Title

DUF1220 copy number is linearly associated with increased cognitive function as measured by total IQ and mathematical aptitude scores

Journal

Human Genetics, 134(1)

ISSN

0340-6717

Authors

Davis, Jonathon M
Searles, Veronica B
Anderson, Nathan
et al.

Publication Date

2015

DOI

10.1007/s00439-014-1489-2

Peer reviewed

HHS Public Access

Author manuscript

Hum Genet. Author manuscript; available in PMC 2018 April 13.

Published in final edited form as:

Hum Genet. 2015 January ; 134(1): 67–75. doi:10.1007/s00439-014-1489-2.

DUF1220 copy number is linearly associated with increased cognitive function as measured by total IQ and mathematical aptitude scores

Jonathon M. Davis,
Department of Biochemistry and Molecular Genetics and Human Medical Genetics, Medical Scientist Training and Neuroscience Programs, University of Colorado School of Medicine, Anschutz Medical Campus, RC1-S, L18-10125, 12801 East 17th Ave, Mailstop 8101, P.O. Box 6511, Aurora, CO 80045, USA

Veronica B. Searles,
Department of Biochemistry and Molecular Genetics and Human Medical Genetics, Medical Scientist Training and Neuroscience Programs, University of Colorado School of Medicine, Anschutz Medical Campus, RC1-S, L18-10125, 12801 East 17th Ave, Mailstop 8101, P.O. Box 6511, Aurora, CO 80045, USA

Nathan Anderson,
Department of Biochemistry and Molecular Genetics and Human Medical Genetics, Medical Scientist Training and Neuroscience Programs, University of Colorado School of Medicine, Anschutz Medical Campus, RC1-S, L18-10125, 12801 East 17th Ave, Mailstop 8101, P.O. Box 6511, Aurora, CO 80045, USA

Jonathon Keeney,
Department of Biochemistry and Molecular Genetics and Human Medical Genetics, Medical Scientist Training and Neuroscience Programs, University of Colorado School of Medicine, Anschutz Medical Campus, RC1-S, L18-10125, 12801 East 17th Ave, Mailstop 8101, P.O. Box 6511, Aurora, CO 80045, USA

Armin Raznahan,
National Institute of Mental Health, Child Psychiatry Branch, 10 Center Drive, Building 10, Room 4C110, Bethesda, MD 20892, USA

L. John Horwood,
Christchurch Health and Development Study, Department of Psychological Medicine, University of Otago, Christchurch, PO Box 4345, Christchurch 8140, New Zealand

David M. Fergusson,
Christchurch Health and Development Study, Department of Psychological Medicine, University of Otago, Christchurch, PO Box 4345, Christchurch 8140, New Zealand

Martin A. Kennedy,

Correspondence to: James M. Sikela.

J. M. Davis and V. B. Searles contributed equally to this study.

Conflict of interest JMS is founder and shareholder of GATC Science, LLC, a biotech company focused on genomics.

Department of Pathology, Gene Structure and Function Laboratory, University of Otago, Christchurch, PO Box 4345, Christchurch 8140, New Zealand

Jay Giedd, and
National Institute of Mental Health, Child Psychiatry Branch, 10 Center Drive, Building 10, Room 4C110, Bethesda, MD 20892, USA

James M. Sikela
Department of Biochemistry and Molecular Genetics and Human Medical Genetics, Medical Scientist Training and Neuroscience Programs, University of Colorado School of Medicine, Anschutz Medical Campus, RC1-S, L18-10125, 12801 East 17th Ave, Mailstop 8101, P.O. Box 6511, Aurora, CO 80045, USA

Abstract

DUF1220 protein domains exhibit the greatest human lineage-specific copy number expansion of any protein-coding sequence in the genome, and variation in DUF1220 copy number has been linked to both brain size in humans and brain evolution among primates. Given these findings, we examined associations between DUF1220 subtypes CON1 and CON2 and cognitive aptitude. We identified a linear association between CON2 copy number and cognitive function in two independent populations of European descent. In North American males, an increase in CON2 copy number corresponded with an increase in WISC IQ (R^2 = 0.13, p = 0.02), which may be driven by males aged 6–11 (R^2 = 0.42, p = 0.003). We utilized ddPCR in a subset as a confirmatory measurement. This group had 26–33 copies of CON2 with a mean of 29, and each copy increase of CON2 was associated with a 3.3-point increase in WISC IQ (R^2 = 0.22, p = 0.045). In individuals from New Zealand, an increase in CON2 copy number was associated with an increase in math aptitude ability (R^2 = 0.10 p = 0.018). These were not confounded by brain size. To our knowledge, this is the first study to report a replicated association between copy number of a gene coding sequence and cognitive aptitude. Remarkably, dosage variations involving DUF1220 sequences have now been linked to human brain expansion, autism severity and cognitive aptitude, suggesting that such processes may be genetically and mechanistically inter-related. The findings presented here warrant expanded investigations in larger, well-characterized cohorts.

Introduction

Despite ample evidence that suggests cognitive function is a highly heritable trait (Devlin et al. 1997), few genetic association studies have been able to identify replicable markers associated with cognitive measures such as IQ (Payton 2009). Studies that have been undertaken have predominantly been genome-wide association studies (GWAS) and have suffered from a lack of reproducibility and small effect (Chabris et al. 2012; Rietveld et al. 2013). Recently, a large-scale study utilizing both training and test data sets (bolstering the reproducibility of the findings) identified two significant SNPs with a cognitive effect. Variation in these SNPs explained 2 % of the variation in cognitive function, and 0.02 % of the variation in educational attainment (Rietveld et al. 2013). Other recent findings have suggested a substantial portion of cognitive ability could be due to the additive effects of many SNPs of small individual effect sizes (Trzaskowski et al. 2014). These results together

suggest that single gene contributions to IQ are minimal, even though IQ exhibits a strong heritability component. While cognitive functioning likely results from a complex interplay of environmental, epigenetic and genetic factors, it is also possible that additional single gene variants that confer large cognitive effects reside in previously unexplored regions of the genome.

