ARTICLE

An estimation of the prevalence of genomic disorders using chromosomal microarray data

Madelyn A. Gillentine[1,2] · Philip J. Lupo[3] · Pawel Stankiewicz[1] · Christian P. Schaaf [1,2,4,5,6]

Abstract

Multiple genomic disorders result from recurrent deletions or duplications between low copy repeat (LCR) clusters, mediated by nonallelic homologous recombination. These copy number variants (CNVs) often exhibit variable expressivity and/or incomplete penetrance. However, the population prevalence of many genomic disorders has not been estimated accurately. A subset of genomic disorders similarly characterized by CNVs between LCRs have been studied epidemiologically, including Williams-Beuren syndrome (7q11.23), Smith-Magenis syndrome (17p11.2), velocardiofacial syndrome (22q11.21), Prader-Willi/Angelman syndromes (15q11.2q12), 17q12 deletion syndrome, and Charcot-Marie-Tooth neuropathy type 1/hereditary neuropathy with liability to pressure palsy (*PMP22*, 17q11.2). We have generated a method to estimate prevalence of highly penetrant genomic disorders by (1) leveraging epidemiological data for genomic disorders with previously reported prevalence estimates, (2) obtaining chromosomal microarray data on genomic disorders from a large medical genetics clinic; and (3) utilizing these in a linear regression model to determine the prevalence of this syndromic copy number change among the general population. Using our algorithm, the prevalence for five clinically relevant recurrent genomic disorders: 1q21.1 microdeletion (1/6882 live births) and microduplication syndromes (1/6309), 15q13.3 microdeletion syndrome (1/5525), and 16p11.2 microdeletion (1/3021) and microduplication syndromes (1/4216), were determined. These findings will inform epidemiological strategies for evaluating those conditions, and our method may be useful to evaluate the prevalence of other highly penetrant genomic disorders.

Introduction

Nonallelic homologous recombination (NAHR), mediated by low copy repeats (LCRs), is a known mechanism of copy number changes for a variety of genomic disorders. Individually, these disorders are rare but cumulatively affect a large portion of the population. In addition, certain genomic disorders affect specific populations at higher frequencies, particularly for neuropsychiatric disorders. Although the enrichment of copy number variants (CNVs) has been noted in affected populations, population-wide prevalence estimates have not been determined for a large proportion of genomic disorders. Estimates that have been proposed are likely to be inaccurate as they are often based on older or incomplete data. For example, it has been estimated that the prevalence of 15q13.3 microdeletion syndrome (OMIM 612001) is 1 in 40,000 individuals, based on the prevalence of intellectual disability in the population at the time of its original discovery [1]. This is likely an underestimate, as 15q13.3 microdeletion syndrome, like many CNV-associated syndromes, exhibits variable expressivity. However, there has not been an easy method to assess the true prevalence of such disorders.

Recurrent CNVs mediated by NAHR occur across the genome, including those causative of Williams-Beuren

✉ Christian P. Schaaf
  schaaf@bcm.edu

1 Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

2 Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, USA

3 Department of Pediatrics-Oncology, Baylor College of Medicine, Houston, TX, USA

4 Institute of Human Genetics, University Hospital Cologne, Cologne, Germany

5 Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany

6 Center for Rare Diseases, University Hospital Cologne, Cologne, Germany

syndrome (WBS, 7q11.23 deletion, OMIM 612547), Smith-Magenis syndrome (SMS, 17p11.2 deletion, OMIM 182290), velocardiofacial syndrome (VCFS, 22q11.21 deletion, OMIM 192430), Prader-Willi/Angelman syndromes (PWS/AS, 15q11.1–15q11.2 deletion, OMIM 176270, and OMIM 105830), 17q12 deletion syndrome, and Charcot-Marie-Tooth Neuropathy type 1 (CMT1A)/hereditary neuropathy with liability to pressure palsy (HNPP, 17q11.2 deletion spanning *PMP22*. The population prevalence of these disorders has been studied epidemiologically. Here, we have taken advantage of the fact that the genomic mechanisms of disease are the same between multiple recurrent CNV loci, although occurring at varying frequencies owing to the structure of the region. As per the American College of Medical Genetics (ACMG) guidelines, chromosomal microarray analysis (CMA) is the first tier test for patients with intellectual disability/developmental delay (ID/DD), autism spectrum disorder (ASD), and multiple congenital abnormalities (MCA), allowing for an exhaustive cohort [2]. We have utilized reported epidemiological data for genomic disorders with known prevalence to determine the population prevalence of a subset of recurrent genomic disorders with clinical relevance: 1q21.1 microdeletion and microduplication syndromes, 15q13.3 microdeletion syndrome, and 16p11.2 microdeletion and microduplication syndromes (Fig. 1).

