Dual copy number variants involving 16p11 and 6q22 in a case of childhood apraxia of speech and pervasive developmental disorder

In this issue, Raca et al. present two cases of childhood apraxia of speech (CAS) arising from microdeletions of chromosome 16p11.2. They propose that comprehensive phenotypic profiling may assist in the delineation and classification of such cases. To complement this study, we would like to report on a third, unrelated, child who presents with CAS and a chromosome 16p11.2 heterozygous deletion. We use genetic data from this child and his family to illustrate how comprehensive genetic profiling may also assist in the characterisation of 16p11.2 microdeletion syndrome.

A number of chromosome 16p11.2 aberrations have been reported in the recent literature, including gross rearrangements and sub-microscopic (<1 Mb) deletions and duplications with incomplete penetrance and variable expressivity and in a heterozygous form. In general, microdeletions appear to be more penetrant than their respective duplications. The ‘typical’ 16p11.2 deletion encompasses 539 kb (from chromosome position 29.5–30.1 Mb, GRCh37/hg19) and 24 genes, but a smaller adjacent distal or ‘atypical’ deletion (between chromosome positions 28.7 and 28.95 Mb, GRCh37/hg19) has also been reported, as have novel anomalies outside of these specified regions. Individuals have been described with deletions spanning both these regions and families have been observed to carry both rearrangement types. It has been suggested that proximal rearrangements may be associated with developmental impairments and distal variations correlated with altered body mass index, although developmental delays and speech and language impairments appear to be a common feature of individuals with various anomalies across this chromosome band.

Screens of clinical cohorts indicate that chromosome 16p11.2 rearrangements are observed at a frequency of 0.3–0.7% in patients with various developmental impairments including autistic disorder (eg, Weiss et al.11), developmental delay (eg, Shinawi et al.3), epilepsy12 and schizophrenia (eg, McCarthy et al.13). Deletions are also observed in apparently healthy individuals, at a similar frequency to clinical cohorts, and while the majority of cases appear to be de novo, inherited imbalances of this region are not uncommon.

Recently there has been a drive to define a core clinical phenotype of the 16p11.2 microdeletion syndrome. In a retrospective screen of 9773 individuals referred for microarray testing, Rosenfeld et al. found that 77 carried chromosome 16p11.2 anomalies (45 deletions, 32 duplications, 0.78%). A detailed review of 18 patients found that the most consistent clinical manifestations in these individuals were intellectual impairment and speech and language delays. These findings were supported by a similar study that included 7400 patients who had undergone array comparative genomic hybridisation (array-CGH) testing in a clinical context, 45 of whom carried 16p11.2 anomalies (27 deletions, 18 duplications, 0.6%). Phenotypic characterisation of 27 individuals also found that all had speech and language delays and cognitive impairment. Other predominant features of 16p11.2 syndrome include dysmorphism, macrocephaly and autistic disorders. However, all of these features have been disputed and it is likely that ascertainment bias will affect the conclusions of many studies, particularly those that focus upon single cases. Thus, the characterisation of the relationships between genetic aberration and clinical presentation is ongoing and will require further, more refined, studies with detailed investigations of this chromosome region and consistent phenotyping of affected individuals.

The child described here was originally assessed for the presence of FOXP2 (OMIM #605317) mutations and rearrangements, as part of an earlier screening project, as disruptions of this gene have been implicated in rare cases of severe speech and language disorder. Although no FOXP2 mutations were identified in the child, we discovered a deletion of chromosome 16p11.2. The child is a second-born male child of unrelated and healthy parents of European (Caucasian) descent. His early development was normal until the age of 1.5 years, when there was evidence of a social withdrawal. He was referred following concerns regarding his speech and language development and received a diagnosis of developmental verbal dyspraxia, also known as CAS, and pervasive developmental disorder not otherwise specified (PDD-NOS) according to DMS-IV. On assessment with the Vineland Adaptive Behaviour Scales, he showed a remarkable impairment in language and communication. At age 14, he scored at an age equivalent of 3 years and 3 months in the communication domain, with a major impairment in the expressive subdomain. He also displayed stereotypic movements and behavioural disturbances with self-aggressive episodes. He did not have hearing or ocular problems and had normal height and weight. The patient showed macrocephaly and peculiar facial features, such as heavy eyebrows with mild synophrys, down-slanting palpebral fissures, hypertelorism, short philtrum, carp shaped mouth and full lips. He had brachydactyly and single crease bilaterally. His parents were clinically normal with no family history of speech delay, autistic disorders or mental retardation.

