Enrichment of small pathogenic deletions at chromosome 9p24.3 and 9q34.3 involving DOCK8, KANK1, EHMT1 genes identified by using high-resolution oligonucleotide-single nucleotide polymorphism array analysis

Jia-Chi Wang *, Loretta W. Mahon, Leslie P. Ross, Arturo Anguiano, Renius Owen and Fatih Z. Boyar

Abstract
Background: High-resolution oligo-SNP array allowed the identification of extremely small pathogenic deletions at numerous clinically relevant regions. In our clinical practice, we found that small pathogenic deletions were frequently encountered at chromosome 9p and 9q terminal regions.

Results: A review of 531 cases with reportable copy number changes on chromosome 9 revealed 142 pathogenic copy number variants (CNVs): 104 losses, 31 gains, 7 complex chromosomal rearrangements. Of 104 pathogenic losses, 57 were less than 1 Mb in size, enriched at 9p24.3 and 9q34.3 regions, involving the DOCK8, KANK1, EHMT1 genes. The remaining 47 cases were due to interstitial or terminal deletions larger than 1 Mb or unbalanced translocations. The small pathogenic deletions of DOCK8, KANK1 and EHMT1 genes were more prevalent than small pathogenic deletions of NRXN1, DMD, SHANK3 genes and were only second to the 16p11.2 deletion syndrome, 593-kb (OMIM #611913).

Conclusions: This study corroborated comprehensive genotype-phenotype large scale studies at 9p24.3 and 9q24.3 regions for a better understanding of the pathogenicity caused by haploinsufficiency of the DOCK8, KANK1 and EHMT1 genes.

Trial registration number: None; it is not a clinical trial, and the cases were retrospectively collected and analyzed.

Keywords: Small pathogenic deletions, High resolution oligonucleotide-single nucleotide polymorphism array analysis, Haploinsufficiency, Homozygous deletions

Background
Chromosomal microarray analysis (CMA) has been widely utilized for the genome-wide screening of microdeletion and microduplication syndromes [1]. The sizes of well-known microdeletion and microduplication syndromes were usually larger than 1 Mb, such as 1.4 Mb for Williams-Beuren syndrome (OMIM #194050) or 2.8 Mb for DiGeorge syndrome (OMIM #188400). Small (<1 Mb) pathogenic deletions at regions which were not well characterized were frequently encountered during our daily clinical practice, for instance, the chromosomal regions at 9p24.3 and 9q34.3.

High-resolution oligo-SNP array is able to reveal a variety of chromosomal disorders including uniparental disomy or extremely small pathogenic deletions which would be missed by low-resolution oligonucleotide CMA. Our and other previous studies showed the cases with uniparental disomy were relatively limited in number on chromosome 9 as compared to chromosome 15, 11 and 7 [2, 3]. In contrast, small pathogenic deletions...
were frequently encountered at chromosome 9p24.3 and 9q34.3 by using high-resolution oligo-SNP array in postnatal studies. Research endeavors have been significantly prioritized to specific genes such as NRXN1 and SHANK3 in the past [4, 5]. To the best of our knowledge, only four cases with small deletions of 192 kb, 225 kb, 465 kb and 518 kb in size at 9p24.3 involving the DOCK8 and/or KANK1 gene [6–8], and a case of 40 kb deletion in the EHMT1 gene at 9q34.3 [9] have been documented.

The purpose of this study is to evaluate 1): the incidence of small (< 1 Mb) pathogenic deletions in postnatal specimens, 2): whether the small pathogenic deletions at 9p24.3 and 9q34.3 constituted a significant proportion of small deletions, 3): what proportion of deletions on chromosome 9 was caused by small pathogenic deletions at 9p24.3 and 9q34.3, 4): the efficacy of identifying extremely small homozygous pathogenic deletions using high-resolution oligo-SNP array.

Results
The incidence of small pathogenic deletions in postnatal specimens studied by high-resolution oligo-SNP array
Approximately 38,000 postnatal specimens were studied by high-resolution oligo-SNP array in our laboratory from 2011 through 2015. Of these, we reported approximately 13,000 (34 %) pathogenic variants or variants of uncertain clinical significance (VOUS). The detection rate was consistent with our previous study [2]. Of the 13,000 variants, a total of 373 recurrent (at least 3 cases) small pathogenic losses were identified (Fig. 1). The 16p11.2 deletion syndrome, 593-kb (OMIM #611913) is the most common small pathogenic loss (107 cases). The remaining involved NRXN1 (35 cases), DMD (31 cases), DOCK8 and/or KANK1 (30 cases), and other chromosomal regions such as 16p11.2 (OMIM #613444, 220-KB), 16p12.2 (OMIM #136570, 520-KB), 22q13.33 (OMIM #606232).