Materials and methods

To develop our algorithm, individuals with SMS, WBS, VCFS, PWS/AS, 17q12 deletion syndrome, and HNPP were identified by review of CMA results performed in the Baylor Genetics (BG) Laboratories, which included 54,407 unique cases. Cases of 1q21 microdeletions and microduplications, 15q13.3 microdeletions, and 16p11.2 microdeletions and microduplications were identified from the same cohort of CMA results. Asymptomatic parental arrays carrying a CNV ($n = 54$) were removed (Table 1). The population prevalence estimates for SMS, WBS, VCFS, PWS/AS, 17q12 deletion syndrome, and HNPP were obtained from previous publications (Table 2) [3–8]. We determined that the reported population prevalence of each genomic disorder correlated with the number of cases in the BG Laboratories ($Y = 0.102 \times -0.001775$; $R^2 = 0.8267$, $p = 0.012$, 95% CI: 0.037–0.9902). We used linear regression analysis in GraphPad Prism, regressing the known prevalence of each genomic disorder and the BG Laboratories number of cases to generate a β-coefficient used to extrapolate the prevalence of 1q21.1 microdeletion and microduplication syndromes, 15q13.3 microdeletion syndrome, and 16p11.2 microdeletion and microduplication syndromes.

Results

We identified 1q21.1 microdeletions in 1/625 (0.16%, $n = 87$) of BG Laboratories clinical CMA samples, whereas the reciprocal microduplications for 1/579 (0.17%, $n = 94$) samples (Table 3). 15q13.3 microdeletions account for 1/513 BG Laboratories clinical CMA samples (0.19%, $n = 106$), with varying frequencies depending on the breakpoints (BPs) utilized (BP3/BP5: 1/4946; BP4/BP5: 1/606; distal-CHRNA7-LCR/BP5: 1/10,881 (Fig. 2). 15q13.3 duplications were not assessed, as they have considerably lower penetrance and occur at similar frequency among clinical CMA samples and in the general population [9]. 16q11.2 microdeletions account for 1/292 (0.34% $n = 186$) of BG Laboratories clinical CMA samples, with the reciprocal duplications accounting for 1/400 (0.25%, $n = 136$) clinical samples. Using our model, we estimate that the population prevalence of each syndrome among live births is the following: 1q21.1 microdeletion syndrome, 0.015% (1/6882); 1q21.1 microduplication syndrome, 0.016% (1/6309); 15q13.3 microdeletion syndrome, 0.018% (1/5525; BP3/BP5: 1/151,515; BP4/BP5: 1/8417; distal-CHRNA7-LCR/BP5: 1/71,429); 16p11.2 microdeletion syndrome, 0.03% (1/3021); and 16p11.2 microduplication syndrome, 0.023% (1/4216) (Fig. 2).

Discussion

In this study, we have developed a model to estimate the population prevalence of highly penetrant recurrent genomic disorders, for which there is currently sparse data owing to the rarity of such disorders. The utilization of CMA data to determine prevalence is a new approach that may be beneficial to apply to additional CNV-associated syndromes. We have utilized from rare (SMS) to more prevalent (VCFS) genomic disorders, likely based on the homology of their LCRs, allowing our regression model to be applied to a range of recurrent, highly penetrant CNVs.