Peripheral blood samples were collected from the proband and his parents, and DNA extracted according to standard procedures. To identify genomic imbalances, DNA samples were hybridised to Agilent 244K and Agilent 4 × 44K arrays (Agilent Technologies Inc., Santa Clara, CA, USA) for the proband and his parents, respectively. The array-CGH was completed as part of an assessment of 36 children with specific language impairment, PDD-NOS and autism spectrum disorders (ASD). Image data were extracted using Agilent Feature Extraction software version 8.5 (Agilent Technologies Inc.) and analysed using Agilent Cgh Analytics software version 3.4 (z-score method setting) (Agilent Technologies Inc.). The reference genomic DNA samples used throughout the study were from the same consented individuals, one male and one female. We estimate that the mean resolution of the Agilent 244 K arrays is ~40 kb.
We identified a de novo chromosome 16p11.2 deletion in the proband’s sample (see Figure 1). The minimal region affected by this deletion spans from chromosome position 29,652,999 to 30,199,351 (GRCh37/hg19) and encompasses 28 Refseq genes, thus coinciding with the ‘proximal’ type reported in the literature. The presence of speech and language abnormalities, macrocephaly and PDD in this child coincides with previously reported core phenotypes of chromosome 16p deletions. The CAS diagnosis of this case provides further support to the findings of Raca et al. and their theory as to the importance of proximal chromosome 16p11.2 abnormalities in CAS. It would thus be of interest to fully assess the CAS in this child using the Madison Speech Assessment Protocol and other relevant speech batteries suggested by Raca et al.

Using the array-CGH data, we catalogued all observed imbalances that spanned four or more consecutive oligonucleotide probes with values outside the log10 Cy-dye threshold ratios for the proband. We excluded any region that had been observed repeatedly either in control data deposited in the Database of Genomic Variants (DGV) or within our own sample sets, and small imbalances that mapped to regions without noted reference genes or mRNAs. This approach allowed us to identify an additional novel duplication of chromosome 6q22.31, which occurred both in the proband and in his clinically normal mother. This duplication has a minimal region from chromosome position 123,527,545 to 124,311,813 (GRCh37/hg19) and does not overlap significantly with any known CNVs in the DGV. The duplication covers two genes: the entire coding region of TRDN (OMIM #603283) (triadin), a ryanodine-sensitive calcium channel expressed in cardiac and skeletal muscle, and the first exon of NKAIN2 (OMIM #609758) (sodium/potassium-transporting ATPase subunit beta-1-interacting protein 2 isoform 2), a transmembrane protein. Truncation of NKAIN2 has been described in patients with developmental delay and complex neurological impairment. All other events found in the patient overlapped with those reported in the DGV. A full list of events can be found in Supplementary Table 1.