Approximately 12 % of the recurrent small pathogenic copy number losses were caused by deletions of DOCK8 and KANK1 at 9p24.3 and EHMT1 at 9q34.3
The total number of deletions involving DOCK8 and/or KANK1 genes at 9p24.3 (30 cases, Fig. 2) and EHMT1 gene at 9q34.3 (16 cases, Fig. 3a) constituted approximately 12 % (46/373 cases) of the recurrent small pathogenic deletions from all the chromosomes.

Small (<1 Mb) and relatively small (1–1.5 Mb) deletions constituted 59 % of pathogenic copy number losses on chromosome 9
Of 13,000 cases with reported copy number variants (CNVs), 531 cases were from chromosome 9, and 142 were pathogenic, including 104 losses, 31 gains, 7 complex chromosomal rearrangements with both losses and gains (Table 1 and Additional file 1: Figure S1). Of 104 pathogenic losses, 57 were smaller than 1 Mb in size. Of 57 cases, 46 were located at 9p24.3 and 9q34.3 regions, involving the DOCK8, KANK1, EHMT1 genes. The size ranged from 22 kb (case 54 with EHMT1 deletion) to
790 kb (case 35), with an average of 384 kb (Table 1 and Additional file 2: Table S1). The remaining 47 cases were either resulted from interstitial deletions over 1 Mb in size (18 cases), terminal deletions over 1 Mb (14 cases), or unbalanced translocations (15 cases) that could be as small as 880 kb in size in a t(9;11)(q34.3;p15.4)(case 47, Additional file 2: Table S2). Very interestingly, of these 47 cases, four had relatively small deletions (1–1.5 Mb in size): a 1,048-kb deletion (case 10, Additional file 2: Table S2) involving the *PTCH1* gene (OMIM 601309) and diagnosis of Gorlin syndrome (OMIM #109400), a 1,349-kb deletion (case 15) involving the *STXBP1* gene and diagnosis of early infantile epileptic encephalopathy (OMIM #612164), and two 9q34.3 deletions of 1,218 kb and 1,268 kb (case 31 and 32). The small (57 cases) and relative small (4 cases) deletions established 59 % (61/104) of pathogenic copy number losses on chromosome 9.

**The 9p24.3 (0–2.2 Mb from 9p telomere) and 9q34.3 (137.4–141.2 Mb from 9p telomere) were two hot spots of pathogenic copy number losses on chromosome 9.**

The 9p24.3 and 9q34.3 regions were two hot spots of copy number losses on chromosome 9. A total of 35 cases majorly involved the 9p24.3 region: 30 small pathogenic deletions (case 1–30, Additional file 2: Table S1), one interstitial deletion (case 1, Additional file 2: Table S2), two terminal deletions (case 19 and 20), and two unbalanced translocations (case 33 and 35). Additionally, 26 cases were localized to 9q34.3 region: 17 small pathogenic deletions (case 41–57, Additional file 2: Table S1), one interstitial deletion (case 18, Additional file 2:
Table S2), 5 large terminal deletions (case 28–32) and 3 unbalanced translocations (case 45–47). Overall, close to 60% (61/104, 59%) of deletions were localized to the very ends of chromosome 9 at band 9p24.3 and 9q34.3.

**Rarely encountered extremely small homozygous pathogenic deletions were discovered in two cases**

Extremely small homozygous pathogenic deletions were identified in two cases: 1): a new born baby girl who presented with a metabolic disorder (abnormal reflexes, hypotonia, seizures, and elevated glycine) was revealed to contain a 25-kb homozygous deletion in the *GLDC* gene, which gave rise to autosomal recessive glycine encephalopathy (nonketotic hyperglycinemia; OMIM #605899). Besides that, a 50-kb heterozygous deletion was also found in the 5’ region of the *GLDC* gene. Parental study showed the mother was a carrier of a 25-kb heterozygous deletion and the father was a carrier of a 75-kb heterozygous deletion of the *GLDC* gene. The 25-kb maternally inherited deletion was located within the 75-kb paternally inherited deletion, and therefore inheritance of abnormal allele from both parents led to a 25-kb homozygous and a 50-kb heterozygous deletion in the proband (Fig. 4); 2): a 1-year-old boy was found to have a 74-kb homozygous deletion of the *CDK5RAP2* gene in a region of homozygosity (Additional file 3: Figure S2) which
led to autosomal recessive primary microcephaly-3 (OMIM #604804). The presence of multiple large regions of homozygosity (a total of 421 Mb) implied these two parents were closely related in blood. The proband inherited the heterozygous abnormal allele with deletion of the CDK5RAP2 gene from both parents.

