Exonic Deletions in AUTS2 Cause a Syndromic Form of Intellectual Disability and Suggest a Critical Role for the C Terminus

Gea Beunders,1,37 Els Voorhoeve,1,37 Christelle Golzio,2 Luba M. Pardo,1 Jill A. Rosenfeld,3 Michael E. Talkowski,4,5 Ingrid Simonic,6 Anath C. Lionel,7,8 Sarah Vergult,9 Robert E. Pyatt,10,11 Jiddeke van de Kamp,1 Aggie Nieuwint,1 Marjan M. Weiss,1 Patrizia Rizzu,1 Lucilla E.N.I. Verwer,1 Rosalina M.L. van Spaendonk,1 Yiping Shen,4,12,13 Bai-lin Wu,12,14 Tingting Yu,12,13 Yongguo Yu,12,13 Colby Chiang,4 James F. Gusella,4,5 Amelia M. Lindgren,15,16 Cynthia S. Morton,5,15,16 Ellen van Binsbergen,17 Saskia Bulk,17 Els van Rossem,18 Olivier Vanakker,9 Ruth Armstrong,6 Sook-Mi Park,9 Lynn Greenhalgh,19 Una Maye,19 Nicholas J. Neill,3 Kristin M. Abbott,6 Susan Sell,20 Roger Ladda,20 Darren M. Farber,21 Patricia I. Bader,22 Tom Cushing,23 Joanne M. Drautz,23 Laura Konczal,24 Patricia Nash,25 Emily de Los Reyes,26 Melissa T. Carter,27 Elizabeth Hopkins,28 Christian R. Marshall,7,8 Lucy R. Osborne,29 Karen W. Gripp,28 Devon Lamb Thrush,10,30 Sayaka Hashimoto,30 Julie M. Gastier-Foster,10,11 Caroline Astbury,10,11 Bauke Ylstra,31 Hanne Meijers-Heijboer,1 Danielle Posthuma,1,32,33 Björn Menten,9 Geert Mortier,34 Stephen W. Scherer,7,8 Evan E. Eichler,35 Santhosh Girirajan,35,36 Nicholas Katsanis,2 Alexander J. Groffen,1,37 and Erik A. Sistemas1,*

Genomic rearrangements involving AUTS2 (7q11.22) are associated with autism and intellectual disability (ID), although evidence for causality is limited. By combining the results of diagnostic testing of 49,684 individuals, we identified 24 microdeletions that affect at least one exon of AUTS2, as well as one translocation and one inversion each with a breakpoint within the AUTS2 locus. Comparison of 17 well-characterized individuals enabled identification of a variable syndromic phenotype including ID, autism, short stature, microcephaly, cerebral palsy, and facial dysmorphisms. The dysmorphic features were more pronounced in persons with 3′ AUTS2 deletions. This part of the gene is shown to encode a C-terminal isoform (with an alternative transcription start site) expressed in the human brain. Consistent with our genetic data, suppression of auts2 in zebrafish embryos caused microcephaly that could be rescued by either the full-length or the C-terminal isoform of AUTS2. Our observations demonstrate a causal role of AUTS2 in neurocognitive disorders, establish a hitherto unappreciated syndromic phenotype at this locus, and show how transcriptional complexity can underpin human pathology. The zebrafish model provides a valuable tool for investigating the etiology of AUTS2 syndrome and facilitating gene-function analysis in the future.

