European Journal of Medical Genetics 55 (2012) 586–598

Contents lists available at SciVerse ScienceDirect

European Journal of Medical Genetics

journal homepage: http://www.elsevier.com/locate/ejmg

Original article

Interpretation of clinical relevance of X-chromosome copy number variations identified in a large cohort of individuals with cognitive disorders and/or congenital anomalies

Marjolein H. Willemsen*, Nicole de Leeuw, Arjan P.M. de Brouwer, Rolph Pfundt, Jayne Y. Hehir-Kwa, Helger G. Yntema, Willy M. Nillesen, Bert B.A. de Vries, Hans van Bokhoven, Tjitske Kleefstra

849 Department of Human Genetics, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 22 January 2012
Received in revised form
5 May 2012
Accepted 5 May 2012
Available online 14 July 2012

Keywords:
X-linked
Intellectual disability
Autism
X-chromosome
Copy number variation
Clinical interpretation

ABSTRACT

Genome-wide array studies are now routinely being used in the evaluation of patients with cognitive disorders (CD) and/or congenital anomalies (CA). Therefore, inevitably each clinician is confronted with the challenging task of the interpretation of copy number variations detected by genome-wide array platforms in a diagnostic setting. Clinical interpretation of autosomal copy number variations is already challenging, but assessment of the clinical relevance of copy number variations of the X-chromosome is even more complex. This study provides an overview of the X-Chromosome copy number variations that we have identified by genome-wide array analysis in a large cohort of 4407 male and female patients. We have made an interpretation of the clinical relevance of each of these copy number variations based on well-defined criteria and previous reports in literature and databases. The prevalence of X-chromosome copy number variations in this cohort was 57/4407 (~1.3%), of which 15 (0.3%) were interpreted as (likely) pathogenic.

1. Introduction

Both intellectual disability (ID) and autism spectrum disorders (ASD) are relatively frequent in the population, with an estimated prevalence of 2–3% and 1% respectively [1,2]. The prevalence of both disorders is higher in males than in females [2,3,4]. ID and ASD are clinically and genetically heterogeneous. Genetic factors are expected to play a causal role in a substantial part of the cases [5–10]. A large proportion of patients with ID have ASD as well, suggesting that shared underlying genetic and biological defects may be involved [11].

In the recent years genome wide array technologies including array comparative genomic hybridization (array-CGH) and single nucleotide polymorphism oligonucleotide arrays (SNP arrays) have become routine tools in the clinical evaluation of patients with ID and/or congenital anomalies (CA) and are increasingly used in the evaluation of other neuropsychiatric disorders as ASD, Attention Deficit Hyperactivity Disorder (ADHD) and schizophrenia as well. This has led to the identification of several causative submicroscopic chromosomal aberrations/copy number variations (CNVs), including X chromosomal aberrations, that are too small in size to be detected by routine cytogenetic and molecular cytogenetic techniques. Among previously reported genome-wide array studies in unselected cohorts of patients with ID, CA and/or dysmorphic features, the reported detection rate of X-chromosome CNVs (X-CNVs) varied from 0% to about 40%, but in the majority of the studies prevalence ranged from 0 to 5%. These figures are influenced by the low total number of patients in some studies, differences in patient selection, the use of different array platforms with variable probe coverage on the X-chromosome and different detection limits. Some studies only reported pathogenic X-CNVs [12–33]. Studies using X-chromosome specific array-CGH reported a higher detection rate of X-CNVs which is also partly explained by the selection of a cohort of predominantly male patients with the suspicion of an X-linked mode of inheritance [34–40]. Studying the X-chromosome aberrations detected by X-chromosome specific arrays has led to the identification of novel X-linked ID genes, like *ZNF674* in non-syndromic X-linked ID [41]. Such studies have also identified deletions and duplications of chromosomal regions that contain known X-linked ID genes, with duplications of the Xq28 region encompassing the *MECP2* gene being the most frequently identified causative X-CNV [42].

Genome-wide array studies in cohorts of patients with ASD revealed several associated CNVs in this group of patients as well,

* Corresponding author. Tel.: +31 24 3613946; fax: +31 24 3668753.
E-mail addresses: m.willemsen@gen.umcn.nl, m.willemsen@antrg.umcn.nl (M.H. Willemsen).

1769-7212/$ – see front matter © 2012 Elsevier Masson SAS. All rights reserved.
http://dx.doi.org/10.1016/j.ejmg.2012.05.001

including X-CNVs. Clinical relevant CNVs are detected in about 5–8% of ASD patients [10,43] Interestingly, several of these CNVs are present both in cohorts of patients with ID, and in cohorts of patients with ASD. In addition, such aberrations have also been identified in patients with other neuropsychiatric disorders like Attention Deficit Hyperactivity Disorder (ADHD) and schizophrenia. Well-known examples are 15q11q13 duplications, 22q11.21 deletions and 16p11.2 deletions and duplications [43]. The study of CNVs in ASD revealed a number of ASD related genes as well, such as *NRXN1, CNTN4, NLGN1* and *ASTN2* [44]. Most of these genes are associated with ID as well. This is another indication for involvement of similar underlying molecular and biological pathways in the etiology of variable cognitive disorders (CD), as ID/developmental delay (DD), ASD and other neuropsychiatric disorders.