Thus we hypothesise that the inherited chromosome 6q22.31 duplication may compound the presence of the de novo 16p11.2 deletion, leading to the observed clinical phenotype in this patient. As most researchers focus solely on the chromosome 16p abnormality, or choose to exclude inherited CNVs, most cases of ‘dual CNV disorder’ such as this will have been missed in the literature. In an attempt to identify similar cases, we performed a PubMed search for ‘16p11.2’ that matched 130 articles (August 1990–February 2012). Fifty of these manuscripts described the characterisation of 16p11.2 anomalies, of which only nine explicitly reported information regarding concurrent CNVs. When limiting our search to cases with typical proximal (29.5–30.1 Mb) 16p11.2 anomalies, we were unable to identify any 16p11.2 cases reported to co-occur with NKAIN2 CNVs. Across the nine studies available, the only regions that were consistently reported across multiple studies as secondary CNVs in 16p11.2 patients were 15q11.2 (Prader–Willi syndrome region, found in 2 of 31 16p11.2 patients studied in Bachmann-Gagescu et al. and 1 of 427 autistic individuals studied in Marshall et al.), 15q13.2 (found in 2 of 138 16p11.2 duplication carriers studied in Jacquemont et al. and 1 of 427 autistic individuals studied in Marshall et al.) and 22q11.2 (DiGeorge syndrome critical region, found in 1 of 31 16p11.2 patients studied in Bachmann-Gagescu et al. and 1 of 36 autistic individuals studied in Davis et al.). Interestingly, these recurrent secondary CNVs align with regions known to be involved in autism and developmental delays.

When we widened our search to include novel chromosome 16p11.2 CNVs outside of the typical region (28.0–31.4 Mb), we did find one study that documented cases with co-occurring chromosome 16p11.2 and 6q22.31 abnormalities. This study, by Sanders et al., investigated 1124 individuals with autism and their unaffected family members (2248 parents and 872 sibs) and identified several recurrent copy number events associated with autism, including rearrangements of 16p11.2, both within and outside the ‘typical’ region. As part of their Supplementary data, the authors published full lists of all high-confidence CNVs found in samples passing quality control. These
The patient we describe in this letter carries the typical 16p11.2 anomalies were intronic to NKAIN1–4 transmembrane protein with four homologues (NKAIN1–4). TRDN patient encompasses all of the coding regions of the gene content. The 6q22.31 duplicated region in our case may represent a very rare CNV with little independent effect, but we cannot rule out a modifying role in combination with the 16p11.2 loss, particularly in the co-occurrence of a further 6q22.31 duplication, both of which are distinct from those described above. The latter does not overlap significantly with noted DGV variants. However, further mining of the Sanders et al. Supplementary data identified three healthy individuals carrying apparently identical 6q22.31 duplications (a father and son and another father) (Supplementary Table S8). Thus the 6q22.31 duplication in our case may represent a very rare CNV with little independent effect, but we cannot rule out a modifying role in combination with the 16p11.2 loss, particularly in view of the gene content. The 6q22.31 duplicated region in our patient encompasses all of the coding regions of TRDN and the first exon of NKAIN2. TRDN codes for a muscle-specific protein, deletion of which leads to cardiac arrhythmia. Interestingly, it is the only gene in common with the smaller 6q22.31 anomalies were intronic to NKAIN1–4 transmembrane protein with four homologues (NKAIN1–4). NKAIN2 is a transmembrane protein with four homologues (NKAIN1–4), all of which are highly conserved and have brain-specific expression. included lists of rare CNVs that did not overlap more than 50% with a CNV present at >1% frequency in the DGV. Using these Supplementary data, we were able to identify seven autistic probands who carried concurrent 16p11.2 (five duplications and two deletions) and NKAIN2 (six duplications and one deletion) anomalies, all inherited from healthy parents (Table 1). However, all of these events were small (<50 kb) and none overlapped with those observed in our patient. No 16p11.2 events were found to be concurrent with TRDN CNVs. Furthermore, the chromosome 16p11.2 anomalies identified by Sanders et al. in these concurrent cases were all outside of the typical region and, on review of the DGV, we noted that those involving the SULT1A1 (OMIM #171150) gene overlapped significantly with regions of common variation (Table 1). Similarly, the NKAIN2 anomalies were intronic to the RefSeq NKAIN2 consensus sequence, or overlapped with common CNVs in the DGV (Table 1). However, it is worth noting that we observed a mRNA, BC035062, that is annotated to include an exon contained within the common 'intronic' duplication, suggesting these events may affect a splice variant.

