nature genetics

© 2007 Nature Publishing Group http://www.nature.com/naturegenetics

Discovery of a previously unrecognized microdeletion syndrome of 16p11.2–p12.2

Blake C Ballif[1], Sara A Hornor[2], Elizabeth Jenkins[3], Suneeta Madan-Khetarpal[3], Urvashi Surti[4,5], Kelly E Jackson[6], Alexander Asamoah[6], Pamela L Brock[6], Gordon C Gowans[6], Robert L Conway[7], John M Graham, Jr[7], Livija Medne[8], Elaine H Zackai[8], Tamim H Shaikh[8], Joel Geoghegan[9], Rebecca R Selzer[9], Peggy S Eis[9], Bassem A Bejjani[1,2,10] & Lisa G Shaffer[1,2]

We have identified a recurrent *de novo* pericentromeric deletion in 16p11.2–p12.2 in four individuals with developmental disabilities by microarray-based comparative genomic hybridization analysis. The identification of common clinical features in these four individuals along with the characterization of complex segmental duplications flanking the deletion regions suggests that nonallelic homologous recombination mediated these rearrangements and that deletions in 16p11.2–p12.2 constitute a previously undescribed syndrome.

Pericentromeric regions of the genome are structurally complex regions adjacent to the centromeres that are enriched for repetitive sequence elements and segmental duplications[1]. This abundance of segmental duplications seems to have made the pericentromeric regions susceptible to deletion or rearrangement[2]. We screened 8,789 consecutive patients with developmental disabilities whose clinical specimens were submitted to our laboratory for analysis with the SignatureChip targeted microarray[3] (**Supplementary Methods** online), which includes a minimum of three to six overlapping BAC clones at the most proximal end of the pericentromeric region for each chromosome arm (excluding the short arms of the acrocentric chromosomes). Four individuals had recurrent deletions in 16p11.2–p12.2 (subjects 1, 2, 3 and 4) (**Fig. 1**). Parental analyses in three of the four subjects (1, 3 and 4) demonstrated that these are *de novo* chromosome abnormalities. Parents were unavailable for testing for subject 2. To clarify the sizes of the deletions further, we analyzed all four subjects with a high-density BAC-based microarray spanning ∼5 Mb of the most proximal unique sequence adjacent to the centromere for all 43 unique pericentromeric regions of the human genome[4] (**Fig. 1a**). Because all four of these abnormalities extended beyond the ∼5-Mb coverage of this pericentromeric array, we characterized the full extent of each abnormality using NimbleGen whole-genome oligonucleotide arrays (for subjects 1, 2 and 4) and/or Affymetrix 250K SNP arrays (for subjects 3 and 4) (**Fig. 1c,d**). FISH analysis using BAC clones that map to the various breakpoint regions confirmed the results of the whole-genome arrays (**Fig. 1a**). Notably, all four deletions in 16p11.2–p12.2 shared the same distal breakpoint, located ∼21.4 Mb from the 16p telomere. However, the proximal breakpoints were ∼28.5 Mb (subject 2), ∼29.3 Mb (subjects 3 and 4) and ∼30.1 Mb (subject 1) from the 16p telomere, resulting in overall deletion sizes of ∼7.1 Mb, ∼7.9 Mb and ∼8.7 Mb, respectively. Computational analysis of the 16p11.2–p12.2 region using the University of California Santa Cruz (UCSC) genome browser (http://genome.ucsc.edu) and the Human Genome Segmental Duplication Database (http://projects.tcag.ca/humandup/) identified a complex arrangement of segmental duplications, some of which directly flanked the deletion breakpoints (**Fig. 1e** and **Supplementary Note** online).