1Department of Clinical Genetics, VU University Medical Center, Amsterdam 1007 MB, The Netherlands; 2Center for Human Disease Modeling, Department of Cell Biology, Duke University Medical Center, Durham, NC 27710, USA; 3Signature Genomic Laboratories, Perkin Elmer, Spokane, WA 99207, USA; 4Center for Human Genetic Research, Massachusetts General Hospital, affiliated with Departments of Genetics and Neurology, Harvard Medical School, Harvard University, Boston, MA 02114, USA; 5Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02142, USA; 6East Anglian Medical Genetics Service, Addenbrooke’s Hospital, Cambridge University Hospitals, National Health Service Foundation Trust, Cambridge, CB2 0QQ, UK; 7The Centre for Applied Genomics and Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 1L7, Canada; 8Department of Molecular Genetics and the McLaughlin Centre, University of Toronto, Toronto, ON M5S 1A1, Canada; 9Center for Medical Genetics, University Hospital Ghent, Ghent 9000, Belgium; 10Pathology and Laboratory Medicine, Nationwide Children’s Hospital, Columbus, OH 43205, USA; 11Department of Pathology, The Ohio State University, Columbus, OH 43210, USA; 12Department of Laboratory Medicine, Boston Children’s Hospital, Boston, MA 02114, USA; 13Shanghai Children’s Medical Center, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China; 14Children’s Hospital of Philadelphia and Institute of Biomedical Science, Fudan University, Shanghai 200032, China; 15Department of Pathology, Brigham and Women’s Hospital, Boston, MA 02115, USA; 16Department of Obstetrics, Gynecology, and Reproductive Biology, Harvard Medical School, Boston, MA 02115, USA; 17Department of Medical Genetics, VU University Medical Center, Amsterdam 1007 MB, The Netherlands; 18Onze-Lieve-Vrouwe Ziekenhuis, Aalst 9300, Belgium; 19Clinical Genetics, Royal Liverpool Children’s Hospital, Eaton Road, Alder Hey, Liverpool L12 2AP, Great Britain; 20Penn State Milton S. Hershey Medical Center, Hershey, PA 17033, USA; 21Department of Neurology, University of Louisville, Louisville, KY 40202, USA; 22Northeast Indiana Genetic Counseling Center, Ft. Wayne, IN 46804, USA; 23Pediatric Genetics Division, Department of Pediatrics, University of New Mexico, Albuquerque, NM 87131, USA; 24University Hospitals, Case Western Reserve University, Cleveland, OH 44106, USA; 25Department of Behavioral Pediatrics, Nationwide Children’s Hospital, Columbus, OH 43205, USA; 26Department of Genetics and Neurology, The Ohio State University, Columbus, OH 43210, USA; 27Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada; 28Division of Medical Genetics, Alfred I. duPont Hospital for Children, Wilmington, DE 19803, USA; 29Departments of Medicine and Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada; 30Department of Pediatrics, The Ohio State University, Columbus, OH 43210, USA; 31Department of Pathology, VU University Medical Center, Amsterdam 1007 MB, The Netherlands; 32Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam 1081 HV, The Netherlands; 33Department of Child and Adolescent Psychiatry, Erasmus University Rotterdam, Rotterdam 3000 CB, The Netherlands; 34Department of Medical Genetics, Antwerp University, Edegem 2650, Belgium; 35Department of Genome Sciences and Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA; 36Department of Biochemistry and Molecular Biology Department of Anthropology, Pennsylvania State University, Pennsylvania, PA 16803, USA

*Correspondence: e.sistemas@vumc.nl
http://dx.doi.org/10.1016/j.ajhg.2012.12.011. ©2013 by The American Society of Human Genetics. All rights reserved.
Introduction

Neurodevelopmental disorders, including intellectual disability (ID) and autism, have a strong genetic component, but only a few of the underlying genes have been identified. Candidate-gene discovery has accelerated in recent years by the implementation of high-resolution genomic arrays. However, detected copy-number variants (CNVs) often either encompass multiple genes or are too rare to provide causal evidence for a particular candidate transcript. Autism susceptibility candidate 2 (AUTS2), located on 7q11.22 (MIM 607270), represents such an ID candidate with inconclusive evidence for causality.

AUTS2 was first identified as a candidate for neurocognitive defects because a translocation-breakpoint analysis in twins with autism, developmental delay, and epilepsy showed that one of the breakpoints disrupted AUTS2.1 Besides the twins, seven additional cases have now been reported to have a disrupted AUTS2 coding region: four individuals with a translocation breakpoint,2,3 one with an inversion breakpoint disrupting AUTS2,2,4 and two with intragenic deletions.2,5 These individuals manifested ID and developmental delay (all nine), dysmorphic features (six), autism (four), and skeletal abnormalities (three). This overview does not include persons with intronic deletions in AUTS2 because the functional significance of such intronic variation is unclear.6

Complicating the candidacy of this locus, some of the genomic rearrangements affecting AUTS2 disrupt other genes as well. A combination of cytogenetic and sequencing studies suggested that CNTNAP2 (7q35) might be causal in an individual with a 7q inversion disrupting AUTS2 and CNTNAP2 (MIM 604569);4 likewise, for three larger multigenic de novo deletions (in the DECIPHER database) encompassing AUTS2, it is unclear whether the disruption of AUTS2 alone drives the phenotype.7 The data presented by Nagamani et al.5 on two individuals with intragenic deletions suggest that deletions in AUTS2 alone might be pathogenic. However, the number of affected individuals was too small to exclude the role of other genes or to delineate a phenotype.5 Here, we present direct evidence from both clinical and genetic data and animal studies for the causal relation of AUTS2 with an ID syndrome and delineate the associated phenotype. Furthermore, we provide evidence that functional elements in the C terminus of AUTS2 are major contributors to both the neurodevelopmental and craniofacial phenotypes of individuals with C-terminal deletions or rearrangements at this locus.