Perhaps the strongest biologic predictor of cognitive functioning is brain size, such that the sizes and connectivity of specific regions demonstrate replicable associations with cognitive function that vary by age and sex (Haier et al. 2005; Shaw et al. 2006; Jung and Haier 2007; Supekar et al. 2013). Complementing this is the fact that increased brain size is a feature of phylogenetic proximity to humans within the primate lineage, likely conferring an evolutionary advantage via increased cognitive functioning.

Considering the importance of regional brain size in cognitive aptitude and the lack of reproducible genetic candidates, novel genomic candidates for IQ would likely be sequences that have been implicated in brain size and/or brain evolution and sequences that have been previously unexamined in genetic studies of human cognition. DUF1220 protein domains are plausible candidates for influencing IQ under these guidelines, given (1) their association with both brain evolution and brain size (Popesco et al. 2006; Dumas et al. 2012; Keeney et al. 2014a), (2) that they are protein-coding sequences that exhibit extremely broad ranges in copy number (Davis, et al. 2014), and (3) that they have not been examined in previous genetic association studies of IQ including GWAS and copy number variation (CNV) scans.

Found within the *NBPF* gene family, DUF1220 sequences have undergone the most extreme human lineage-specific copy number expansion of any protein-coding region in the genome (Vandepoele et al. 2005; O'Bleness et al. 2012). Humans have approximately 290 haploid copies of DUF1220 that can be subdivided into 6 clades defined by sequence similarity (CON1-3 and HLS1-3) (O'Bleness et al. 2012, 2014). DUF1220 copy number (dosage) has been implicated in normal and pathological variation in human brain size and in neuron number across primate lineages (Dumas et al. 2012). DUF1220 copy number also follows a broad Gaussian distribution in the human population (Davis et al. 2014), suggesting that variation in DUF1220 could confer a wide range of phenotypic effects. Additionally, there is evidence suggesting that strong selection pressures have acted on DUF1220 sequences throughout human evolution. This evidence includes the association between DUF1220 copy number and the evolutionary expansion of the human brain (O'Bleness et al. 2012; Dumas et al. 2012), the rapidity with which DUF1220 copy number increased in the human genome and the signatures of positive selection found in DUF1220 coding regions (Popesco et al. 2006). These properties, together with our recent research implicating DUF1220 domains in driving neural stem cell proliferation (Keeney et al. 2014b) and in influencing symptom severity in Autism Spectrum Disorder (Davis et al. 2014), make DUF1220 domains attractive candidates for influencing cognitive functioning.

Based on this evidence, this study aimed to investigate the association between DUF1220 copy number variation and cognitive aptitude. We hypothesized that increased copy number of DUF1220 CON1 and CON2 subtypes (clades), specifically due to their associations with increased gray matter in healthy individuals (Dumas et al. 2012), would be associated with

increased cognitive function. To investigate this possibility we tested for associations between DUF1220 clade-specific copy number variation and cognitive aptitude in North Americans and replicated our analysis in New Zealanders. We also explored associations between specific brain region measurements and IQ in North Americans. Finally, we explored associations between CON1/CON2 copy number variation and specific brain region measurements based on our previous work demonstrating associations between these clades and total gray matter volume (Dumas et al. 2012). Our findings include a positive association between DUF1220 CON2 copy number and cognitive ability in two separate populations, and evidence for associations between specific brain region measurements and cognitive ability. Given these promising findings we call for expansion of this research in large well-characterized populations.

Methods

Ethics statement

All subjects were consenting participants in collaborating investigations, and all data were deidentified. The Colorado Multiple Institutional Review Board approved this research.

Populations

We utilized two separate populations consisting of individuals of European descent: (a) from North America and (b) from New Zealand. The North American population included unrelated non-Hispanic white individuals selected for brain size extremes through a procedure identifying residual values from a regression controlling for sex and age from a population of more than 600. This included 59 individuals (41 males and 18 females) whose ages ranged from 6 to 22. To determine reproducibility of our results, we gathered array comparative genomic hybridization (arrayCGH)-based copy number data on 75 individuals from New Zealand, 51 of whom were of European decent and selected for extremes in birth head circumference from the Christchurch Health and Development Study (Fergusson and Horwood 2001) (CHDS) cohort. The CHDS is a longitudinal study of a birth cohort of 1,265 children born in the Christchurch urban area in 1977. Participants have been examined on 23 occasions through age 35, with a focus on understanding psychosocial development, health and well-being. Cognitive measures were gathered in children aged 8–13 years in the CHDS. ArrayCGH values for CON1 and CON2 were sent to blinded statisticians to identify associations with cognitive function tests. Brain and head circumference size extremes were originally selected for association studies examining DUF1220, head circumference and brain size (Dumas et al. 2012). Although selected for size extremes, these populations were not selected for IQ metrics and IQ was normally distributed in these populations. Importantly, to account for potential confounding due to this selection strategy, statistical analysis utilized a brain size adjustment discussed further below.

Assay methods

ArrayCGH

ArrayCGH was performed as described in Dumas et al. (2012). Briefly, Oxford Gene Technology (OGT) performed arrayCGH utilizing custom-designed 1q21 arrays composed

of probes corresponding to both unique and non-unique sequences with ~75 % of the probes located on chromosome 1. The non-unique probes have sequences that are found at multiple loci in this region and the majority of them are located within the highly duplicated DUF1220-domain-encoding regions. The original design of the microarray probe sequences was based on the hg18 2006 assembly, which was converted to the hg19 build through BLAT alignment. The $\log_2$ ratios for the non-unique array probes were assigned to each newly mapped non-unique probe locus. The $\log_2$ ratios for probes that mapped to the same location were averaged. Dye bias was normalized within the arrays utilizing mean values from probes not aligning with chromosome 1, X or Y. Normalization was also conducted across arrays utilizing single locus probes. Throughout this report the $\log_2$ ratios are referred to as copy ratios, meaning the ratio of florescence of the test DNA to the common reference DNA. For example, a copy ratio of zero would indicate that both the test and reference DNA had the same amount of DUF1220 copies, and a copy ratio of one would indicate that the test has twice as many copies. A positive value indicated that the test DNA had more DUF1220 copies than the reference, and a negative value indicated the test sample had fewer copies.

