

## DATA REPORT

## Characteristics of rare and private deletions identified in phenotypically normal individuals

Keiko Shimojima<sup>1,2</sup> and Toshiyuki Yamamoto<sup>1,2</sup>

Genomic copy number variations (CNVs) identified through chromosomal microarray testing must be validated to confirm whether they are pathogenically and functionally relevant to their respective clinical features. Although larger deletions have a higher probability to be pathogenic, this is not always true. Phenotypically normal individuals showed five CNV deletions larger than 1.5 Mb. The genes related to autosomal dominant trait were absent within these CNV deletions.

*Human Genome Variation* (2017) 4, 17037; doi:10.1038/hgv.2017.37; published online 14 September 2017

Chromosomal microarray testing (CMT) has been established as a diagnostic tool for detecting genomic copy number variations (CNVs). In 2010, CMT was suggested as a first-tier screening for patients with developmental delay (DD) due to unknown etiology.<sup>1</sup> The identified CNVs should be validated through several steps.<sup>2</sup> First, the identified CNVs are checked for any relationship to known chromosomal disorders or recurrently identified genomic deletion/duplication syndromes. If the identified CNVs do not correlate with any known syndromes, further analyzes must be performed via PUBMED (<https://www.ncbi.nlm.nih.gov/pubmed>), CNV databases such as dbVAR (<https://www.ncbi.nlm.nih.gov/dbvar/>), and CNV morbidity map of DD (<https://www.ncbi.nlm.nih.gov/dbvar/studies/nstd100/>) to evaluate whether the same CNVs have been previously reported. If the pathogenicity of CNVs is difficult to define, the test laboratories can deposit the genotype/phenotype data to online databases such as DECIPHER (<https://decipher.sanger.ac.uk/>) in order to curate the pathogeni-

city of those CNVs gradually. However, the publicly available information on genomic variants, especially rare/unique variants, may not always be accurate. Thus, the pathogenicity of genomic variants should be reviewed periodically.

If the same CNVs cannot be found in the databases or the literature, parental examination is recommended to evaluate whether the CNVs were inherited or *de novo*.<sup>3</sup> However, *de novo* occurrences *per se* cannot be regarded as evidence to support pathogenicity. In contrast, inherited CNVs are more complicated due to phenotypic variability, incomplete penetrance, and a recessive mode of inheritance. The most frequently observed 22q11.2 deletions and duplications are often shared among relatives possibly due to incomplete penetrance.<sup>4,5</sup> Thus, a clear classification of whether the CNVs are pathogenic or benign is not always easy in such cases.

At present, we have analyzed ~2,000 samples from patients and their relatives associated with DD for the possible presence of

**Table 1.** Summary of the deletions larger than 1.5 Mb identified in healthy individuals

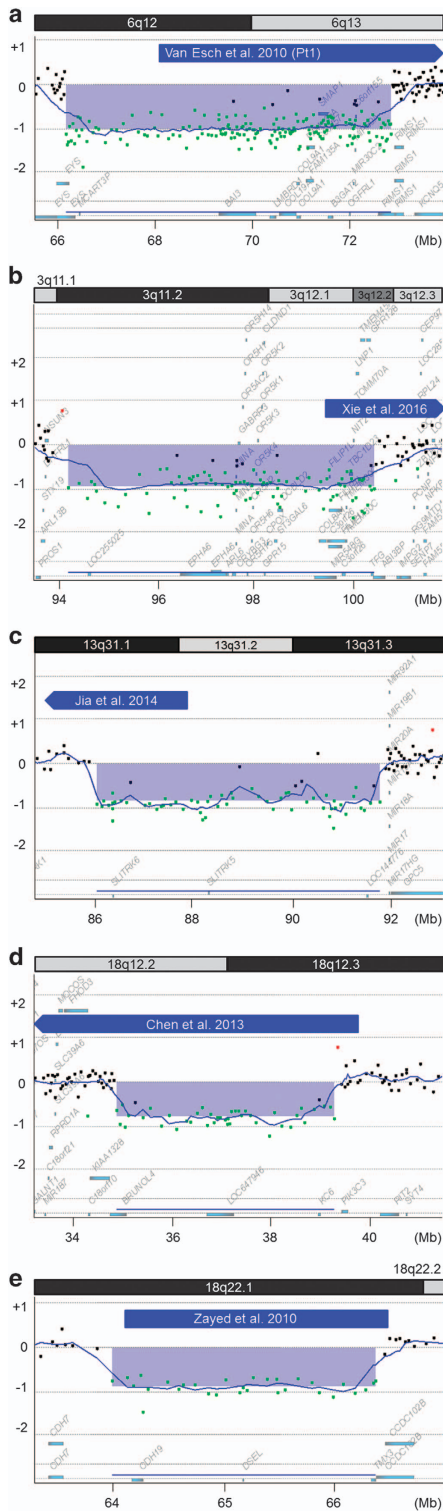
Subject	Chromosome	Band	Start <sup>a</sup>	End <sup>a</sup>	Size (Mb)	RefSeq gene numbers included	OMIM gene numbers included <sup>b</sup>	Numbers of the gene related to haploinsufficiency scores <sup>c</sup> (Autosomal dominant trait)	Notes
1	chr6	q12–q13	66,231,424	72,860,605	6.63	16	4 (2)	1 (0)	Paternally derived
2	chr3	q11.2–q12.2	94,258,997	100,487,464	6.23	36	3 (1)	1 (0)	Paternally derived
3	chr13	q31.1–q31.3	86,046,386	91,816,350	5.77	18	1 (0)	0 (0)	Paternally derived
4	chr18	q12.2–q12.3	34,959,559	39,389,665	4.43	4	0	0 (0)	Maternally derived
5	chr18	q22.1	63,993,067	66,381,438	2.39	5	0	0 (0)	Identified in a normal individual
ex	chr22	q11.21	18,919,942	21,440,514	2.52	60	12 (1)	5 (1)	