Misalignment of segmental duplications in meiosis followed by nonallelic homologous recombination (NAHR) can generate microdeletions, microduplications and inversions of the intervening genomic sequence, depending on the orientation of the duplicated segments[2,5–7]. The locations of the breakpoints in these four subjects with respect to the location and orientation of the segmental duplications in the 16p11.2–p12.2 region suggest that NAHR mediated these rearrangements (**Fig. 1e**). Although the proximal breakpoint in subject 2 seems to be atypical, in that it does not seem to have a paired segmental duplication at the distal breakpoint, this is not without precedent. For example, some of the more rare rearrangements of 17p11.2 associated with Smith-Magenis syndrome do not have breakpoints that fall within the typical paired segmental duplications and may not be associated with known genomic architectural features[8].

The clinical features of the four subjects with microdeletions in 16p11.2–p12.2 include distinct facial features, including flat facies, downslanting palpebral fissures, low-set and malformed ears and eye anomalies (**Supplementary Table 1** online and **Fig. 1f–i**). Other commonly described features include orofacial clefting, heart defects, frequent ear infections with potential hearing loss, short stature, minor hand and foot anomalies, feeding difficulties, hypotonia and cognitive and developmental delays (**Supplementary Table 1**).

To our knowledge, only one other individual has been reported in the literature with a deletion in the 16p11.2 region. This individual, identified in ref. 9 using conventional CGH, was a 5-month-old male proband with distinct craniofacial features, including flat facies, microretrognathia, blepharophimosis, hypoplasic alae nasi and absent

[1]Signature Genomic Laboratories, Spokane, Washington 99202, USA. [2]Health Research and Education Center, Washington State University, Spokane, Washington 99210, USA. [3]Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [4]Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA. [5]Magee-Womens Hospital, Pittsburgh, Pennsylvania 15213, USA. [6]Weisskopf Child Evaluation Center, Louisville, Kentucky 40202, USA. [7]Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. [8]The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [9]NimbleGen Systems, Madison, Wisconsin 53711, USA. [10]Sacred Heart Medical Center, Spokane, Washington 99204, USA. Correspondence should be addressed to L.G.S. (lshaffer@wsu.edu).

Received 16 February; accepted 20 June; published online 19 August 2007; doi:10.1038/ng2107

© 2007 Nature Publishing Group http://www.nature.com/naturegenetics

a Subject 1

b Subject 5

Log2 ratio of means

(Mb)