The ability to estimate the population prevalence of genetic disorders is useful and beneficial epidemiologically as well as in the clinic. With the often wide range of clinical manifestations, and in the absence of pathognomonic clinical features, a substantial number of individuals with these CNVs are not identified. This was highlighted by Turner et al. (2008) [10], who measured de novo genomic disorder rates based on sperm NAHR, who determined that, whereas the dominant disorders, resulting from reciprocal deletions and duplications between LCRs at 17p11.2 (HNPP and CMT1A, respectively) are diagnosed 1:4, it should really be closer to 2:1, but owing to milder phenotypes, the deletion is likely underdiagnosed. We believe that our methodology supports a similar conclusion, and that our estimates may
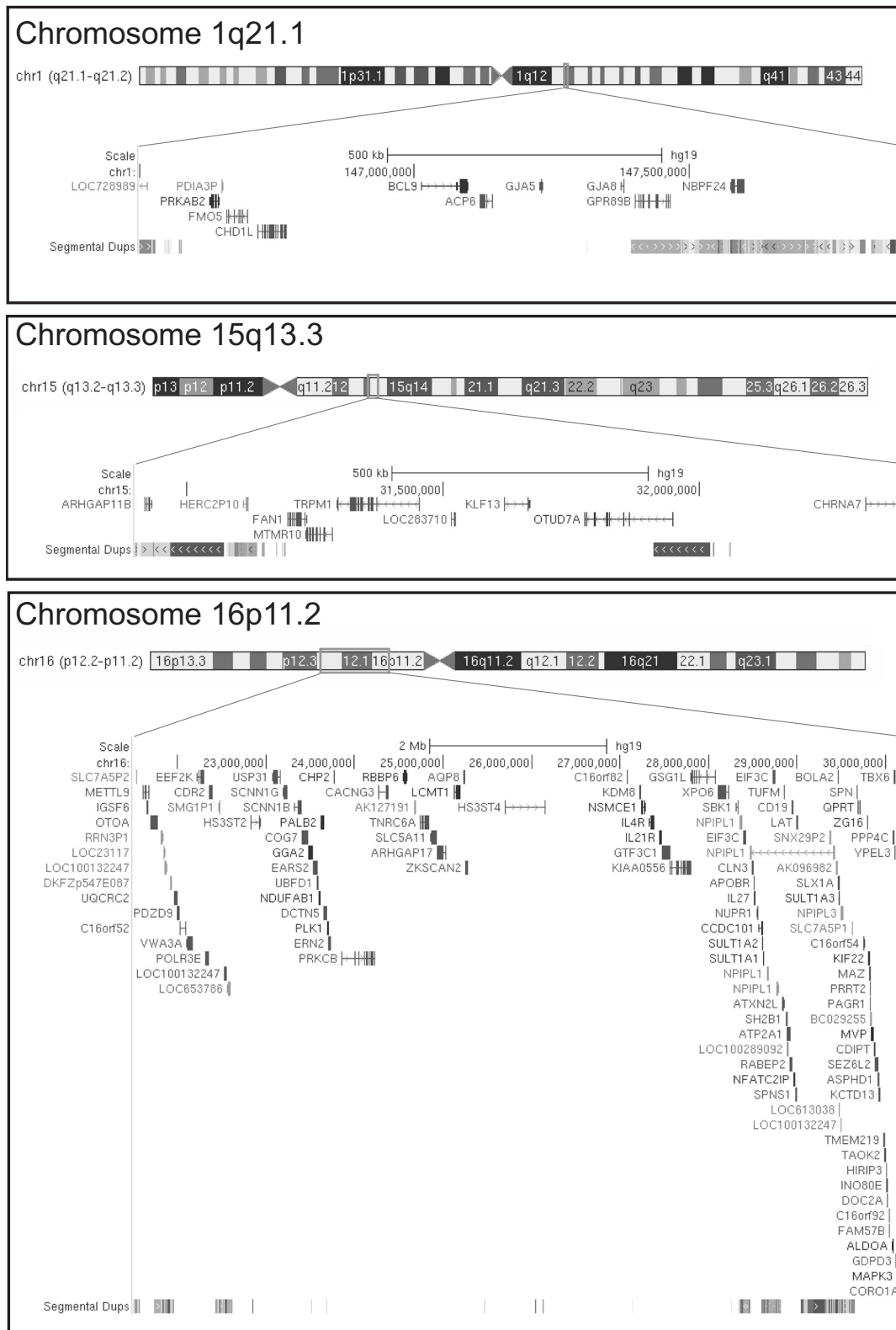
Fig. 1 DECIPHER coordinates genomic disorders at chromosomes 1q21.1, 15q13.3, and 16p11.2. Region is highlighted in red box along the chromosome. For 15q13.3 microdeletions, larger deletions spanning breakpoint 3 to breakpoint 5 were included, but only breakpoint 4 to breakpoint 5 is shown. Adapted from UCSC Genome Browser