The patient we describe in this letter carries the typical 16p11.2 loss co-occurring with a further 6q22.31 duplication, both of which are distinct from those described above. The latter does not overlap significantly with noted DGV variants. However, further mining of the Sanders et al. Supplementary data identified three healthy individuals carrying apparently identical 6q22.31 duplications (a father and son and another father) (Supplementary Table S8). Thus the 6q22.31 duplication in our case may represent a very rare CNV with little independent effect, but we cannot rule out a modifying role in combination with the 16p11.2 loss, particularly in view of the gene content. The 6q22.31 duplicated region in our patient encompasses all of the coding regions of TRDN and the first exon of NKAIN2. TRDN codes for a muscle-specific protein, deletion of which leads to cardiac arrhythmia. Although primarily expressed in cardiac tissue, this gene is also expressed in skeletal muscle, where it is involved in the regulation of resting calcium levels. NKAIN2 is a transmembrane protein with four homologues (NKAIN1–4), all of which are highly conserved and have brain-specific expression. Interestingly, it is the only gene in common with the smaller 6q22.31 CNV regions described above. The cellular functions of the NKAIN proteins are unknown but they have been shown to localise and interact with the plasma membrane protein ATP1B1. Drosophila dNKAIN mutants show decreased co-ordination and temperature-sensitive paralysis. Microdeletions in NKAIN2 have previously been reported as rare events contributing to the risk of schizophrenia and Attention

*European Journal of Human Genetics*

## Table 1 Probands with Concurant 16p11.2 and 6q22.31 CNVs from the Sanders et al.\(^{(29)}\) Study (a) Chromosome 16p11.2 anomalies and (b) Concurrent chromosome 6q22.31 anomalies

| Individual       | 16p11.2 start (hg18) | 16p11.2 end (hg18) | 16p11.2 size | 16p11.2 State | 16p11.2 Genes | Intronic/Exonic? | DGV frequency range | Average DGV frequency |
|------------------|----------------------|-------------------|--------------|---------------|---------------|-----------------|---------------------|-----------------------|
| (a) Our patient  | 29560 500            | 30106 852         | 546352       | Deletion      | SULT1A1       | Exonic, 30 genes | 0.00–0.00            | 0.00                  |
| 11009.p1         | 28521 466            | 28528 253         | 6787         | Duplication   | SULT1A1       | Exons 1–7 (of 8) | 0.03–0.30            | 0.17                  |
| 11087.p1         | 28522 302            | 28528 253         | 5951         | Duplication   | SULT1A1       | Exons 1–7 (of 8) | 0.03–0.30            | 0.17                  |
| 11996.p1         | 28522 302            | 28528 253         | 5951         | Duplication   | SULT1A1       | Exons 1–7 (of 8) | 0.03–0.30            | 0.17                  |
| 12961.p1         | 28527 444            | 28528 253         | 5951         | Duplication   | SULT1A1       | Exons 1–7 (of 8) | 0.03–0.30            | 0.17                  |

