reported statistical local minimums to narrow the association in reference to a region of nominal significance, including SNPs residing within 1 Mb of each other (**Supplementary Fig. 10**). The resulting nominally significant CNVRs were excluded if they met any of the following criteria: (i) residing on telomere- or centromere-proximal cytobands; (ii) located in a 'peninsula' of common CNVs arising from variation in boundary truncation of CNV calling (**Supplementary Fig. 11**); (iii) genomic regions with extremes of GC content, which produces hybridization bias; or (iv) samples contributing to multiple CNVRs. See ParseCNV for details (see URLs). We statistically adjusted for the relatedness of cases using permutation (1,000×). Three lines of evidence were considered to establish statistical significance: independent replication with $P < 0.05$, permutation significance of the observations and no control-enriched significance. We used DAVID (Database for Annotation, Visualization and Integrated Discovery)⁵⁸ to assess the significance of the functional annotation clustering of independently associated CNV results into InterPro categories.

Network analysis. We used Cytoscape software¹⁸ to identify 228 genes within two degrees of relation to eight *GRM* genes based on the merged human interactome. We clustered this network of genes into 17 distinct modular clusters based solely on network topology using the ClusterViz plugin for the software using the FAG-EC algorithm with default parameters. Component genes in each of the 17 modules were submitted to DAVID⁵⁸ to assess the significance of functional enrichment using *Homo sapiens* Gene Ontology annotations.

CNV validation by qRT-PCR. Universal Probe Library (Roche) probes were selected using the ProbeFinder v2.41 software (Roche). qRT-PCR was performed on an ABI 7500 Real Time PCR Instrument or on an ABI Prism 7900HT Sequence Detection System (Applied Biosystems). Each sample was analyzed in quadruplicate either in 25 μ l of reaction mixture (250 nM probe, 900 nM of each primer, the Fast Start TaqMan Probe Master from Roche and 10 ng of genomic DNA) or in 10 μ l of reaction mixture (100 nM probe, 200 nM of each primer, 1× Platinum Quantitative PCR SuperMix-Uracil-DNA-Glycosylase with ROX from Invitrogen and 25 ng of genomic DNA). The results were evaluated using Sequence Detection Software v2.2.1 (Applied Biosystems). Further data analysis was performed using the $\Delta\Delta C_T$ method. Reference genes, chosen from *COBL*, *GUSB* and *SNCA*, were included based on the minimal coefficient of variation, and data was then normalized by setting a normal control to a value of 1.

PennCNV-Affy workflow adapted to Perlegen 600K data. CNV calling on Perlegen used a similar algorithm to the Illumina arrays but had additional preprocessing. To perform data normalization and signal extraction from the raw final report files generated in the genotyping experiments, we first reformatted data from dbGaP into the format produced by the following Affymetrix Power Tools: birdseed.calls.txt, birdseed.confidences.txt and quant-norm.pm-only.med-polish.expr.summary.txt (**Supplementary Table 11**). We calculated the \log_{10} of the X and Y values provided in the sample-based report files. For each SNP marker, the allele-specific intensity for the AA, AB and BB genotypes on all genotyped samples was used to construct three canonical genotype clusters in the polar coordinates θ and R. Once canonical genotype clusters were constructed, signal intensity values for each SNP were transformed into LRR and BAF values. For details, see URLs.

To optimize the Hidden Markov Model, we used the reference file hh550.hmm and ran "train" in PennCNV in one batch of thirty samples with the lowest s.d. of the LRR value followed by two batches that included random representative samples. We also created a population B allele frequency definition file specifically adapted to the Perlegen data. This allowed for CNV calls to be made in 1,887 (642 cases and 1,245 parents) out of 2,789 Perlegen 600K samples available. Although the global s.d. of the LRR value was below 0.2 for the majority (84%) of the samples, the intensity data was noisy in regions of called CNVs and showed subpopulations of SNPs that were unable to differentiate the deletion signal, perhaps as a result of PCR saturation during lab processing. Nevertheless, deletion and duplication features were detected with confirmation of homozygote and AAB and ABB genotypes (**Supplementary Figs. 6 and 7**). Lastly, Perlegen CNV calls were screened for overlap with the 12 loci associated with ADHD based on the CHOP Illumina data. To ensure that each detected CNV was a true DNA feature, we validated each CNV using qRT-PCR (**Supplementary Fig. 5**).

Permutation to adjust significance for relatedness. For the initial Fisher's exact test, related individuals were not controlled for, as our primary objective was to detect CNVs in multiple samples regardless of relatedness. CNVRs passing this initial screen were scored for statistical significance based on a permuted P value, which permutes case and control labels randomly for all samples, with the condition that related individuals have same label. Based on the number of samples with the CNVR being calculated in randomly assigned 'cases' and 'controls', a Fisher's exact test P value was assigned. The number of hypothetical scenarios with significance equal or greater to this P value provides the permuted P value, which corrects for relatedness. The Fisher's exact test P value and the counts of cases and controls with each CNVR are provided in **Table 1** for transparency.

Study criteria for inclusion in IMAGE. The probands were required to have combined subtype ADHD, were children aged 6–17 years (inclusive), had one or more sibling(s) in same age range, had both parents available or one parent and two or more siblings available to provide a DNA sample, an intelligence quotient above 70, were free of single-gene disorders associated with ADHD, were free of neurological disease and damage, were living at home with at least one biological parent and one full sibling and did not meet the criteria for autism or Asperger's syndrome.

Study criteria for inclusion in IMAGE II. The probands were required to have ADHD according to the standards of the current diagnostic and statistical manual of mental health (DSM-IV-TR) and had a semi-structured diagnostic interview based on KSADS-PL or Kinder-DIPS, the Child Behavior Checklist, the Conners parent and teacher Scales or the German Teachers Report on children aged 6–18 years. Probands also had an intelligence quotient above 70, a birth weight >2,000 g, no major medical events during their mother's pregnancy and no drug abuse in their mother during pregnancy. They were also free of single-gene disorders known to be associated with ADHD and free of neurological disease and damage, and they did not meet the criteria for autism, Asperger's syndrome, schizophrenia, bipolar disorder, primary major depressive episode, anxiety disorder or Tourette's Syndrome.

Controls for IMAGE II. The control subjects used were drawn from Affymetrix 6.0 genotyped subjects from the NIMH genetics repository collected through

the United States nationally representative survey panel. Participants were screened for psychosis and bipolar disorder. Control participants were not screened for ADHD. Control participants gave written consent for their biological materials to be used for medical research at the discretion of NIMH. Controls were genotyped using the Affymetrix 6.0 array at the Broad Institute National Center for Genotyping and Analysis. Genotype calls were made with the BIRDSEED program, a module of the BIRDSUITE package.

Ethics Statement. This research was approved by the institutional review board of the Children's Hospital of Philadelphia. The appropriate informed consent was obtained for all sample donors.

57. Diskin, S.J. *et al.* Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res.* **36**, e126 (2008).
58. Dennis, G. Jr. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, P3 (2003).