Subjects and Methods

Subjects
Routine diagnostic array comparative genomic hybridization (CGH) was performed for ID and/or multiple congenital anomalies (MCAs) for a total of 49,684 individuals across ten diagnostic centers in The Netherlands, Belgium, Great Britain, the United States, and Canada (each center used their standard diagnostic platform; in total, six analogous platforms were used). In some of these individuals, karyotyping was also performed. From this cohort, we selected all individuals with a deletion involving AUTS2, as well as one person with a translocation and another person with an inversion in which one of the breakpoints was in AUTS2. To map the region further and to delineate the associated phenotypes, we obtained peripheral-blood samples and collected clinical information through either medical letters or a data sheet filled in by the referring physicians with approval of the local medical ethical committee. Results were confirmed with different methods (high-density array, multiplex ligation-dependent probe amplification [MLPA], and fluorescence in situ hybridization [FISH]) depending on the laboratory (see Tables S1, S2, and S3, available online). Exact breakpoint delineation of the translocation with one breakpoint in 7q11.22 was performed with FISH, and the inversion was characterized with whole-genome sequencing, as previously described.8–10 Informed consent was obtained from parents or caregivers as appropriate, and specific consent for publishing photographs was obtained from all individuals whose photographs are shown here. Institutional approval of the local medical ethical committee was obtained as well. Individuals with a confirmed exonic deletion or a genomic rearrangement involving AUTS2 and available clinical data were included for phenotypic studies.

Controls
To assess the frequency of AUTS2 deletions within a large general population, we analyzed CNV data of 16,784 subjects from several control groups. A total of 4,783 DNA samples from the Wellcome Trust Case Control Consortium 2 (WTCCC2) were analyzed with a SNP array. This control group included individuals who had been nationally ascertained and regarded as healthy from the 1958 Birth Cohort and the UK Blood Service collection (October 26, 2011).11 Further control CNV data from 8,329 cell-line- and blood-derived controls were obtained primarily from genome-wide association studies of nonneurological phenotypes. Because these included 2,090 controls from the UK Blood Service collection, this set added only 6,239 unique controls. Although these data were not ascertained specifically for neurological disorders, they consist of adult individuals who provided informed consent as described previously.12 In addition, publicly available data from HapMap phase 3 (October 26, 2011), which consists of 1,056 healthy controls from 11 different populations, were checked for deletions involving AUTS2.13 CNV data were available from the following four control sets: The Ottawa Heart Institute (OHI) controls (n = 1,234) from Canada, POPGEN controls (n = 1,123) from Germany, SÄGE controls (n = 1,287) from the United States, and the Low-Lands-Consortium controls (n = 981) from The Netherlands.13–15 See Table S4 for details on all cohorts and the array platforms used. The array platforms used for controls have the same or a comparable resolution as the platforms used for cases. The number of deletions found in the cases was compared to that in the controls with a Fisher’s exact test.

Genotype-Phenotype Correlations
We received clinical data from 17 individuals and 4 family members carrying an exonic AUTS2 disruption. We used these individuals to identify features that occurred in at least two unrelated individuals, i.e., features with a minimal frequency of 10%.
A recent systematic review of Oeseburg et al. showed that in a general ID cohort, the most frequent additional health conditions (epilepsy and cerebral palsy) are as frequent as 20%, but the remainder of the comorbid clinical features (including autism and a congenital malformation in general) are seen in less than 10%. Therefore, a frequency of 10% for a specific feature in this AUTS2 cohort is an enrichment compared to ID cases in general. These recurrent features were scored for all individuals and family members carrying the familial deletion, and asymmetrically occurring features were counted as positive. The sum of positive features was counted for each individual and was defined as his or her individual AUTS2 syndrome severity score.

Because deletions or genomic rearrangements affecting the 3′ end of the AUTS2 coding sequence seem to be associated with a more severe phenotype, persons with exonic deletions were categorized in two groups depending on whether the deletion disrupted the highly conserved AUTS2 segment (containing exons 9–19) that is also encoded by the alternative 3′ transcript (see Alternative Transcription Start Sites below and the Results). We used a Kolmogorov-Smirnov test to test whether the corresponding AUTS2 syndrome severity scores for these two groups differed significantly.

Alternative Transcription Start Sites

To search for an explanation for the observed genotype-phenotype trend, we first determined the evolutionary conservation of human AUTS2 exonic sequences. We used the following species for comparison: gorilla (Gorilla gorilla; gorGor3), macaque (Macaca mulatta; Mmul_1), dog (Canis familiaris; BroadD2), cow (Bos taurus; Btau_4.0), pig (Sus scrofa; Sscrofa9), mouse (Mus musculus; NCBI37), chicken (Gallus gallus; Washuc2), clawed frog (Xenopus tropicalis; JG4_2), and zebrafish (Danio rerio; ZB9). Accession numbers of protein sequences are ENSGGOPO0000011519, ENSMMPUP00000023254, ENSBTAP00000002697, ENSSSCPO0000008253, ENSCAFPO00000016549, ENSMUSP00000062515, ENSGALP0000001729, ENSXETPO0000007747, and ENSDARP00000073379. Two different methods were used. We first aligned the predicted protein of the longest isoform in humans to the predicted amino acid sequences of the orthologous species by using MUSCLE v.3.8 software. For that purpose, we downloaded sequences from the latest builds from Ensembl. Then, to detect similarity in nonannotated or noncoding genomic DNA, we used the tblastn algorithm (see Web Resources) with the human amino acid sequence as query. The degree of homology was calculated as the percentage of identical amino acids.