| (b) Individual   | 6q22.31 start (hg18) | 6q22.31 end (hg18) | 6q22.31 size | 6q22.31 State | 6q22.31 Genes | Intronic/Exonic? | DGV frequency range | Average DGV frequency |
|------------------|----------------------|-------------------|--------------|---------------|---------------|-----------------|---------------------|-----------------------|
| Our patient      | 123581 324           | 124201 824        | 620500       | Duplication   | TRDN, NKAIN2  | Exons 1–41 (of 41), | 0.00–0.00            | 0.00                  |
| 11009.p1         | 124477 640           | 124510 591        | 32951        | Duplication   | NKAIN2        | Exon 1 (of 6)    | 0.001–0.029           | 0.01                  |
| 11087.p1         | 124477 640           | 124510 591        | 32951        | Duplication   | NKAIN2        | Exon 2 of BC035062 mRNA | 0.001–0.029           | 0.01                  |
| 11096.p1         | 124479 205           | 124510 591        | 31386        | Duplication   | NKAIN2        | Exon 2 of BC035062 mRNA | 0.001–0.029           | 0.01                  |
| 11229.p1         | 124480 321           | 124510 591        | 30270        | Duplication   | NKAIN2        | Exon 2 of BC035062 mRNA | 0.001–0.029           | 0.01                  |
| 11246.p1         | 124477 640           | 124510 591        | 32951        | Duplication   | NKAIN2        | Exon 2 of BC035062 mRNA | 0.001–0.029           | 0.01                  |
| 11996.p1         | 124959 283           | 124961 396        | 2113         | Duplication   | NKAIN2        | Exon 2 of BC035062 mRNA | 0.001–0.029           | 0.01                  |
| 12961.p1         | 124477 640           | 124510 591        | 32951        | Duplication   | NKAIN2        | Exon 2 of BC035062 mRNA | 0.001–0.029           | 0.01                  |

Abbreviation: DGV, database of genomic variants.

*frequency DGV gives the average frequency of deletions or duplications (as appropriate) in all DGV studies including at least 30 European individuals in the DGV.

4DGV frequency range gives the frequency range of deletions or duplications (as appropriate) in all DGV studies including at least 30 European individuals in the DGV.
Deficit/ Hyperactivity Disorder (ADHD),38 and variants within this gene have been associated with neuroticism,39

Taking all of this information into consideration, we believe that the clinical presentation of chromosome 16p11.2 deletion cases may be modulated by the presence of additional genomic imbalances, such as the inherited duplication of chromosome 6q22.31 observed in our case. Researchers of developmental disorders have proposed a dual CNV model at other loci,33,40–42 as well as compound heterozygotes with a CNV-mediated deletion of one allele and non-synonymous mutation of the other (mixed genomic disorders).33,44

The genetic background, of course, extends beyond CNVs and, as genetic technologies advance, we predict that a whole-genome view will allow the elucidation of many combinatorial factors. For example, a recent study extended the dual CNV model to incorporate rare point mutations across common functional pathways, where an ASD proband was identified with both a de novo mutation of FOXP1 (OMIM #605515) and an inherited mutation of CNTNAP2 (OMIM #604569).45 The validity of this model and the significance of concurrent CNVs can only be tested by the consistent and detailed description of CNV cohorts in a whole-genome context. This is especially true for studies such as ours, which involve only a single patient. We would therefore urge researchers characterising chromosome abnormalities to consider, and to explicitly report, the anomalies in the context of whole genome copy number variation and genomic cataloguing. Advances in genetic technology mean that there is no longer a need to consider genomic imbalances in isolation, particularly in case reports. We suggest that the capture of complete genomic contexts, alongside detailed phenotypic profiling, will allow us to develop a better understanding of the variability of the chromosome 16p11.2 phenotype and may assist in the delineation of a core clinical phenotype.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We would like to thank the family for their participation in the study, Sonja Vernes for her assistance with DNA preparations and Jane Hurst for her support of this project. We would also like to thank the Sanders group for their comprehensive, publically available data. This work was supported by the NIHR Biomedical Research Centre, Oxford with funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health. The project was also supported by the Wellcome Trust [090532/Z/09/Z] and the MRC [G1000569/1]. Simon Fisher was supported by the Simons Foundation Fellowship and a Junior Research Fellow at St John's College, Oxford. Jenny Taylor is funded by the NIHR Biomedical Research Centre.

