Open access • Journal Article • DOI:10.1038/NG.1013

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

Josephine Elia, Joseph T. Glessner, Kai Wang, Nagahide Takahashi ...+66 more authors

Institutions: University of Pennsylvania, Children's Hospital of Philadelphia, Icahn School of Medicine at Mount Sinai, University of California, Los Angeles ...+25 more institutions

Published on: 01 Jan 2012 - Nature Genetics (Nature Publishing Group)

Topics: Copy-number variation and Single-nucleotide polymorphism

Genome Wide Copy Number Variation Study Associates Metabotropic Glutamate Receptor Genes with Attention Deficit Hyperactivity Disorder

Josephine Elia, M.D.[1,2*], Joseph T. Glessner, M.S.[3*], Kai Wang, Ph.D.[3], Nagahide Takahashi, M.D., Ph.D.[4], Corina J. Shtir, Ph.D.[5], Patrick M.A. Sleiman, Ph.D.[3], Haitao Zhang, Ph.D.[3], Cecilia E. Kim, B.S.[3], Reid Robison, Ph.D.[6], Gholson J. Lyon, M.D.,Ph.D.[6], James H. Flory, M.D., Ph.D.[3], Jonathan P. Bradfield, B.S.[3], Marcin Imielinski, M.D.,Ph.D.[3], Cuiping Hou, B.S.[3], Edward C. Frackelton, B.S.[3], Rosetta M. Chiavacci, B.S.N.[3], Takeshi Sakurai, Ph.D.[4], Cara Rabin, B.A.[7], Frank A. Middleton, Ph.D.[8], Kelly A. Thomas, B.S.[3], Maria Garris, B.S.[3], Frank Mentch, Ph.D.[3], Christine M. Freitag, M.D.[9], Hans-Christoph Steinhausen, M.D.,Ph.D.[10,11,12], Alexandre A. Todorov, Ph.D.[13], Andreas Reif, M.D.[14], Aribert Rothenberger, M.D.,Ph.D.[15], Barbara Franke, Ph.D.[16], Eric O. Mick, Sc.D.[17], Herbert Roeyers, Ph.D.[18], Jan Buitelaar, M.D. [19], Klaus-Peter Lesch, M.D.[14], Tobias Banaschewski, M.D.,Ph.D.[20], Richard P. Ebstein, Ph.D.[21], Fernando Mulas, M.D.[22], Robert D. Oades, Ph.D.[23], Joseph Sergeant, Ph.D.[24], Edmund Sonuga-Barke, Ph.D.[25,26,27], Tobias J. Renner, M.D.[28], Marcel Romanos, M.D.[28], Jasmin Romanos, M.D.[28], Andreas Warnke, M.D.[28], Susanne Walitza, M.D.[28,29], Jobst Meyer, Ph.D.[30], Haukur Pálmason, M.D.[30], Christiane Seitz, M.D.[31], Sandra K. Loo, Ph.D.[32], Susan L. Smalley, Ph.D.[32], Joseph Biederman, M.D.[17], Lindsey Kent, Ph.D.[33], Philip Asherson, M.D.,Ph.D.[29], Richard Anney, Ph.D.[34], J. William Gaynor, M.D. [35], Philip Shaw, Ph.D.[7], Marcella Devoto, Ph.D.[36,37,38,39], Struan F.A. Grant, Ph.D.[3,36,37], Peter White, Ph.D.[40,41], Joseph D. Buxbaum, Ph.D.[4], Judith L. Rapoport, M.D.[7], Nigel M. Williams, Ph.D.[42], Stanley F. Nelson, M.D.[5], Stephen V. Faraone, Ph.D.[8], Hakon Hakonarson, M.D.,Ph.D.[3,36,37,43]

[1] *Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, 19104, USA;*
[2] *Department of Child and Adolescent Psychiatry, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA;*
[3] *Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA;*
[4] *Laboratory of Molecular Neuropsychiatry, Department of Psychiatry, Mount Sinai School of Medicine, New York, NY 10029 USA;*
[5] *Department of Human Genetics and Neuroscience, University of California - Los Angeles, Los Angeles, CA 90095, USA;*
[6] *Department of Psychiatry, University of Utah, Salt Lake City, Utah, 84112;*
[7] *Child Psychiatry Branch, National Institute of Mental Health, Bethesda, MD, 20892, USA;*
[8] *Departments of Psychiatry and of Neuroscience and Physiology, State University of New York Upstate Medical University, Syracuse, New York, NY 13210, USA;*
[9] *Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, JW Goethe University, 60528 Frankfurt am Main, Germany*
[10] *Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, CH – 8008, Switzerland;*
[11] *Aalborg Psychiatric Hospital, Aarhus University Hospital, Denmark*
[12] *Institute of Psychology, University of Basel, Switzerland*
[13] *Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri 63110, USA;*
[14] *ADHD Clinical Research Network, Unit of Molecular Psychiatry, Department of Psychiatry, Psychosomatics, and Psychotherapy, University of Wuerzburg, Wuerzburg, 97080 Germany;*
[15] *Child and Adolescent Psychiatry, University of Goettingen, Germany 37073;*

[16] *Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, 6500, The Netherlands;*
[17] *Department of Psychiatry, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts 02114, USA;*
[18] *Ghent University, Dunantlaan, Ghent, B-9000 Belgium;*
[19] Radboud University Nijmegen Medical Centre, Donders Institute for Brain Cognition and Behaviour, Department of Cognitive Neuroscience, 6500 HB Nijmegen, The Netherlands;
[20] *Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany;*
[21] *Geha MHC, Petach-Tikva, Israel;*
[22]*Department of Neuropaediatrics, La Fe University Hospital, Faculty of Medicine, Valencia, Spain*
[23] *Clinic for Child and Adolescent Psychiatry, University of Duisburg-Essen, Essen, Germany*
[24] *Vrije Universiteit, De Boelelaan, Amsterdam, The Netherlands;*
[25] *MRC Social Genetic and Developmental Psychiatry, Kings's College London, SE5 8AF, UK;*
[26] *School of Psychology, Institute for Disorder on Impulse and Attention, University of Southampton, Highfield, Southampton SO17 1BJ, UK;*
[27] *Child Study Center, New York University, New York, NY 10016, USA;*
[28]*Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, 97080 Germany;*
[29] *Department of Child and Adolescent Psychiatry, University of Zuerich, Zuerich, 8032 Switzerland;*
[30] *Institute of Psychobiology, Department of Neurobehavioral Genetics, University of Trier, Trier, Germany;*
[31] *Department of Child and Adolescent Psychiatry, Saarland University, Homburg, Germany*
[32] *Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience & Human Behavior, UCLA, Los Angeles, California 90095 USA;*
[33] *Bute Medical School, St Andrews, Scotland, UK;*
[34] *Department of Psychiatry, Trinity Centre for Health Sciences, St. James's Hospital, Dublin 8, Ireland;*
[35] Division of Cardiothoracic Surgery, The Children's Hospital of Philadelphia, Philadelphia, Pa.
[36] *Division of Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA;*
[37] *Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, 19104, USA;*
[38] *Dipartimento di Medicina Sperimentale, University La Sapienza, 00185 Rome, Italy;*
[39] *Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA;*
[40] *Center for Biomedical Informatics, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA;*
[41] *Division of Oncology, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA;*
[42] *Department of Psychological Medicine, School of Medicine, Cardiff University, Cardiff, Wales, CF10 3xQ, UK;*
[43] *Division of Pulmonary Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA;*

*equal contribution

Background

Attention-Deficit, Hyperactivity Disorder (ADHD) is a common, heritable neuropsychiatric disorder of unknown etiology. Recently, we reported an enrichment of rare variants in genes involved in learning, behavior, synaptic transmission and central nervous system development in autism [1], suggesting that rare inherited structural variants could also play a role in the etiology of ADHD, a related neuropsychiatric disorder.

Methods

We performed a whole-genome CNV study in a cohort of 1,013 ADHD cases and 4,105 healthy children of European ancestry who were genotyped with 550,000 SNP markers. Positive findings were evaluated in multiple independent cohorts, totaling 2,493 ADHD cases and 9,222 controls of European ancestry, with respective case-control cohorts genotyped on matched platforms.

Results

CNVs impacting metabotropic glutamate receptor genes were significantly enriched across all independent cohorts (P= 2.1x10^{-9}). Among them, deletions in *GRM5* (glutamate receptor, metabotropic 5) occurred in ten cases across three independent cohorts and in only one control subject (P=1.36x10^{-6}). In addition, deletions in *GRM7* occurred in six cases and *GRM8* in eight cases, both with a control frequency of zero. *GRM1* was duplicated in eight cases, a frequency notably enriched above controls. Observed variants were experimentally validated using quantitative PCR.

Conclusions

We have identified several rare recurrent CNVs that are overrepresented in multiple independent ADHD cohorts that impact genes involved in glutamatergic neurotransmission, an important mediator for the developing brain and normal brain function. These results suggest that

variations involving glutamatergic gene networks of the brain contribute to the genetic susceptibility to ADHD.

Attention Deficit Hyperactivity Disorder (ADHD) is a common neuropsychiatric disorder with heritability estimates ranging from 30 to 90% [2-4]. Most neurodevelopmental disorders have been resistant to the genome wide association (GWA) approach, although recent progress has been made in autism [1,5]. GWA studies have been reported in ADHD utilizing a cohort of 958 parent-child trios recruited through the International Multicentre ADHD Genetics (IMAGE) study. Results of these studies did not report any association at genome-wide significance level [6,7]. Using quantitative measures of ADHD, Lasky-Su and colleagues recently reported nominal evidence from a PBAT analysis of tagging SNPs located at *CDH13* (rs6565113) and *GFOD1* (rs552655) [8]. A SNP in strong linkage disequilibrium with rs6565113 impacting *CDH13* was also reported in a GWA study of an independent sample of ADHD adults [9].

We performed a two-stage whole-genome CNV study in a large cohort of ADHD cases and healthy children of European ancestry who were genotyped with the Illumina Infinium HumanHap550K BeadChip. Positive findings were evaluated in multiple independent replication cohorts of similar size also of European ancestry.

Study Participants: The discovery cohort included a total of 1,013 ADHD cases of Northern European descent genotyped at CHOP. This consisted of 664 cases without parents and 349 cases from complete trios recruited at CHOP (Supplementary Tables 1 and 2). To address

replication, we accessed the IMAGE cohorts which are a part of the Genetic Association Information Network (GAIN). There were 624 IMAGE samples that met quality control criteria for the study. Access to these genotypes and intensity data for IMAGE was provided through the database of Genotypes and Phenotypes (dbGaP) (http://www.ncbi.nlm.nih.gov/gap). The PUWMa consortium contributed 864 ADHD cases and 1,258 parents. The IMAGE II consortium contributed 787 ADHD cases and 898 unrelated controls. Furthermore, 128 cases recruited at the NIMH and 90 cases recruited at The University of Utah also served for replication. The DNA samples from CHOP, NIMH, and Utah cohorts were genotyped using the Illumina Infinium HumanHap550K BeadChip at CHOP. The IMAGE cohort was genotyped using the Perlegen 600K platform. The PUWMa cohort was genotyped on the Illumina 1M BeadChip. The IMAGE II cohort was genotyped on the Affymetrix 5.0 array. To manage differences in CNV detection between arrays we used controls genotyped on platforms matching the case platforms, including: 4,105 Illumina 550k from CHOP, 3,297 Perlegen 600k from GAIN psoriasis and depression projects, 3,469 Illumina 1M from PUWMa parents and SAGE, and 2,456 Affymetrix 5.0 and 6.0 controls from the NIMH genetics repository and AGRE parents.

Results

CNV size and number in cases and controls: To search for novel CNVs we analyzed the 1,013 CHOP cases as a discovery cohort in comparison with 4,105 control children, all of whom were of European ancestry. Data from the IMAGE, PUWMa, IMAGE II, NIMH, and Utah cohorts were used for replication, together with an independent control cohort of 9,222 genotyped on the same platforms. Thus, the control CNV frequency is robustly characterized in multiple large independent cohorts, based on the Illumina, Perlegen, and Affymetrix platforms. We note that of

the 2,713 (934 cases) IMAGE samples available in dbGaP, 1,886 (624 cases) met strictly established data quality thresholds for CNVs (see Methods).

The PennCNV software was used to produce CNV calls for cases and controls as previously described [10]. The CNV frequency of the subjects who met quality standards, which included removing substantial outliers in the count CNV call quality metric that deviated exponentially from the distribution of the majority of the cohort, resulted in 93% of subjects having 8-45 CNV calls (Supplementary Figure 1). We called four different copy number states, including 3,172 homozygous deletions (copy number, or CN =0), 27,810 hemizygous deletions (CN =1), 14,806 one copy duplications (CN =3), and 581 two copy duplications (CN =4). Supplementary Figure 2 shows an example of raw Illumina data as viewed in the BeadStudio software and the resulting CNV call. The CNV calls spanned from 3 to 598 SNPs, with an average of 14 SNPs per CNV call, with the largest CNV of 2.2 Mb and an average CNV size of 62 kb. Variable probe coverage allows for detection of CNVs down to a small physical size, provided at least 3 SNPs are present, and the CNVs were experimentally validated using qPCR.

Control individuals examined also had 93% of subjects with 8-45 CNV calls (Supplementary Figure 1). Among the CNV calls, we identified 4,471 homozygous deletions (CN =0), 49,726 hemizygous deletions (CN =1), 27,032 one copy duplications (CN =3), and 1,480 two copy duplications (CN =4). The CNV calls spanned from 3 to 708 SNPs, with an average of 12.8 SNPs per CNV call, with the largest CNV of 2.9 Mb and an average CNV size of 53.6 kb.

SNP association testing: We performed GWA analysis on the discovery cohort, however, we did not detect any single SNP genotype association signals that met statistical criteria for genome-wide significance ($p<5x10^{-8}$) (See Supplement: Analysis of Genotype Call Genome-Wide Association and Tables 3-5). However, we did observe evidence of replication of several terminal exon SNPs within the *GFOD1* gene in the CHOP families, using TDT (*P*-value range = $8x10^{-4}$ - $1x10^{-2}$, for rs1866863, rs9370020, rs2254292, and rs2439565). We additionally report observed significance for other SNPs reported previously[9,11] with converging evidence in Supplementary Table 6.

Segment-based comparative analysis of CNVs: To identify novel genomic loci harboring CNVs potentially contributing to ADHD, we applied a segment-based scoring approach that scans the genome for consecutive SNPs with more frequent copy number changes in cases compared to controls as we have previously described [1,10] . The genomic span for these consecutive SNPs delineates common copy number variation regions, or CNVRs. In the CHOP cohort, we identified 10 CNVRs that were observed in multiple cases but not in controls, as well as 2 CNVRs that had higher frequency in cases compared to controls (Table 1). To ensure reliability of our CNV detection method, we experimentally validated all CNVRs using quantitative PCR (qPCR), a method commonly used for independent validation of CNVs (Supplementary Figure 3). Thus, we have applied a separate validation technique on all the CNVs reported to ensure positive confirmation. Using this approach, we have identified and experimentally validated a total of 12 CNV loci that were either observed in ADHD cases only or overrepresented in the ADHD cases that we subsequently took forward for replication in independent study cohorts.

Replication analysis was performed in five independent cohorts, including ADHD subjects from IMAGE, PUWMa, IMAGE II, NIMH, and Utah. Based on the 10 case-specific CNVs from the discovery cohort, 3 were also exclusive to replication cohort cases, notably *GRM7, GRM8 and NEGR1*, with resulting combined *P*-values of 3.52×10^{-6} and 8.14×10^{-5}, for *GRM8* and *GRM7*, respectively (Table 1A). A third GRM gene, *GRM5*, was observed in 9 ADHD cases (9/3,506) and one control case (1/13,327) with resulting $P=1.36 \times 10^{-6}$ (Table 1A). GRM1 was observed in 8 cases and 2 controls $P=1.05 \times 10^{-4}$. Thus, these 4 GRM genes were impacted by deletions that associated with ADHD and replicated successfully in the independent ADHD cohorts (Table 1A and Table 2), whereas the other CNV loci were also observed to be enriched in the ADHD cases, albeit at nominally significant P values (Table 1B and Supplementary Table 2). Figure 1 shows the CNV deletions observed at the *GRM5* locus (9 cases vs 1 control), using UCSC Genome Browser [12] with Build 36 of the human genome. Experimental validation of IMAGE CNVs, using qPCR, together with Raw BAF and LRR plots are shown in Supplementary Figures 4-6.

Taken together, we have uncovered four genes directly impacted by CNVRs in multiple independent cohorts that belong to the metabotropic glutamate receptor gene family (InterPro category "GPCR, family 3, metabotropic glutamate receptor"; $P= 2.1 \times 10^{-9}$). Examining all *GRM* family genes, we also detected single CHOP cases with deletion of *GRM2* and *GRM6* not observed in controls that we also detected in single *GRM2* and *GRM6* IMAGE II cases, respectively. We additionally evaluated the significance of the GRM genes, using SNP genotype TDT in the same cohort and the best support was observed for *GRM7,* $P=8.35 \times 10^{-5}$ (Supplementary Table 7). Analyis was also performed to addres familybased CNV statistics

based on transmission disequilibrium and *de novo* events in the family-based subset of 311 CHOP families and 422 IMAGE families (Supplementary Tables 8 and 9).