Abbreviation: ex, a representative example of typical 22q11.2 deletion.

<sup>a</sup>Genomic positions referred to build19.

<sup>b</sup>The numbers in parentheses indicate the numbers of the genes related to autosomal dominant trait.

<sup>c</sup>This is checked through ClinGen Dosage Sensitivity Curation Page (<https://www.ncbi.nlm.nih.gov/projects/dbvar/clingen/>).



**Figure 1.** Chromosomal microarray testing results. The identified deletions are shown in Gene View created by the Agilent Genomic Workbench (Agilent Technologies). (a) subject 1, (b) subject 2, (c) subject 3, (d) subject 4, and (e) subject 5. Previously reported deletions that overlapped with the deletions reported in this study are depicted by blue rectangles or pentagons for comparison. Directions suggested by the pentagons indicate large deletions beyond the edge. All genomic positions are unified to the hg19 assembly.

CNVs. Through this analysis, definite pathogenic CNVs were detected in 17% of these samples.<sup>4</sup> The CNVs that are frequently observed and already indexed in existing databases can be easily interpreted; however, rare/private CNVs are not, and hence cannot be cross-referenced in such databases. Benign CNVs could also be primarily detected in healthy parents in some cases. Furthermore, the pathogenicity of CNVs is related to their size. Larger CNVs are typically suspected to be pathogenic, whereas smaller CNVs are usually shared within the families and are considered benign.<sup>2,6,7</sup> However, there are exceptional cases of large but benign CNVs.

In this study, we analyzed the characteristics of large CNVs that were identified in healthy individuals to better understand their role in the genome. For this purpose, unique CNVs larger than 1.5 Mb that were only identified in a single family or healthy individuals were selected, since most CNVs larger than 1.5 Mb are usually pathogenic (affected versus control = 52/1,782; odds ratio = 20.3) and most benign CNVs are less than 1.5 Mb.<sup>6</sup> The criteria used for the selection of CNVs include (1) larger than 1.5 Mb, (2) identified in healthy individuals, (3) unique and not repeatedly detected in our laboratory, and (4) previously not reported as pathogenic. This study was performed in accordance with the Declaration of Helsinki, and permission was obtained from the ethical committee of our institution. Blood samples were collected from patients and their families after obtaining written informed consent. CMT was then performed using the Agilent microarray 60 K/105 K (Agilent Technologies, Santa Clara, CA, USA) as described previously.<sup>8</sup> CNVs were detected using the Agilent Genomic Workbench (Agilent Technologies), and the CNV sizes were measured using the report function. Upon identification of unique CNVs, a trio analysis via microarray or fluorescence *in situ* hybridization (FISH) was performed using the parental samples. The identified CNVs were then indexed into a database to be evaluated for their frequency and cross-referenced with the GRCh37/hg19 genome assembly via the UCSC genome browser (<https://genome.ucsc.edu/>).