Dianne F Newbury1,9, Francesca Mari2,3,9, Elham Sadigih Akha1,4, Kay D MacDermot1, Roberto Canitano4, Anthony P Monaco1, Jenny C Taylor1,14, Alessandra Renieri1,2, Simon E Fisher1,2,8, and Samantha JL Knight1,14

1Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 2Medical Genetics, University of Siena, Siena, Italy; 3Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy; 4NIHR Biomedical Research Centre, Oxford, UK; 5Kennedy Galton Centre (North West Thames Regional

Genetic service), Imperial College, London, UK; 6Child Neuropsychiatry, Azienda Ospedaliera Universitaria Senese, Siena, Italy; 7Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; 8Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands E-mail: sknight@well.ox.ac.uk

*These authors contributed equally to this work.
Do regulatory regions matter in FOXG1 duplications?

European Journal of Human Genetics (2013) 21, 365–366; doi:10.1038/ejhg.2012.142; published online 4 July 2012

Duplications of FOXG1 gene at 14q12 have been reported in patients with infantile spasms and developmental delay of variable severity.1,2,3 FOXG1 encodes the forkhead protein G1, a brain-specific transcriptional repressor, regulating corticogenesis in the developing brain and neuronal stem cell self-renewal in the postnatal brain.4 Recently, Amor et al.5 reported on this journal an interstitial duplication of ~88 kb at 14q12 in a father–son pair with hemipalial microsoma and normal neurocognitive phenotype. The duplication contains only two polypeptide-encoding genes, FOXG1 and CHFP23, suggesting that FOXG1 duplication may be benign or at least incompletely penetrant. That makes the involvement of FOXG1 duplication in the pathogenesis of the neurocognitive impairment and epilepsy controversial. As also discussed by Brunetti-Pierri et al.,6 we feel that this statement needs special caution.

Functional consequences of chromosomal microduplication and microdeletion rely on the final gene dosage, which is strongly influenced by the location of the breakpoint. In this context, the understanding of the contribution of regulatory sequences in gene transcription is critical to understand the relationship between CNVs and human diseases. With this purpose, the Encyclopedia of DNA Elements (ENCODE) project has recently performed a systematic analysis of transcriptional regulation in different human cell lines, providing new understanding about transcription start sites, including their relationship with specific regulatory sequences and histone modification and features of chromatin accessibility.7,8 Interestingly, analysis of histone modifications from the ENCODE project revealed the presence of a putative regulatory element upstream FOXG1 gene between 28188 and 28217 kb (UCSC genome browser, NCBI Build 36/hg18) (Figure 1). This conserved region localizes about 130 kb upstream FOXG1 gene and contains histone modifications typical of enhancers of gene transcription (e.g., histone H3 and Lysine 4 monomethylation) in eight different human cell lines. Analysis of regulatory potential scores, comparing frequencies of short alignment patterns between known regulatory elements and neutral DNA,9 also disclose two additional putative elements typical of cis-regulatory modules within this region (Figure 1). Moreover, it contains a DNaes hypersensitive site (DHS). DHSs reflect genomic regions thought to be enriched for regulatory information and many DHSs reside at or near transcription start site. Notably, no other polypeptide-encoding genes or non-coding RNAs and pseudogenes are present in the region, suggesting that this regulatory element might regulate FOXG1 transcription. Analysis of duplication breakpoints previously reported on 14q12 revealed that duplications associated with an epileptic phenotype localizes uniquely upstream this regulatory element, whereas downstream duplications were identified only in the cases without seizures (Figure 1). On the basis of this finding, we suggest that FOXG1 duplication including this putative regulatory region allows the efficient transcription of the supernumerary copy of FOXG1 gene, resulting in an effective increase in FOXG1 expression and, thereby, in brain hyperexcitability. In contrast, duplications starting downstream this putative regulatory site do not allow efficient transcription of FOXG1, which may underlie the lack of neurological phenotype in the case reported by Amor et al.8 Even if the functional relevance of this putative long-range regulatory element on FOXG1 transcription deserves to be experimentally verified, it provides an interesting clue to dissect