Discussion

At present, there is a notable paucity of genome wide association studies in ADHD, and no study has reported CNVs that are significantly associated with ADHD. As such, our study represents the first large-scale, unbiased two-stage genome-wide scanning of CNVs in ADHD. The genes from the metabotropic glutamate receptor family (*GRM5*, *GRM7, GRM8* and *GRM1*) are for the first time shown to be impacted by CNVs that associatewith ADHD and observed to replicate in multiple independent case control data sets.

Metabotropic glutamate receptors (GRMs or mGluRs) are a class of G-protein-coupled receptors that possess a seven transmembrane region involved in the modulation of excitatory synaptic transmission in the nervous system [13]. There are three receptor groups based on sequence homology, putative signal transduction mechanisms, and pharmacologic properties [14]. *GRM5* and *GRM1* are members of Group I expressed particularly in the basal ganglia and cerebellum [15], relevant brain areas for ADHD. These receptors have been shown to activate phospholipase C and it has been postulated they may play a role in addiction, anxiety and behavioral disorders [16]. *GRM7* and *GRM8* are members of Group III which is linked to the inhibition of the cyclic AMP cascade. *GRM7* has been linked with anxiety [17] and is the most highly conserved of all mGluR subtypes across different mammalian species [18].

Evidence for glutamatergic involvement in ADHD is arising from diverse fields. While association studies investigating variants in glutamatergic receptors and transporters have reported mixed results [19-22] a genome-wide association study investigating response to the methylphenidate in ADHD children detected nominal evidence for association of several SNPs including SNPs within *GRM7* (rs3792452) [23]. *GRIN2A* was reportedly associated with ADHD in a genetic linkage study[19] and *GRIN2B* was associated by TDT [24]. Magnetic resonance spectroscopy studies have shown increased glutamatergic tone in frontal and striatal brain regions of ADHD subjects[25-27] which normalizes with stimulants and atomoxetine [28]. The *SLC6A3-KO* (DAT-KO) mouse, an ADHD animal model, remains responsive to methylphenidate in spite of the lack of a dopamine transporter[29] and hyperactivity in these mice can be increased by NMDA-receptor blockers and suppressed by drugs that increase glutamatergic transmission[30]. Increased midbrain *SLC6A3* and *DRD4* expression were reported in rats where glutamate transporter increases were found in the striatum [31] suggesting that decreases in dopamine may alter glutamate signaling. Also, glutamate receptor subunit gene (*GRIN2A*) disruption increased DA and serotonin metabolism in the frontal cortex and striatum of mice, and increased locomotor activity that was reduced by dopamine or serotonin receptor antagonists [32]. Moreover, dysregulated expression of genes in glutametergic pathways has been observed in the SHR [33-36] and in the PCB exposed rat model of ADHD [35].

Apart from the GRM family of genes, we have detected association of 8 other loci with ADHD, five which directly impact genes (Table 1B). Among those are genes with intriguing biology with respect to ADHD. *DPP6* has been previously associated with Amyotrophic Lateral Sclerosis (ALS) in a genome wide association studies [37,38], and CNVs impacting *DPP6* have

been reported in relation with autism[39]. *DPP6* and *CTNNA2* (although our association does not directly impact *CTNNA2*) have been implicated by earlier ADHD SNP genotype GWAS [9]. *NLN* is an interesting candidate responsible for metabolic inactivation of neural peptides, such as Neuropeptide Y (NPY) which has previously been implicated in ADHD [44,45]. *SLC7A10* has been shown to play a role in the modulation of glutamatergic transmission through mobilization of D-serine at the glutamatergic synapse. *LARP7* is important for snRNP integrity, a protein complex responsible for post transcriptional splicing. *NEGR1* encodes a neural cell adhesion molecule and a trans-neural growth-promoting factor in regenerative axon sprouting and neuronal growth in the mammalian brain. Interestingly, this neuronal gene was recently associated with obesity [40].

In the CHOP discovery cohort, Family 230 is impacted with both *GRM5* deletion inherited from the mother and *NEGR1* duplication inherited from the father in all three ADHD cases in the family. In spite of superior IQ levels these 3 children had severe impairment. These were the only CNV regions observed in all three familial cases and not observed in controls. Assessment of the mother using an adult ADHD Self-Report Scale[41] indicated a likelihood of ADHD.

There are seven CNVRs presented that directly disrupt the respective gene in these regions (including *GRM5*, *GRM7*, GRM8, *GRM1, DPP6, NEGR1* and *LARP7*) while the remainder are annotated with the closest gene (Table 1A and B). Further functional studies will be needed to fully characterize the function of the associated genes in relation with the ADHD phenotypes. Thus, our unbiased approach to assess the entire genome in multiple independent cohorts has revealed CNVs in novel genes that have not previously been studied for any potential biological or physiological impact on the brain in ADHD and await further characterization.

In conclusion, using a two-stage genome-wide association approach for high-resolution CNV detection, we have identified 12 loci demonstrating enrichment of CNVs in ADHD cases as compared to controls, and successfully replicated 4 of them using independent data sets of ADHD cases and healthy controls genotyped on three different platforms matched for cases and controls. Four of the genes affected belong to the metabotropic glutamate receptor family. The enrichment of genes within this molecular system suggests novel susceptibility mechanisms for ADHD, and will spur assessment of additional variations, including structural variations and single-base changes in candidate genes within these molecular networks. Our results call for functional expression assays to assess the biological effects of CNVs in these candidate genes.

Methods

Illumina Infinium assay for CNV Discovery

We performed high-throughput, genome-wide SNP genotyping, using the InfiniumII HumanHap550 BeadChip technology (Illumina San Diego CA), at the Center for Applied Genomics at CHOP. The genotype data content together with the intensity data provided by the genotyping array provides high confidence for CNV calls. Importantly, the simultaneous analysis of intensity data and genotype data in the same experimental setting establishes a highly accurate definition for normal diploid states and any deviation thereof. To call CNVs, we used the PennCNV algorithm, which combines multiple sources of information, including Log R Ratio (LRR) and B Allele Frequency (BAF) at each SNP marker, along with SNP spacing and population frequency of the B allele to generate CNV calls. The replication case and control

cohorts utilized genome-wide SNP genotyping using the Perlegen 600K, Illumina 1M, and Affymetrix 5.0 arrays. Raw X and Y values were normalized with log(10) and clustered to establish BAF and LRR with PennCNV-Affy protocol (Supplementary Methods and Table 10). Rare recurrent CNVs were the focus of our study.

CNV quality control

We calculated Quality Control (QC) measures on our HumanHap550 GWAS data based on statistical distributions to exclude poor quality DNA samples and false positive CNVs. The first threshold is the percentage of attempted SNPs which were successfully genotyped. Only samples with call rate > 98% were included. The genome wide intensity signal must have as little noise as possible. Only samples with the standard deviation (SD) of normalized intensity (LRR) < 0.35 were included. All samples must have Caucasian ethnicity based on principle components analysis (Supplementary Figure 7) and all other samples were excluded. Furthermore, case and control matching was insured by calculating a genomic inflation factor (GIF=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 accurate inference of copy number variations [42]. Only samples where the wave factor of LRR to wave model ranged between -0.5<x<0.6 were accepted. If the count of CNV calls made by PennCNV exceeds 70 (Supplementary Fig 1), the DNA quality is usually poor. Thus, only samples with CNV call count < 70 were included. Any duplicate samples (such as monozygotic twins) had one sample excluded. Supplementary Table 11 provides the number of samples excluded for each quality control measure.

Statistical analysis of CNVs

CNV frequency between cases and controls was evaluated at each SNP using Fisher's exact test. We only considered loci that were nominally significant between cases and controls ($p<0.05$) where cases in the CHOP discovery cohort had the same variation, replicated in IMAGE, PUWMa, or IMAGE II or were not observed in any of the control subjects, and validated with an independent method. We report 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 Figure 8). Resulting nominally significant CNVRs were excluded if they met any of the following criteria: i) residing on telomere or centromere proximal cytobands; ii) arising in a "peninsula" of common CNV arising from variation in boundary truncation of CNV calling (Supplementary Figure 3); iii) genomic regions with extremes in GC content which produces hybridization bias; or iv) samples contributing to multiple CNVRs. We statistically adjusted for relatedness of cases with permutation (1000x). Three lines of evidence establish statistical significance: independent replication $p<0.05$, permutation of observations, and no loci observed with control enriched significance. We used DAVID (Database for Annotation, Visualization, and Integrated Discovery) [43] to assess the significance of functional annotation clustering of independently associated CNV results into InterPro categories.

References

1. Glessner JT, Wang K, Cai G, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature 2009;459:569-573.
2. Derks EM, Hudziak JJ, Dolan CV, van Beijsterveldt TC, Verhulst FC, Boomsma DI. Genetic and environmental influences on the relation between attention problems and attention deficit hyperactivity disorder. *Behav Genet* **38**(1), 11-23 (2008).
3. Wood AC, Rijsdijk F, Saudino KJ, Asherson P, Kuntsi J. High heritability for a composite index of children's activity level measures. *Behav Genet* **38**(3), 266-276 (2008).
4. Haberstick BC, Timberlake D, Hopfer CJ, Lessem JM, Ehringer MA, Hewitt JK. Genetic and environmental contributions to retrospectively reported DSM-IV childhood attention deficit hyperactivity disorder. *Psychol Med* **38**,1057-1066 (2008).
5.Wang K, Zhang H, Ma D, et al., Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459**, 528-533 (2009).
6. Franke B, Neale BM, Faraone SV, Genome-wide association studies in ADHD. *Hum Genet*. doi: 10.1007/s00439-009-0663-4 (2009).
7. Neale BM, Lasky-Su J, Anney R, *et al*. Genome-wide association scan of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 1337-1344. (2008).
8. Lasky-Su J, Neale BM, Franke B *et al*., Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **147B**(8), 1345-1354 (2008).
9. Lesch KP, Timmesfeld N, Renner TJ, 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).
10. Wang K, Li M, Hadley D, *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).
11. Zhou K, Dempfle A, Arcos-Burgos M *et al*. Meta-analysis of genome-wide linkage scans of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet*. **147B**(8), 1392-8 (2008).
12. Kent WJ, Sugnet CW, Furey TS, *et al*. The human genome browser at UCSC. *Genome Res.* **12**(6), 996-1006 (2002).
13. Taniura, H.., Sanada, N., Kuramoto, N., Yoneda, Y. Metabotropic Glutamate Receptor Family Gene in Dictyostelium discoideum. *J. Biol. Chem.*, **281**(18), 12336-12343, (2006).
14. Conn, P. J. & Pin, J.. Phamacology and Functions of Metabotropic Glutamate Receptors. *Annu. Rev. Pharmacol. Toxicol*. **37**, 205–37 (1997).
15. Berthele A, Platzer S, Laurie DJ, et al. Expression of metabotropic glutamate receptor subtype mRNA (mGluR1-8) in human cerebellum. *NeuroReport* **10**(18), 3861-3867 (1999).
16. Koob, G. F., Sanna, P.P., Bloom, F. E., Neuroscience of Addiction. *Neuron*, **21**, 467–476, (1998).
17. Cryan JF, Kelly PH, Neijt HC, Sansig G, Flor PJ, van Der Putten H. Antidepressant and anxiolytic-like effects in mice lacking the group III metabotropic glutamate receptor mGluR7. *European Journal of Neuroscience*. **17**(11):2409-2417, (2003).

18. Makoff, A., Pillinga, C., Harrington, K., Emson, P. Human metabotropic glutamate receptor type 7: Molecular cloning and mRNA distribution in the CNS. *Molecular Brain Research*. **40**(1), 165-170 (1996).
19. Turic D, Langley K, Mills S, 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).
20. Mick, E. and Faraone, S.V., Genetics of attention deficit hyperactivity disorder. *Child Adolesc Psychiatr Clin N Am* **17**, 261-284, vii-viii. (2008).
21. Turic D, Langley K, Williams H, et al. A family based study implicates solute carrier family 1-member 3 (SLC1A3) gene in attention-deficit/hyperactivity disorder. Biol Psychiatry. 2005 Jun 1;57(11):1461-6.
22. Elia J, Capasso M, Zaheer Z, 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*. PMID: 19352218 (2009).
23. Mick E, Neale B, Middleton FA, McGough JJ, Faraone SV. 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).
24. Dorval KM, Wigg KG, Crosbie J, *et al*., Association of the glutamate receptor subunit gene GRIN2B with attention-deficit/hyperactivity disorder. *Genes Brain Behav.* **6**(5) (2007).
25. Jin Z, Zang YF, Zeng YW, Zhang L, Wang YF. Striatal neuronal loss or dysfunction and choline rise in children with attention-deficit hyperactivity disorder: a 1H-magnetic resonance spectroscopy study. *Neurosci Lett* **315**, 45-48 (2001).
26. MacMaster FP, Carrey N,. Sparkes S, Kusumakar V. Proton spectroscopy in medication-free pediatric attention-deficit/hyperactivity disorder. *Biol Psychiatry* **53**:184-187. (2003).
27. Courvoisie H, Hooper SR, 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).
28. Carrey N, MacMaster FP, Fogel J, et al. Metabolite changes resulting from treatment in children with ADHD: a 1H-MRS study. *Clin Neuropharmacol* **26**,218-221 (2003).
29. 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* **283**, 397-401 (1999).
30. 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 U S A* **98**,11047-11054 (2001).
31. 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).
32. Miyamoto K, Nakanishi H, Moriguchi S, 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).
33. 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).
34. Russell VA. 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).
35. DasBanerjee T, Middleton FA, Berger DF, Lombardo JP, Sagvolden T, Faraone SV. A comparison of molecular alterations in environmental and genetic rat models of ADHD: a pilot study. Am J Med Genet B Neuropsychiatr Genet 147B, 1554-63 (2008).
36. Sagvolden T, Johansen EB, Wøien G, et al. The spontaneously hypertensive rat model of ADHD--the importance of selecting the appropriate reference strain. *Neuropharmacology* **57**, 619-26 (2009).
37. Del Bo R, Ghezzi S, Corti S, *et. al*., DPP6 gene variability confers increased risk of developing sporadic amyotrophic lateral sclerosis in Italian patients. *Journal of Neurology, Neurosurgery & Psychiatry*. **79**(9), 1085 (2008).
38. Cronin S, Tomik B, Bradley DG, Slowik A, Hardiman O. Screening for replication of genome-wide SNP associations in sporadic ALS. *European Journal of Human Genetics* **17**, 213–218 (2009).
39. Marshall CR, Noor A, Vincent JB, *et al*. Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 477-88 (2008).
40. Renström F, Payne F, Nordström A, et al. Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden *Human Molecular Genetics* **18**(8):1489-1496 (2009).
41. Kessler RC, Adler LA, Barkley R, *et al*., Patterns and predictors of attention-deficit/hyperactivity disorder persistence into adulthood: results from the national comorbidity survey replication. *Biological Psychiatry* **57**(11)**,** 1442-1451 (2005).
42. Diskin S, Li M, Hou C, *et al*., Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Research*. **36**(19) (2008).
43. Dennis G Jr, Sherman BT, Hosack DA, *et al*., DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*. **4**(9), (2003).
44. Lesch KP, Selch S, Renner TJ, *et al.* Genome-wide copy number variation analysis in ADHD: association with neuropeptide Y gene dosage in an extended pedigree. *Mol Psychiatry*, Epub ahead of print (2010).
45. Oades RD, Daniels R, Rascher W. Plasma neuropeptide Y levels, monoamine metabolism, electrolyte excretion, and drinking behavior in children with attention-deficit hyperactivity-disorder (ADHD). *Psychiat. Res*., **80**, 77-186. (1998).

Acknowledgements

We gratefully thank all the ADHD children and their families who participated in this study and all the control subjects who donated blood samples to Children's Hospital of Philadelphia (CHOP) for genetic research purposes. We thank the technical staff in the Center for Applied Genomics, Children's Hospital of Philadelphia for generating the genotypes used in this study and the medical assistants, 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. Hermannsson and A. Krisbjörnsson for their software design and contribution.

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 NIH grant R01MH62873 to S. V. Faraone and the genotyping of samples was provided through the Genetic Association Information Network (GAIN). The dataset used for the analyses of IMAGE described in this manuscript were obtained from the database of Genotype and Phenotype (dbGaP) found at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number ADHD IMAGE: phs000016.v2.p2, Depression: phs000020.v2.p1, and Psoriasis: phs000019.v1.p1. Samples and associated phenotype data were accessed via an authorized data access request by Hakon Hakonarson and Struan F. A. Grant. Data collection for the PUWMa sample was supported by NIH grants to S.V. Faraone, J.Biederman, S. Smalley, and R.Todd/A.Todorov. GWAS genotyping of PUWMa samples was completed by Genizon and provided through a grant for genotyping services to S.V. Faraone. PUWMa consortium represents a Pfizer funded collaboration among UCLA, Washington University and Massachussetts General Hospital. We thank Michael C. O'Donovan, Michael Gill, Michael J. Owen, Peter A. Holmans, Anita Thapar, Benjamin M. Neale, and Ana Miranda for contributing DNA samples and phenotypes to the study and editing the manuscript. We thank Gail DePalma, Theresia Töpner, Allan Guiney, and Hua Zhang for providing samples for qPCR CNV validation.