Digital droplet polymerase chain reaction (ddPCR)

Digital droplet polymerase chain reaction (ddPCR), a third-generation PCR protocol (Hindson et al. 2011), was utilized in a subset of 17 individuals as a confirmatory measurement. Genomic DNA was first digested with a restriction enzyme then diluted to 2 ng/μl and added to a PCR mix containing: primers to the target sequence and to a reference sequence of known copy number, Bio-Rad droplet PCR master mix and fluorescent probes specific to the target and reference. Primer and probe sequences were as follows: CON1: Left—′AATGTGCCATCACTTGTTCAAATAG′, Right—′GACTTTGTCTTCCTCAAATGTGATTTT′, Hyb—′CATGGCCCTTATGACTCCAACCAGCC′; RPP30: Left—′GATTTGGACCTGCGAGCG′, Right—′G CGGCTGTCTCCACAAGT′, Hyb—′TTCTGACCTGAAG GCTCTGCGC′; CON2: Left—′CTGGCTCATCAGGA ATCTGC′, Right—′CGAATAACCTTCATCCCAGGAC′, Hyb—′AAGAGGAGGAAGTCCCCCAGGA′. Oil droplets were then generated using 20 μl PCR mix and 70 μl oil on the QX100 Droplet Generator. These droplets were subjected to a standard thermocycler protocol producing over 14,000 independent PCR reactions per sample. Copy number of the target sequence was then estimated by comparing the ratio of the target to the reference. Each sample was run in triplicate to confirm results and the copy estimates were then merged to produce a final copy number for each sample. A Pearson's correlation of 0.65 was identified between ddPCR and arrayCGH measurements suggesting moderate agreement. The imperfect correlation between measurements suggests a level of measurement error. To avoid bias and ensure random error, and thus underestimated associations, all assays were randomized and blinded. The ddPCR assay was found to be highly reproducible (Pearson's $r > 0.90$). Results derived from both ddPCR and arrayCGH are presented in Table 1.

Brain size

Brain region measurements from the North American cohort were estimated using MRI volumetric and surface area measurements and are described further in (Raznahan et al.

2011). Adult head circumference (age ~35 years) and birth head circumference was obtained from individuals participating in the New Zealand cohort.

Cognitive measurements

Wechsler Intelligence Scale for Children (WISC)—The WISC is a clinical test administered by psychologists to assess intelligence quotient in children (Weschler 1974). This study utilized Total WISC IQ, WISC Verbal IQ and WISC Performance IQ scores in the analysis.

Progressive Achievement Test (PAT)—The PAT is a set of seven multiple choice tests used by teachers to assess student achievement and understanding as defined by the New Zealand curriculum (Reid and Huges 1974; Elley and Reid 1969). PAT math and verbal scores were utilized for analysis in the study.

Test of Scholastic Abilities (TOSCA)—The TOSCA is a group test of student scholastic ability that is administered to students by teachers in a group setting, with answers recorded individually (Reid et al. 1981). Total scholastic ability was utilized for analysis in this study.

Statistical methods

This study utilized two independent analyses to test for dosage associations between DUF1220 CON1/CON2 subtypes and cognitive aptitude. First, the association between DUF1220 CON1/CON2 and WISC IQ was explored in the North American population. Second, findings from the North American population were explored in the New Zealand population. Linear regression was used to test associations among CON1, CON2 and cognitive aptitude while exploring adjustments for sex and age. Due to previous reports of relationships between brain regions and IQ that differed by sex, we hypothesized the interaction of CON1/2 by sex would be important. When the interaction was significant the analysis was in turn stratified by sex, and we present results from the North American population in a reduced univariate model stratified by sex. We followed this analysis with an adjustment for the brain region most strongly predictive of IQ, bilateral temporal surface area (BTSA) (Table 1), and an adjustment for right frontal surface area (RFSA), the brain region exhibiting the strongest association with CON1 and CON2 (Supplemental Tables 2 and 3). We also investigated adjustments for BTSA, RFSA and total gray matter volume simultaneously. As there were only two primary comparisons (CON1 vs. IQ and CON2 vs. IQ) and a replication analysis in an independently selected and assayed population, results from this step of the analysis are presented without correction for multiple comparisons. Finally, due to reports that IQ profiles change with brain development over time (Shaw et al. 2006), and the differences between sexes mentioned above, we followed CON2 analyses by exploring higher level interaction terms of CON2 × sex × age. Analyses were subsequently stratified due to the significant interaction term.

Importantly, replication of the association between CON2 and cognitive ability was explored in a separate cohort of individuals from New Zealand. The analysis in the New Zealand cohort explored adjustments for adult head circumference (as well as age and sex) as

specific MRI data were not available. The associations tested in the New Zealand group are presented univariately from a reduced model. These are corrected for multiple comparisons with the Bonferroni method due to the additional comparisons analyzed (six outcomes tested with one primary exposure: CON2 vs. 6 cognitive assessments).

We also explored associations between MRI-derived brain region size measurements and IQ in the North American population. These were tested with linear regression. Despite the selection strategy, only total gray matter volume was bimodal. The other regions examined did not depart substantively from a normal distribution and residual diagnostics did not suggest deviations from model assumptions. This analysis included 54 comparisons and findings with significant or trending p values ($p < 0.01$) are displayed, and those meeting Bonferroni criteria ($p < 0.0009$) are presented in bold in Table 1.

This project also sought to expand on our previous findings that CON1 and CON2 are associated with total gray matter volume in healthy individuals (Dumas et al. 2012), and to identify regions that may be important confounders due to the selection strategy. We explored associations between both CON1 and CON2 copy number and specific brain regions in this same population to further identify regions that may be uniquely associated with CON1 and CON2 variation. As these analyses are a follow-up study to previous findings (Dumas et al. 2012) uncorrected p values are presented along with the Bonferroni critical p value (Supplement Tables 2 and 3).