Second, we searched for putative alternative transcription start sites (TSSs) that were associated with a shorter 3′ isoform in the human brain. We used mRNA from the caudate nucleus and the medial frontal gyrus from one donor provided by the Dutch Brain Bank and performed a replication experiment by the same procedure on a mRNA sample from the medial frontal gyrus of a second donor. Rapid amplification of 5′ cDNA ends (5′ RACE) was performed with the Ambion FirstChoice RLM-RACE kit according to the manufacturer’s instructions. Nested PCR amplification was performed with 5′-ATGCTTCTGGCCTGAATGGCT-3′ as the outer AUTS2-specific reverse primer and 5′-GGAGAAGAGACTGTGCCG-3′ as the inner primer (Figures S1A and S1B).

Knockdown and Rescue Experiments in Zebrafish Embryos

To investigate the role of AUTS2 in the regulation of head size, neuronal development, and morphology in general, we performed zebrafish knockdown experiments. Zebrafish (Danio rerio) embryos were raised and maintained as described. Splice-blocker morpholinos (MOs) against the AUTS2 ortholog auts2 were designed and obtained from Gene Tools (Table S5). We injected 1 nl of diluted MOs (4.5 ng for the 5′ MO targeting the exon 2 donor splice and 6 ng for the 3′ MO targeting the exon 10 donor splice) and/or 100 pg of mRNA into wild-type zebrafish embryos at the 1- to 2-cell stage (n = 50–100 embryos per injection dose) and performed RT-PCR to measure efficiency of the splice blocking. Injected embryos were scored visually at 3 days postfertilization (dpf) and classified as normal or microcephalic on the basis of the relative head size compared with that of age-matched controls from the same clutch. For rescue experiments, the human wild-type mRNAs (full-length or short transcript [GenBank accession numbers JQ670866 and JQ670867, respectively]) were cloned into the pCS2 vector and transcribed in vitro with the SP6 Message Machine kit (Ambion); 100 pg of the human wild-type mRNAs were coinjected with the MOs. All experiments were repeated three times and evaluated statistically with a Student’s t test. Aclian-blue staining of cartilaginous structures was performed for investigating the morphology of the head. Zebrafish embryos were fixed with 4% paraformaldehyde (PFA), and the cartilage structures were visualized by Aclian-blue staining according to an established protocol. Further, whole-mount immunostainings with either HuC/D (postmitotic neurons) or phosphohistone H3 (proliferating cells) were performed for investigating neuronal development and head-size regulation at a cellular level. Embryos were fixed in 4% PFA overnight and stored in 100% methanol at −20°C. After rehydration in PBS, PFA-fixed embryos were washed in immunofluorescence (IF) buffer (0.1% Tween-20 and 1% BSA in PBS 1×) for 10 min at room temperature. The embryos were incubated in the blocking buffer (10% FBS and 1% BSA in PBS 1×) for 1 hr at room temperature. After two washes in IF buffer for 10 min each, embryos were incubated in the first antibody solution, 1:750 anti-histone H3 (ser10)-R (sc-8656-R, Santa Cruz) or 1:1,000 anti-HuC/D (A21271, Invitrogen), in blocking solution overnight at 4°C. After two washes in IF buffer for 10 min each, embryos were incubated in the secondary antibody solution, 1:1,000 Alexa Fluor donkey anti-rabbit IgG and Alexa Fluor goat anti-mouse IgG (A21207, A11001, Invitrogen), in blocking solution for 1 hr at room temperature. Staining was quantified by the counting of positive cells in defined regions of the head and with ImageJ software.

Results

Genotypes

To assess the candidacy of AUTS2 in cognitive impairment in humans, we examined the AUTS2 region in 49,684 individuals with ID and/or MCAs by using array CGH and/or karyotyping. We identified 44 deletions encompassing at least part of AUTS2 and a maximum of two other genes (WBSCR17 and CALN1 [MIM 607176]), and conventional karyotyping revealed one translocation and one inversion each with one breakpoint in AUTS2 (Table S1). AUTS2-encompassing duplications found in this cohort were not included in this study because the functional relevance of these lesions is unclear. Twenty-four deletions were found to include at least one AUTS2 exon, whereas another 17 did not. For the remaining three deletions, it was
unclear whether they included an exon because of the limited resolution of the array. For these three individuals, we had no consent to perform further studies.