The study is supported by an Institutional Development Award to the Center for Applied Genomics from the Children's Hospital of Philadelphia (HH), which funded all CHOP genome-wide genotyping for this study. This work was additionally supported in part by National Institutes of Health grant K23MH066275 (JE), University of Pennsylvania CTR grant UL1-RR-024134 (JE) and by a Research Development Award from the Cotswold Foundation (HH & SFAG). NT, T Sakurai, and JDB 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. Lesch and A. Reif, ME 1923/5-1, ME

1923/5-3 to C. Freitag and J. Meyer) and the Bundesministerium für Bildung und Forschung (BMBF 01GV0605 to K.P. Lesch).

Author Contributions

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

Author Information

The authors other than M.R. declare no competing financial interests. During the last 3 years Marcel Romanos has been in the speakers' bureau for Janssen-Cilag. In previous years he has been on the speakers' bureau for MEDICE.

Correspondence and requests for materials should be addressed to H.H. (hakonarson@email.chop.edu).

Table 1. Novel CNVRs Over-represented in ADHD Patients

A) Loci Significantly Associated with ADHD

CNVR	CHOP Cases n=1013	IMAGE Cases n=624	PUWMa Cases n=864	IMAGE II Cases n=787	NIMH/ Utah Cases n=218	CHOP Controls n=4105	Additional Controls n=9222	Combined P-value	Type	Gene
chr11:88269449-88351661	4	5	1	0	0	0	1	$1.36x10^{-6}$	Del	*GRM5*
chr7:126525124-126536202	3	3	2	0	0	0	0	$3.52x10^{-6}$	Del	*GRM8*
chr3:7183953-7197236	4	2	0	0	0	0	0	$8.14x10^{-5}$	Del	*GRM7*
chr6:146657076-146694047	5	0	0	1	2	2	0	$1.05x10^{-4}$	Dup	*GRM1*

B) ADHD Loci with Nominal Significance

CNVR	CHOP Cases n=1013	IMAGE Cases n=624	PUWMa Cases n=864	IMAGE II Cases n=787	NIMH/ Utah Cases n=218	CHOP Controls n=4105	Additional Controls n=9222	Combined P-value	Type	Gene
chr1:72317292-72328395	4	0	1	0	0	0	0	$3.91x10^{-4}$	Dup	*NEGR1*
chr7:153495598-153564827	5	0	2	0	1	0	2	$4.08x10^{-4}$	Dup	*DPP6*
chr5:65027976-65046520	4	1	0	1	0	0	1	$4.68x10^{-4}$	Del	*SGTB/ NLN*
chr1:56053497-56064495	2	3	0	1	0	0	2	$1.54x10^{-3}$	Del	*USP24**
chr19:38427720-38444834	5	1	1	0	0	2	3	$4.95x10^{-3}$	Del	*SLC7A10**
chr3:1844168-1859889	4	0	2	1	0	0	6	$8.81x10^{-3}$	Del	*CNTN4**
chr2:81419297-81446082	2	1	0	1	0	0	3	$3.83x10^{-2}$	Dup	*CTNNA2**
chr4:113772340-113788584	2	1	1	0	0	0	3	$3.83x10^{-2}$	Dup	*LARP7*

Rare variants that were recurrent and observed to be enriched among ADHD cases relative to control frequencies and detected in multiple independent cohorts are reported. All GRM genes are directly impacted by the CNVR. Regions listed represent the optimal overlap of cases and significance with respect to controls as described in the Methods and Supplementary Fig. 8. The closest gene is listed for each CNVR locus since it is most likely to be impacted. For detailed counts from each contributing project see Supplementary Table 12. *No gene directly impacted so closest proximal gene listed.

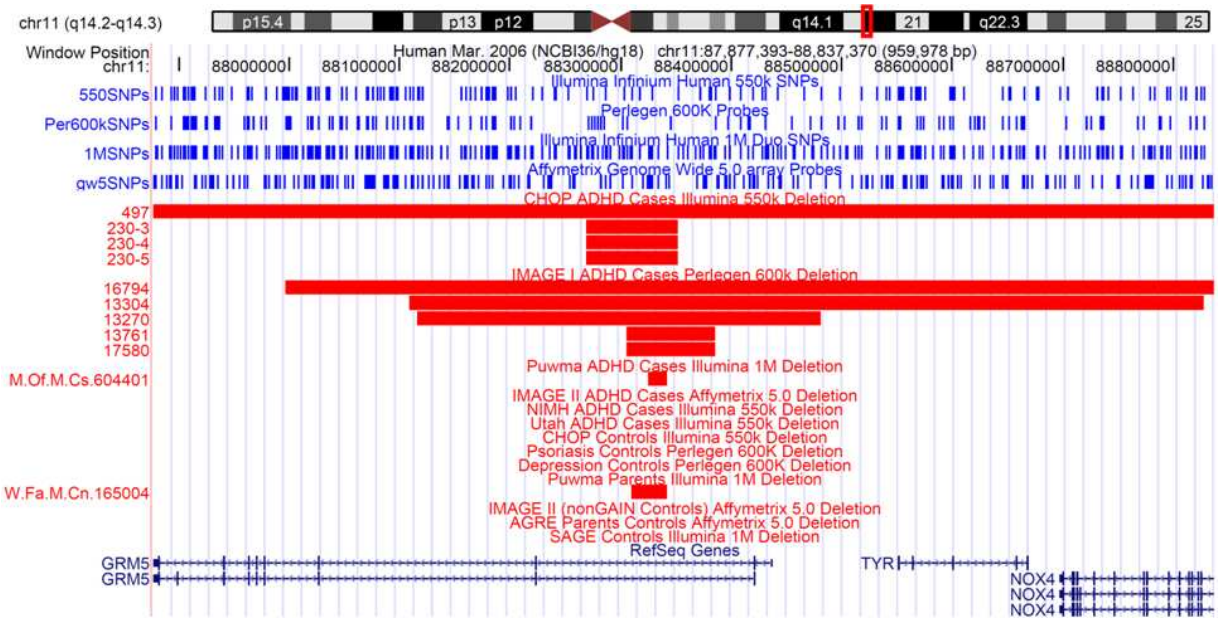
Table 2. Discovery, Replication, and Combined Significance of CNV Regions.

CNVR	Discovery P-value	Replication P-value	Combined P-value	Permuted Discovery P-value	Permuted Replication P-value	Permuted Combined P-value	Type	Gene
chr11:88269449-88351661	**1.53×10^{-3}**	**5.29×10^{-4}**	**1.36×10^{-6}**	**0.025**	**0.001**	**0.002**	**Del**	***GRM5***
chr7:126441593-126621501	**7.74×10^{-3}**	**4.35×10^{-4}**	**3.52×10^{-6}**	**0.013**	**<0.001**	**<0.001**	**Del**	***GRM8***
chr3:7183953-7197236	**1.53×10^{-3}**	**4.53×10^{-2}**	**8.14×10^{-5}**	**0.011**	**0.039**	**<0.001**	**Del**	***GRM7***
chr6:146657076-146694047	**4.42×10^{-3}**	**9.63×10^{-3}**	**1.05×10^{-4}**	**0.006**	**<0.001**	**<0.001**	**Dup**	***GRM1***
chr1:72317292-72328395	1.53×10^{-3}	2.13×10^{-1}	3.91×10^{-4}	0.036	0.213	0.011	Dup	*NEGR1*
chr7:153495598-153564827	1.53×10^{-3}	6.82×10^{-2}	4.08×10^{-4}	<0.001	0.058	<0.001	Dup	*DPP6*
chr5:65027976-65046520	1.53×10^{-3}	1.17×10^{-1}	4.68×10^{-4}	0.003	0.108	0.001	Del	*SGTB/ NLN*
chr1:56053497-56064495	3.91×10^{-2}	2.12×10^{-2}	1.54×10^{-3}	0.035	0.024	<0.001	Del	*USP24*
chr19:38427720-38444834	4.42×10^{-3}	2.89×10^{-1}	4.95×10^{-3}	0.002	0.262	0.007	Del	*SLC7A10*
chr3:1844168-1859889	1.53×10^{-3}	4.12×10^{-1}	8.81×10^{-3}	0.008	0.416	0.015	Del	*CNTN4*
chr2:81419297-81446082	3.91×10^{-2}	2.89×10^{-1}	3.83×10^{-2}	0.046	0.294	0.032	Dup	*CTNNA2*
chr4:113772340-113788584	3.91×10^{-2}	2.89×10^{-1}	3.83×10^{-2}	0.033	0.288	0.042	Dup	*LARP7*

The top 4 most significant loci are shown in bold.

Figure 1. Deletion Directly Impacting *GRM5* Exclusive to ADHD Cases and Replicating in IMAGE and PUWMa

Three CHOP ADHD case hemizygous deletions in *GRM5* replicated by 2 deletions and 3 larger deletions found in IMAGE and 1 PUWMa deletion. SNP coverage of the Illumina 550k, Perlegen 600k, Illumina 1M, and Affymetrix 5.0 arrays are shown as vertical blue lines.



Supplementary Information

CNV Calls and Review of Significant Loci

No additional "CNV burden" was observed in cases vs. controls, rather the distribution of calls made was highly comparable (Supplementary Figure 1). We established CNV call reliability in Illumina and Perlegen data by observing Mendelian patterns of inheritance. Trios were first verified by genotype inheritance and analyzed to establish the quality of CNV calls from both Illumina and Perlegen platforms based on observed inheritance. Based on all CNV calls called in trios from the Illumina CHOP data, 8,647 CNVs observed in offspring were inherited from a parent while 437 CNVs were putatively *de novo* which is a *de novo* rate of 4.811%. Based on all CNV calls called in trios from the Perlegen IMAGE data, 1,862 CNVs observed in offspring were inherited from a parent while 505 CNVs were putatively *de novo* which is a *de novo* rate of 21.335%. 51 IMAGE cases, 22 deletion loci, and 5 duplication loci had multiple *de novo* events due to low data quality and were excluded as outliers; once excluded, 785 CNVs were inherited and 63 were denovo which lowered the observed denovo rate to an acceptable level of 7.429%. Based on CNVs observed in parents from Illumina CHOP data, 9,305 CNVs were passed to the child while 7,432 CNVs were not inherited resulting in a 55.595% inheritance rate. Based on all CNVs observed in parents from Perlegen IMAGE data, 2,114 CNVs were passed to the child while 3,789 CNVs were not inherited resulting in a 35.812% inheritance rate. We excluded 65 parent samples that were outliers with 20 or greater CNVs not inherited to offspring and filtering these samples out resulted in 1,204 CNVs were passed to the child while 1,221 were not inherited resulting in a 49.650% inheritance rate which established confidence in this CNV call set.

It is intractable to review all PennCNV calls and wasteful to exclude CNVs smaller than a size threshold. Instead, we statistically score the loci based on all CNVs detected and review these nominally associated CNVR loci for appropriate overlap, signal quality, and Mendelian inheritance. As in Table 1, all reported loci show at least one case with the CNV inherited from a parent, in cases where both parents were available.

In total, there are 3,506 cases and 13,327 controls, representing greater than a three-fold abundance of control samples to robustly define CNVs to be absent or at a lower frequency than case samples. Although the number of CNVs detected per sample was as high as 70, there are actually inferred normal diploid (CN=2) calls which make every sample equivalent. These CNVs are very rare and thus the number of observed CNV calls will vary between samples.

CNV validation by quantitative PCR (QPCR)

Universal Probe Library (UPL; Roche, Indianapolis, IN) probes were selected using the ProbeFinder v2.41 software (Roche, Indianapolis, IN). Quantitative PCR was performed on an ABI 7500 Real Time PCR Instrument or on an ABI Prism™ 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA). Each sample was analyzed in quadruplicate either in 25 ul reaction mixture (250 nM probe, 900 nM each primer, Fast Start TaqMan Probe Master from Roche, and 10 ng genomic DNA) or in 10 ul reaction mixture (100 nM probe, 200 nM each primer, 1x Platinum Quantitative PCR SuperMix-Uracil-DNA-Glycosylase (UDG) with ROX from Invitrogen, and 25 ng genomic DNA).

The values were evaluated using Sequence Detection Software v2.2.1 (Applied Biosystems, CA). Data analysis was further 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 then data was normalized by setting a normal control to a value of 1.

PennCNV-Affy Workflow Adapted to Perlegen 600K Data

The CNV calling on Perlegen platform used a highly similar algorithm to those used on the Illumina arrays, but the signal pre-processing steps differ. Unlike the Illumina platform, where normalized signal intensities (Log R Ratio and B Allele Frequency) can be exported directly from the BeadStudio software, these signal intensity measures in the Perlegen 600K platform need to be calculated from the collection of genotyped samples based on raw X and Y values. To perform data normalization and signal extraction from raw final report files generated in genotyping experiments, we first reformatted data from dbGaP into the format produced by Affymetrix Power Tools: birdseed.calls.txt, birdseed.confidences.txt, and quant-norm.pm-only.med-polish.expr.summary.txt (Supplementary Table 10). The X and Y values provided in the sample based report files from dbGaP were reduced to a more finite range by taking the logarithm base 10. For each SNP marker, we then relied on the allele-specific signal intensity for the AA, AB and BB genotypes on all genotyped samples to construct three canonical genotype clusters in polar coordinates theta and R, similar to the Illumina clustering generation approach. The "-conf 2" option was included in running generate_affy_geno_cluster.pl since 1 was coded as the best score. Once the canonical genotype clusters were

constructed, we then transformed the signal intensity values for each SNP to Log R Ratio (LRR) and B Allele Frequency (BAF) values using normalize_affy_geno_cluster.pl. For more technical details, see http://www.openbioinformatics.org/penncnv/penncnv_tutorial_affy_gw6.html.

To optimize the Hidden Markov Model (HMM), we used the baseline reference file hh550.hmm and ran "-train" in PennCNV in three successive batches of thirty. The first training used the samples with the lowest standard deviation of LRR while the other two runs, using the file created as a new reference, included more random representative samples. We also created definition files providing inter-SNP distance and population b-allele frequency to further inform CNV calling specifically adapted to the observed 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 standard deviation of LRR was below 0.2 for the majority (84%) of samples, the intensity data was notably noisier in regions of called CNV and often showed a subpopulation of SNPs unable to differentiate a deletion signal, perhaps due to PCR saturation during the lab processing. Nevertheless, the deletion and duplication features were still detected with confirmation of homozygote and AAB/ABB genotypes respectively shown for the same SNPs (Supplementary Figure 5 and 6).

Lastly, Perlegen CNV calls were screened for overlap with the 11 loci associated based on the CHOP Illumina data. The SNP level data underlying each CNV call was reviewed to ensure clean signal quality. To ensure that each detected CNV was a true DNA feature and not in any way an artifact of the Perlegen 600K array used or our

bioinformatics manipulations of the data, we validated each CNV with qPCR at an independent lab (Supplementary Figure 4).

Permutation to Adjust Significance for Relatedness

For initial Fisher's exact test, related individuals are not controlled for since our primary objective is to detect CNVs in multiple samples regardless of relatedness. CNVRs passing this initial screen are scored for statistical significance based on a permuted P-value which permutes case and control labels randomly of all samples with the condition that related individuals must have the same label. Each unrelated individual is assigned a case or control label and their related sibling is assigned the 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 is assigned. The number of hypothetical scenarios with significance equal or greater (lower P-value) provides the permuted P-value which corrects for relatedness. The Fisher's exact test P-value and counts of cases and controls with each CNVR are provided for transparency.

Analysis of Genotype Call Genome-Wide Association

Full scale genotype genome-wide association was performed and the genomic inflation factor (GIF) was at an acceptable level (GIF=1.02409). We also checked pairwise population concordance to check for and filter out cryptic relatedness which could give rise to rare CNVs specific to ultra-stratified subpopulations of Europe. We performed Transmission Disequilibrium Test (TDT) statistic using Plink on 397 ADHD cases with both parents on the CHOP Illumina HumanHap550 genotype data (Supplementary Table

3). The top result with more than one significant SNP in a region was chr4p12 P(rs1018199)=2.71x10^{-5} and P(rs11724347)=6.19x10^{-5} which impacts *TEC*. We also performed a case:control genotype genome-wide association on 735 cases and 2,298 controls using the same Illumina data set (Supplementary Table 4). The strongest signal was chr19p12 P(rs2081051)=4.60x10^{-6} and P(rs399686)=4.72x10^{-6} residing between *ZNF66* and *ZNF85*. Lastly, 623 ADHD cases with both parents on the IMAGE Perlegen 600K data were analyzed with TDT statistic (Supplementary Table 5). The most significant signal was chr5q23.1 P(rs17144308)= 9.70x10^{-6} and P(rs2043053)=3.36x10^{-5} which is 237 kb from the closest proximal gene *DTWD2*. Taken together, SNPs residing around exon 4 of contactin 3 (*CNTN3*) appear to replicate most consistently between Illumina and Perlegen ADHD TDT statistics. SNP rs12488030 is common to both platforms P=2.51x10^{-3} Illumina and P=4.97x10^{-3} Perlegen. There are two supporting SNPs in close proximity also showing significance Illumina: P(rs4073942)= 2.78x10^{-3} and P(rs9869828)=8.61x10^{-3} in addition Perlegen: P(rs11915713) =1.86x10^{-5} and P(rs7372975) =7.59x10^{-5}

Study Criteria for inclusion in IMAGE

Proband diagnosis: combined subtype ADHD.