Results

DUF1220 CON1 and CON2 copy number association with IQ

Age-appropriate full-scale WISC IQ was modeled in a multivariate linear regression with CON2 copy number as the primary exposure variable including age and sex as covariates in the North American group. The interaction of CON2 × sex was significant ($p = 0.038$), suggesting a more pronounced effect in males. After model reduction and stratification, and using two independent assay methods (arrayCGH and ddPCR), increasing CON2 dosage (e.g.; copy number) was significantly associated with increasing IQ in males (arrayCGH-based $R^2 = 0.13$, $p = 0.02$ ddPCR-based $R^2 = 0.22$, $p = 0.045$) (Table 2), but not in females ($p = 0.54$, data not shown). CON1 copy number, meanwhile, was not associated with IQ ($p = 0.18$). In the subset of males with the ddPCR assay ($N = 17$) CON2 ranged from 26 to 33 copies with a mean of 29. In this group each additional copy of CON2 was associated with an IQ increase of 3.3 points (SE = 1.52, $p = 0.045$).

Brain size associations with IQ in North American individuals of European descent

Each region measured by MRI was tested for volumetric or surface area association with IQ. The volumes and surface areas of a number of gray matter regions were associated with total age-appropriate WISC IQ. Notably, increased bilateral temporal surface area was associated with progressively increased WISC IQ ($p = 0.0006$, Bonferroni critical p value = 0.0009) (Fig. 1).

DUF1220 CON1 and CON2 copy number association with brain region size in North American individuals of European descent

Post-hoc analysis of MRI-measured volumes or surface areas identified eight brain regions marginally associated with CON1 copy number ($p < 0.05$, Supplemental Table 2). Two regions were marginally associated with CON2 ($p < 0.05$, Supplemental Table 3). In each, increased size corresponded with increased CON1 and CON2 dosage. Perhaps most interestingly, right frontal surface area displayed the strongest association of the regions tested with both CON1 and CON2 copy number. For each array-CGH-based 0.01 copy ratio increase of CON1, right frontal surface area increased 130 mm^2 ($p = 0.03$). For each array-CGH-based 0.01 copy ratio increase of CON2, right frontal surface area increased 150 mm^2 ($p = 0.04$ Supplemental Table 1). There were no CON1 or CON2 associations with white matter regions or with gyrification index.

Multivariate analysis including adjustment for size of brain regions and exploration of higher order interactions with CON2

Due to the selection strategy employed for this group, which utilized individuals with brain size extremes, the results may be overestimates. This is plausible given that the R^2 value of 0.13 (detected in North American males with the arrayCGH assay) in this study is substantially larger than previously reported in a large GWAS examining educational attainment and cognitive function (Rietveld et al. 2013). To address this, and to examine effects of CON2 copy number that may be independent of brain size, we employed models adjusting for brain size (and age, as age is an important component of brain size). CON2 remained significant in both models ($p < 0.05$) and the adjusted R^2 increased proportionately, suggesting little overlap of the variance explained by BTSA and CON2 (CON2 $R^2 = 0.13$, BTSA adjusted $R^2 = 0.14$, combined adjusted $R^2 = 0.22$). The beta estimate of CON2, BTSA adjusted, did decrease 16.2 %, however, suggesting a complex etiologic mechanism between CON2 dosage, brain region size and IQ. Adjusting for RFSA yielded a less pronounced change in beta, with a CON2 beta decrease of 10.3 %. RFSA also was not a significant predictor in this model. These effects suggest that the CON2 IQ association was not substantially confounded by RFSA nor does CON2 share a similar complex brain volume IQ relationship with RFSA as it may with BTSA. Simultaneous adjustments of BTSA, RFSA and total gray matter volume yielded similar results. CON2 remained significant ($p = 0.025$) and the beta estimate changed minimally (<10 %).

We further investigated higher level interactions of CON2 dosage with sex and age on IQ in the North American cohort due to findings suggesting that developing regions of the brain are important when examining IQ over time (Shaw et al. 2006). The three-level interaction of CON2 × age × sex was highly significant ($p = 0.0004$) suggesting that these relationships follow a complex pattern affected by sex and age, such that with increasing age the correlation between CON2 and IQ in males appears to decrease. For ease of interpretation, follow-up analysis was stratified by sex and age, and age was categorized as young or old based on the mean age of the sample. Further, the mean age of this sample (10.86) corresponds with findings that suggested that cortex thickening in the brightest children peaked by 11 or 12 (Shaw et al. 2006). In males younger or equal to the mean age ($n = 19$), increasing CON2 dosage was strongly associated with increasing WISC IQ ($R^2 = 0.42$, $p =$

0.003, Fig. 2). For each 0.01 copy ratio increase in arrayCGH measured CON2, IQ increased on average 1.91 points (SE = 0.54). CON2 remained strongly significant in this group after adjustment for BTSA and RFSA. This association was not detected in older males (p = 0.96).

DUF1220 CON2 with IQ: replication of association in individuals from New Zealand

As described, to test whether the association between CON2 copy number and IQ could be replicated in an independent cohort, we followed the North American analysis with an exploration of a second cohort of 51 individuals selected for birth brain size extremes (as determined by birth head circumference) but not for IQ metrics or adult brain size. A significant association was identified in this cohort between CON2 dosage and Progressive Achievement Mathematics (a measure of mathematical aptitude; R^2 = 0.10, Bonferroni corrected p = 0.018, Table 2) and a positive trend was seen between CON2 copy number and WISC verbal IQ R^2 = 0.05, Bonferroni corrected $p < 0.098$) (Table 2). Birth head circumference was not significant and did not affect the association of CON2 dosage with Progressive Achievement Mathematics. The interaction CON2 × sex was not significant.

Discussion

In pursuit of further elucidating the function of DUF1220, a protein domain implicated in human brain evolution, autism severity, and disorders of brain size, this study examined associations among copy number variations in subtypes of DUF1220 and cognitive aptitude. Results demonstrate a linear association between CON2 clade copy number and IQ, and an association between specific regional brain surface areas and IQ.