Overall, in our cohort of 49,684 affected individuals, we identified 24 persons (0.05%) harboring deletions disrupting the coding sequence. To assess the significance of this observation, we analyzed 16,784 controls from 12 cohorts by using arrays with high-density coverage of the AUTS2 locus (Table S4). Although nine deletions were found, none of them disrupted an AUTS2 exon (Table S6). The difference between exonic deletions in the cases (24/49,651) and those in controls (0/16,784) was highly significant (p = 0.00092), suggesting that exonic disruptions of AUTS2 give rise to a highly penetrant phenotype in humans. This is supported by CNV data from the latest version of the Database of Genomic Variants (August 25, 2012), wherein none of the array-based studies show CNVs that disrupt an exon, and by the fact that none of the 24 probands with an exonic AUTS2 deletion had a rare de novo CNV at another locus (Table S1).

We were able to obtain phenotypic data from 15 out of 24 probands with an exonic AUTS2 deletion (cases 1–15), from the inversion case (16) and the translocation case (17) with one breakpoint in AUTS2, as well as from four family members carrying the familial AUTS2 deletion. In these 17 probands, MLPA, FISH, high-density array CGH, and breakpoint sequencing confirmed the aberrations and further delineated the breakpoints. (Figure 1 and Table S1).

In total, 21 individuals from 17 families were included in our genotype-phenotype study (Table S7). In 8 cases (5, 8, 10, 12, 14, 15, 16, and 17) out of 11 probands in which both parents were available for genetic testing, the AUTS2 aberrations occurred de novo; the other three probands inherited the AUTS2 deletion from an unaffected parent (case 1) or an affected parent (cases 4 and 6). In six probands (cases 2, 3, 7, 9, 11, and 13), the inheritance status of the AUTS2 deletion could not be fully resolved because one or both parents were unavailable for testing.

Of the ten individuals with an intragenic deletion (not including the first and last exon), four probands (cases 6, 7, 8, and 11) carried a deletion predicted to cause a frameshift, whereas the other six individuals (cases 1, 2, 3, 4, 9, and 10) carried in-frame deletions. Finally, in case 14, the deletion also included one downstream gene (WBSCR17), and in cases 13 and 15, the deletions also affected two downstream genes (WBSCR17 and CALN1) (Figure 1, Table S7, and Figure S2).

**Phenotypes**

Next, we asked whether there were any recurrent phenotypic features associated with AUTS2 disruptions. All 17 probands from whom detailed clinical data were available had ID and/or developmental delay; this had been the reason for diagnostic testing. One of the parents (the mother of case 4) carrying an AUTS2 deletion had learning difficulties, one (the mother of case 6) had mild ID, and one (the father of case 1) had normal intelligence. Seven
probands (cases 2, 5, 9, 12, 13, 16, and 17) were diagnosed with autism spectrum disorder or showed autistic behavior. In addition to the expected neurocognitive defects, we also observed a constellation of other recurrent clinical features in individuals with exonic deletions. These included microcephaly (14 individuals), short stature (12), feeding difficulties (10), hypotonia (8), and cerebral palsy (9). We also found recurrent dysmorphic features: hypertelorism (10), ptosis (6), short palpebral fissures (8), epicanthal folds (7), a short and/or upturned philtrum (8), micrognathia (7), and a narrow mouth (12). Less frequent features were skeletal abnormalities including (signs of) arthrogryposis (3), umbilical or inguinal hernia (2), and heart defects (3) (Figure 2 and Table 1). The striking phenotypic complexity and variable size and position of the CNVs prompted us to evaluate the clinical information from the 17 probands and 4 family members carrying the familial AUTS2 deletion included in this study to derive pathology scores on the basis of simple, objective criteria; we summarized these as the “AUTS2 syndrome severity score” (the maximum score is 32). Even though this paradigm is a crude approximation of the phenotypic diversity at this locus, we nonetheless observed dichotomization of phenotypes. Cases and family members 1–4, (all with 5’ in-frame deletions) scored significantly lower (median AUTS2 syndrome severity score = 5) than did cases (5–17) and family members with deletions of downstream exons, whole-gene deletions, or exons 1–4 deletions including the initiation codon (median AUTS2 syndrome severity score = 12) (Figures 1 and 3 and Table S7). This difference

**Figure 2. Facial Characteristics of Cases with an AUTS2 Aberration**

(A) Case 1 at age 3 years shows no dysmorphic features.
(B and L) Front (B) and side (L) views of case 4 at age 2.5 years show a repaired cleft lip, mild proptosis, and short and mild upslanting palpebral fissures.
(C) The mother of case 4 shows a repaired cleft lip, ptosis, and retrognathia.
(D) Case 5 at age 3 years shows highly arched eyebrows, mild downslanting palpebral fissures, epicanthal folds, and a short philtrum.
(E, F, M, and N) Front (E) and side (M) views of case 6 at age 6 years. She is hyperteloric and has ptosis and downslanting palpebral fissures, a short philtrum, and a narrow mouth similarly to her brother, shown in (F) and (N) at the age of 10 years.
(G and O) Front (G) and side (O) views of case 9 at age 32 years show hypertelorism, proptosis, upslanting palpebral fissures, a short upturned philtrum, and a narrow mouth.
(H) Case 10 at age 2 years shows a prominent nasal tip, anteverted nares, and a short philtrum.
(I and P) Front (I) and side (P) views of case 13 at age 5.5 years show hypertelorism, ptosis, a broad nasal bridge, a short and upturned philtrum, and a narrow mouth.
(J, K, and Q) Case 15 at age 1 year (J) and 4.8 years (K and Q) shows a broad nasal bridge, short palpebral fissures and a short philtrum, and a narrow mouth. See also Table 1 and Table S7.
was significant regardless of the inclusion or exclusion of affected family members (p = 0.001 or p = 0.011, respectively).