As a result, five deletions were confirmed to fulfill our criteria (Table 1). Clinical information of the subjects is summarized in Supplementary Information. The number of RefSeq (<https://www.ncbi.nlm.nih.gov/refseq/>) and OMIM (<http://omim.org/>) genes included in the deleted regions was evaluated. The dosage effects (haploinsufficiency scores) were also checked through ClinGen Dosage Sensitivity Curation Page (<https://www.ncbi.nlm.nih.gov/projects/dbvar/clingen/>). The largest deletion identified via the 60 K array (the other four deletions were identified through the 60 K array) was a 6.63-Mb deletion in 6q12–q13 (Figure 1a). This deletion partially overlaps with a previously reported pathogenic deletion.<sup>9</sup> The second largest deletion identified was a 6.23-Mb deletion in 3q11.2–q12.2 (Figure 1b). Although similar deletions in 3q12 have been previously reported, only a partial overlap was observed with this interstitial deletion.<sup>10</sup> The third largest deletion was a 5.77-Mb deletion in 13q31.1–q31.3 (Figure 1c). Previously reported deletions in 13q were fairly large; hence, smaller deletions were thought to be quite rare.<sup>11,12</sup> Additionally, a previously reported deletion was very similar to this deletion; however, it had been identified in a fetus through non-invasive prenatal testing.<sup>13</sup> The fourth was a 4.43-Mb deletion in 18q12.2–q12.3 (Figure 1d). An overlapping deletion has been previously reported in a fetus with unknown pathogenicity.<sup>14</sup> The fifth largest deletion was a 2.39-Mb deletion in 18q22.1 (Figure 1e). A similar CNV has been reported in a patient with a late-presenting diaphragmatic hernia and microphthalmia; however, the CNV was maternally inherited.<sup>15</sup>

The five deletions larger than 1.5 Mb identified in this study were from healthy individuals. The fifth deletion in the 18q22.1 region was first identified in a healthy individual, whereas the other identified deletions were shared between the patients with DD and their parents. We cannot rule out the possibility that these deletions may be related to any clinical feature with incomplete

penetrance or autosomal recessive inheritance. Even if the pathogenicity of these deletions cannot be clearly determined, it is still important to index these findings into the public databases.

The most characteristic feature among these five CNVs was the sparse distribution of genes in the corresponding regions. The human genome is approximately 3,000 Mb in size and the OMIM database contains gene descriptions for 15,394 genes (<https://omim.org/statistics/entry>). Hence, we expected an average density of 5 OMIM genes per 1 Mb of genomic region. As such, typical deletions in 22q11.21 include 12 OMIM genes within a 2.52-Mb region (Table 1). However, the density of OMIM genes within the deleted regions identified in this study was smaller than the expected average. Furthermore, all five deletions identified in this study do not include genes with haploinsufficiency scores in association with autosomal dominance trait (evaluated by the ClinGen Dosage Sensitivity Curation Page; Table 1; Supplementary Information). These results were finally evaluated when this manuscript was submitted. Because most of the genes within the relevant regions have not been curated yet and database will be updated periodically, these results may change in the future. In contrast, the typical deletions in 22q11.21 include a gene associated with haploinsufficiency scores as the autosomal dominant trait (Table 1). These findings suggest that we may not be able to evaluate observed CNVs without first checking OMIM genes and their haploinsufficiency scores, even if those CNVs are larger than 1.5 Mb.

As mentioned above, two deletions similar to those identified in this study (deletions in 13q31.1–q31.3 and 18q12.2–q12.3) were primarily identified through prenatal screening. Such rare CNVs identified via prenatal screening often present diagnostic challenges owing to the limited availability of clinical information about the fetus. Therefore, we suggest that indexing these potentially benign but rare CNVs would be beneficial for researchers to allow for evaluation of CNVs with unknown significance. However, compiling a complete list of such potentially benign CNVs would be perpetual because of the large number of benign CNVs being continuously reported. This study also reaffirmed that trio analysis is necessary for the evaluation of individual rare and private CNVs.

## HGV DATABASE

The relevant data from this Data Report are hosted at the Human Genome Variation Database at <http://dx.doi.org/10.6084/m9.figshare.hgv.1399>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1405>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1408>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1411>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1414>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1417>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1420>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1423>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1426>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1429>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1432>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1435>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1438>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1441>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1444>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1447>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1450>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1453>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1456>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1459>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1462>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1465>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1468>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1471>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1474>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1477>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1480>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1483>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1486>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1489>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1492>; [http://dx.doi.org/10.6084/m9.](http://dx.doi.org/10.6084/m9.figshare.hgv.1495)

[figshare.hgv.1498](http://dx.doi.org/10.6084/m9.figshare.hgv.1501); <http://dx.doi.org/10.6084/m9.figshare.hgv.1501>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1504>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1507>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1510>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1513>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1516>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1519>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1522>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1525>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1528>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1531>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1534>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1537>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1540>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1543>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1546>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1549>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1552>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1555>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1558>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1561>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1564>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1567>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1570>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1573>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1576>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1579>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1582>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1585>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1588>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1591>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1594>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1597>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1600>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1603>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1606>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1609>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1612>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1615>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1618>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1621>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1624>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1627>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1630>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1633>; <http://dx.doi.org/10.6084/m9.figshare.hgv.1636>.