Whole-genome scanning technologies have also identified a large number of submicroscopic CNVs that could not be directly associated with CD and/or CA. Furthermore, many CNVs, either on the autosomes or the X-chromosome, are also present in the general population, giving rise to difficulties in the interpretation of these CNVs [45–47]. X-CNVs form a special group, because dosage effects likely differ in males and females and causative CNVs in a male patient can be inherited from a healthy mother.

In total, various studies reported over 100 X-CNVs in control individuals, which are collected in databases (Databases of Genomic Variants: http://projects.tcag.ca/variation; [37]).

In the present study we provide an overview of X-CNVs identified by genome wide array analysis in a cohort of 4407 individuals with various clinical presentations of CD, and/or CA, including an interpretation of the clinical relevance. Our aim was to provide a resource for clinicians and laboratory specialist for the interpretation of X-CNVs that are picked up by genome wide array platforms in a routine diagnostic setting.

2. Patient data

2.1. Ascertainment of patients

Our data are derived from genome-wide copy number profiling performed among a cohort of 4407 individuals. All individuals had been referred between January 2003 and August 2010 to our diagnostic center for the evaluation of cognitive disorders (CD) -including unexplained ID/DD, neuropsychiatric disorders- and/or CA. Neuropsychiatric disorders included mainly ADHD and ADHD. In the majority the indication for genome-wide array analysis was DD/ID (with or without other neuropsychiatric disorders and/or CA). In a minority of the patients genome-wide array analysis was performed because of CA or behavior problems *without* presence of DD/ID. Two third of the total cohort had an age between 1 and 18 years. The male to female ratio was 1.2:1.

3. Methods

3.1. Genome-wide copy number profiling

Genome-wide copy number profiling was either performed by the Agilent 32k BAC array (669 patients) or the Affymetrix 250k SNP array (3738 patients) analysis platform. Since 2009, genome-wide 250k SNP array analysis has replaced routine cytogenetic chromosome studies as the first line diagnostic test for patients with CD and/or CA in our laboratory.

Genome-wide 32k BAC array analysis was performed as previously described [14]. Samples for genome-wide 250k SNP array analysis were processed in accordance with the standard Affymetrix GeneChip protocol (Affymetrix inc, Santa Clara, California, USA) and analyzed using the Copy Number Analyzer for GeneChip (CNAG) v2.0 software package [48]. CNVs were mapped according to the USCS genome browser build March 2006.

3.2. Segregation analysis

Segregation of the detected CNVs was tested in the parents if available. In case of affected family members, segregation of the CNV with the disease was further examined. Segregation was tested by Multiplex Ligation-dependent Probe Amplification (MLPA), Fluorescent In Situ Hybridization (FISH), quantitative Polymerase Chain Reaction (qPCR) experiments or 250k SNP array analysis.

3.3. X-chromosome inactivation (XCI) analysis

XCI analysis was performed whenever an X-CNV inherited from a healthy (or obviously milder affected) mother was identified. XCI patterns were determined using a widely used method based on the presence of a polymorphic repeat within the 5′ end of the *Androgen Receptor (AR) gene*, as described by Allen et al.[49].

A sample was considered to have skewed X-inactivation if the same X-chromosome was inactivated in at least 90% of the cells [50].

3.4. Assessment of clinical relevance

X-CNVs were classified into three different categories reflecting their presumed clinical relevance. The classification categories include:

1) (likely) pathogenic CNVs
2) CNVs with unknown clinical relevance
3) (likely) non-pathogenic CNVs

Table 1 shows a schematic overview of the criteria that were considered to assess clinical relevance and to categorize the X-CNVs into these three classification categories. Classifications of X-CNVs in each individual patient in this cohort were based on careful consideration of the total combination of the criteria for assessment of clinical relevance as mentioned in Table 1. None of the criteria was absolute and each criterion was considered in perspective to the other criteria, the currently available information in literature and databases, and the phenotype of the patient.

4. Results

4.1. General overview

In total 57 X-CNVs were analyzed. Tables 2–4 summarize (likely) pathogenic X-CNVs, CNVs with unknown clinical relevance and (likely) non-pathogenic CNVs, respectively. Concise information about the phenotype is included as well. X-CNVs that had initially been identified by conventional karyotyping and were further delineated by SNP array analysis, microscopically visible X-CNVs, and X-CNV gains of the pseudo-autosomal region (PAR) indicating an XXY or XYY karyotype, were excluded from the analysis. The total number of X-CNVs comprised 57 in 4407 (1.3%), including 22 losses and 35 gains. CNV size ranged from 10 kb to 8.5 Mb. Twenty-six percent of these (15/57, 7 gains, 8 losses) were classified as pathogenic, and 42% (24/57) were classified as non-pathogenic (16 gains, 8 losses). In 18 cases we could not make a decision about the pathogenicity (12 gains, 6 losses). This gives a detection rate of approximately 0.3% pathogenic X-CNVs in our cohort, that noteworthy consists of patients not preselected on the