Table 1 Number of parental CMA cases removed for each genomic disorder

Syndrome	Parental cases removed
SMS	0
WBS	0
PWS/AS	0
VCFS	0
HNPP	1
1q21.1 microdeletion syndrome	1
1q21.1 microduplication	7
15q13.3 microdeletion syndrome	28
16p11.2 microdeletion syndrome	5
16p11.2 microduplication syndrome	12

SMS Smith-Magenis syndrome, *WBS* Williams-Beuren syndrome, *PWS/AS* Prader-Willi syndrome/Angelman syndrome, *VCFS* velocardiofacial syndrome, *HNPP* hereditary neuropathy with liability to pressure palsies, *CMA* chromosomal microarray

Table 2 Known prevalence of recurrent deletion syndromes and percent of cases in Baylor Genetics Laboratories

Syndrome	Prevalence of live births	Percent of CMA cases
SMS	0.004–0.006% (1/25,000–1/15,000) [8]	0.08
WBS	0.013–0.005% (1/7,500–1/20,000) [6]	0.18
PWS/AS	PWS 0.006% AS 0.005–0.008% (combined (0.011–0.014%)) (1/9,090–1/7,142) [3, 4]	0.23
VCFS	0.0625–0.05% (1/1,600–1/2,000) [7]	0.51
17q12 deletion syndrome	0.007% (1/14,500) [5]	0.079
HNPP	0.016% (1/6,250) [5]	0.059

SMS Smith-Magenis syndrome, *WBS* Williams-Beuren syndrome, *PWS/AS* Prader-Willi syndrome/Angelman syndrome, *VCFS* velocardiofacial syndrome, *CMA* chromosomal microarray

allow physicians the opportunity to better recognize these underdiganosed disorders. Our approach covers individuals with a wide range of phenotypes, allowing us to avoid missing individuals with diagnoses that may be more rare in these microdeletion syndromes or misreported. As genes within these regions may be potential drug targets, understanding their prevalence may impact the speed at which specific therapeutics are developed. Furthermore, with accurate prevalence measures, physicians have better tools to recognize the syndromes and recommend CMA testing for patients.

1q21.1 microdeletion and microduplication syndromes

We estimate that pathogenic 1q21.1 deletions and duplications (DECIPHER coordinates: chr1: 146.5–147.9 Mb, hg19) occur in 1/6882 and 1/6309 live births, respectively. However, these CNVs are incompletely penetrant, so it is likely that the true frequency in the population is higher than our prediction. 1q21.1 microdeletion and microduplication syndromes result from an ~800 kb CNV spanning minimally seven genes, none of which has been targeted therapeutically [11]. The reciprocal genomic disorders are characterized by mild to moderate ID/DD, ASD (more so for duplication probands), dysmorphic features, cardiac defects, microcephaly or macrocephaly (for deletion or duplication probands, respectively), and multiple psychiatric diagnoses, including those that are later onset, such as schizophrenia [12]. However, there is considerable variable expressivity of phenotypes among probands. Previously, Mefford et al. (2008) reported that deletions at 1q21.1 occur in 0.5% of probands with ID/DD, ASD, or MCA. More recent studies have found 1q21.1

microdeletions to occur between 0.02% and 0.27%, depending on the clinical cohort, with higher frequency among ID/DD/MCA and schizophrenia individuals [13]. The duplications have been noted to occur in 0.14% to 0.26% of varying neuropsychiatric cohorts, with the highest frequency among probands with ASD. Although the differing frequencies of these CNVs, although not considerably different from our findings, highlight that our method may produce differing results dependent on the cohort assessed, they may also reflect the changes in CMA assessment over the last decade, with more probands having varying phenotypes being tested. Our estimated prevalence suggests that physicians may want to consider CMA analysis for individuals with phenotypes observed with 1q21.1 CNVs who may not typically be assessed by CMA, such as cardiac defects and adult onset psychiatric disorders.