**Discussion**

The subtelomeric region such as 1p36 is known to be gene-rich and prone to have deletions, supported by a study with a large cohort of over 5,000 cases [10]. However, as compared to 1p36, 22q13, 4p16, 5p15, very limited cases at the ends of chromosome 9p and 9q were established [10]. When we reviewed the profile of copy number losses from our database of 38,000 postnatal cases studied by using high-resolution oligo-SNP array and sorted it based on chromosomal regions, we discovered that all the cases with 1p36 deletions were over 1 Mb; 20 cases were 1–3 Mb, eight cases were 3–5 Mb, eight cases were 5–10 Mb and five cases were 10–
20 Mb in size (unpublished data). In contrast, 61 of 104 pathogenic deletions on chromosome 9 were either smaller than 1 Mb (57 cases) or between 1 and 1.5 Mb (4 cases). This finding demonstrated that the size of copy number losses conspicuously varied between chromosomal regions. In the clinical practice, the concept of vigilant selecting appropriate methods to characterize the genomic losses at different regions becomes essentially important. For instance, FISH analysis using subtelomeric or locus-specific probe may be appropriate to identify 1p36 microdeletions, but may miss cases with small deletions on chromosome 9.

Extremely small intragenic deletion of the EHMT1 gene was only reported in one case: a 40-kb intragenic deletion of the EHMT1 gene with uncertain phenotypic consequence [9]. A more recent update on Kleefstra syndrome exhibited that the 16 newly diagnosed 9q34.3 deletions were all larger than 200 kb [11]. In our cohort, we identified a total of 24 cases with 9q34.3 deletions: 16 with small (<1 Mb) deletions involving the EHMT1 gene (case 42–57, Additional file 2: Table S1), 5 with terminal deletion of 9q (case 28–32, Additional file 2: Table S2), and 3 with small (880–1001 kb in size) 9q34.3 deletions due to unbalanced translocations (case 45–47, Additional file 2: Table S2). Remarkably, we brought about three extremely small (22 kb, 39 kb and 40 kb in size) intragenic deletions of the EHMT1 gene and all were clustered at the 3′ end of the gene (Fig. 3, b–d). The 22-kb deletion (Fig. 3b) was identified in a 32-year-old female with intellectual disability; whereas, the 40-kb deletion (Fig. 3c) was found in a 5-year-old girl and the 39-kb deletion (Fig. 3d) was discovered in a 1-year-old girl with typical features of 9q34.3 deletion including developmental delay, speech and motor delay, and hypotonia [9]. In addition, a 26-year-old male with a 165-kb deletion involving EHMT1 and CACNA1B gene (case 57, Additional file 2: Table S1) also presented clinical features which were typical for 9q34.3 deletion syndromes including mental retardation, developmental delay, speech delay, motor delay, learning disability, autism spectrum disorder, asymmetry of temporal lobe, localized polymicrogyria, loping gait, and scoliosis [11].

In contrast to the EHMT1 gene of which haploinsufficiency score was much better established by ClinGen
Conclusions
This study demonstrated 1): the incidence of recurrent small pathogenic deletions (<1 Mb) was approximately 3% (373/13,000) of all reported CNVs; 2): the small pathogenic deletions at 9p24.3 and 9q34.3 constituted 12% of small pathogenic deletions from all the chromosomes, 3): 59% of pathogenic deletions on chromosome 9 were due to small (<1 Mb) or relatively small (1–1.5 Mb) pathogenic deletions; 4): 81% (46/57) of small pathogenic deletions were enriched at 9p24.3 and 9q34.3 regions involving the DOCK8, KANK1 and EHMT1 genes; 5): high-resolution oligo-SNP array was capable of identifying homozygous deletions as small as 25 kb in size.

Methods
Patients
Patients with a broad range of clinical indications including intellectual disability, developmental delay, multiple congenital anomalies, dysmorphic features and pervasive developmental disorders were referred to our laboratory for oligo-SNP array studies. The data for this study were compiled from de-identified results of 38,000 consecutive patient specimens referred to our laboratory for constitutional oligo-SNP array study from 2011 to 2015. The patients were majorly from general population in the United States, with <5% from Mexico and other countries.