c Subject 1 Subject 4 Subject 2 Subject 5

d Subject 3

20 Mb 21 Mb 22 Mb 23 Mb 24 Mb 25 Mb 26 Mb 27 Mb 28 Mb 29 Mb 30 Mb 31 Mb

e Chromosome band 16p12.2 16p12.1 16p11.2

338J22 141O15 713L8 5.7 Mb 15I10 264B17 74E23

TEL CEN

A1 B1 A2 A3 B2 C1 C2 D1 C3 C4 D2 C5 A4 A5 A6

Subject 1 8.7 Mb del

Subject 4 7.9 Mb del

Subject 3 7.9 Mb del

Subject 2 7.1 Mb del

Subject 5 1.1 Mb dup, 5.7 Mb trp

Copy number–variant regions

OTOA *SCNN1G* *SCNN1B* *COG7* *PRKCB1* *SLC5A11* *IL21R* *IL4R* *JMJD5* *CLN3* *ATP2A1* *ALDOA* *TBX6*

f Subject 1 g Subject 2 h Subject 3 i Subject 3 j Subject 5

Figure 1 Analysis of individuals with copy-number imbalances of 16p11.2–p12.2. (**a**) Pericentromeric array CGH profile and FISH confirmation (arrow) of a >4.99-Mb deletion in 16p11.2–p12.2 in subject 1. For the array CGH plots, clones are ordered on the *x*-axis according to physical mapping positions. The top plot shows a normal chromosome 16; the bottom plot shows the abnormal chromosome 16. (**b**) Pericentromeric array CGH profile and metaphase and interphase (inset) FISH confirmation (arrows) of a *de novo* triplication of ~4 Mb of 16p11.2–p12.1 and duplication of ~1 Mb of 16p12.1–p12.2 in subject 5. Microarray plots are arranged as in **a**. (**c**) NimbleGen whole-genome oligonucleotide array CGH profiles for subjects 1, 2, 4 and 5. (**d**) Affymetrix 250K SNP array profile for subject 3. (**e**) Schematic of the 16p11.2—p12.2 region with a summary of the abnormalities identified in five subjects. A simplified interpretation of the segmental duplications located in 16p11.2–p12.2 (see **Supplementary Note**) is shown with relative orientation of each duplication (arrows). The block of segmental duplications shown in green does not share identity with the 16p11.2–p12.2 region. Some of the RP11 BAC clones used for FISH confirmation of abnormalities are shown as black dots along the chromosome. Red bars indicate deleted regions for each subject. Light and dark purple bars for subject 5 indicate regions of duplication and triplication, respectively. Orange bars indicate regions of copy number variation based on the Database of Genomic Variants. The locations of select genes from >100 known genes in the region are shown. (**f**) Subject 1 at 13 years of age. Subject 1 has had multiple reconstructive surgeries for cleft lip and palate, multiple mandibular distraction, tracheostomy and cholesteatoma. (**g**) Subject 2 at 7 years of age. (**h**,**i**) Subject 3 at 3 years of age. (**j**) Subject 5 at 10 years of age.

nasal bridge, low-set and malformed ears and glossoptosis with hypoplastic palate. The proband was also described as having heart defects (tetralogy of Fallot with pulmonary atresia), eye abnormalities and other features that were similar to those described for the four subjects identified here (**Supplementary Table 1**). Although the precise breakpoints for the deletion in the individual described in ref. 9 were not determined, based on the brief clinical description, this individual probably contains a similar microdeletion in 16p11.2–p12.2.

Among the 8,789 patients screened in our laboratory, we also identified an individual with a complex *de novo* pericentromeric abnormality (subject 5) involving a duplication and triplication of the same region of 16p11.2–p12.2 that is deleted in the four microdeletion subjects. Because this abnormality extended beyond the ~5-Mb coverage of the pericentromeric microarray (**Fig. 1b**), we performed whole-genome oligonucleotide array CGH (NimbleGen) to refine the breakpoints further (**Fig. 1c**). The breakpoints in this subject also cluster at segmental duplications (**Fig. 1e**). By this analysis, the distal duplication was determined to be ~1.1 Mb in size, with the same distal breakpoint as all four individuals with 16p11.2–p12.2 microdeletions. The triplicated segment is ~5.7 Mb in size and is also flanked by segmental duplications. We confirmed these duplication and triplication breakpoints by FISH using BAC clones mapping within these regions of 16p11.2–p12.2 (**Fig. 1b** and **Supplementary Table 2** online). The clinical features of subject 5 are listed in **Supplementary Table 1**, and facial features are illustrated in **Figure 1j**.

© 2007 Nature Publishing Group http://www.nature.com/naturegenetics

Although we have not identified an exact reciprocal duplication product of one of the deletions in this region, two subjects with tandem duplications of 16p11.2–p12.2 have been reported and may constitute the reciprocal duplication product of this microdeletion syndrome[10]. Given the complex distribution of segmental duplications within 16p11.2–p12.2 (ref. 11), we anticipate that other deletions, duplications, triplications and inversions will be observed in this region of 16p11.2–p12.2. Indeed, the complex duplication and triplication observed here is most likely mediated by NAHR, and inversions in the distal region of 16p12.1 have already been reported in the Database of Genomic Variants (http://projects.tcag.ca/variation/).

Although **Supplementary Table 1** reports the major clinical findings in our subjects with 16p11.2–p12.2 imbalances, genotype-phenotype correlations are more difficult to establish for large microdeletions encompassing a substantial number of genes. Of the 104 RefSeq genes located within the largest deleted region of 16p11.2–p12.2, there are no obvious candidates for all of the features of these syndromes, nor are any known to be dosage sensitive. However, at least 13 genes (**Supplementary Table 3** online) are known to be associated with various genetic diseases in the autosomal recessive state or have functions potentially relevant to the clinical features of these patients (**Fig. 1e**).