Detection of a C-Terminal AUTS2 Isoform
The apparent dependence of severity scores on CNV location prompted us to evaluate the evolutionary conservation of each AUTS2 exon (Figures S3 and S4); conservation was especially high in the 3' gene region. Given the fact that the ENSEMBL annotation of the AUTS2 sequences predicts the presence of several splice isoforms, we next looked for the presence of alternative isoforms in human brain mRNA. Using 5' RACE, we identified a short 3' AUTS2 mRNA variant starting in the middle of exon 9, depicted in Figure 4. All transcripts detected employed the same start site (see also Figures S1 C and S1D). The reading frame of the short transcript is identical to that of the full-length AUTS2 transcript and is predicted to encode a polypeptide of 697 amino acids instead of the 1,259 amino acids of the full-length protein. The evolutionary conservation from humans to zebrafish suggests an important biological function for AUTS2 and, together with the shorter transcript, gave us the opportunity to analyze the function of the C terminus of AUTS2 in a zebrafish model.

In Vivo Analysis of AUTS2 in Zebrafish Embryos
Taken together, our CNV mapping data, our RACE analyses, and the strong correlation between phenotypic severity and position of the deletion suggest that the 3' end of the AUTS2 locus contains major functional elements that are encoded by both the full-length transcript and the shorter C-terminal isoform. Microcephaly is one of the most consistent clinical features in our cases (14/20; Table 1). We therefore asked whether AUTS2,
particularly the shorter C-terminal isoform, might be involved in the regulation of head size. Given that we have shown recently how head-size evaluations in zebrafish embryos can serve as a surrogate for the evaluation of candidate genes for neurocognitive traits,\textsuperscript{21} we decided to create a zebrafish morphant for \textit{auts2}. Using reciprocal BLAST, we identified a single \textit{Danio rerio AUTS2} ortholog (\textit{auts2} on chromosome 10; 62\% amino acid identity and 72\% similarity with the long isoform of \textit{AUTS2}) (Figure 4A). We were able to detect endogenous \textit{auts2} message by RT-PCR as early as embryonic 5-somite stage by using both 5' and 3' primer sets (data not shown). Next, we designed two splice-blocking morpholinos (sb-MOs): a 5' MO targeting the splice donor site of exon 2 and a 3' MO targeting the splice donor site of exon 10 (the 5' and 3' MOs were chosen to suppress the full-length transcript only and both \textit{auts2} transcripts [if present], respectively; see Figure 4A and Figure S5). RT experiments demonstrated that both sb-MOs affected correct splicing of the \textit{auts2} transcript (Figure S5). Masked scoring of embryos at 3 dpf showed a reproducible microcephaly phenotype—53\% and 48\% for 5' and 3' sb-MOs, respectively (Figures 5A and 5B)—that was concomitant with the efficiency of splice blocking of the two sb-MOs, as established by RT-PCR (Figure S5). The phenotype was unlikely to be driven by overall developmental delay; morphants had a normal appearance with regard to their pigment cells, there was no apparent pathology in other internal organs, such as the heart or the swim bladder, and their body length was indistinguishable from that of control embryos from the same clutch (Figure 5C). The phenotype was specific; the observed microcephaly caused by the two sb-MOs could be rescued efficiently with coinjection of wild-type human full-length mRNA (GenBank JQ670866) (Figures 5A and 5B). Strikingly, microcephalic embryos could also be rescued with the human short \textit{AUTS2} isoform (GenBank JQ670867) in a manner indistinguishable from that with the full-length form, indicating that the observed phenotype is driven by sequences in exons 9–19. We also observed another recurrent dysmorphic feature in knockout zebrafish morphants: micrognathia and retrognathia. To quantify this defect, we stained embryos injected with either a 5' or 3' sb-MO at 5 dpf with Alcian blue and performed quantitative morphometric analysis of the lower jaw. We observed a significant reduction of the distance between the Meckel and ceratohyal cartilages, indicating a reduced lower-jaw size comparable to the micrognathia and retrognathia seen in individuals with an \textit{AUTS2} disruption (Figures 5D and 5E).

