

DATA NOTE

Open Access



# A comprehensive list of human microdeletion and microduplication syndromes

Alyssa S. Wetzel\*  and Benjamin W. Darbro

## Abstract

**Objective:** The phenotypic spectrum of human microdeletion and microduplication syndromes (MMS) is heterogeneous but often involves intellectual disability, autism spectrum disorders, dysmorphic features and/or multiple congenital anomalies. While the common recurrent copy number variants (CNVs) which underlie these MMS have been well-studied, the expansion of clinical genomic testing has led to the identification of many rare non-recurrent MMS. To date, hundreds of unique MMS have been reported in the medical literature, and no single resource exists which compiles all these MMS in one location. This comprehensive list of MMS will aid further study of CNV disorders as well as serve as a resource for clinical laboratories performing diagnostic CNV testing.

**Data description:** Here we provide a comprehensive list of MMS which have been reported in the medical literature to date. This list is sorted by genomic location, and for each MMS, we provide a list of publications for referral, as well as the consensus coordinates, representative region, shortest regions of overlap (SRO), and/or subregions where applicable.

**Keywords:** Copy number variant, Microdeletion, Microduplication, Microduplication/microdeletion disorders, Genomic disorders

## Objective

While the field of genetics is over 150 years old, it has only been within the last twenty years that we have begun to understand the role of copy number variants (CNVs) in human disease [1]. In the early 1990s, the underlying microdeletion was identified for several clinical syndromes [2–6]. These discoveries led to research on the surrounding genomic architecture and mechanism of formation for these microdeletions which in turn led to the discovery of the predicted reciprocal microduplications and their associated clinical syndromes. Then, the completion of the human genome project and development of more cost-effective sequencing technologies in the early 2000s paved the way for the expansion of the genotype-first approach to identifying genomic disorders [2, 3, 7].

The spectrum of CNVs which have been identified has been further expanded since the American College of Medical Genetics and Genomics formally recommended the chromosomal microarray as a first-line diagnostic test. Cooper et al. [8] leveraged both patient and control population data to identify regions of the human genome which are enriched for CNVs in patient population identifying 59 putative pathogenic regions. While 14 of these putative regions were novel and/or weakly supported at the time, several have since been decidedly associated with human disorders. In 2014, Nevado et al. [9] reported that nearly 100 *new* CNV disorders had been reported in a five-year period, further underscoring the breadth and depth of the role of CNVs in human disease. Unlike prior reviews of microdeletion and microduplication syndromes (MMS) [9–11], this comprehensive list provides both genomic coordinates and a list of representative publications for phenotypic analysis and/or further reading. Thus, this list serves as a valuable resource in the interpretation of

\*Correspondence: alyssa-wetzel@uiowa.edu

Stead Family Department of Pediatrics, University of Iowa Hospitals and Clinics, Iowa City, IA 52242, USA



clinically identified CNVs or for use in downstream variant prioritization or classification pipelines.

### Data description

A non-redundant list of microduplication and microdeletion syndromes (MMS) was obtained from the following databases and publications: Clinical Genome Resource (ClinGen; [www.clinicalgenome.org](http://www.clinicalgenome.org)) [12] via UCSC's Table Browser (iscaCuratedPathogenic;  $n = 61$ ) [13, 14], Database of genomic variation and Phenotype in Humans using Ensembl Resources (DECIPHER; <https://decipher.sanger.ac.uk/>;  $n = 66$ ) [15], Online Mendelian Inheritance in Man (OMIM; [www.omim.org](http://www.omim.org)) [16] via the Gene2Map file (keywords: contiguous gene, chromosome (exclude: open reading frame), deletion, duplication, triplication, and quadruplication;  $n = 115$ ), Nevado et al. [9] ( $n = 99$ ), Table 2 of Bentacur [10] ( $n = 39$ ), Supplemental Tables 1 and 2 of Kaminsky et al. [17] ( $n = 41$ ), Additional File 5 of Marcinkowska et al. [18] ( $n = 20$ ) and Table 1 of Wiese et al. [11] ( $n = 193$ ).