The smallest region of overlap (SRO) for the four microdeletions in 16p11.2–p12.2 is ~7.1 Mb, equivalent in size to the deletion of subject 2. However, analysis of the Database of Genomic Variants suggests that structural and copy number variants (CNVs) exist within this region of 16p11.2–p12.2. These variants consist of inversions and copy number gains and losses identified in the apparently normal population. The locations of these CNVs relative to the SRO are shown in **Figure 1e**. Assuming that regions of common copy number variation do not harbor dosage-sensitive genes[6], they may be useful in refining the location of dosage-sensitive genes responsible for the phenotypic features of microdeletions in 16p11.2–p12.2. Of the 104 known RefSeq genes located within the 8.7-Mb region defined by the largest of the 16p11.2–p12.2 deletions identified in this study, only 33 are not located within regions of copy number variation. However, refinement of the SRO based on CNVs may be problematic, given the relatively recent discovery of the prevalence of CNVs and our immature understanding of them[12–14].

We have established the identity of a previously unknown microdeletion syndrome of 16p11.2–p12.2 by analyzing individuals with mental retardation, developmental delay or dysmorphic features with array CGH. The screening of additional individuals using targeted BAC arrays and/or higher-resolution whole-genome oligonucleotide arrays is likely to uncover other new microdeletion and microduplication syndromes, thus adding to our growing knowledge of the cytogenetic basis of developmental disabilities.

For the subjects with 16p11.2–p12.2 abnormalities described here, we obtained informed consent to perform high-resolution molecular cytogenetic testing and to publish photographs, using a consent form approved by the Washington State University Institutional Review Board (for subjects 1, 2, 3 and 5) or by the Children's Hospital of Philadelphia Institutional Review Board (for subject 4).

Microarray data can be found at ArrayExpress under accession codes E-MEXP-1148 (Affymetrix platform) and E-TABM-286 (NimbleGen platform).

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

ACKNOWLEDGMENTS
The authors thank all of the families who participated in this study. We also recognize A. Theisen (Signature Genomic Laboratories) for his editorial assistance. This work was partly funded by a grant from the US National Institutes of Health (NIH) (HD047280-02) to L.G.S. (Washington State University). We appreciate the support from SHARE's Childhood Disability Center, the Steven Spielberg Pediatric Research Center, the Cedars-Sinai Burns and Allen Research Institute, the Skeletal Dysplasias NIH/National Institute of Child Health and Human Development Program (grant HD22657-11) and the Medical Genetics NIH/National Institute of General Medical Sciences Training Program (grant GM08243).

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

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. She, X. *et al. Nature* **430**, 857–864 (2004).
2. Stankiewicz, P. & Lupski, J.R. *Trends Genet.* **18**, 74–82 (2002).
3. Bejjani, B.A. *et al. Am. J. Med. Genet. A.* **134**, 259–267 (2005).
4. Ballif, B.C. *et al. Genet. Med.* **9**, 150–162 (2007).
5. Koolen, D.A. *et al. Nat. Genet.* **38**, 999–1001 (2006).
6. Sharp, A.J. *et al. Nat. Genet.* **38**, 1038–1042 (2006).
7. Shaw-Smith, C. *et al. Nat. Genet.* **38**, 1032–1037 (2006).
8. Stankiewicz, P. *et al. Am. J. Hum. Genet.* **72**, 1101–1116 (2003).
9. Hernando, C. *et al. J. Med. Genet.* **39**, E24 (2002).
10. Finelli, P. *et al. J. Med. Genet.* **41**, e90 (2004).
11. Martin, J. *et al. Nature* **432**, 988–994 (2004).
12. Iafrate, A.J. *et al. Nat. Genet.* **36**, 949–951 (2004).
13. Redon, R. *et al. Nature* **444**, 444–454 (2006).
14. Sebat, J. *et al. Science* **305**, 525–528 (2004).