To probe the underlying cause(s) of the microcephalic phenotype further, we stained embryos at 2 dpf with antibodies against phosphohistone H3, an M phase marker, and HuC/D, a marker of postmitotic neurons.\textsuperscript{22} This time point was selected because it precedes the development of microcephaly and, as such, allowed us to evaluate the forebrain prior to gross anatomical defects. We observed a striking reduction in phosphohistone-H3- and HuC/D-positive cells in embryos injected with either the 5' or the 3' MO, as well as loss of bilateral symmetry in HuC/D protein levels, indicating that the microcephaly phenotype is caused by disturbed neuronal proliferation. Both phenotypes could be rescued with the 3' human mRNA (Figure 6).

**Discussion**

Our studies of 49,684 individuals with ID and/or MCAs revealed deletions in \textit{AUTS2} in 44 individuals; at least 24 of these deletions involved exons. In contrast, we only found nine \textit{AUTS2} deletions, none of which were exonic, in 16,784 controls, strongly indicating that intragenic \textit{AUTS2} deletions that disrupt at least a portion of the coding sequence are a recurrent cause of neurodevelopmental defects in humans. The frequency of exonic deletions that we found was 1 in 2,000 cases, comparable with some of the recurrent deletions described by Cooper et al.\textsuperscript{12} such as the 10q23 deletion (\textit{NRG3} [MIM 605533] and \textit{GRID1} [MIM 610659]) and deletions causing Sotos syndrome (MIM 117550) or Rubinstein-Taybi syndrome (MIM 180849). This observed frequency is likely to be an

---

**Figure 3. Scatter Plot of the AUTS2 Syndrome Severity Score for Disruptions Affecting the N or C Terminus of AUTS2**

Scatter plot of the AUTS2 syndrome severity score for disruptions that affect the highly conserved amino acid sequence block encoded by exons 9–19 (yes) and the deletions not affecting this amino acid sequence (no) (see also Table S7 and Figure 4). The numbers refer to case numbers. The following abbreviations are used: f, father of patient x; m, mother of patient x; and s, sibling of patient x (see Table S7). The AUTS2 syndrome severity scores between these groups of cases differ significantly (p = 0.001, Kolmogorov-Smirnov Z test).
underestimate because smaller deletions (single-exon deletions and small indels within exons) and nonsense mutations are likely to cause AUTS2 syndrome and are missed with the techniques used here.

The individuals (cases 1–17) with an AUTS2 aberration affecting the coding sequence studied here, together with previously reported cases, allowed us to delineate recurrent phenotypic features (ID, autism, microcephaly, mild short stature, feeding difficulties, hypotonia, cerebral palsy, and dysmorphic features) of the AUTS2 syndrome. Only 1 (the father of case 1) of the 21 persons studied in detail did not have any features of the AUTS2 syndrome, indicating a penetrance of around 95%. Although the phenotype of the AUTS2 syndrome is variable and the features are sometimes subtle, there are other examples where reverse genomics have shown variable phenotypes associated with the same locus.23,24 Several lines of evidence support the causality of AUTS2 deletions for this broad phenotypic spectrum; these are (1) the significant enrichment of exonic deletions in cases, (2) the fact that auts2 zebrafish morphants show microcephaly and smaller lower-jaw size comparable to the human phenotype (these aberrant phenotypes can be fully rescued by both full-length and size comparable to the human phenotype (these aberrant morphants show microcephaly and smaller lower-jaw size). This might suggest that the C-terminal part of the protein contains the crucial region for the observed dysmorphology. It is uncertain whether the shorter 3' transcript is expressed at sufficiently high levels to explain the milder phenotype in humans with in-frame 5' deletions. The milder phenotype might well be explained by the fact that AUTS2 alleles with these deletions can still be transcribed and can thus result in a protein that contains the important C-terminal sequences.

In aggregate, our data indicate that AUTS2 deletions, particularly when they involve the C terminus, give rise to a highly penetrant syndrome that includes neurocognitive defects. Our data highlight transcriptional complexity at the AUTS2 locus and show that careful genomic, genetic, and functional dissection of such complexity can offer both clinical and mechanistic insights. Although little is known about the function(s) of AUTS2 or its isoforms, a role in neurodevelopment is suggested by the reduction of postmitotic neurons and loss of bilateral symmetry, which both might be driven by neurogenesis and/or migration defects in the zebrafish auts2 morphants. The zebrafish model can be of great value for further studies of AUTS2 function and can be helpful for defining the pathogenicity of specific genomic disruptions.

In conclusion, detailed analysis of the AUTS2 locus allowed us to delineate a hitherto undescribed microdeletion
syndrome occurring with a frequency that approximates the frequency of deletions causing Sotos syndrome or Riebstein-Taby syndrome. This AUTS2 syndrome has presumably remained undescribed because (1) the specific characteristics of the resulting phenotype are subtle, (2) the severity of the syndrome is highly variable, and (3) the penetrance is dependent on location and type of deletion.