15q13.3 microdeletion syndrome

15q13.3 microdeletion syndrome (DECIPHER coordinates: chr15: 30.9–32.4 Mb, hg19) is characterized by ID/DD, seizures, and ASD among other variably expressive neuropsychiatric phenotypes [14]. Remarkably, we have shown that its prevalence is over sevenfold higher than previously estimated. The assessed 1/5525 live births carrying a 15q13.3 microdeletion is considerably higher than the previous suggestions of 1/40,000, and is comparable to that of SMS, WBS, and PWS/AS. As the previous estimate was based on the prevalence of ID only, it is not surprising that our number is higher, as 15q13.3 microdeletion syndrome manifests with multiple neuropsychiatric phenotypes. Incomplete penetrance of 15q13.3 deletions (estimated to be 40–80% penetrant) must also be taken into account. As we have removed asymptomatic individuals from our analysis, our value of 1/5525 live births is representative of the prevalence of pathogenic 15q13.3 microdeletions in the

Table 3 Genomic disorder cases in Baylor Genetics (BG) Laboratories and their estimated population prevalence

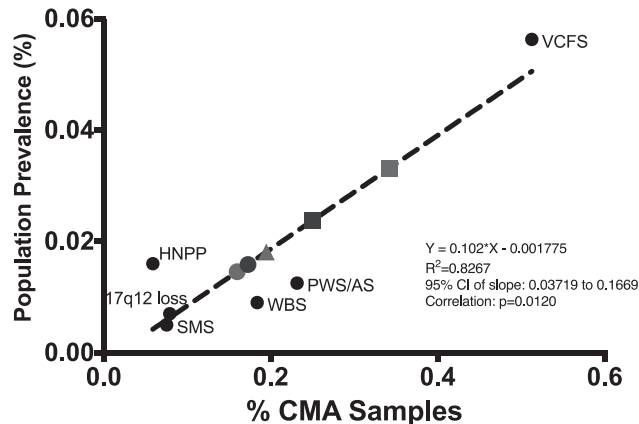| Genomic disorder	Percent of CMA cases (total $n = 54{,}407$)	Population prevalence estimate
1q21.1 microdeletion syndrome	0.16% (n = 87)	0.015% (1/6882)
1q21.1 microduplication syndrome	0.17% ($n = 94$)	0.016% (1/6309)
15q13.3 microdeletion syndrome	0.19% ($n = 106$)	0.018% (1/5525)
16p11.2 microdeletion syndrome	0.34% ($n = 186$)	0.03% (1/3021)
16p12.2 microduplication syndrome	0.25 ($n = 136$)	0.023% (1/4216)

CMA chromosomal microarray

Fig. 2 Linear regressions used to determine prevalence of highly penetrance genomic disorders. Percent of CMA cases at the Baylor Genetics (BG) Laboratories was plotted against known population prevalence. The prevalence of 1q21.1 microdeletion syndrome (red circle), 1q21.1 microduplication syndrome (blue circle), 15q13.3 microdeletion syndrome (red triangle), 16p11.2 microdeletion syndrome (red square), and 16p11.2 microduplication syndrome (blue square) were extrapolated based on the prevalence of Smith-Magenis syndrome (SMS), Williams-Beuren syndrome (WBS), Prader-Willi syndrome/Angelman syndrome (PWS/AS), and Velocardiofacial syndrome (VCFS).

population, suggesting that the disorder may be underdiagnosed.

Studies of the genomic structure of this region between populations suggest that there may be population differences of the prevalence of 15q13.3 CNVs, as certain structures are more frequent in differing groups. However, these were not considered in this study. These genomic structures include both a duplication adjacent to *CHRNA7*, found to be fixed in the population, and a polymorphic inversion (the γ inversion) between BPs 4 and 5 at 15q13.3 that predispose to deletions in offspring [15]. The haplotype frequency of the γ inversion polymorphism has been estimated at 6%. Thus, our estimates may be slightly decreased from the real pathogenic 15q13.3 microdeletion prevalence, but without the knowledge of how many of those 6% actually result in pathogenic deletions, it is impossible to consider that in our calculations. The genomic structure of the region also plays a role in the prevalence of each size of

15q13.3 microdeletion, with the largest deletions spanning from BPs 3 to 5 being the rarest, due to increased distance between the LCRs and decreased homology between the two LCRs. Deletions from BPs 4 to 5 are the most prevalent, owing to the over 99% homology between the two LCRs.