Our results demonstrate that increased CON2 dosage (copy number) is associated with increased WISC IQ and increased scores in Progressive Achievement Mathematics, with a more pronounced association exhibited in males. To our knowledge, the magnitude of this association (R^2 = 0.42 in young males) is greater than that reported in any previous study of a single candidate genetic or genomic predictor of IQ. While further research is needed to definitively confirm this association in a larger, more generalizable sample, this is an important finding, particularly given that the variation in IQ explained by CON2 is substantially higher than recently reported (Rietveld et al. 2013). Additionally, the linear nature of this relationship, with increased copies of CON2 correlating with increased IQ and math aptitude scores, suggests an underlying mechanism related to DUF1220 protein domain dosage. Of note, controlling for brain volumes and sizes did not eliminate the CON2-IQ association, implying there is an additional mechanism through which CON2 copy number may influence cognitive ability. Further, the strong association detected in younger males that explained an exceptional amount of variance suggests a critical time period where the effect of CON2 dosage is important. This is a small sample, however, and further investigations in large cohorts will be necessary to confirm these effects. Previous evidence does, though, suggest that cognitive performance may be related to different brain regions depending on sex and age (Haier et al. 2005). The evidence we present suggests a complex interplay of cognitive performance with brain growth, DUF1220 CON2 copy number, sex and age, and warrants expanded investigations.

Analyses of MRI-measured brain volumes and surface areas demonstrate that left temporal surface area and bilateral temporal surface area are associated with IQ. The observed associations between IQ test performance and MRI measures might be explained by different neuron number profiles within these brain regions that in turn effect size differences. Regions of the genome implicated in controlling neuron number, such as sequences encoding DUF1220 protein domains, are therefore plausible candidates for generating intellectual diversity such as the differences observed on the WISC IQ test performance.

We previously reported that the copy number of CON1 and CON2 was linearly associated with total gray matter volume (Dumas et al. 2012). To further explore this relationship we examined specific brain regions to clarify the potential role of CON1 and CON2 in brain size. While not significant after correction for multiple comparisons, these results suggest specific roles for DUF1220 dosage in region-specific brain size. The link with surface area also suggests a role for DUF1220 in neural stem cell proliferation, given that such proliferation is thought to be the major contributor to surface area (Rakic 2000). This possible link is supported by recent research in cell culture models (Keeney et al. 2014b). The link with cortical thickness, though, suggests that an additional mechanism may be involved through which DUF1220 may mediate changes in brain size and cognition. Further research is necessary to determine what molecular pathways may be involved, but given the association between DUF1220 copy number expansion and increased neuron number in primates (Dumas et al. 2012; Keeney et al. 2014b), a plausible mechanism could involve additional rounds of neurogenic cell divisions per cortical layer.

A mechanism of differential brain size involving differences in neuron number underlying IQ test performance would also be consistent with the features of primate brain expansion, which are that an increase in cerebral size within this lineage is almost linearly proportional to neuron number (unlike any other lineages surveyed thus far) (Herculano-Houzel 2009). These scaling rules indicate that an increase in brain size within primates is almost exclusively due to the addition of more neurons, rather than a change in neuronal size or density. Consistent with this, DUF1220 copy number shows a dramatic increase that is unique to the primate order (O'Bleness et al. 2012) and is strongly correlated with brain neuron number among primate species (Dumas et al. 2012; Keeney et al. 2014b). In addition, as referenced above it has been recently shown that DUF1220 domains can function as drivers of proliferation in cultured neural stem cells (Keeney et al. 2014b). These findings are consistent with the possibility that increases in DUF1220 dosage may be resulting in the production of more neurons, providing a plausible explanation for the association between DUF1220 copy number and brain size and, in turn, a component of the observed association with cognitive performance (Keeney et al. 2014a).

The potential link between DUF1220 subtypes and regional surface areas, thickness and volumes also suggests possible mechanisms by which DUF1220 may influence disease phenotypes. Recent research examining DUF1220 in individuals with ASD has shown that copy number increases of DUF1220 correlate with increasing severity of the primary symptoms of ASD in a similar dose response fashion as presented in this report with IQ (Davis et al. 2014). Such findings suggest that increased DUF1220 copies are beneficial in

otherwise healthy individuals but deleterious in individuals with ASD. Given the suggested link between DUF1220 and neuron number, a similar mechanism could be an important factor in individuals with ASD as individuals with ASD often exhibit increased brain size and neuron number (Courchesne et al. 2011). Investigations that aim to elucidate the mechanism behind the link between increased brain size/neuron number and autism symptomology have the potential to explain a large portion of the condition.

This study has limitations that need to be considered while interpreting the results. The selection strategy that utilized brain size extremes may have led to an overestimate of the strength of association between CON2 and IQ. However, adjusting for brain volume in a multivariate model suggests this estimate is not inflated substantially. The adjustment did decrease the beta estimate of CON2, which suggests a complex mechanism of aptitude etiology in which CON2-brain volume relationships partly account for cognitive ability, but are complemented by an independent mechanism of CON2 dosage effects on IQ.

The associations presented appear to be driven by younger males and may highlight the importance of CON2 and brain growth in this narrow age range. Conversely, however, a recent report suggested that the same genes contribute to cognitive ability from childhood into adolescence with increased heritability over time (Trzaskowski et al. 2014). This report examined the additive contributions of thousands of SNPs of low individual effect related to cognitive function, while omitting DUF1220 domains, and found a large cumulative effect such that these SNPs together explained 45 % of the inter-individual variation in cognitive ability. This is a notable contrast to the findings presented in this report, where we suggest the dosage of a specific protein domain subtype, DUF1220 CON2, may be highly important in young males with a similarly important explanatory effect ($R^2 = 0.42$). Taken together, these findings suggest that the genetic underpinnings of cognitive ability are complex and necessitate examinations of highly duplicated copy number polymorphic sequences as well as more traditional SNP-based approaches.

Future investigations of DUF1220 CON2 in larger, well-characterized populations with longitudinal brain growth and longitudinal IQ measurements will aid in determining the exact mechanisms mediating the associations among CON2 dosage, regional brain size, and cognitive aptitude. Meanwhile, measurement of higher copy number DUF1220 clades such as HLS1-3 has been challenging given the technologies currently available. It is noteworthy that these are the DUF1220 clades that show the most dramatic human-specific increase in copy number (O'Bleness et al. 2012). As more precise techniques become available, future research should also explore the relationship of these higher copy number clades with cognitive features.

It should also be noted that the cognitive associations identified were not perfectly replicated between cohorts. This discrepancy could be due to a number of factors, including regional differences in measurement methods and interpretations of performance. Given the heterogeneity between populations presented in this report it is not surprising that there would be subtle differences in associations between CON2 and cognitive ability. Further, the age range in the New Zealand cohort when IQ metrics were made is slightly older (8–13) than the group showing the strongest association in the North American sample (6–11).