The non-redundant list was curated to exclude disorders which were only described in a single patient (e.g. 3p11.2-p12.1 [9, 19]) or family and disorders which are not associated with congenital disease (e.g. 8p11 myeloproliferative syndrome; OMIM 613523). CNV disorders were retained even if only the duplication or deletion has been observed or associated with human disease. When possible, coordinates for the CNV disorders were obtained from either DECIPHER or ClinGen (accessed August 2021). For all remaining MMS without consensus coordinates, CNV coordinates for each reported patient were identified as were the coordinates for any defined shortest region of overlaps (SRO) or minimal regions (Dataset 2). All coordinates were converted from their initial assembly into hg19 coordinates using UCSC's Liftover Tool [20]. These patient and minimal regions were utilized to define the representative CNV interval and when applicable minimal region(s) of overlap. Some MMS were divided into multiple interval regions to account for: recurrent CNVs with multiple break-points (e.g. 15q11.2 BP1–2 vs 15q11.2 BP1–3 vs 15q11.2 BP2–3), size differences due to the presence of segmental duplications within a CNV call (1q21.1 TAR Susceptibility Locus), and to account for multiple minimal CNV regions (e.g. 1q24-q25). Further details on the representative and minimal intervals for each CNV disorder can be found in Datafile 1. BED file versions of Dataset 1 are available for hg19 (Dataset 3) and GRCh38 (Dataset 4) coordinates.

We identified 192 recurrent and non-recurrent Microdeletion and Microduplication Syndromes (Dataset 1) which were divided up into 320 individual CNV intervals and 141 non-overlapping CNV regions. Of

these intervals, 25% ( $n = 80$ ) had consensus coordinates from ClinGen and/or DECIPHER. Coordinates for the remaining CNV intervals ( $n = 240$ ) were determined manually by leveraging over 2500 patient CNVs, defined SRO(s) or minimal region(s) from the medical literature (Datafile 1; Dataset 2).

Label	Name of Data Set/File	File Type/ Extension	Data repository and identifier (DOI or accession number)
<b>Dataset 1</b>	<i>Microdeletion and Microduplication Syndromes</i>	MS Excel file (.xlsx)	Zenodo ( <a href="https://doi.org/10.5281/zenodo.6975072">https://doi.org/10.5281/zenodo.6975072</a> ) [21]
<b>Dataset 2</b>	<i>MMS CNVs Reported in Medical Literature</i>	MS Excel file (.xlsx)	Zenodo ( <a href="https://doi.org/10.5281/zenodo.6975072">https://doi.org/10.5281/zenodo.6975072</a> ) [21]
<b>Dataset 3</b>	<i>MMS hg19</i>	Browser Extensible Data (.bed)	Zenodo ( <a href="https://doi.org/10.5281/zenodo.6975072">https://doi.org/10.5281/zenodo.6975072</a> ) [21]
<b>Dataset 4</b>	<i>MMS GRCh38</i>	Browser Extensible Data (.bed)	Zenodo ( <a href="https://doi.org/10.5281/zenodo.6975072">https://doi.org/10.5281/zenodo.6975072</a> ) [21]
<b>Datafile 1</b>	<i>CNV Coordinate Determination by Chromosome</i>	MS Word file (.doc)	Zenodo ( <a href="https://doi.org/10.5281/zenodo.6975072">https://doi.org/10.5281/zenodo.6975072</a> ) [21]
<b>Datafile 2</b>	<i>Reference List for Microdeletion and Microduplication Syndromes</i>	MS Word file (.doc)	Zenodo ( <a href="https://doi.org/10.5281/zenodo.6975072">https://doi.org/10.5281/zenodo.6975072</a> ) [21]

### Limitations

We chose to limit the scope of this MMS list to the characterization of the representative genomic coordinates rather than conducting a thorough phenotypic review of each reported MMS. Therefore, researchers and clinicians will need to refer to each MMS' reference list or other databases such as OMIM to collect this information. This list of MMS was restricted to those described in at least two unrelated patients, and relied on databases and prior reviews of MMS. Therefore, there is the potential that a MMS has been missed and a future iteration(s) of this comprehensive MMS list will likely be needed.