## ACKNOWLEDGEMENTS

We would like to express our gratitude to the participants to this study and their attending medical doctors for their cooperation. This work was supported by the Practical Research Project for Rare/Intractable Diseases from Japan Agency for Medical Research and development (AMED), a Grant-in-Aid for Scientific Research from Health Labor Sciences Research Grants from the Ministry of Health, Labor and Welfare, Japan, and JSPS KAKENHI Grant Number 15K09631. We certify that we have read the journal's position regarding issues pertaining to ethical publication, and affirm that this report is consistent with the guidelines.

## COMPETING INTERESTS

The authors declare no conflict of interest.

## PUBLISHER'S NOTE

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP *et al*. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010; **86**: 749–764.
- Buysse K, Delle Chiaie B, Van Coster R, Loeys B, De Paepe A, Mortier G *et al*. Challenges for CNV interpretation in clinical molecular karyotyping: lessons learned from a 1001 sample experience. *Eur J Med Genet* 2009; **52**: 398–403.
- ACOG. Committee Opinion No. 581: the use of chromosomal microarray analysis in prenatal diagnosis. *Obstet Gynecol* 2013; **122**: 1374–1377.

- 4 Yamamoto T, Shimojima K, Shimada S, Yokochi K, Yoshitomi S, Yanagihara K *et al.* Clinical impacts of genomic copy number gains at Xq28. *Hum Genome Var* 2014; **1**: 14001.
- 5 Ou Z, Berg JS, Yonath H, Enciso VB, Miller DT, Picker J *et al.* Microduplications of 22q11.2 are frequently inherited and are associated with variable phenotypes. *Genet Med* 2008; **10**: 267–277.
- 6 Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* 2011; **43**: 838–846.
- 7 Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P *et al.* Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 2014; **46**: 1063–1071.
- 8 Shimojima K, Okamoto N, Suzuki Y, Saito M, Mori M, Yamagata T *et al.* Subtelomeric deletions of 1q43q44 and severe brain impairment associated with delayed myelination. *J Hum Genet* 2012; **57**: 593–600.
- 9 Van Esch H, Rosser EM, Janssens S, Van Ingelghem I, Loeys B, Menten B. Developmental delay and connective tissue disorder in four patients sharing a common microdeletion at 6q13-14. *J Med Genet* 2010; **47**: 717–720.
- 10 Xie L, Luo X, Yang J, Wang J, Nie C, Wang Z. A Chinese patient with Toriello-Carey syndrome and an interstitial deletion of 3q. *Am J Med Genet A* 2016; **173**: 721–726.
- 11 Poreau B, Lin S, Bosson C, Dieterich K, Satre V, Devillard F *et al.* 13q31.1 microdeletion: A prenatal case report with macrocephaly and macroglossia. *Eur J Med Genet*. 2015; **58**: 526–530.
- 12 Alp MY, Cebi AH, Seyhan S, Cansu A, Aydin H, Ikbali M. 22.5 MB deletion of 13q31.1-q34 associated with HPE, DWM, and HSCR: A case report and redefining the smallest deleted regions. *Genet Couns* 2016; **27**: 43–49.
- 13 Jia Y, Zhao H, Shi D, Peng W, Xie L, Wang W *et al.* Genetic effects of a 13q31.1 microdeletion detected by noninvasive prenatal testing (NIPT). *Int J Clin Exp Pathol* 2014; **7**: 7003–7011.
- 14 Chen CP, Huang MC, Chen YY, Chern SR, Wu PS, Chen YT *et al.* Prenatal diagnosis of de novo interstitial deletions involving 5q23.1-q23.3 and 18q12.1-q12.3 by array CGH using uncultured amniocytes in a pregnancy with fetal interrupted aortic arch and atrial septal defect. *Gene* 2013; **531**: 496–501.
- 15 Zayed H, Chao R, Moshrefi A, Lopezjimenez N, Delaney A, Chen J *et al.* A maternally inherited chromosome 18q22.1 deletion in a male with late-presenting diaphragmatic hernia and microphthalmia-evaluation of DSEL as a candidate gene for the diaphragmatic defect. *Am J Med Genet A* 2010; **152A**: 916–923.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

© The Author(s) 2017

Supplementary Information for this article can be found on the Human Genome Variation website (<http://www.nature.com/hgv>)