Children aged 6-17 years (inclusive).

One or more sibling(s) in the same age range.

Both parents available to provide DNA sample or one parent available plus two or more siblings.

IQ above 70.

Free of single-gene disorders known to be associated with ADHD (e.g. fragile-X, phenylketonuria, hypercalcaemia, thyroid hormone resistance).
Free of neurological disease and damage (e.g. hemiplegia and other cerebral palsies, epilepsy, hydrocephalus, post- encephalitic syndromes, psychosis, sensorimotor handicaps).
Living at home with at least one biological parent and one full sibling.
Not meeting criteria for autism or Asperger's syndrome.

Study Criteria for inclusion in IMAGE II

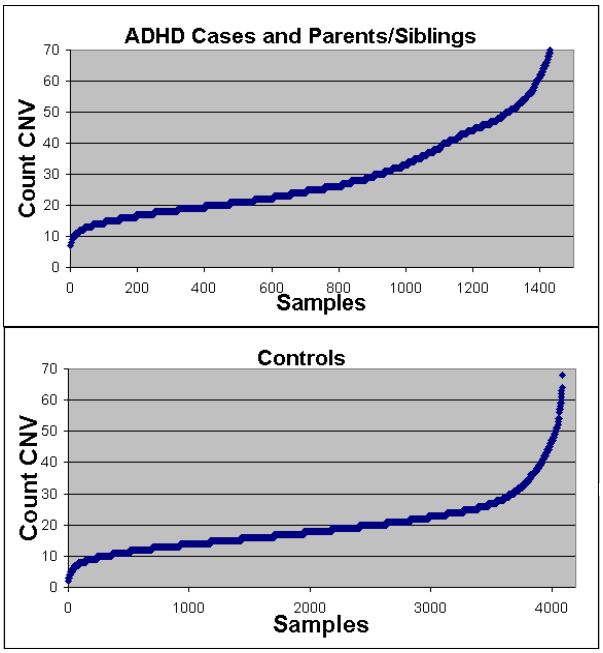
Proband diagnosis: ADHD according to DSM-IV-TR
Semi-structured diagnostic interview: KSADS-PL or Kinder -DIPS
Child Behavior Checklist, Conners parent and teacher Scales or German Teachers Report on ADHD symptoms according to DSM-IV
Children aged 6-18 years (index patients older than 8 years).
IQ above 70; birth weight > 2000 g; no major medical events during pregnancy; no drug abuse in mother during pregnancy
Free of single-gene disorders known to be associated with ADHD (e.g. fragile-X, phenylketonuria, hypercalcaemia, thyroid hormone resistance).
Free of neurological disease and damage (e.g. hemiplegia and other cerebral palsies, epilepsy, hydrocephalus, post- encephalitic syndromes, motor neuron disorder etc.).
Not meeting criteria for autism or Asperger's syndrome, schizophrenia, bipolar disorder, primary major depressive episode, and anxiety disorder, Tourette’s Syndrome, .

Controls for IMAGE II

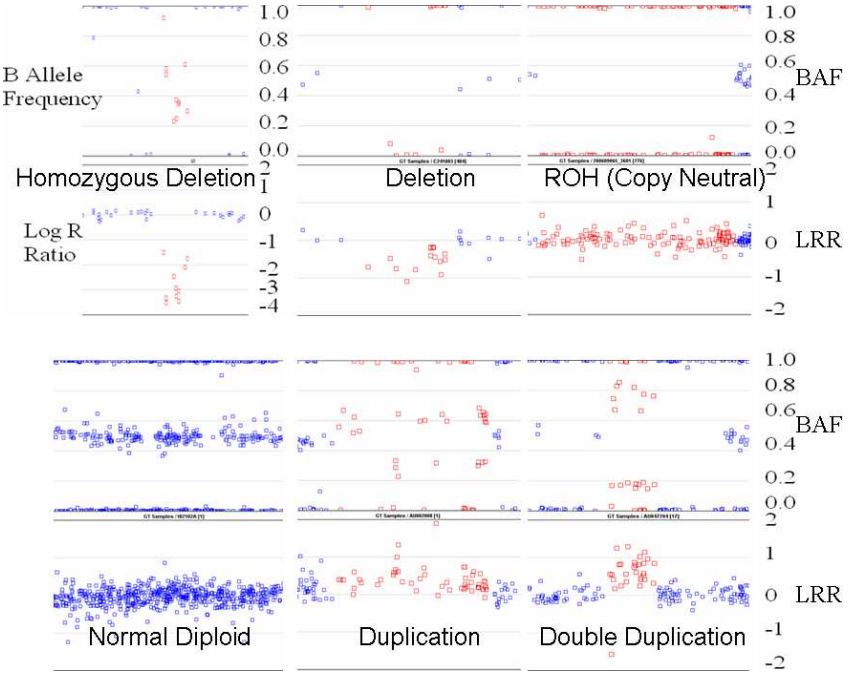
The control subjects used were drawn from Affymetrix 6.0 genotyped subjects from the NIMH genetics repository. They had been collected through a US Nationally representative survey panel (of approximately 60,000 adult individuals at any one time, with constant turnover) ascertained via random digit dialing. Participants were screened for psychosis and bipolar disorder. Control participants were not screened for ADHD. A blood sample was collected via a US national phlebotomy service. 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.

Supplementary Figures:

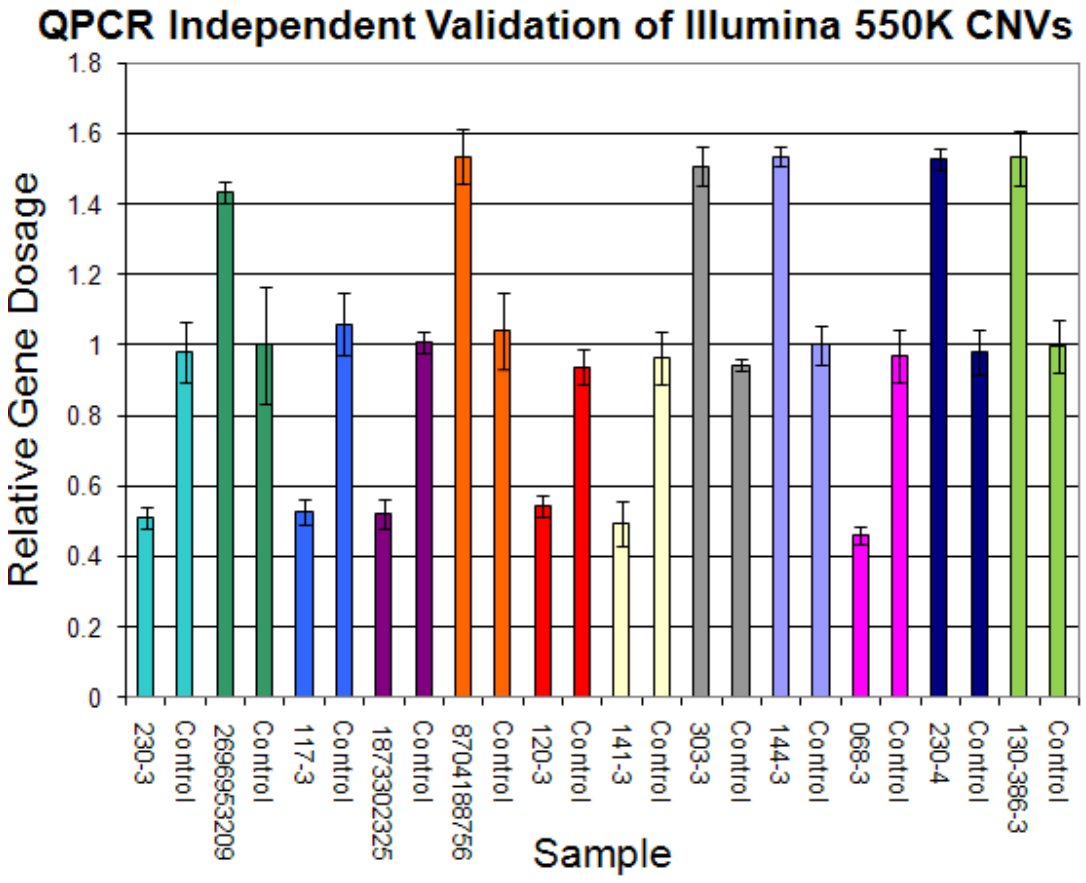
Supplementary Figure 1. Distribution of CNV calls per individual cases (top panel) vs controls (bottom panel).



Supplementary Figure 2. Examples of CNV observance based on B-allele frequency (BAF) and Log R Ratio (LRR).



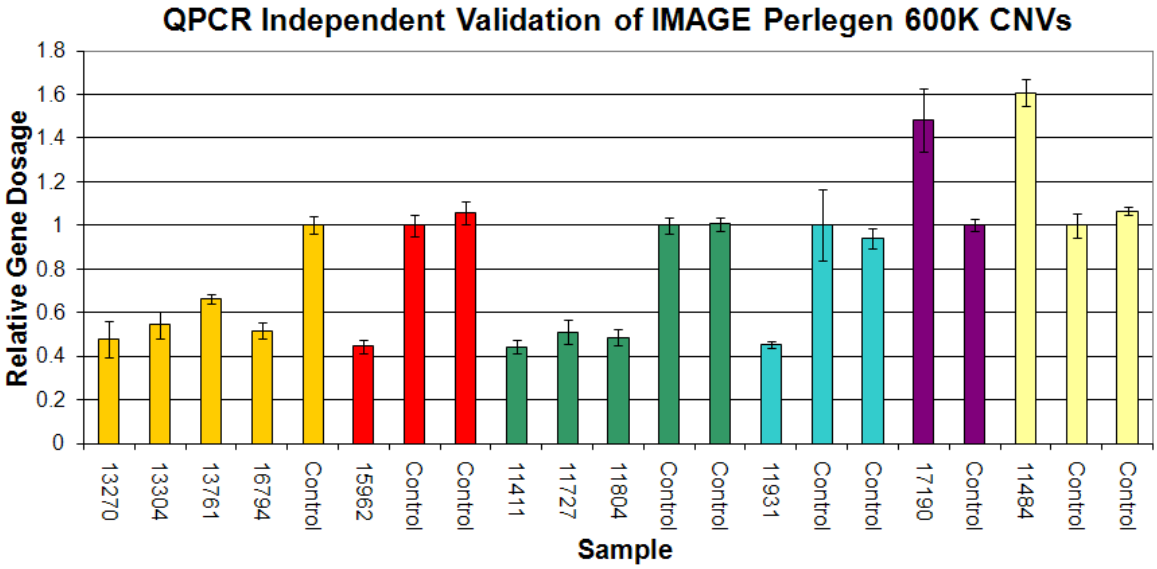
Supplementary Figure 3. CHOP Illumina Human Hap550 Independent Validation Using qPCR.



11q14.3 GRM5 Del; 7q36.2 DPP6 Dup; 5q12.3 SGTB/NLN Del; 1p32.3 USP24 Del; 7q31.33 GRM8 Del; 19q13.11 SLC7A10 Del; 3p26.3 CNTN4 Del; 4q25 LARP7 Dup; 2p12 CTNNA2 Dup; 3p26.1 GRM7 Del; 1p31.1 NEGR1 Dup; 6q24.3 GRM1 Dup

Fluorescent probe-based qPCR assays using Roche Universal probe were designed to validate every candidate CNV with a completely independent test (representative series shown for each locus in case and control pairs). Error bars denote the standard deviation of quadruplicate runs. Del, deletion; Dup,duplication.

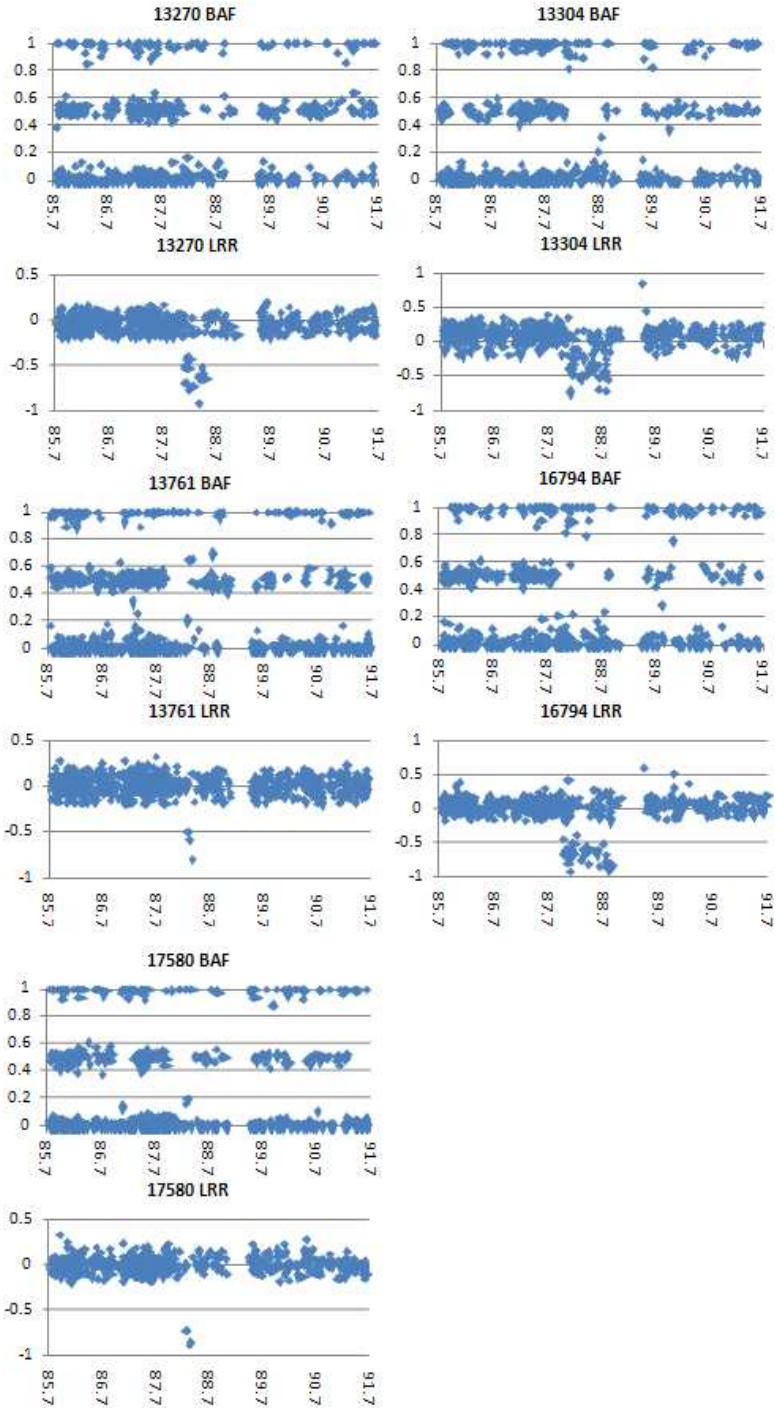
Supplementary Figure 4. IMAGE Perlegen 600K Independent Validation Using qPCR.



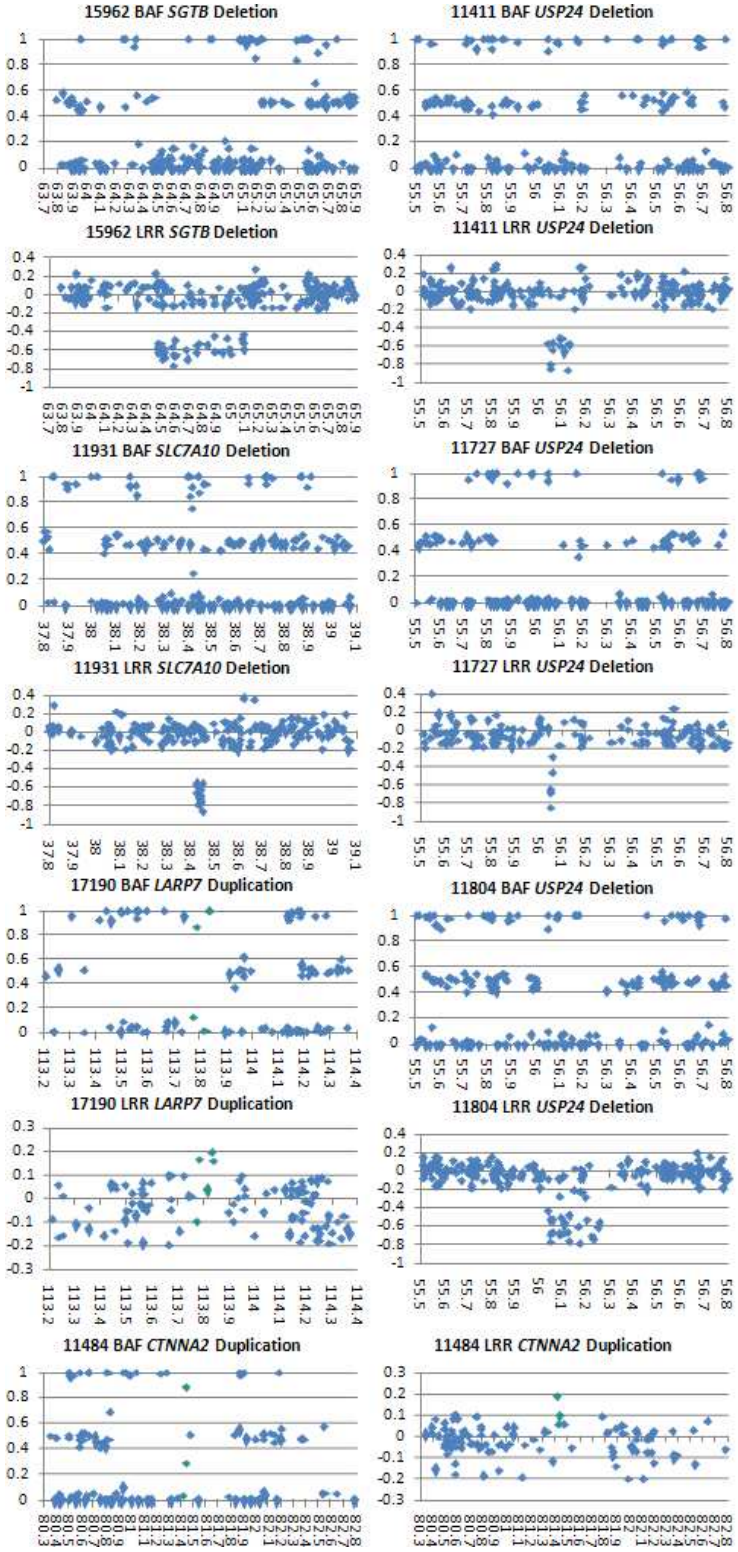
11q14.3 GRM5 Del; 5q12.3 SGTB/NLN Del; 1p32.3 USP24 Del; 19q13.11 SLC7A10 Del; 4q25 LARP7 Dup; 2p12 CTNNA2 Dup. Samples in IMAGE/NIMH not available: 17580 (11q14.3), 1135 (7q36.2).