Oligonucleotide-single nucleotide polymorphism array analysis platforms and threshold setting
Oligo-SNP array analysis was performed on either Human SNP Array 6.0 (in 2011) or CytoScan® HD (2012–2015)(Affymetrix, Santa Clara, CA), using genomic DNA extracted from whole blood. The Human SNP Array 6.0 has 1.8 million genetic markers, including about 906,600 SNPs and 946,000 probes for the detection of CNVs. The CytoScan® HD has more than 2.67 million probes, including 1.9 million non-polymorphic copy number probes and 750,000 SNP probes. The overall resolutions are approximately 1.7 kb for Human SNP Array 6.0 and 1.15 kb for CytoScan® HD. For chromosome 9, the probes for Human SNP Array 6.0 covered: 9p(chr9:37,747–47,217,164) and 9q(chr9:65,596,318–141,091,382); for CytoScan®: 9p (chr9:129,40,784,142, chr9:43,400,082–44,900,526) and 9q (chr9:66,837,485–141,025,328). Genomic coordinates were based upon genome build 37/hg19 (2009). Hybridization, data extraction, and analysis were performed as per manufacturers’ protocols. The Affymetrix® Chromosome Analysis Suite (ChAS) Software version 2.0 was used for data analysis, review, and reporting. For genome-wide screening, thresholds were set at >200 kb for gains and >50 kb for losses. For cytogenetically relevant regions, thresholds were set at >100 kb for gains and >20 kb for losses. Benign CNVs that are documented in the database of genomic variations (http://dgv.tcag.ca/dgv/app/home?ref=GRCh37/hg19) and present in the general population were excluded from reporting.
Additional file 1: Figure S1. Profile of abnormal CNVs and VOUS in chromosome 9 from 531 cases. A total of 142 cases with pathogenic CNVs, and 389 cases with VOUS were identified. (JPG 52 kb)

Additional file 2: Table S1. List of the cases with small (<1 Mb) pathogenic copy number loss on chromosome 9 without involvement of other chromosomes (N = 57). Table S2. List of the remaining cases with copy number loss on chromosome 9 (N = 47). Table S3. Interpretation, references and ClinGen evaluation of haplinsufficiency score of selected cytogenetically relevant genes on chromosome 9. (DOXX 43 kb)

Additional file 3: Figure S2. A. The finding of multiple large regions of homozygosity (421 Mb) was due to closely parental relatedness. B. An approximately 74-kb homozygous deletion involved multiple exons of the CDK5RAP2 gene in a region of homozygosity at 9q33.1-q34.11 (chr9:118,503,864-132,425,233), which caused autosomal recessive primary microcephaly-3 (OMIM #604804). (JPG 60 kb)

Abbreviations
CMA: Chromosomal microarray analysis; CNV: Copy number variant; FISH: Fluorescence in situ hybridization; ISCN: International System of Cytogenetic Nomenclature; Oligo-SNP array: Oligonucleotide-single nucleotide polymorphism array; UPD: Uniparental disomy; VOUS: Variants of uncertain clinical significance

Acknowledgements
The authors acknowledge that Andrew Hellman and Jeff Radcliff provided valuable assistance in manuscript revisions.

Funding
Not applicable.

Availability of data and materials
The raw data and materials used for this study have been de-identified and were freely available to any scientist wishing to use them without breaching participant confidentiality.

Authors' contributions
J-CW – reviewing, verification, analysis and interpretation of the data, drafting and revising the article, final approval of the version to be published. LWM – conception and design, interpretation of data. AA – conception and design, interpretation of data. RO – analysis of data. FZB – interpretation of data. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable since this manuscript did not contain any individual’s personal data in any form.

Ethics approval and consent to participate
This manuscript is a retrospective case report that does not require ethics committee approval at our institution. Consent was considered not required because the cases reported here were anonymous and de-identified. Received: 24 August 2016 Accepted: 1 November 2016 Published online: 15 November 2016