15q13.3 CNVs encompass the *CHRNA7* gene, encoding for the α7 nicotinic acetylcholine receptor. The α7 nicotinic acetylcholine receptor, important for signal transduction via calcium signaling in the brain, is a potential drug target, with agonists and positive allosteric modulators in development and testing. Our higher prevalence estimate is suggestive that therapeutics could reach a broader population than previously anticipated. These therapeutics may also be beneficial for other α7 nicotinic acetylcholine receptor implicated diseases, including 15q13.3 microduplications, Alzheimer Disease, and Parkinson Disease.

16p11.2 microdeletion and microduplication syndromes

Deletions at 16p11.2 are characterized by varying levels of ID/DD, high incidence of ASD (accounting for up to 1% of all ASD cases), and language delay, as well as other features including obesity and neuropsychiatric disorders. The reciprocal duplications have a similar range of phenotypes, although a lower estimated prevalence [13]. Over 29 genes are encompassed by these deletions, with *KCTD13* suggested to have a role in at least a subset of phenotypes. Unlike 1q21.1 CNVs and 15q13.3 microdeletions, deletions at 16p11.2 (chr16: 21.5–30.2 Mb, hg19) have been assessed previously epidemiologically. For Icelandic population, it has been suggested that 3.5 of 10,000 (1/2857) individuals carry 16p11.2 deletions [16], which is not significantly different from our estimate of 1/3021 live births. This supports the notion that, while the number of individuals with each CNV varies from where the data were obtained, it is likely that similar frequencies would be identified, assuming similar ascertainment. With our estimates, and those from previous publications, 16p11.2 deletions and duplications are among the more common pathogenic CNVs.

Control populations

None of the analyzed CNV exhibit complete penetrance. Studies have been done to determine the rate of CNVs among control populations. Of note, the prevalence of the microdeletions we have chosen have been estimated to be from 0.02 to 0.04% in "control" populations, suggesting that the penetrance of these deletions is either considerably lower than previously estimated, or the deletions may not be pathogenic, which is very unlikely [13, 17]. However, studies reporting these values do not deeply phenotype their control populations. As these disorders, like many genomic disorders, are known to have variable expressivity, it is possible that a subset of the "controls" had subclinical neuropsychiatric conditions, such as borderline ID. Furthermore, individuals reported in "control" populations carrying CNVs associated with neurodevelopmental disorders have been found to have mild neuropsychiatric phenotypes [18, 19].

Limitations

The best way to determine the population prevalence of genomic disorders would be formal epidemiological studies in large populations. However, this is not feasible for several reasons: (1) the rarity of these disorders makes it difficult to find them in many populations; (2) many of genomic disorders exhibit variable expressivity, so it is difficult to identify them by phenotypes; and (3) many of these genomic disorders look similar to each other, making molecular diagnoses necessary for accurate estimates. With our model, we can make an estimate; however, it is based only on those subjects who were able to receive a molecular diagnosis, meaning that the model likely represents an underestimate. In addition, our number of cases is based on the BG Laboratories CMA data only, and of course may vary based on other institutions available data.

All of the genomic disorders utilized for our regression and those we estimate the prevalence for are mediated by NAHR. However, some of these disorders, such as AS/PWS, do not result from NAHR-mediated CNVs alone, but also from uniparental disomy or single gene disruptions. In addition, certain disorders may be diagnosed using a method other than clinical CMA, such as fluorescent *in situ* hybridization, whole exome or genome sequencing, or other indicative clinical tests. As the clinical CMA data do not include reports from all assays, we likely have an underestimate of the prevalence for these disorders. The vast majority (if not all) of the CNVs reported in this manuscript occur owing to NAHR and therefore have recurrent BPs. We believe our approach is useful for genomic disorders occurring owing to similar mechanisms, but may not be as accurate for non-recurrent CNVs.

Incompletely penetrant CNVs present another challenge for our model. For CNVs that are primarily inherited, the reportedly asymptomatic parents must be (a) removed, as was done here, or (b) clinically phenotyped. In addition, incomplete penetrance makes this approach challenging for lowly penetrant CNVs, particularly with milder phenotypes, as they may have a much higher prevalence among the general population. An example of this is 15q13.3 duplications, which have been estimated to occur in 1/123 (0.81%) samples submitted for CMA and 0.55–0.62% in control individuals, a non-significant difference [9, 13]. Although the number of cases for these duplications could be put into our model and produce an estimate, the resulting value may be difficult to interpret. This would only estimate the population prevalence of individuals who carry those CNVs who had a severe enough phenotype and access to CMA testing, which is a considerable number of variables. For some CNVs, this information may be valuable, but is likely a vast underestimate for how often these copy changes occur in the general population.