Although a subtle difference, this may further suggest a critical age range where CON2 dosage is important in cognitive functioning. The fact that such an association existed in both groups, though, is a substantial strength of this report. It is also worth noting that the full New Zealand cohort of over 600 samples had previously been analyzed by SNP-arrays for CNVs that could be genetic predictors of IQ and no loci were identified as significant (Bagshaw et al. 2013). Importantly, DUF1220 was not among the loci examined. This is consistent with our previous point that conventional genome-wide CNV studies, both SNP-based and arrayCGH-based, typically overlook highly duplicated, though potentially important regions of the genome.

The findings reported here, together with previous studies, suggest that increases in copy number of sequences from the same protein domain family (DUF1220) are involved in promoting human brain expansion, increasing autism severity and improving cognitive capacity. Such intriguing associations call for additional investigations that include assessment of the role of other DUF1220 clades in these conditions, more precise determination of how DUF1220 copy number is being altered at the genome level, and examination of DUF1220 copy number in large, well-characterized populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding for this work was provided by NIH R01 MH081203 (JMS), Colorado Clinical and Translational Science Institute TL1 TR001081 (VBS), and a Graduate Assistantship from the Coleman Institute for Cognitive Disabilities (JK). CHDS was funded by Grants from the Health Research Council of New Zealand, the National Child Health Research Foundation, the Canterbury Medical Research Foundation, the New Zealand Lottery Grants Board, the Marsden Fund, and the James Hume Bequest Fund. We thank Allison Miller for preparation of DNA samples.

References

Bagshaw ATM, Horwood LJ, Liu Y, et al. No effect of genome-wide copy number variation on measures of intelligence in a New Zealand birth cohort. PLoS One. 2013; 8:e55208.doi: 10.1371/journal.pone.0055208 [PubMed: 23383111]

Chabris CF, Hebert BM, Benjamin DJ, et al. Most reported genetic associations with general intelligence are probably false positives. Psychol Sci. 2012; 23:1314–1323. DOI: 10.1177/0956797611435528 [PubMed: 23012269]

Courchesne E, Mouton PR, Calhoun ME, et al. Neuron number and size in prefrontal cortex of children with autism. JAMA. 2011; 306:2001–2010. DOI: 10.1001/jama.2011.1638 [PubMed: 22068992]

Davis JM, Searles VB, Anderson N, et al. DUF1220 dosage is linearly associated with increasing severity of the three primary symptoms of autism. PLoS Genet. 2014; 10:e1004241.doi: 10.1371/journal.pgen.1004241 [PubMed: 24651471]

Devlin B, Daniels M, Roeder K. The heritability of IQ. Nature. 1997; 388:468–471. DOI: 10.1038/41319 [PubMed: 9242404]

Dumas LJ, O'Bleness MS, Davis JM, et al. DUF1220-domain copy number implicated in human brain-size pathology and evolution. Am J Hum Genet. 2012; 91:444–454. DOI: 10.1016/j.ajhg.2012.07.016 [PubMed: 22901949]

Elley, WB., Reid, NA. Progressive achievement tests: teacher manual: reading comprehension, reading vocabulary. New Zealand Council for Educational Research; Wellington: 1969.

Fergusson DM, Horwood LJ. The Christchurch Health and Development Study: review of findings on child and adolescent mental health. Aust NZ J Psychiatry. 2001; 35:287–296.

Haier RJ, Jung RE, Yeo RA, et al. The neuroanatomy of general intelligence: sex matters. NeuroImage. 2005; 25:320–327. DOI: 10.1016/j.neuroimage.2004.11.019 [PubMed: 15734366]

Herculano-Houzel S. The human brain in numbers: a linearly scaled-up primate brain. Front Hum Neurosci. 2009; doi: 10.3389/neuro.09.031.2009

Hindson BJ, Ness KD, Masquelier DA, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem. 2011; 83:8604–8610. DOI: 10.1021/ac202028g [PubMed: 22035192]

Jung RE, Haier RJ. The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. Behav Brain Sci. 2007; 30:135–154. DOI: 10.1017/S0140525X07001185 [PubMed: 17655784]

Keeney JG, Dumas L, Sikela JM. The case for DUF1220 domain dosage as a primary contributor to anthropoid brain expansion. Front Hum Neurosci. 2014a; 8:427.doi: 10.3389/fnhum.2014.00427 [PubMed: 25009482]

Keeney JG, Davis JM, Siegenthaler J, et al. DUF1220 protein domains drive proliferation in human neural stem cells and are associated with increased cortical volume in anthropoid primates. Brain Struct Funct. 2014b; doi: 10.1007/s00429-014-0814-9

O'Bleness M, Searles VB, Dickens CM, et al. Finished sequence and assembly of the DUF1220-rich 1q21 region using a haploid human genome. BMC Genom. 2014; 15:387.doi: 10.1186/1471-2164-15-387

O'Bleness MS, Dickens CM, Dumas LJ, et al. Evolutionary history and genome organization of DUF1220 protein domains. G3 (Bethesda). 2012; 2:977–986. DOI: 10.1534/g3.112.003061 [PubMed: 22973535]

Payton A. The Impact of Genetic Research on our Understanding of Normal Cognitive Ageing: 1995 to 2009. Neuropsychol Rev. 2009; 19:451–477. DOI: 10.1007/s11065-009-9116-z [PubMed: 19768548]

Popesco MC, Maclaren EJ, Hopkins J, et al. Human lineage-specific amplification, selection, and neuronal expression of DUF1220 domains. Science. 2006; 313:1304–1307. DOI: 10.1126/science.1127980 [PubMed: 16946073]

Rakic P. Radial unit hypothesis of neocortical expansion. Novartis Found Symp. 2000; 228:30–42. (discussion 42–52). [PubMed: 10929315]

Raznahan A, Shaw P, Lalonde F, et al. How does your cortex grow? J Neurosci. 2011; 31(19):7174–7177. DOI: 10.1523/JNEUROSCI.0054-11.2011 [PubMed: 21562281]

Reid, NA., Huges, DC. Progressive Achievement Tests: Teachers Manual, Mathamatics. New Zeland Council for Educational Research; Wellington: 1974.