References
1. Watson CT, Marques-Bonet T, Sharp AJ, Mefford HC. The genetics of microdeletion and microduplication syndromes: an update. Annu Rev Genomics Hum Genet. 2014;15:215–44.
2. Wang JC, Ross L, Mahon LW, Owen R, Hemmat M, Wang BT, El Naggar M, Kopita KA, Randolph LM, Chase JM, Matas Aguilea MJ, Siles JL, Church JA, Hauser N, Shen JJ, Jones MC, Wierenga KJ, Jiang Z, Haddadin M, Boyar FZ, Anguiano A, Strom CM, Sahoo T. Regions of homozygosity identified by oligonucleotide SNP arrays: evaluating the incidence and clinical utility. Eur J Hum Genet. 2015;23:663–71.
3. Lieht T. Cytogenetic contribution to uniparental disomy (UPD). Mol Cytogenet. 2010;3:8.
4. Bene F, Bruno DL, Eriksson M, van Ravenswaaij-Arts C, Stark Z, Dijkhuizen T, Gerkes E, Gimelli S, Ganesamoorthy D, Thuresson AC, Labalme A, Till M, Bilan F, Pasquier L, Kitzis A, Duboucq C, Rossi M, Bottani A, Gagnebin M, Sanlaville D, Gilbert-Dussardier B, Guipponi M, van Haeringen A, Krieck M, Ruivenkamp C, Antonanakis SE, Anderlid BM, Slater HR, Schoumans J. Molecular and clinical characterization of 25 individuals with exonic deletions of NROB1 and comprehensive review of the literature. Am J Med Genet B Neuropsychiatr Genet. 2013;162B:388–403.
5. Soonya L, Kolevzon A, Zweifach J, Lim T, Doby Y, Schwartz L, Frank Y, Wang AT, Cai G, Parkhomenko E, Halpern D, Groddberg D, Angarita B, Willner JP, Yang A, Canitano R, Chaplin W, Betancor C, Bubbaum JD. Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency. Mol Autism. 2013;4:18.
6. Tassano E, Accogli A, Pavanello M, Bruno C, Capra V, Gimelli G, Cucco C. Interstitial 9p24.3 deletion involving only DOK8 and KANK1 genes in two patients with non-overlapping phenotypic traits. Eur J Med Genet. 2016;59:20–5.
7. Vanzo RJ, Martin MM, Sanado MR, South ST. Familial KANK1 deletion that does not follow expected imprinting pattern. Eur J Med Genet. 2013;56:256–9.
8. Lerer I, Sagi M, Meiner V, Cohen T, Zlotogora J, Abelliovich D. Deletion of the ANKRD15 gene at 9q24.3 causes parent-of-origin-dependent inheritance of familial cerebral palsy. Hum Mol Genet. 2005;14:9911–20.
9. Kleefstra T, van Zelst-Stams WA, Nillesen WM, Corinier-Daive V, Hougue G, Foulds N, van Dooren M, Willemsen MH, Pfundt R, Turner A, Wilson M, McLaughran J, Rauch A, Zener K, Adam MP, Innes M, Davies C, Lopez AZ, Casalone R, Weber A, Brueton LA, Navarro AD, Bralo MP, Venselaar H, Stegmann SP, Ytretrea HG, van Bokhoven H, Brunner HG. Further clinical and molecular delineation of the 9q subtelomeric deletion syndrome supports a major contribution of EHMT1 haplinsufficiency to the core phenotype. J Med Genet. 2009;46:598–606.
10. Shao L, Shaw CA, Lu XY, Sahoo T, Bacino CA, Lalani SR, Staniekwicz P, Yatsenko SA, Li Y, Neill S, Purisely AN, Chinnault AC, Patel A, Beaudet AL, Lupski JR, Cheung SW. Identification of chromosome abnormalities in subtelomeric regions by microarray analysis: a study of 5,380 cases. Am J Med Genet A. 2008;146A:2242–51.
11. Willemsen MH, Vufo-van Silfhout AT, Nillesen WM, Wissink-Lindhout WM, van Bokhoven H, Philip N, Berry-Kravis EM, Kin K, van Ravenswaaij-Arts CM, Delle Chiaie B, Innes AM, Hougue G, Kosonen T, Crenner K, Fannemel M, Stray-Pedersen A, Reedon W, Ignatius J, Lachlan K, Mircher C, van den Enden PT, Mastebroek M, Cohn-Hakke PE, Ytretrea HG, Drunat S, Kleefstra T. Update on Kleefstra Syndrome. Mol Syndromol. 2012;2:202–12.
12. Griggs BL, Ladd S, Saul RA, DuPont BR, Skrabatka AK. Deletion of the KANK1 gene in a region of homozygosity at 9q33.1-q34.11 (chr9:118,503,864-132,425,233), which caused autosomal recessive primary microcephaly-3 (OMIM #604804). (JPG 60 kb)