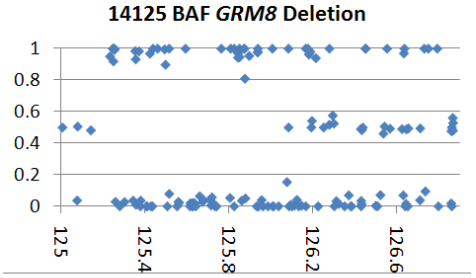
Fluorescent probe-based qPCR assays using Roche Universal probe were designed to validate every candidate CNV with a completely independent test (11 of the 14 IMAGE samples with replicating CNV calls for the loci reported were available for validation and all validated in comparison with control pairs; the other 3 loci were visually validated – see Supplem Figures 5 and 6). Error bars denote the standard deviation of quadruplicate runs. Del, deletion; Dup,duplication.

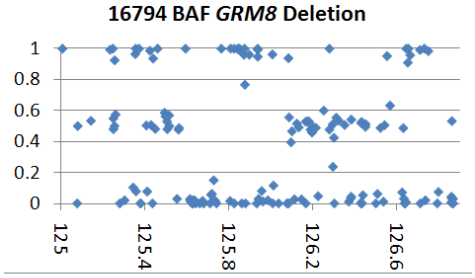
Supplementary Figure 5. Normalized SNP Level Perlegen 600K Data. The X axis shows base pair position in Megabases on chromosome 11. Raw SNP Level Data Showing GRM5 Deletion in five samples from IMAGE Perlegen 600K Data Normalized by Adapted PennCNV-Affy Protocol.

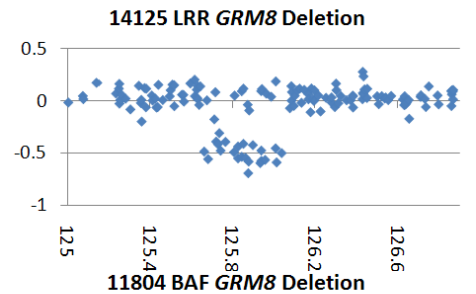


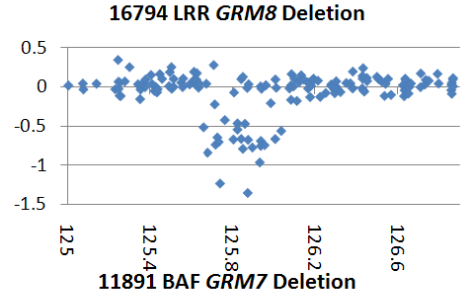
Supplementary Figure 6A Full SNP-Level Normalized Perlegen 600K Data.

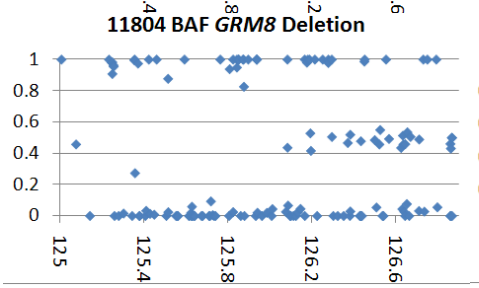


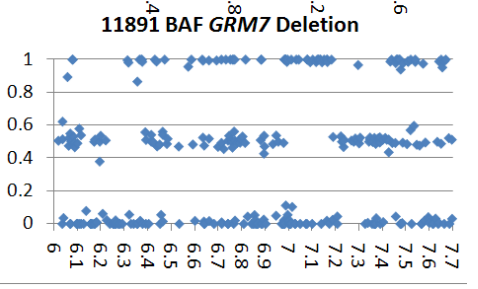

14125 BAF GRM8 Deletion
1
0.8
0.6
0.4
0.2
0
125
125.4
125.8
126.2
126.6

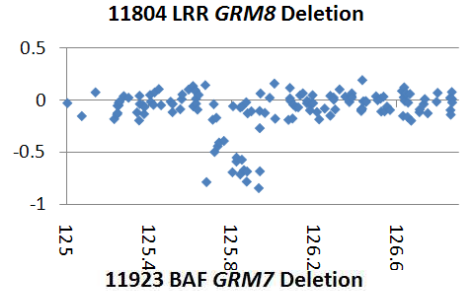

16794 BAF GRM8 Deletion
1
0.8
0.6
0.4
0.2
0
125
125.4
125.8
126.2
126.6

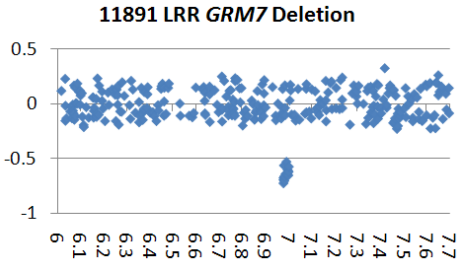

14125 LRR GRM8 Deletion
0.5
0
-0.5
-1
125
125.4
125.8
126.2
126.6

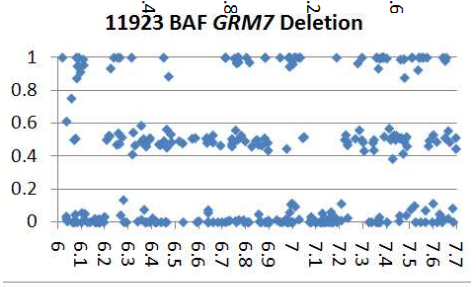

16794 LRR GRM8 Deletion
0.5
0
-0.5
-1
-1.5
125
125.4
125.8
126.2
126.6

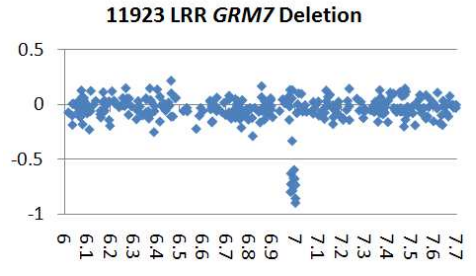

11804 BAF GRM8 Deletion
1
0.8
0.6
0.4
0.2
0
125
125.4
125.8
126.2
126.6


11891 BAF GRM7 Deletion
1
0.8
0.6
0.4
0.2
0
6
6.1
6.2
6.3
6.4
6.5
6.6
6.7
6.8
6.9
7
7.1
7.2
7.3
7.4
7.5
7.6
7.7

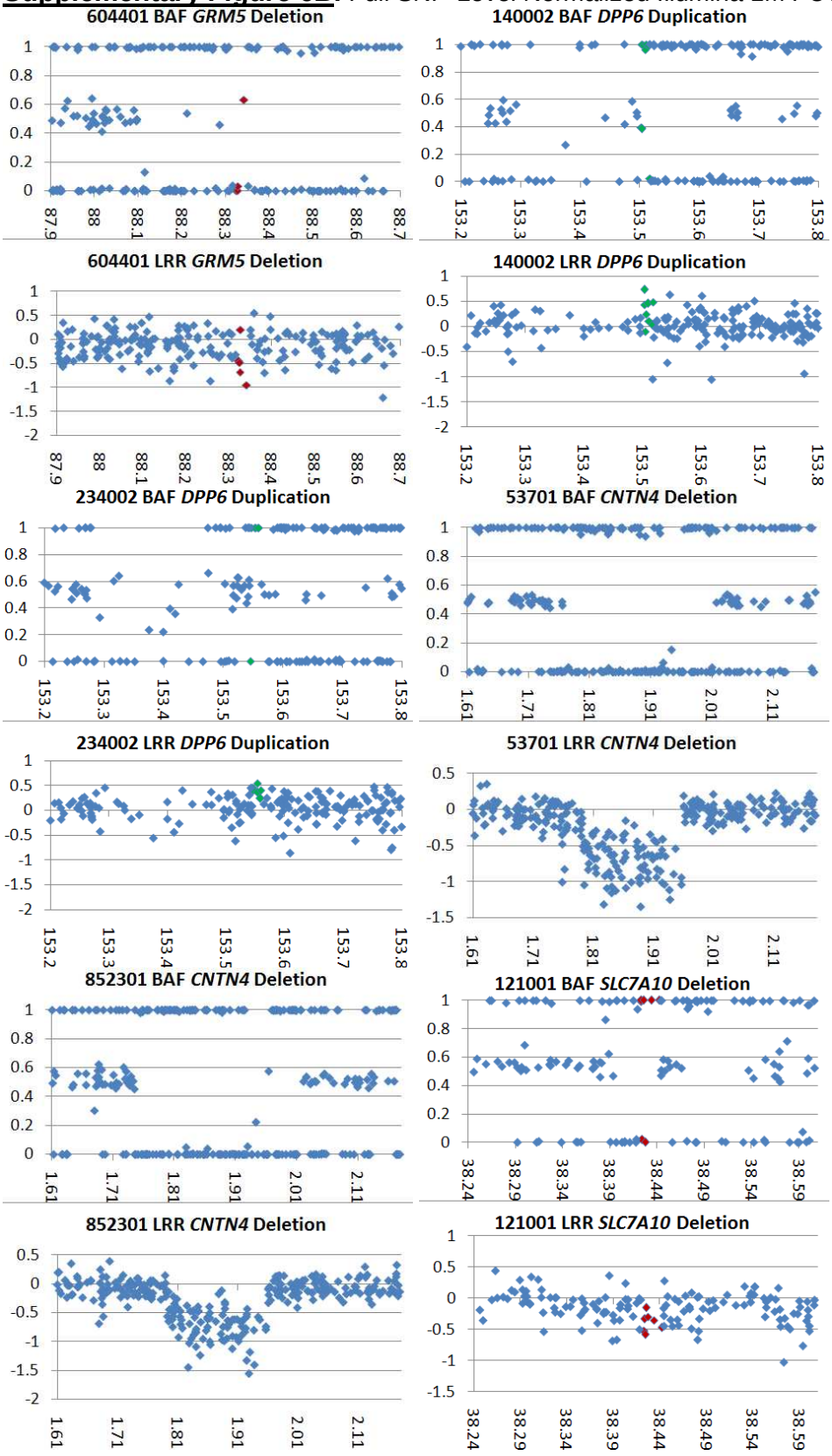

11804 LRR GRM8 Deletion
0.5
0
-0.5
-1
125
125.4
125.8
126.2
126.6

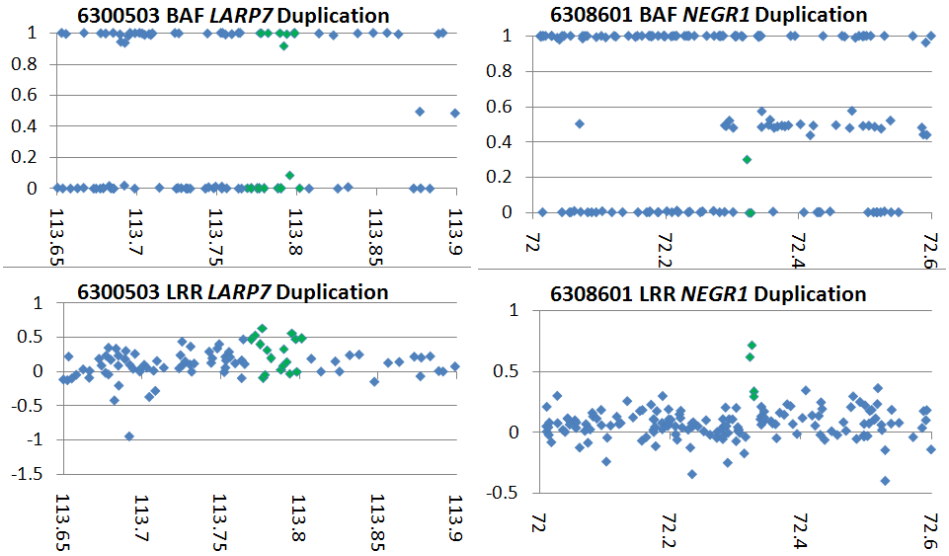

11891 LRR GRM7 Deletion
0.5
0
-0.5
-1
6
6.1
6.2
6.3
6.4
6.5
6.6
6.7
6.8
6.9
7
7.1
7.2
7.3
7.4
7.5
7.6
7.7


11923 BAF GRM7 Deletion
1
0.8
0.6
0.4
0.2
0
6
6.1
6.2
6.3
6.4
6.5
6.6
6.7
6.8
6.9
7
7.1
7.2
7.3
7.4
7.5
7.6
7.7

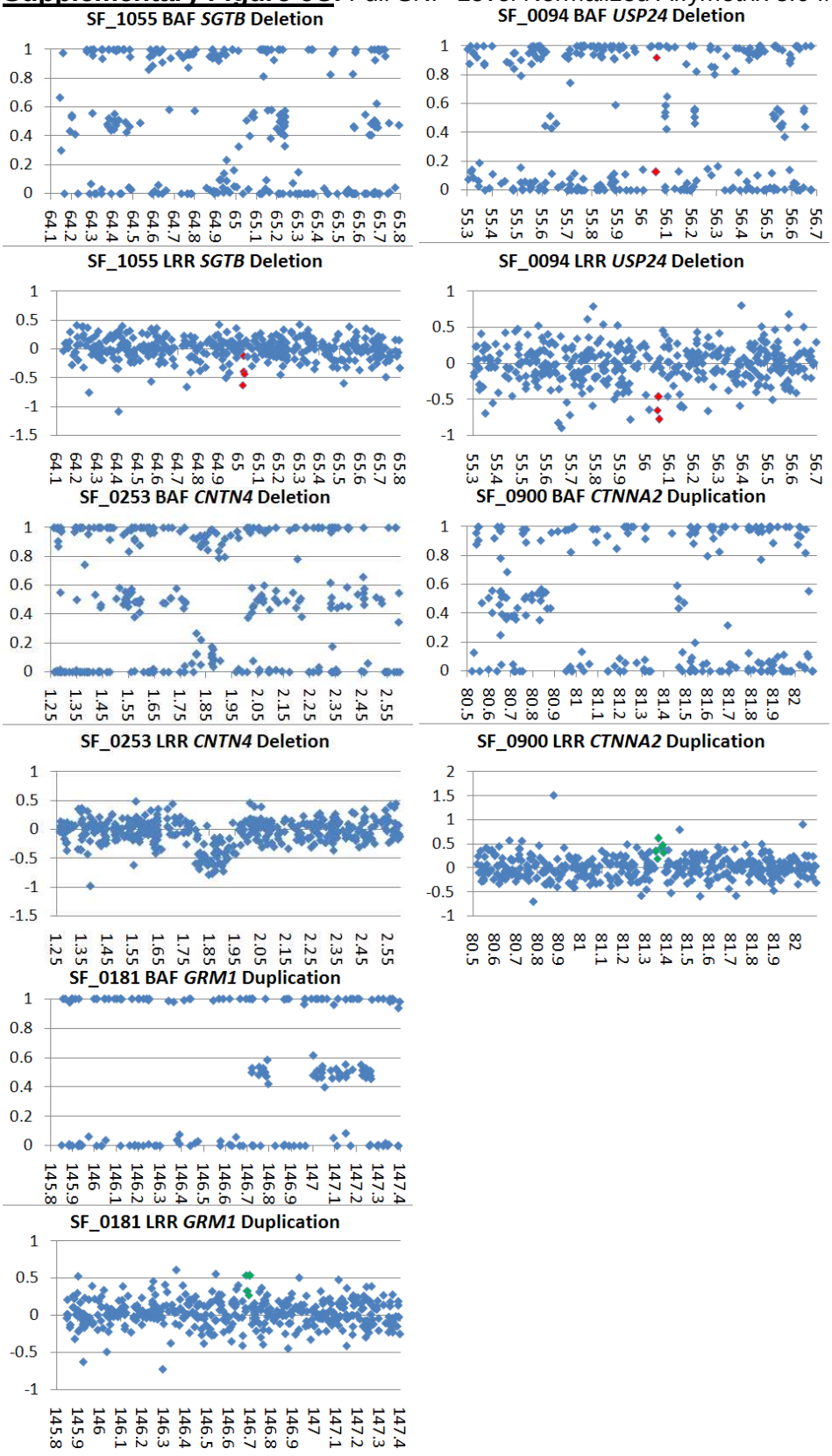

11923 LRR GRM7 Deletion
0.5
0
-0.5
-1
6
6.1
6.2
6.3
6.4
6.5
6.6
6.7
6.8
6.9
7
7.1
7.2
7.3
7.4
7.5
7.6
7.7

Supplementary Figure 6B. Full SNP-Level Normalized Illumina 1M PUWMa Data.