### Abbreviations

CNVs: Copy number variants; MMS: Microdeletion and microduplication syndromes; ClinGen: Clinical Genome Resource; DECIPHER: Database of genomic variation and Phenotype in Humans using Ensembl Resources; OMIM: Online Mendelian Inheritance in Man; SRO: Shortest region of overlap.

### Acknowledgements

This study makes use of data generated by the DECIPHER community. A full list of centres who contributed to the generation of the data is available from <https://deciphergenomics.org/about/stats> and via email from [contact@deciphergenomics.org](mailto:contact@deciphergenomics.org). Funding for the DECIPHER project was provided by Wellcome. The authors would also like to acknowledge the Shivanand R. Patil Cytogenetics and Molecular Laboratory.

**Authors' contributions**

AW compiled the comprehensive list of CNV disorders and drafted the manuscript. BW reviewed the list of CNV disorders. All authors read and approved the final manuscript.

**Authors' information**

Not applicable.

**Funding**

This work was supported by the Stead Family Department of Pediatrics.

**Availability of data and materials**

The data described in Table 1 of this Data Note can be freely and openly accessed on GitHub under aswetzal/MMS (<https://github.com/aswetzal/MMS>) [21].

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 2 December 2021 Accepted: 18 October 2022

Published online: 26 November 2022

**References**

- Rimoin DL, Hirschhorn K. A history of medical genetics in pediatrics. *Pediatr Res*. 2004;56(1):150–9.
- Harel T, Lupski JR. Genomic disorders 20 years on—mechanisms for clinical manifestations. *Clin Genet*. 2018;93(3):439–49.
- Lupski JR. Genomic disorders ten years on. *Genome Med*. 2009;1(4):42.
- De Decker HP, Lawrenson JB. The 22q11.2 deletion: from diversity to a single gene theory. *Genet Med*. 2001;3(1):2–5.
- McDonald-McGinn DM, Sullivan KE, Marino B, Philip N, Swillen A, Vorstman JAS, et al. 22q11.2 deletion syndrome. *Nat Rev Dis Primers*. 2015;1(1):15071.
- Robin NH, Shprintzen RJ. Defining the clinical spectrum of deletion 22q11.2. *J Pediatr*. 2005;147(1):90–6.
- Martin CL, Warburton D. Detection of chromosomal aberrations in clinical practice: from karyotype to genome sequence. *Annu Rev Genomics Hum Genet*. 2015;16:309–26.
- 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(9):838–46.
- Nevado J, Mergener R, Palomares-Bralo M, Souza KR, Vallespín E, Mena R, et al. New microdeletion and microduplication syndromes: a comprehensive review. *Genet Mol Biol*. 2014;37(1 Suppl):210–9.
- Betancur C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res*. 2011;1380:42–77.
- Weise A, Mrasek K, Klein E, Mulatinho M, Llerena JC, Hardekopf D, et al. Microdeletion and microduplication syndromes. *J Histochem Cytochem*. 2012;60(5):346–58.
- Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen—the Clinical Genome Resource. *N Engl J Med*. 2015;372(23):2235–42.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996–1006.
- Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, et al. The UCSC table browser data retrieval tool. *Nucleic Acids Res*. 2004;32(Database issue):D493–6.
- Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am J Hum Genet*. 2009;84(4):524–33.
- Online Mendelian Inheritance in Man, OMIM®. In. Baltimore, MD: McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University. <https://www.omim.org/help/faq>.
- Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, et al. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet Med*. 2011;13(9):777–84.
- Marcinkowska M, Szymanski M, Krzyzosiak WJ, Kozlowski P. Copy number variation of microRNA genes in the human genome. *BMC Genomics*. 2011;12:183.
- Gat-Yablonski G, Frumkin-Ben David R, Bar M, Potievsky O, Phillip M, Lazar L. Homozygous microdeletion of the POU1F1, CHMP2B, and VGLL3 genes in chromosome 3—a novel syndrome. *Am J Med Genet A*. 2011;155a(9):2242–6.
- Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, et al. The UCSC genome browser Database: update 2006. *Nucleic Acids Res*. 2006;34(Database issue):D590–8.
- Wetzel AS, Darbro BW. A comprehensive list of human microdeletion and microduplication syndromes. Zenodo. 2022. <https://doi.org/10.5281/zenodo.6975072>.

**Publisher's Note**

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