Reid, NA., Jackson, PF., Gilmore, A., Croft, C. Test of Scholastic Abilities. New Zeland Council for Educational Research; Wellington: 1981.

Rietveld CA, Medland SE, Derringer J, et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science. 2013; 340:1467–1471. DOI: 10.1126/science.1235488 [PubMed: 23722424]

Shaw P, Greenstein D, Lerch J, et al. Intellectual ability and cortical development in children and adolescents. Nature. 2006; 440:676–679. DOI: 10.1038/nature04513 [PubMed: 16572172]

Supekar K, Swigart AG, Tenison C, et al. Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proc Natl Acad Sci. 2013; 110:8230–8235. DOI: 10.1073/pnas.1222154110 [PubMed: 23630286]

Trzaskowski M, Yang J, Visscher PM, Plomin R. DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7–12. Mol Psychiatry. 2014; 19:380–384. DOI: 10.1038/mp.2012.191 [PubMed: 23358157]

Vandepoele K, Van Roy N, Staes K, et al. A novel gene family NBPF: intricate structure generated by gene duplications during primate evolution. Mol Biol Evol. 2005; 22:2265–2274. DOI: 10.1093/molbev/msi222 [PubMed: 16079250]

Weschler, D. Manual for the Wechsler Intelligence Scale for Children - Revised. Psychological Corporation; New York: 1974.

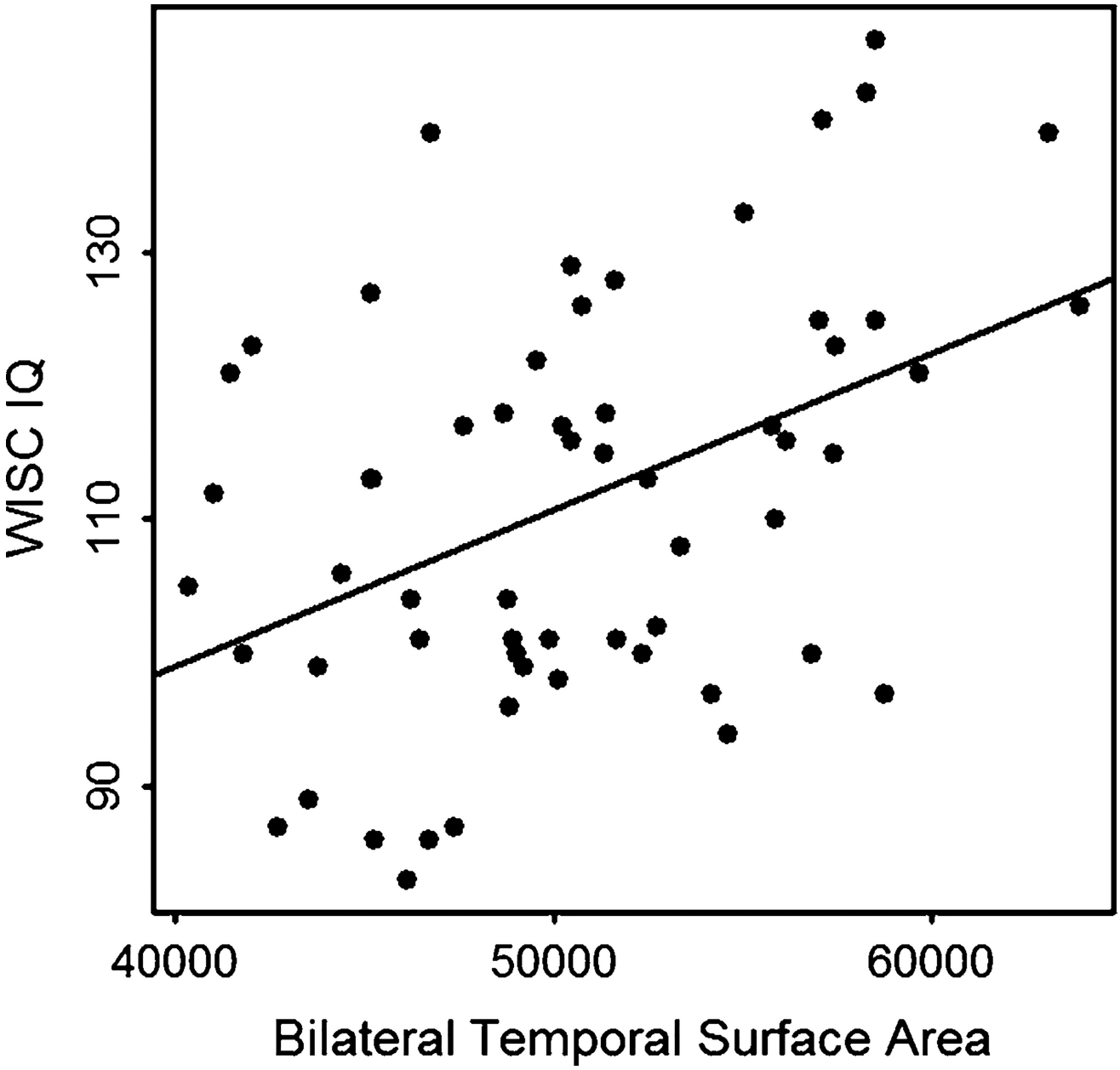

Fig. 1.