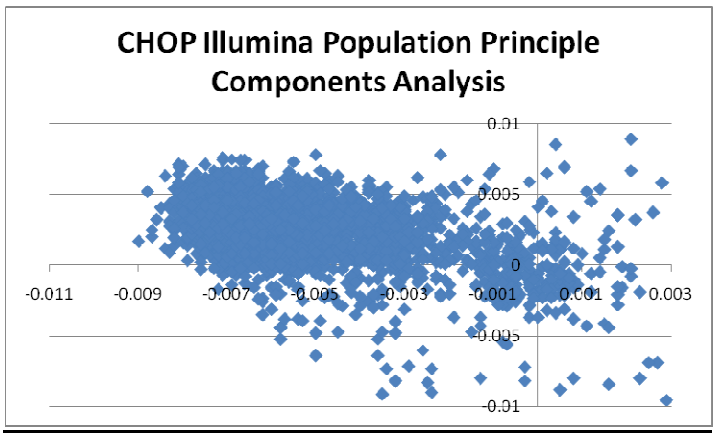


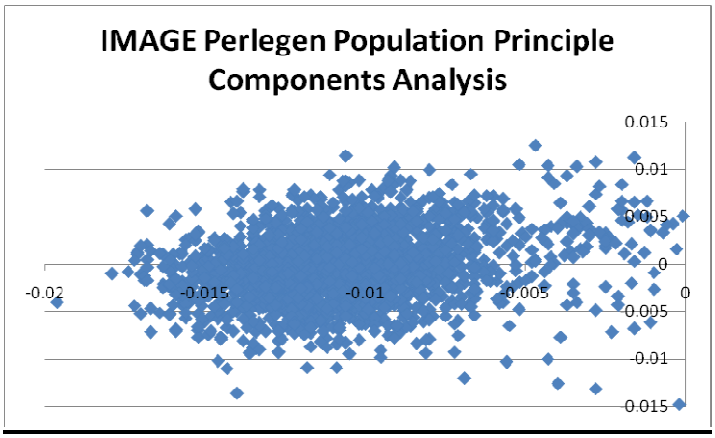

6300503 BAF LARP7 Duplication
1
0.8
0.6
0.4
0.2
0
113.65
113.7
113.75
113.8
113.85
113.9
6308601 BAF NEGR1 Duplication
1
0.8
0.6
0.4
0.2
0
72
72.2
72.4
72.6
6300503 LRR LARP7 Duplication
1
0.5
0
-0.5
-1
-1.5
113.65
113.7
113.75
113.8
113.85
113.9
6308601 LRR NEGR1 Duplication
1
0.5
0
-0.5
72
72.2
72.4
72.6

Supplementary Figure 6C. Full SNP-Level Normalized Affymetrix 5.0 IMAGE II Data.



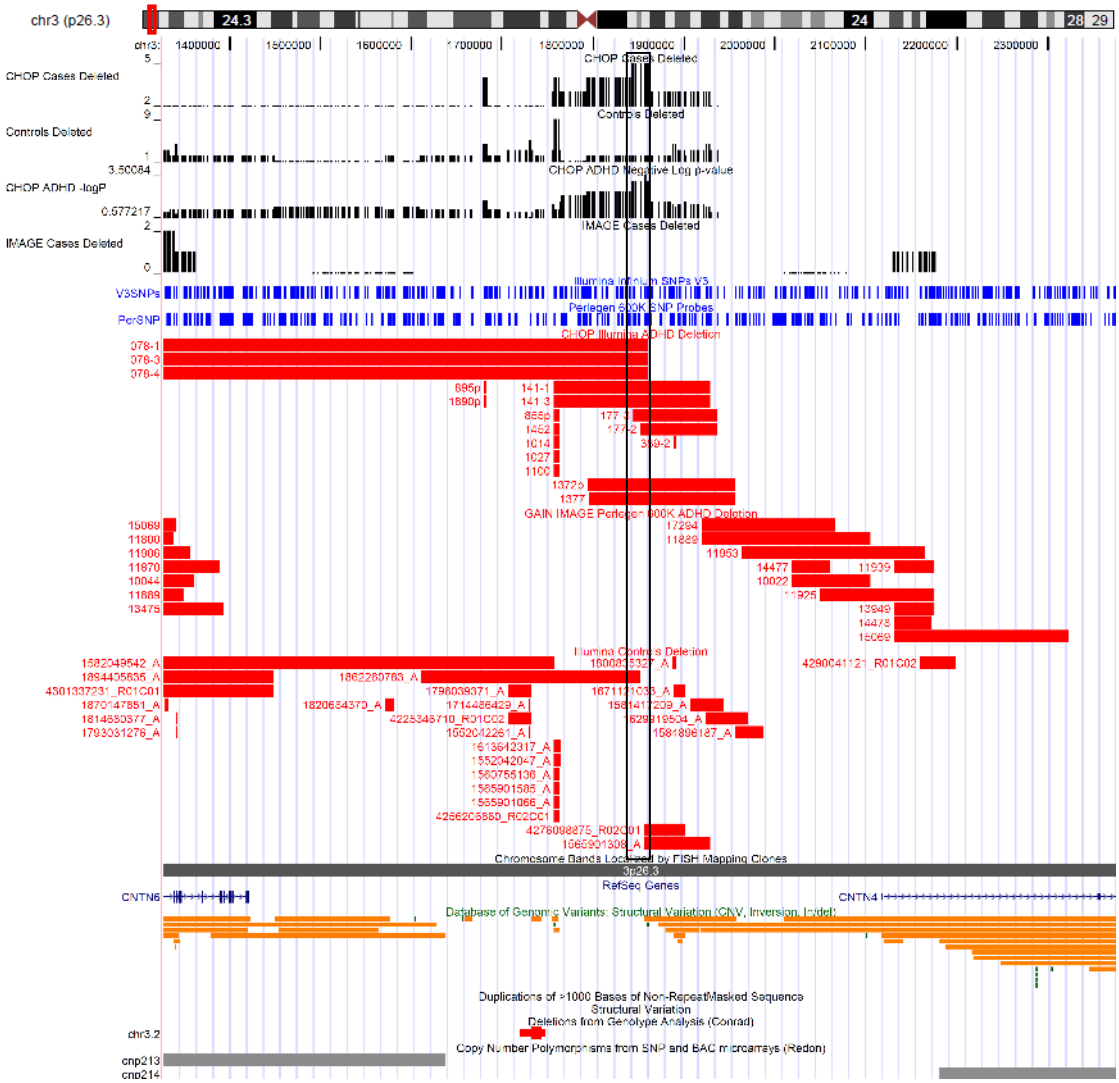
Supplementary Figure 7. Eigenstrat Principle Components Analysis. Cases and Controls were simultaneously analyzed to minimize population substructure in case control CNV association. Samples deviating from the Caucasian cluster shown were removed. The genomic inflation factor (GIF) within Plink was at an acceptable level (GIF=1.02409). We also checked pairwise population concordance to check for and filter out cryptic relatedness which could give rise to rare CNVs specific to ultra-stratified subpopulations of Europe.



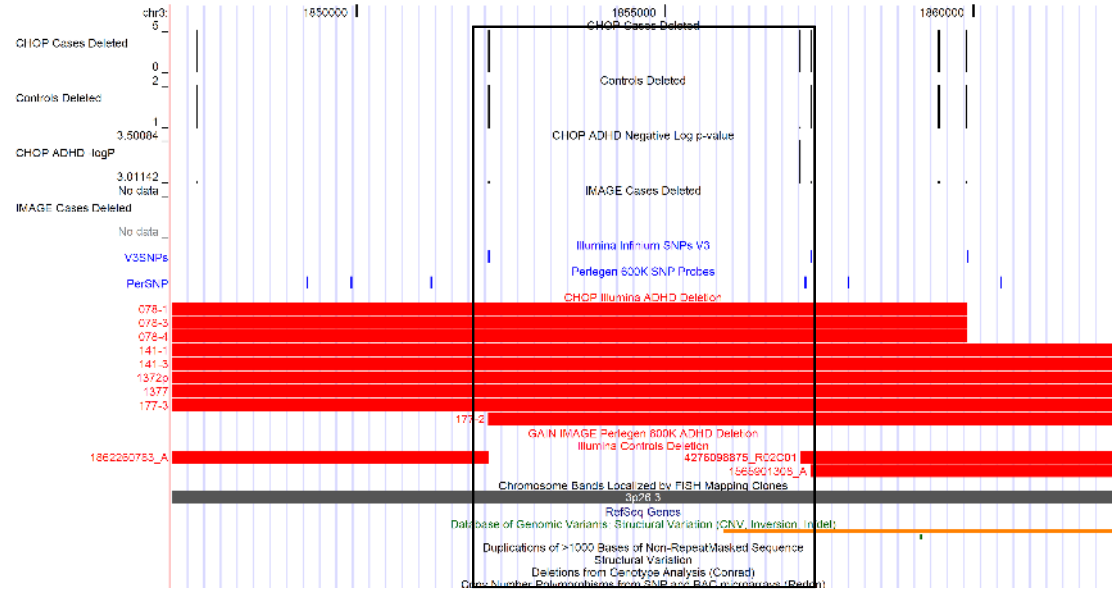


Supplementary Figure 8. Example of the SNP-based statistics applied and the resulting highest significance region Called**,** Examples from chr 3 are shown; A) 1,327,963-2,376,095 and B) 1,847,000-1,862,261. Complex CNV overlap is simplified by producing SNP-based statistics. As seen in plots for cases deleted and controls deleted, each SNP has a specific number of CNVs. The cases and controls are compared with a Fisher's exact test and the negative log p value is shown in the third plot. Regions of significance ranging within a power of ten are reported and the region of highest significance (local minimum p-value) within 1MB is reported. The IMAGE cases deleted plot shows only one case sample #11939 since the remaining red regions 3' are parents.

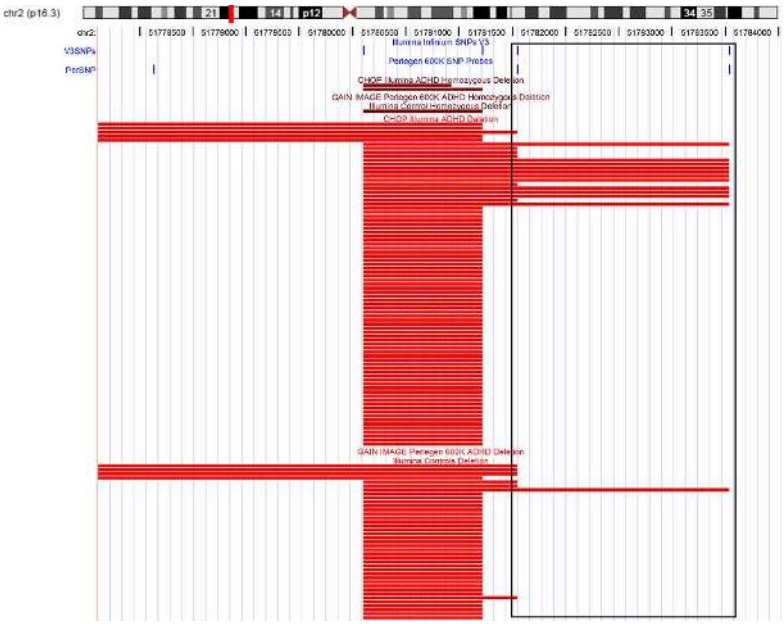
A)



B)



Supplementary Figure 9. CNV peninsula false positive association example. An example from chr 2 is shown (location 51,777,616-51,784,033**).** All significant CNVRs are reviewed for CNV peninsulas indicating uncertainty in boundary truncation.



Supplementary Tables

Supplementary Table 1. Clinical Demographics of Study Participants.

ADHD Cohort	N	ADHD subjects Age range	Ancestry	ADHD ascertainment
CHOP ADHD trios	349	6-18	European	K-SADS-IVR
CHOP ADHD cases	664	6-18	European	Clinical ADHD diagnosis & treatment with ADHD meds; K-SADS-IVR on majority
NIMH ADHD trios	128	6-12	European	DICA; Conners Scales
UTAH cases	90	19-60	European	WRAADDS, WURS, PRS, strict DSM-IV criteria, including age of onset before 7
IMAGE ADHD trios	642	6-17	European	PACS, Conners, SDQ, WISC
IMAGE II ADHD trios	787	5-14	European	K-SADS German version, Kinder-DIPS, Conners parent and teacher scales, WISC, K-ABC
PUWMa trios	864	6-18	European	K-SADS

PACS: Parental Account of Child Symptoms; Conners: Behavioral rating scales; SDQ: Strength and Difficulties Questionnaire; WISC: Wechsler Intelligence Scale for Children (WISC-IV); KSADS-IVR: Schedule for Affective Disorders and Schizophrenia for School-Age Children-IVR; DICA: Diagnostic Interview for Children and Adolescents; Kinder-DIPS: Diagnostic Interview for Psychiatric Disorders in Children, K-ABC: Kaufman-ABC intelligence scale. WRAADDS=Wender-Reimherr Adult Attention Deficit Disorder Scale; WURS=Wender Utah Rating Scale; PRS=Parent Rating Scale.

Supplementary Table 2. K-SADS ADHD Severity of of CHOP Study Participants in Inattentive, Impulsive, and Hyperactive Domains.

Diagnostic Criteria	Score 1	Score 2	Score 3	Score 4
Often Careless	7	40	372	81
Loses Things	18	126	277	79
Difficulty Finishing	16	90	311	83
Listening	10	22	320	148
Concentration*	2	25	337	135
Distracted	1	10	307	182
Organizing	19	79	304	98
Avoiding	19	55	278	148
Forgetful	19	75	290	116
Interrupts	28	73	305	94
Acts Before Thinking	28	112	283	77
Shifts Activities	72	134	247	47
Blurts†	135	82	232	48
Difficulty Waiting Turn	80	172	200	48
Hyperactive	53	127	227	93
Fidgeting	15	47	301	137
Difficulty Staying Seated	45	80	287	88
On the Go	49	89	255	107
Talks Excess	37	77	255	131
Difficulty Playing Quietly	98	120	233	49

*Concentration 1 record missing †Blurts 3 records missing. Scores 1 and 2 means that symptoms are within the normal range while scores 3 and 4 are excessive.

Supplementary Table 3. TDT Analysis of 397 ADHD Cases and Parents from CHOP genotyped on the Illumina HH550 chip.