Figure 5. Suppression of auts2 in Zebrafish Leads to Small Head Size and Craniofacial Defects
(A) Lateral views of representative control embryos and embryos injected with auts2 MOs.
(B) Quantification of microcephaly was performed in embryo batches injected with 4.5 ng 5' MO (targeting exon 2 donor splice) or 6 ng 3' MO (targeting exon 10 donor splice) plus 100 pg wild-type human AUTS2 full-length (FL) or short isoform (3') mRNAs (n = 56–91 embryos per injection). p values are denoted on the bar graph. The following abbreviation is used: NS, nonsignificant.
(C) No significant difference in body length was observed in auts2 morphants and rescued embryos at 2 dpf. Bars represent the average length of 30 embryos, which were scored blindly to injection cocktail. Data are shown as the mean ± SD.
(D) Ventral views of representative control embryos and those injected with auts2 MOs (either a 3' or 5' MO) at 5 dpf. Cartilage structures were visualized by whole-mount Alcian-blue staining at 5 dpf, allowing measurement of the distance between ceratohyal and Meckel's cartilages (red lines).
(E) Averaged distance measurements are presented as the mean ± SEM. The corresponding p values are denoted on the bar graph (two-tailed t test comparisons). The following abbreviations are used: ch, ceratohyal cartilage; and Mk, Meckel's cartilage.
Supplemental Data

Supplemental Data include five figures and seven tables and can be found with this article online at http://www.cell.com/AJHG.

Acknowledgments

The authors would like to thank all patients and their families for their major contribution; without them, we would not have been able to describe this syndrome. We thank Marjolein van de Mespel, Desiree Linders, Stefan Strijbis, Amanda Clarkson, and Carrie Hanscom for their contribution to the laboratory work in this project and Shahrin Pereira for clinical coordination. This study made use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available at www.wtccc.org.uk. Funding for this project was provided by the Wellcome Trust under awards 076113 and 085475, as well as by the National Institutes of Health (GM061354, HD065286, and MH095867), the American Heart Association (11POST7160006), and a Distinguished Brumley Professorship to N.K. Evan Eichler is a Howard Hughes Medical Institute investigator. The authors thank Klaas Kok for making available the control data of The Low-Lands Consortium for this study and for developing the custom design of the 180K Agilent oligo array used in The Netherlands. We acknowledge Soheil Shams from Bio-Discovery for his help with processing the different array data files. J.A.R. and N.J.N. are employees of Signature Genomic Laboratories, a subsidiary of PerkinElmer, and E.E.E. is on the scientific advisory boards for Pacific Biosciences, SynapDx, and DNA nexus.

Received: September 14, 2012
Revised: October 6, 2012
Accepted: December 20, 2012
Published: January 17, 2013

Web Resources

The URLs for the data presented herein are as follows:
DECIPHER, http://decipher.sanger.ac.uk/
DGV (Database of Genomic Variants), http://projects.tcag.ca/cgi-bin/variation/gbrowse/hg18/#search
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
Wellcome Trust Case Control Consortium 2, https://www.wtccc.org.uk/ccc2/

Figure 6. Suppression of auts2 Leads to Reduced HuC/D Protein Levels and Fewer Proliferating Cells

(A) Suppression of auts2 leads to a decrease of HuC/D levels at 2 dpf. Representative photographs (with HuC/D-antibody staining) show the ventral views of a control, an embryo injected with an auts2 MO, and a rescued embryo injected with an auts2 MO plus 3’ human AUTS2 mRNA at 2 dpf. HuC/D levels in the anterior forebrain of the embryo injected with the auts2 MO are considerably lower than those of the control embryo. This defect was rescued significantly by coinjection of full-length (FL) or short isoform (3’) human AUTS2 mRNAs.

(B) Percentage of embryos with normal, bilateral HuC/D protein levels in the anterior forebrain (blue) or decreased and/or unilateral HuC/D protein levels (red) in embryo batches injected with auts2 MOs alone or MOs plus human AUTS2 FL or 3’ mRNA (MO + 3’ mRNA). p values are denoted on the bar graph.

(C) Phosphohistone-H3 staining for proliferating cells in the zebrafish brain at 2 dpf.

(D) Quantification of phosphohistone-H3-staining intensities from 20 embryos each (control embryos or embryos injected with MOs alone or MOs plus 3’ or 5’ human AUTS2 mRNA). Data are represented as the mean ± SEM. The corresponding p values are denoted on the bar graph (two-tailed t test comparisons between MO-injected and rescued embryos).
Accession Numbers

The Gene Expression Omnibus accession numbers for the microarray data, the nucleotide sequence of the full-length human AUTS2 transcript, and the nucleotide sequence of the shorter 3’ alternative human AUTS2 transcript reported in this paper are GSE37657, GQ670866, and GQ670867, respectively.

References