Linear association of bilateral temporal surface area (BTSA) with IQ in North Americans. MRI-based surface area versus age-appropriate total WISC IQ. Increasing BTSA was associated with increasing IQ ($R^2 = 0.19$, $p = 0.0006$)

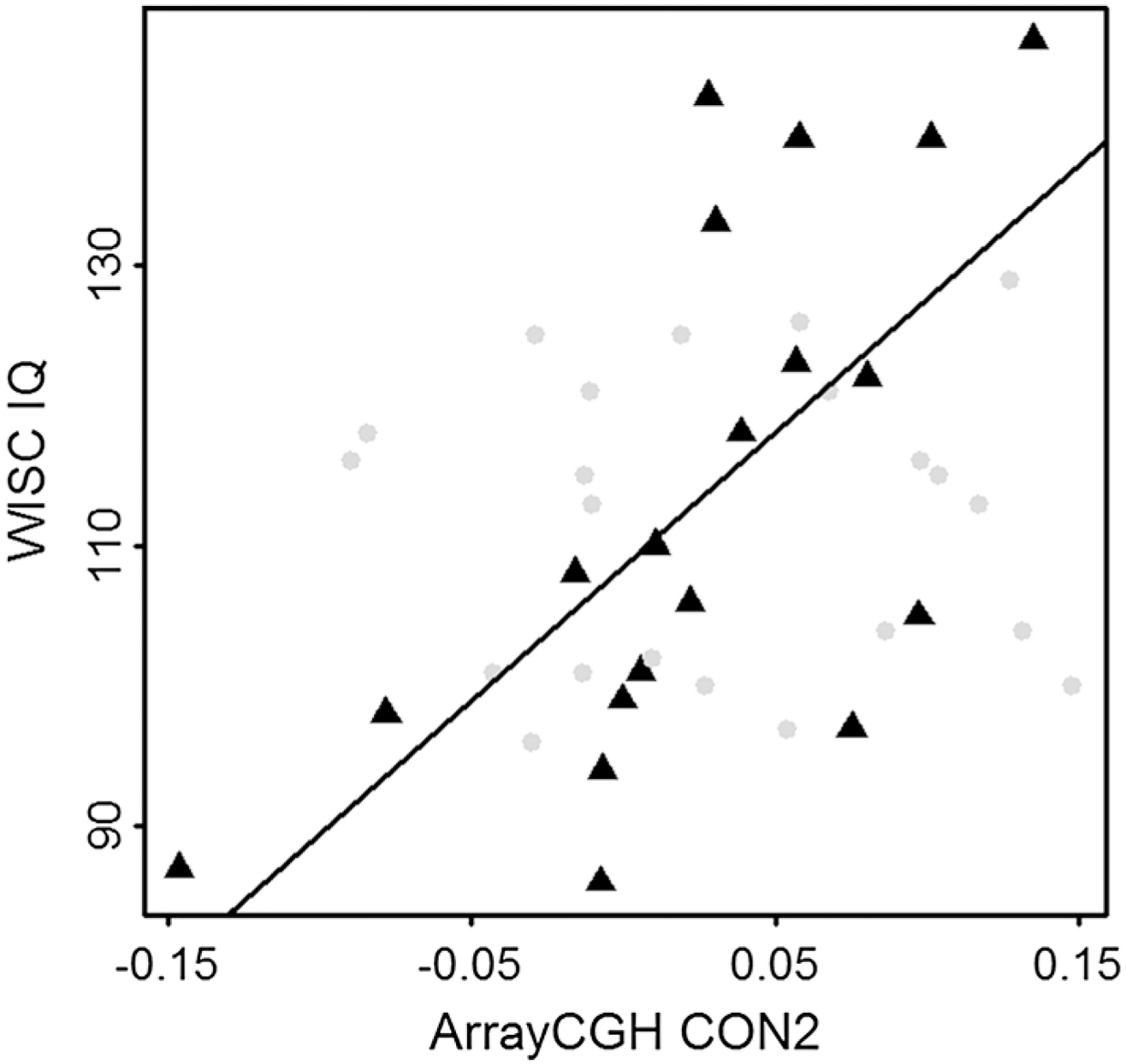

Fig. 2.

Linear association of CON2 with IQ in North American males. The *triangles* display array-based CON2 copy ratio versus WISC IQ in younger males (≤10.9). The *gray dots* display the same but in older males. The *line* is a least squares line of best fit in the younger male group (R^2 = 0.42, p = 0.003, n = 19)

Table 1

Select brain region associations with WISC IQ

Region	Beta	SE	*p* value
Right frontal cortical volume	0.0003	0.0001	0.004
Right parietal cortical volume	0.0006	0.0002	0.004
Right temporal cortical volume	0.0003	0.0001	0.005
Right occipital cortical volume	0.0009	0.0003	0.002
Left frontal cortical volume	0.0003	0.0001	0.003
Left frontal surface area	0.0014	0.0005	0.003
Right temporal surface area	0.002	0.0006	0.001
Left temporal surface area	**0.002**	**0.0006**	**0.0006**
Right temporal surface area	0.002	0.0006	0.001
Right occipital surface area	0.003	0.0010	0.006
Bilateral frontal surface area	0.0007	0.0002	0.007
Bilateral parietal surface area	0.001	0.0004	0.009
Bilateral temporal surface area	**0.001**	**0.0003**	**0.0006**
Right surface area	0.0006	0.0002	0.003
Left surface area	0.0006	0.0002	0.003
Bilateral surface area	0.0003	0.0001	0.002
Bilateral frontal cortical volume	0.00015	0.000048	0.003
Bilateral parietal cortical volume	0.00027	0.00009	0.004
Bilateral temporal cortical volume	0.00017	0.000057	0.004
Bilateral occipital cortical volume	0.00045	0.00013	0.002
Bilateral cortical volume	0.000057	0.000018	0.002

Table 1 includes brain regions that were associated with IQ ($p < 0.01$). Bold highlights those regions meeting Bonferroni criteria (0.05/54 = 0.0009)

Table 2

ArrayCGH-measured CON2 association with cognitive scores in North American and New Zealand cohorts

IQ measure	NIMH*		NZ	
	R^2	*p* value	R^2	*p* value¥
Total WISC IQ	**0.13**	**0.020**	0.03	0.190
Total WISC IQ#	**0.22**	**0.045**	–	–
WISC verbal	–	–	0.05	0.098
WISC performance	–	–	<0.01	0.547
PAT reading	–	–	0.01	0.420
PAT math	**–**	**–**	**0.10**	**0.018**
Scholastic ability	–	–	0.03	0.257

Table 2 directly compares arrayCGH-based results of CON2 dosage from different cohorts and also includes replication of assays (array-CGH vs. ddPCR of CON2) in the case of WISC IQ in the NIMH cohort

*In males,

#Results of ddPCR assay,

¥Bonferroni corrected *p* values