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
18	rs8095193	58834095	1	2	167	92	1.815	21.72	3.16E-06
17	rs4357980	13498634	1	2	99	174	0.569	20.6	5.65E-06
18	rs8091710	72897492	1	2	29	73	0.3973	18.98	1.32E-05
14	rs899116	97495185	1	2	101	172	0.5872	18.47	1.73E-05
13	rs9595945	48099556	1	2	245	160	1.531	17.84	2.40E-05
4	rs1018199	47927632	1	2	35	80	0.4375	17.61	2.71E-05
1	rs3795324	157456184	2	1	91	157	0.5796	17.56	2.78E-05
3	rs6444186	188156541	1	2	81	36	2.25	17.31	3.18E-05
9	rs11144627	75654927	2	1	46	14	3.286	17.07	3.61E-05
8	rs1462011	108104653	1	2	199	125	1.592	16.9	3.94E-05
X	rs5991935	100480088	1	2	22	59	0.3729	16.9	3.94E-05
7	rs1013572	78350227	1	2	63	118	0.5339	16.71	4.35E-05
11	rs952619	20316347	1	2	108	177	0.6102	16.71	4.37E-05
4	rs7689018	85116479	1	2	41	87	0.4713	16.53	4.79E-05
18	rs1943825	69128567	2	1	97	162	0.5988	16.31	5.37E-05
4	rs4696821	8473961	1	2	210	135	1.556	16.3	5.39E-05
18	rs1943823	69131624	2	1	157	237	0.6624	16.24	5.57E-05
4	rs11724347	47923023	1	2	26	64	0.4062	16.04	6.19E-05
1	rs7530899	76950752	2	1	89	151	0.5894	16.02	6.28E-05
18	rs4890560	41457783	1	2	93	156	0.5962	15.94	6.54E-05
6	rs2677099	45527900	1	2	220	144	1.528	15.87	6.79E-05
12	rs11067228	113556980	2	1	231	153	1.51	15.84	6.88E-05
6	rs2790102	45540192	1	2	222	146	1.521	15.7	7.44E-05
1	rs4926757	48961624	1	2	192	122	1.574	15.61	7.80E-05
11	rs17147479	84055504	1	2	137	79	1.734	15.57	7.93E-05
17	rs9913261	12026365	2	1	89	150	0.5933	15.57	7.96E-05
9	rs7041883	135352660	1	2	17	49	0.3469	15.52	8.19E-05
12	rs7309946	103478293	2	1	119	188	0.633	15.51	8.22E-05
7	rs10226468	42907176	2	1	144	219	0.6575	15.5	8.27E-05
5	rs438418	2902436	2	1	78	36	2.167	15.47	8.37E-05
8	rs12682232	108078371	2	1	199	128	1.555	15.42	8.63E-05
X	rs5956634	123092612	2	1	59	110	0.5364	15.39	8.74E-05
7	rs7786719	42850356	1	2	133	205	0.6488	15.34	8.99E-05
6	rs910586	45518290	1	2	221	146	1.514	15.33	9.04E-05
6	rs9395010	44453984	1	2	152	91	1.67	15.31	9.11E-05
14	rs11844273	97489409	1	2	100	163	0.6135	15.09	1.02E-04
2	rs11904235	36288350	1	2	64	27	2.37	15.04	1.05E-04
11	rs487518	131283728	1	2	150	225	0.6667	15	1.08E-04
6	rs6920606	33105652	2	1	164	242	0.6777	14.99	1.08E-04
14	rs2014525	97491178	1	2	109	174	0.6264	14.93	1.12E-04
11	rs7948111	23403649	1	2	65	117	0.5556	14.86	1.16E-04
16	rs12598067	60940038	2	1	65	117	0.5556	14.86	1.16E-04
6	rs9472494	45559814	1	2	223	149	1.497	14.72	1.25E-04
7	rs533486	99085345	2	1	163	240	0.6792	14.71	1.25E-04
8	rs7835921	96345468	1	2	157	96	1.635	14.71	1.26E-04
4	rs827019	8460842	2	1	69	122	0.5656	14.71	1.26E-04

CHR:Chromosome number, SNP:SNP identifier, A1:Minor allele code, A2:Major allele code, T:Transmitted minor allele count, U:Untransmitted allele count, OR:TDT odds ratio, CHISQ:TDT chi-square statistic, P:TDT asymptotic p-value

Supplementary Table 4. Case:Control Analysis of 735 ADHD Cases and 2,298 Unrelated Controls from CHOP genotyped on the Illumina HH550 chip.

CHR	SNP	BP	A1	A2	F_A	F_U	OR	CHISQ	P
18	rs16943400	23086102	1	2	0.02778	0.08875	0.2934	57.53	3.33E-14
3	rs7649108	166136126	1	2	0.3156	0.2497	1.386	24.88	6.11E-07
6	rs9390261	145283744	1	2	0.02585	0.009072	2.899	24.54	7.29E-07
X	rs4609327	37790223	2	1	0.1441	0.08032	1.928	24.48	7.50E-07
X	rs5917547	37803525	2	1	0.1578	0.09074	1.878	24.22	8.59E-07
16	rs2278656	54885245	1	2	0.01443	0.04091	0.3432	22.04	2.67E-06
8	rs17834541	2674349	2	1	0.1083	0.1565	0.6545	21.01	4.56E-06
19	rs2081051	20866811	1	2	0.1382	0.1911	0.6786	21	4.60E-06
19	rs399686	20772798	1	2	0.143	0.1962	0.6833	20.95	4.72E-06
X	rs5917937	39750534	2	1	0.1195	0.06572	1.929	20.93	4.76E-06
19	rs10419820	20943636	2	1	0.1789	0.2357	0.7067	20.9	4.84E-06
X	rs10522011	32517409	1	2	0.05924	0.02509	2.447	19.48	1.02E-05
8	rs11203872	17531028	2	1	0.4342	0.37	1.306	19.34	1.09E-05
X	rs9633179	3535471	2	1	0.1089	0.05969	1.925	19.24	1.15E-05
4	rs10519629	143040375	2	1	0.1864	0.1398	1.409	18.81	1.44E-05
19	rs7253306	20951939	2	1	0.219	0.2759	0.736	18.77	1.48E-05
13	rs9569383	55299477	1	2	0.1415	0.1909	0.6984	18.64	1.58E-05
12	rs12229174	62532933	1	2	0.06054	0.03502	1.776	18.56	1.64E-05
19	rs6511169	20893589	1	2	0.1461	0.1961	0.7014	18.51	1.69E-05
11	rs10833476	21190445	1	2	0.1224	0.08502	1.502	18.48	1.72E-05
2	rs1821659	212064488	2	1	0.3109	0.2527	1.334	18.15	2.05E-05
X	rs2480443	53212284	2	1	0.06525	0.02994	2.262	18.1	2.09E-05
7	rs1486173	45965025	2	1	0.1131	0.07764	1.515	17.91	2.32E-05
15	rs4381545	93039961	2	1	0.2296	0.18	1.358	17.8	2.45E-05
7	rs10265665	96175055	1	2	0.0619	0.0365	1.742	17.79	2.46E-05
10	rs11593585	44391199	1	2	0.1286	0.09093	1.475	17.69	2.60E-05
X	rs4134188	17474194	1	2	0.1016	0.05571	1.917	17.62	2.69E-05
4	rs11131363	63013616	2	1	0.2643	0.212	1.335	17.6	2.72E-05
19	rs1469402	20738115	2	1	0.145	0.1934	0.7075	17.52	2.85E-05
11	rs12279152	133861485	1	2	0.02653	0.01139	2.365	17.43	2.98E-05
X	rs5957334	119125665	2	1	0.06667	0.03136	2.206	17.13	3.49E-05
X	rs6632558	36075450	2	1	0.0812	0.04176	2.028	16.94	3.85E-05
1	rs2057594	117348535	1	2	0.2483	0.1983	1.335	16.89	3.96E-05
8	rs17834523	2672777	1	2	0.09592	0.1367	0.6699	16.84	4.06E-05
7	rs10485959	78702412	2	1	0.3007	0.3595	0.7659	16.83	4.09E-05
X	rs5945330	152438289	2	1	0.08807	0.04698	1.959	16.63	4.55E-05
3	rs16854851	145238402	1	2	0.02381	0.009916	2.435	16.62	4.56E-05
8	rs2237826	17519195	2	1	0.4355	0.376	1.28	16.59	4.65E-05
X	rs16987407	35968032	2	1	0.1041	0.05857	1.868	16.5	4.87E-05
X	rs4089885	22878045	2	1	0.1193	0.07027	1.792	16.47	4.94E-05
1	rs2024766	181385290	2	1	0.5027	0.4424	1.274	16.45	4.99E-05
4	rs9312518	173526549	1	2	0.4639	0.4042	1.276	16.45	5.00E-05
4	rs9997484	173517324	2	1	0.4639	0.4042	1.276	16.45	5.00E-05
17	rs4338847	7870502	1	2	0.3102	0.3679	0.7725	16.35	5.28E-05
12	rs17497206	113000660	2	1	0.1537	0.2011	0.7219	16.33	5.32E-05

CHR:Chromosome, SNP:SNP ID, BP:Physical position (base-pair), A1:Minor allele name (based on whole sample), F_A:Frequency of this allele in cases, F_U:Frequency of this allele in controls, A2:Major allele name, OR:Estimated odds ratio (for A1, i.e. A2 is reference), CHISQ:Basic allelic test chi-square (1df), P:Asymptotic p-value for this test.

Supplementary Table 5. TDT Analysis of 623 ADHD Cases and Parents from IMAGE genotyped on the Perlegen platform.

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
12	rs3782309	26750663	1	2	172	99	1.737	19.66	9.23E-06
5	rs17144308	117965870	2	1	244	352	0.6932	19.57	9.70E-06
2	rs7609261	80530821	2	1	199	297	0.67	19.36	1.08E-05
3	rs1344870	21282405	2	1	16	52	0.3077	19.06	1.27E-05
18	rs7244637	17876224	1	2	134	215	0.6233	18.8	1.45E-05
1	rs3850879	48004718	1	2	226	143	1.58	18.67	1.56E-05
14	rs2295426	58446208	2	1	209	307	0.6808	18.61	1.60E-05
16	rs7204253	5576184	2	1	114	189	0.6032	18.56	1.64E-05
4	rs1378945	25382295	2	1	212	310	0.6839	18.4	1.79E-05
3	rs11915713	74568983	1	2	176	266	0.6617	18.33	1.86E-05
12	rs11830382	41718893	2	1	198	122	1.623	18.05	2.15E-05
12	rs4761641	93525817	2	1	137	215	0.6372	17.28	3.22E-05
5	rs2043053	117958083	2	1	126	201	0.6269	17.2	3.36E-05
18	rs12965880	22313077	1	2	235	333	0.7057	16.91	3.92E-05
9	rs17306197	97862011	1	2	162	96	1.688	16.88	3.97E-05
8	rs17668689	96254526	1	2	216	310	0.6968	16.8	4.16E-05
2	rs4852567	80556890	2	1	206	298	0.6913	16.79	4.17E-05
13	rs1002468	93085569	2	1	287	197	1.457	16.74	4.30E-05
1	rs10873925	77234323	2	1	305	212	1.439	16.73	4.31E-05
16	rs12596741	17345435	1	2	228	324	0.7037	16.7	4.39E-05
9	rs2991298	3284851	2	1	81	142	0.5704	16.69	4.41E-05
14	rs1427324	58434446	1	2	206	297	0.6936	16.46	4.96E-05
10	rs11258682	13951273	1	2	204	130	1.569	16.4	5.14E-05
4	rs10520276	175420068	2	1	216	140	1.543	16.22	5.63E-05
1	rs17375519	179499648	1	2	75	133	0.5639	16.17	5.78E-05
1	rs10800069	163296159	1	2	232	327	0.7095	16.14	5.87E-05
7	rs13340504	75277632	1	2	142	82	1.732	16.07	6.10E-05
2	rs6543239	104056246	2	1	251	349	0.7192	16.01	6.31E-05
2	rs4664452	162762970	1	2	30	6	5	16	6.33E-05
4	rs16889099	13341184	2	1	48	96	0.5	16	6.33E-05
5	rs12520147	2000122	1	2	158	237	0.6667	15.8	7.04E-05
11	rs10400283	23523711	1	2	222	314	0.707	15.79	7.07E-05
4	rs1378946	25382548	1	2	197	284	0.6937	15.74	7.28E-05
3	rs7372975	74602140	2	1	169	250	0.676	15.66	7.59E-05
17	rs11654470	74388926	2	1	82	141	0.5816	15.61	7.79E-05
3	rs9878591	121464488	1	2	107	173	0.6185	15.56	8.01E-05
12	rs1553953	28724544	1	2	76	133	0.5714	15.55	8.06E-05
11	rs7121790	45021541	1	2	171	252	0.6786	15.51	8.20E-05
12	rs1452231	83750252	2	1	223	314	0.7102	15.42	8.60E-05
7	rs194847	103560404	1	2	347	251	1.382	15.41	8.65E-05
2	rs11902138	80565100	1	2	173	254	0.6811	15.37	8.86E-05
16	rs12932714	80320240	1	2	150	226	0.6637	15.36	8.88E-05
1	rs1015144	200004976	2	1	204	291	0.701	15.29	9.22E-05
22	rs6009441	47873456	1	2	107	172	0.6221	15.14	9.97E-05
8	rs4734069	104169047	1	2	275	191	1.44	15.14	9.97E-05
20	rs2024946	61678306	2	1	112	61	1.836	15.03	1.06E-04

CHR:Chromosome number, SNP:SNP identifier, A1:Minor allele code, A2:Major allele code, T:Transmitted minor allele count, U:Untransmitted allele count, OR:TDT odds ratio, CHISQ:TDT chi-square statistic, P:TDT asymptotic p-value

Supplementary Table 6. SNP GWAS Significance of Top Ranked ADHD Associated SNPs Reported by Lesch and Zhou. A) ADHD TDT CHOP Illumina 550k data; B) ADHD Case:Control CHOP Illumina 550k data; C) ADHD IMAGE Perlegen 600k data.

A)

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
2	rs2241685	1896290	1	2	72	62	1.161	0.7463	0.3877
2	rs13395022	79793915	2	1	136	136	1	0	1
2	rs2587695	120038047	1	2	183	197	0.9289	0.5158	0.4726
2	rs2242073	208819551	2	1	108	106	1.019	0.01869	0.8913
2	rs1110998	217169458	1	2	175	159	1.101	0.7665	0.3813
3	rs10510238	2876647	2	1	84	93	0.9032	0.4576	0.4987
3	rs9879164	54040611	2	1	185	198	0.9343	0.4413	0.5065
3	rs2084358	57457928	2	1	182	198	0.9192	0.6737	0.4118
3	rs10490808	59939739	2	1	175	204	0.8578	2.219	0.1363
3	rs10510850	60542142	1	2	90	83	1.084	0.2832	0.5946
4	rs755403	6507714	2	1	195	180	1.083	0.6	0.4386
4	rs10516182	7143981	2	1	155	169	0.9172	0.6049	0.4367
4	rs7697323	7801488	1	2	180	222	0.8108	4.388	0.03619
5	rs173754	65102081	1	2	218	202	1.079	0.6095	0.435
5	rs258082	66166352	1	2	199	205	0.9707	0.08911	0.7653
6	rs160666	2719051	2	1	179	181	0.989	0.01111	0.9161
6	rs2842643	41758714	2	1	180	149	1.208	2.921	0.08744
6	rs3799977	44945334	2	1	209	183	1.142	1.724	0.1891
6	rs8180608	89064414	2	1	178	218	0.8165	4.04	0.04442
6	rs1358601	91532294	1	2	180	181	0.9945	0.00277	0.958
6	rs6921403	154156020	2	1	86	90	0.9556	0.09091	0.763
7	rs2237349	28536203	2	1	176	191	0.9215	0.6131	0.4336
7	rs2002865	154132035	2	1	134	157	0.8535	1.818	0.1776
8	rs6991017	5508780	2	1	127	126	1.008	0.003953	0.9499
8	rs2248529	14657354	1	2	188	190	0.9895	0.01058	0.9181
8	rs4961315	142110882	2	1	186	152	1.224	3.42	0.06441
9	rs2418326	114759028	1	2	141	142	0.993	0.003534	0.9526
9	rs2502731	128056111	2	1	170	178	0.9551	0.1839	0.668
14	rs10483393	31530235	1	2	146	137	1.066	0.2862	0.5927
15	rs2556560	42609135	2	1	169	171	0.9883	0.01176	0.9136
16	rs8060494	78808972	2	1	190	174	1.092	0.7033	0.4017
17	rs4790372	2701606	2	1	163	169	0.9645	0.1084	0.7419
17	rs12453316	69027654	1	2	177	179	0.9888	0.01124	0.9156
19	rs997669	34996323	2	1	201	183	1.098	0.8438	0.3583
20	rs1555322	33312595	1	2	94	79	1.19	1.301	0.2541

B)

CHR	SNP	BP	A1	F_A	F_U	A2	OR	CHISQ	P
2	rs2241685	1896290	1	0.09116	0.09283	2	0.9802	0.03733	0.8468
2	rs13395022	79793915	2	0.2088	0.2095	1	0.9961	0.002865	0.9573
2	rs2587695	120038047	1	0.4973	0.4922	2	1.021	0.1161	0.7333
2	rs2242073	208819551	2	0.1605	0.1568	1	1.029	0.1216	0.7273
2	rs1110998	217169458	1	0.3116	0.2928	2	1.093	1.886	0.1697
3	rs10510238	2876647	2	0.1293	0.1376	1	0.9304	0.6621	0.4158
3	rs9879164	54040611	2	0.4218	0.4359	1	0.9441	0.9084	0.3406
3	rs2084358	57457928	1	0.5184	0.4722	2	1.203	9.597	0.001949
3	rs10490808	59939739	2	0.4068	0.4266	1	0.9218	1.8	0.1797
3	rs10510850	60542142	1	0.1211	0.1116	2	1.097	1.001	0.3172
4	rs755403	6507714	2	0.3985	0.3973	1	1.005	0.007242	0.9322
4	rs10516182	7143981	2	0.2801	0.2954	1	0.9279	1.274	0.259
4	rs7697323	7801488	1	0.3782	0.38	2	0.9927	0.0142	0.9051
5	rs173754	65102081	1	0.4925	0.4915	2	1.004	0.004285	0.9478
5	rs258082	66166352	1	0.4619	0.4521	2	1.04	0.4342	0.5099
6	rs160666	2719051	2	0.2857	0.3025	1	0.9222	1.515	0.2183
6	rs2842643	41758714	2	0.2932	0.2909	1	1.011	0.02797	0.8672
6	rs3799977	44945334	2	0.4306	0.4076	1	1.099	2.452	0.1174
6	rs8180608	89064414	2	0.4101	0.4441	1	0.8703	5.265	0.02176
6	rs1358601	91532294	1	0.3852	0.3846	2	1.003	0.002076	0.9637
6	rs6921403	154156020	2	0.1373	0.1405	1	0.9736	0.09408	0.7591
7	rs2237349	28536203	2	0.4109	0.4082	1	1.011	0.03276	0.8564
7	rs2002865	154132035	2	0.2075	0.217	1	0.9445	0.6065	0.4361
8	rs6991017	5508780	2	0.1891	0.1873	1	1.012	0.02315	0.8791
8	rs2248529	14657354	1	0.3604	0.363	2	0.9888	0.03305	0.8557
8	rs4961315	142110882	2	0.2959	0.2995	1	0.983	0.06846	0.7936
9	rs2418326	114759028	1	0.2534	0.252	2	1.007	0.01179	0.9135
9	rs2502731	128056111	2	0.3626	0.3508	1	1.053	0.6767	0.4107
14	rs10483393	31530235	1	0.2272	0.2203	2	1.041	0.3146	0.5749
15	rs2556560	42609135	2	0.419	0.4215	1	0.9899	0.02811	0.8668
16	rs8060494	78808972	2	0.3215	0.3228	1	0.9943	0.008131	0.9282
17	rs4790372	2701606	2	0.3014	0.3112	1	0.9546	0.5122	0.4742
17	rs12453316	69027654	1	0.3612	0.3662	2	0.9788	0.1193	0.7298
19	rs997669	34996323	2	0.4023	0.3876	1	1.064	1.025	0.3114
20	rs1555322	33312595	1	0.1279	0.1277	2	1.002	0.0004034	0.984