Conclusions

Here, we have utilized clinical CMA data and previously published epidemiological data to generate a model to determine the prevalence of a selection of highly penetrant CNVs. With increasing data available, more accurate measurements can be determined. With more precise measurements, we can better understand genomic disorders at both the genomic and clinical level. This is very useful, as it is likely that many genomic disorders are likely underdiagnosed [20]. Furthermore, with the diagnosis of a genomic disorder, more targeted therapeutics may be available in the future for individuals. Using this model, it is possible to estimate the population prevalence of genomic disorders, allowing for faster identification in probands and potentially contributing to the development of therapeutics.

Compliance with ethical standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Sharp A, Mefford H, Li K, Baker C, Skinner C, Stevenson R, et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. Nat Genet. 2008;40:322–328.

2. Miller D, Adam M, Aradhya S, Biesecker L, Brothman A, Carter N, 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.

3. Õiglane-Shlik E, Talvik T, Žordania R, Põder H, Kahre T, Raukas E, et al. Prevalence of Angelman syndrome and Prader–Willi syndrome in Estonian children: Sister syndromes not equally represented. Am J Med Genet A. 2006;140A:1936–1943.

4. Kyllerman. On the prevalence of Angelman syndrome. Am J Med Genet. 1995;59:405–405.

5. Meretoja P, Silander K, Kalimo H, Aula P, Meretoja A, Savontaus ML. Epidemiology of hereditary neuropathy with liability to pressure palsies (HNPP) in south western Finland. Neuromuscul Disord. 1997;7:529–532.

6. Strømme P, Bjørnstad PG, Ramstad K. Prevalence estimation of Williams syndrome. J Child Neurol. 2002;17:269–271.

7. Shprintzen R. Velo-cardio-facial syndrome: 30 Years of study. Dev Disabil Res Rev. 2008;14:3–10.

8. Greenberg F, Guzzetta V, Montes de Oca-Luna R, Magenis RE, Smith AC, Richter SF, et al. Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). Am J Hum Genet. 1991;49:1207–1218.

9. Gillentine MA, Berry LN, Goin-Kochel RP, Ali MA, Ge J, Guffey D, et al. The cognitive and behavioral phenotypes of individuals with CHRNA7 duplications. J Autism Dev Disord. 2017;47:549–562.

10. Turner D, Miretti M, Rajan D, Fiegler H, Carter N, Blayney M, et al. Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. Nat Genet. 2007;40:90–95.

11. Bernier R, Steinman KJ, Reilly B, Wallace AS, Sherr EH, Pojman N, et al. Clinical phenotype of the recurrent 1q21.1 copy-number variant. Genet Med. 2016;18:341–349.

12. Dolcetti A, Silversides C, Marshall C, Lionel A, Stavropoulos D, Scherer S, et al. 1q21.1 Microduplication expression in adults. Genet Med. 2012;15:282–289.

13. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell. 2012;148:1223–1241.

14. Gillentine MA, Schaaf CP. The human clinical phenotypes of altered CHRNA7 copy number. Biochem Pharmacol. 2015;97:352–362.

15. Antonacci F, Dennis M, Huddleston J, Sudmant P, Steinberg K, Rosenfeld J, et al. Palindromic GOLGA8 core duplicons promote chromosome 15q13.3 microdeletion and evolutionary instability. Nat Genet. 2014;46:1293–1302.

16. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2013;505:361–366.

17. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al. Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK biobank subjects.Biol Psychiatry. 2017;82:103–110.

18. Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. Genet Med. 2013;15:478–481.

19. Männik K, Mägi R, Macé A, Cole B, Guyatt AL, Shihab HA, et al. Copy number variations and cognitive phenotypes in unselected populations. JAMA. 2015;313:2044–2054.

20. Girirajan S, Rosenfeld JA, Coe BP, Parikh S, Friedman N, Goldstein A, et al. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. N Engl J Med. 2012;367:1321–1331.