C)

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
1	rs2281597	34132445	0	2	0	0	NA	NA	NA
1	rs642969	197590139	0	2	0	0	NA	NA	NA
2	rs2587695	120038287	1	2	320	294	1.088	1.101	0.2941
2	rs2242073	208702290	2	1	185	182	1.016	0.02452	0.8756
3	rs10510850	60542142	1	2	109	115	0.9478	0.1607	0.6885
3	rs17233461	125807474	2	1	305	322	0.9472	0.4609	0.4972
4	rs755403	6440543	2	1	296	278	1.065	0.5645	0.4525
4	rs3857174	7089831	2	1	202	217	0.9309	0.537	0.4637
4	rs7697323	7734317	1	2	269	278	0.9676	0.1481	0.7004
5	rs1457720	110998762	2	1	247	260	0.95	0.3333	0.5637
6	rs160666	2719051	2	1	248	262	0.9466	0.3843	0.5353
6	rs3799977	44945334	2	1	302	282	1.071	0.6849	0.4079
6	rs6921403	154105599	2	1	149	150	0.9933	0.003344	0.9539
8	rs6991017	5508780	2	1	193	191	1.01	0.01042	0.9187
9	rs2418326	116719295	1	2	236	210	1.124	1.516	0.2183
9	rs2416606	119862757	2	1	264	262	1.008	0.007605	0.9305
10	rs16928529	72652991	2	1	277	312	0.8878	2.08	0.1493
10	rs11594082	72969259	1	2	126	138	0.913	0.5455	0.4602
10	rs10786284	98125495	0	1	0	0	NA	NA	NA
10	rs515910	105956394	2	1	300	272	1.103	1.371	0.2417
11	rs3893215	17721406	0	2	0	0	NA	NA	NA
11	rs10830468	87604834	0	2	0	0	NA	NA	NA
12	rs4964805	102716954	0	2	0	0	NA	NA	NA
13	rs7995215	93206507	1	2	279	317	0.8801	2.423	0.1196
14	rs2239627	22705999	0	2	0	0	NA	NA	NA
14	rs10483286	24273582	0	2	0	0	NA	NA	NA
16	rs10514604	83003885	0	2	0	0	NA	NA	NA
17	rs2440129	6847295	0	2	0	0	NA	NA	NA

Supplementary Table 7. ADHD Genotype GWAS of Glutamatergic Genes. The most significant SNP genotype association in each of the eight GRM gene regions. A) ADHD TDT CHOP Illumina 550k B) ADHD Case:Control CHOP Illumina 550k C) ADHD IMAGE Perlegen 600k.

A)

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P	Gene
11	rs4237549	88407924	2	1	31	61	0.5082	9.783	0.001762	GRM5
7	rs17864159	126444172	1	2	22	46	0.4783	8.471	0.003609	GRM8
6	rs3887555	34177040	1	2	208	161	1.292	5.986	0.01442	GRM4
7	rs6943762	86047914	2	1	69	99	0.697	5.357	0.02064	GRM3
3	rs7623055	7485891	1	2	151	193	0.7824	5.128	0.02354	GRM7
6	rs362839	146721428	2	1	125	161	0.7764	4.531	0.03328	GRM1
3	rs4687770	51730105	2	1	114	94	1.213	1.923	0.1655	GRM2
5	rs2078183	178357150	2	1	190	210	0.9048	1	0.3173	GRM6

B)

CHR	SNP	BP	A1	F_A	F_U	A2	OR	CHISQ	P	Gene
3	rs7623055	7485891	1	0.3582	0.4129	2	0.7936	15.48	8.35E-05	GRM7
11	rs1354411	88016449	2	0.03643	0.0566	1	0.6302	10.21	0.001396	GRM5
7	rs2283100	126643293	2	0.2281	0.193	1	1.235	9.527	0.002024	GRM8
6	rs1873250	34130718	2	0.2134	0.2455	1	0.8338	7.062	0.007873	GRM4
7	rs10952890	86193151	1	0.02753	0.03917	2	0.6945	4.782	0.02877	GRM3
5	rs2078183	178357150	2	0.4593	0.4897	1	0.8852	4.605	0.03189	GRM6
6	rs1983635	146707365	2	0.316	0.2917	1	1.122	3.515	0.06081	GRM1
3	rs4687592	51630896	1	0.03442	0.04041	2	0.8464	1.191	0.2752	GRM2

C)

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P	Gene
6	rs12206652	34173960	2	1	265	216	1.227	4.992	0.02547	GRM4
11	rs160195	87932621	2	1	302	253	1.194	4.326	0.03753	GRM5
7	rs11563486	126621501	1	2	130	162	0.8025	3.507	0.06112	GRM8
3	rs11717471	7599469	2	1	238	280	0.85	3.405	0.06498	GRM7
6	rs2300620	146745874	2	1	160	133	1.203	2.488	0.1147	GRM1
7	rs1468413	86271589	1	2	190	162	1.173	2.227	0.1356	GRM3
5	rs7725272	178338994	2	1	289	261	1.107	1.425	0.2325	GRM6
3	rs6445959	51747387	2	1	169	153	1.105	0.795	0.3726	GRM2

Supplementary Table 8. ADHD CNV Family Based Transmission Disequilibrium and *de novo* Statistical Tests.

A) Illumina CHOP Deletions Enriched for Inheritance

CNVR	Count SNPs	TDTDel	InhDel	*de novo* Del	ParDel NotInh	Gene	Distance
chr18:74258734-74260996	3	0.001953	9	0	0	*SALL3*	580267
chr7:120092385-120099982	3	0.001953	9	0	0	*KCND2*	0
chr4:92499956-92502794	8	0.001953	9	0	0	*KIAA1680*	0
chr11:69755529-69759313	12	0.007813	7	0	0	*FADD*	24395
chr4:42400885-42403451	15	0.007813	7	0	0	*ATP8A1*	47238
chr5:104463047-104518786	17	0.007813	7	0	0	*NR_000039*	0
chr13:69637654-69666685	18	0.015625	6	0	0	*NR_002717*	25969
chr3:195971510-195982215	5	0.03125	5	1	0	*FAM43A*	80455
chr19:44369918-44376749	3	0.03125	5	1	0	*LOC342897*	2695
chr1:2349841-2356176	4	0.03125	5	1	0	*PEX10*	15971
chr21:45777720-45782727	3	0.03125	5	0	0	*SLC19A1*	0
chr10:67748487-67785209	30	0.03125	5	0	0	*CTNNA3*	0

B) Illumina CHOP Duplications Enriched for Inheritance

CNVR	Count SNPs	TDTDup	InhDup	*de novo* Dup	ParDup NotInh	Gene	Distance
chr20:59015708-59022667	4	0.007813	7	0	0	*CDH4*	238287
chr12:72808323-72832667	5	0.015625	6	0	0	*BC061638*	0
chr6:73021641-73023171	3	0.03125	5	0	0	*RIMS1*	0
chr17:74089903-74106726	9	0.03125	5	0	0	*DNAHL1*	10904
chr1:9243828-9310031	22	0.03125	5	0	0	*H6PD,SPSB1*	0

C) Illumina CHOP Deletions Enriched for *de novo*

CNVR	Count	*de novo*	InhDel	*de novo*	ParDel	Gene	Distance

	SNPs	TDTDel		Del	NotInh		
chr16:87694595-87778383	16	3.02E-05	32	2	21	*AX748415,CDH15,LOC197322*	0
chr18:65358832-65367619	18	3.02E-05	33	2	21	*DOK6*	0
chr12:55902280-55923860	3	0.000367	9	3	19	*NDUFA4L2,NXPH4,SHMT2,STAC3*	0
chr17:71112486-71120734	4	0.001848	12	3	16	*KIAA1783*	0
chr22:38384374-38403731	8	0.018158	4	4	13	*CACNA1I*	0
chr19:15992679-15997923	2	0.025875	15	6	15	*LOC126536*	0

D) Illumina CHOP Duplications Enriched for *de novo*

CNVR	Count SNPs	*de novo* TDTDup	InhDup	*de novo* Dup	ParDup NotInh	Gene	Distance
chr19:59423491-59428132	12	4.85E-09	74	3	38	*LILRB3,LIR-3*	0
chr8:145217675-145247517	4	3.05E-05	19	0	15	*CYC1,MAF1,SHARPIN,hSIPL1A*	0
chr18:64897188-64906488	48	0.000122	9	0	13	*CCDC102B*	23782
chr14:104225150-104339273	35	0.00293	7	1	11	*ADSS,ADSSL1,AKT1,SIVA1*	0
chr9:138606913-138647195	17	0.005371	10	1	10	*AF161442*	15688
chr16:650256-2028586	41	0.015625	8	0	6	*Many*	0
chr20:61642713-61668792	11	0.03125	4	1	7	*C20orf195,PRIC285,SRMS*	0
chr16:87399730-87430019	22	0.03125	7	1	7	*APRT,CDT1,FLJ00319,GALNS*	0
chr16:3553005-3590430	20	0.03125	8	0	5	*BTBD12,NLRC3*	0
chr22:17257787-17355587	60	0.03125	3	0	5	*DGCR6,KIAA1647,PRODH*	0

E) Perlegen IMAGE Deletions Enriched for Inheritance

CNVR	Count SNPs	TDTDel	InhDel	*de novo* Del	ParDel NotInh	Gene	Distance
chr2:180271795-180274556	5	0.003204	2	1	13	*ZNF533*	0
chr14:79919894-79924934	5	0.03125	1	0	7	*BC039670*	0
chr7:19828746-	7	0.041656	4	0	11	*MGC42090*	49005

19840916							

F) Perlegen IMAGE Duplications Enriched for Inheritance

CNVR	Count SNPs	TDTDup	InhDup	*de novo* Dup	ParDup NotInh	Gene	Distance
chr22:17361563-17369020	3	0.015625	6	0	0	*CR623368, KIAA1647*	0
chr15:30088094-30090949	3	0.03125	5	1	0	*CHRNA7*	19069
chr7:71664963-71712086	5	0.03125	5	0	0	*MGC87315*	0

G) Perlegen IMAGE Deletions Enriched for *de novo*

CNVR	Count SNPs	Denovo TDTDel	InhDel	*de novo* Del	ParDel NotInh	Gene	Distance
chr2:180271795-180274923	6	0.000854	2	1	13	*ZNF533*	0
chr10:85445139-85446804	7	0.03125	5	1	7	*GHITM*	442361

H) Perlegen IMAGE Duplications Enriched for *de novo*

CNVR	Count SNPs	Denovo TDTDup	InhDup	*de novo* Dup	ParDup NotInh	Gene	Distance
chr12:31276361-31285014	9	6.87E-05	15	1	17	*OVOS2*	26006
chr10:47089854-47154881	31	6.87E-05	11	1	17	*AK057316*	0
chr7:140018-162903	13	0.005371	10	1	10	*AL137655*	23529
chr8:2437197-2492653	23	0.03125	4	1	7	*BC045738*	0
chr6:168234697-168295618	13	0.043945	5	2	8	*FLJ00181*	9639

Supplementary Table 9. ADHD CNV Family Based Transmission Disequilibrium and *de novo* Statistical Tests.

CNVR (hg18/B36/ Mar2006)	Type	TDTDel	TDTDup	*de novo* TDTDel	*de novo* TDTDup	InhDel	*de novo* Del	ParDel NotInh	InhDup	*de novo* Dup	ParDup NotInh
chr7:126441593-126621501	Del	1	1	1	1	0	0	0	0	0	0
chr11:88269449-88351661	Del	0.125	1	0	1	3	0	0	0	0	0
chr3:7183953-7197236	Del	0.25	1	1	1	2	0	0	0	0	0
chr6:146657076-146694047	Dup	1	1	1	1	0	0	0	0	0	0
chr7:153495598-153564827	Dup	0.205	1	0.016	1	4	0	6	0	0	0
chr5:65027976-65046520	Del	1	0.5	1	1	0	0	0	1	0	0
chr1:56053497-56064495	Del	1	1	1	1	0	0	0	0	0	0
chr1:72317292-72328395	Dup	1	1	1	1	0	0	0	0	0	0
chr19:38427720-38444834	Del	0.183	1	0.004	1	6	0	8	0	0	0
chr3:1844168-1859889	Del	0.063	1	0	1	4	0	0	0	0	0
chr2:81419297-81446082	Dup	1	0.5	1	1	0	0	0	1	0	0
chr4:113772340-113788584	Dup	0.375	1	0.5	1	2	0	1	0	0	0

Supplementary Table 10. Perlegen Data Reformatted File Samples to match Affymetrix Power Tools output format.

A) Genotype Calls File (0=AA,1=AB,2=BB,-1=NoCall).

probeset_id	10009	10010	10021	10022
SNP_rs10000023	1	1	2	1
SNP_rs10000030	1	0	0	1
SNP_rs10000037	0	0	1	1
SNP_rs10000068	2	2	2	2

B) Genotype Calls Confidence Scores (All set to 1).

probeset_id	10009	10010	10021	10022
SNP_rs10000023	1	1	1	1
SNP_rs10000030	1	1	1	1
SNP_rs10000037	1	1	1	1
SNP_rs10000068	1	1	1	1

C) Intensity Summary (-A=log10(X), -B=log10(Y) (X and Y value from dbGaP Single Sample Final Report files).

probeset_id	10009	10010	10021	10022
SNP_rs10000023-A	2.85	2.78	2.07	2.89
SNP_rs10000023-B	2.86	2.84	2.98	2.96
SNP_rs10000030-A	2.9	2.99	2.95	3.02
SNP_rs10000030-B	2.91	2.4	2.38	3.05

Supplementary Table 11. Sample exclusion based on quality control measures.

Exclusion Criteria	CHOP	Control
Call Rate < 98%	170	271
SD LRR > 0.35	73	124
Ethnicity non-Caucasian	71	48
Wave Factor -0.5>X>0.6	251	1040
Count CNVs > 70	197	237
Monozygotic Twin	31	38

Samples excluded based on Quality Control (QC) measures on our HumanHap550 GWAS data based on statistical distributions to exclude poor quality DNA samples and false positive CNVs.

Supplementary Table 12. Sample Source Contributions to Impacting CNV Loci.

CNVR	CHOP Cases	CHOP Controls	NIMH cases	Utah cases	IMAGE cases	Per Psoriasis Control	Per Depression Control	PUW Ma Cases	PUW Ma Parents	IMAGE II Cases	IMAGE II Controls	SAGE Illumina 1M Controls	AGRE Affy 5.0 Parents Controls	Type	Gene
chr11:88269449-88351661	4	0	0	0	5	0	0	1	1	0	0	0	0	Del	*GRM5*
chr7:126441593-126621501	3	0	0	0	3	0	0	2	0	0	0	0	0	Del	*GRM8*
chr3:7183953-7197236	4	0	0	0	2	0	0	0	0	0	0	0	0	Del	*GRM7*
chr6:146657076-146694047	5	2	1	1	0	0	0	0	0	1	0	0	0	Dup	*GRM1*
chr1:72317292-72328395	4	0	0	0	0	0	0	1	0	0	0	0	0	Dup	*NEGR1*
chr7:153495598-153564827	5	0	1	0	0	0	0	2	0	0	1	0	1	Dup	*DPP6*
chr5:65027976-65046520	4	0	0	0	1	0	0	0	0	1	1	0	0	Del	*SGTB/NLN*
chr1:56053497-56064495	2	0	0	0	3	0	0	0	0	1	0	0	2	Del	*USP24*
chr19:38427720-38444834	5	2	0	0	1	0	0	1	3	0	0	0	0	Del	*SLC7A10*
chr3:1844168-1859889	4	0	0	0	0	0	0	2	2 (inh)	1	1	4	1	Del	*CNTN4*
chr2:81419297-81446082	2	0	0	0	1	0	3	0	0	1	0	0	0	Dup	*CTNNA2*
chr4:113772340-113788584	2	0	0	0	1	0	0	1	1	0	0	1	1	Dup	*LARP7*