Research in Autism Spectrum Disorders 23 (2016) 145–151

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: http://ees.elsevier.com/RASD/default.asp

Detection of small copy number variations (CNVs) in autism spectrum disorder (ASD) by custom array comparative genomic hybridization (aCGH)

Eloisa S. Moreira[a], Isabela M.W. Silva[a], Naila Lourenço[a], Danielle P. Moreira[a], Cintia M. Ribeiro[a], Ana Luiza B. Martins[b], Karina Griesi-Oliveira[a], Monize Lazar[a], Silvia S. Costa[a], Michel S. Naslavsky[a], Kátia M. Rocha[a], Meire Aguena[a], Agnes C. Fett-Conte[c], Mayana Zatz[a], Carla Rosenberg[a], Elaine C. Zachi[a], Débora R. Bertola[a], Estevão Vadasz[d], Maria Rita Passos-Bueno[a,*]

[a] Departamento de Genética e Biologia Evolutiva, Centro de Estudos do Genoma Humano, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil
[b] Departamento de Biologia, Universidade Estadual Júlio Mesquita Filho, São José do Rio Preto, São Paulo, Brazil
[c] Faculdade de Medicina de São José do Rio Preto, Departamento de Biologia Molecular, São José do Rio Preto, São Paulo, Brazil
[d] Instituto de Psiquiatria do Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 10 May 2015
Received in revised form 13 November 2015
Accepted 27 December 2015
Available online 6 January 2016

Keywords:
Autism
Copy number variation
Comparative genomic hybridization
Neurodevelopmental disorder
MBD2
SLC17A6

ABSTRACT

Autism spectrum disorder (ASD) has a strong genetic basis and advances in genomic scanning methods have resulted in the identification of the underlying alterations in about 30% of the cases. The overwhelming majority of these alterations are either sequencing variants or large copy number variations (CNVs). In this pilot study, we tested whether the use of a customized array comparative genomic hybridization (aCGH), targeting exons of 269 ASD candidate genes, would allow the identification of small potentially pathogenic CNVs (<100 Kb). We detected 10 rare, potentially pathogenic CNVs in nine out of 98 patients with idiopathic ASD, and none of 200 Brazilian controls. Two out of five CNVs identified among the non-syndromic cases, involving the genes *MBD2* and *SLC17A6*, were smaller than 100 Kb. In a subsequent screening of other 407 patients and 350 non-affected controls for CNVs involving *SLC17A6*, a gene without previous documentation in the literature of involvement with neurodevelopmental disorders, we found intragenic duplications in another proband but also in five controls. Of note, a commercial 500 K SNP-array did not detect the smallest gains in *SLC17A6*. Our results suggest that small CNVs contribute to the etiology of ASD and that customized CGH array has significant potential to improve the sensitivity for detecting this class of alterations.

Abbreviations: aCGH, array comparative genomic hybridization; CNV, copy number variation; MLPA, multiplex ligation probe amplification; NGS, next generation sequencing.

* Corresponding author at: Departamento de Genetica e Biologia Evolutiva, Instituto de Biociencias, Rua do Matao, 277, Sala 200, Cidade Universitaria, CEP: 05508-090 Sao Paulo, SP, Brazil. Fax: +55 11 3091 7419.
E-mail address: passos@ib.usp.br (M.R. Passos-Bueno).

http://dx.doi.org/10.1016/j.rasd.2015.12.012
1750-9467/

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by significant and persistent deficits in communication and social interaction, and by restricted and repetitive patterns of behavior, interests or activities (American Psychiatric Association, 2013). ASD is considered a common disorder with a strong genetic component, affecting one in 68 individuals and presenting a gender ratio of 4.2 boys–1 girl (Centers for Disease Control and Prevention, 2014; Devlin & Scherer, 2012; Fombonne, 2009; Gillberg, Cederlund, Lamberg, & Zeijlon, 2006).

High-throughput genomic approaches, as array comparative genomic hybridization (aCGH) and next generation sequencing (NGS), have revolutionized the knowledge on ASD genetics, unraveling a number of rare copy number variations (CNVs) and sequencing variants for which there is strong evidence for a causal role in the phenotype (Betancur, 2011; Hoischen, Krumm, & Eichler, 2014). Nevertheless, in more than 70% of the ASD patients, the underlying mutations remain to be identified (Betancur, 2011; Buxbaum et al., 2014). We have therefore raised the hypothesis that small CNVs (between 1 and 100 Kb), below the resolution of most of the commercial microarrays and still poorly detected by NGS, could contribute to the phenotype in a significant proportion of cases.

As a first step to address this question, we have performed a pilot study using a customized aCGH with high density of probes targeting exons of 269 ASD candidate genes, in order to screen for small potentially pathogenic CNVs among Brazilian patients with ASD.

2. Materials and methods

2.1. Patients

A total of 505 Brazilian patients with idiopathic ASD were included in the study. While 434 patients were referred to the Human Genome and Stem-Cell Research Center (HUG-CELL), Sao Paulo (most ascertained at the Institute of Psychiatry, University of Sao Paulo), 71 were ascertained at the *Hospital de Base*, Faculty of Medicine of Sao Jose do Rio Preto. ASD was diagnosed based on the fourth or fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM) by psychiatrists at the above-mentioned institutions, and evaluated following previously standardized criteria (Griesi-Oliveira et al., 2012; Moreira et al., 2014; Orabona et al., 2009). In most cases children's behavior was characterized by score questions of Childhood Autism Rating Scale (CARS) and/or by an interview adapted from the Autism Diagnostic Interview-Revised (ADI-R) (Griesi-Oliveira et al., 2012). At the HUG-Cell, the intelligence quotient (IQ) of ASD patients has been systematically assessed in the last two years by means of the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) and of the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III) (Nascimento & Figueiredo, 2002). For some patients, particularly for those who present severe verbal and interaction difficulties, non-verbal IQ calculations were based on the Raven's matrices standard (Campos, 2004).

Patients presenting dysmorphic features and/or congenital malformations were classified as syndromic. Patients with clinically recognizable, autism-related syndromes and metabolic disorders, such as Rett, Prader-Willi, Angelman, Williams, DiGeorge, Beckwith-Wiedmann, and Smith-Lemli-Opit syndromes, were not included in the present study. Exposition to teratogenic drugs or infections during pregnancy was also an exclusion criterion. Finally, all patients were negative for CNVs at 15q11-13, 16p11 and 22q13 through multiplex ligation probe amplification (MLPA) analysis with the SALSA MLPA KIT P343-B1 AUTISM-1 (MRC-Holland, Amsterdam, The Netherlands; Moreira et al., 2014), and all boys were negative for *FMR1* expansion (Haddad, Mingroni-Netto, Vianna-Morgante, & Pena, 1996).

Patients at HUG-CELL were 10.9 ± 7.8 years old at ascertainment, about 23% of them were syndromic, and 56 (13%) were probands of familial cases. The male–female ratio was 3.2–1, which may be accounted for by a higher proportion of patients with intellectual disability (ID) as compared to other studies (roughly 77% versus 55%) with reported male-female ratios closer to 4 to 1 (Centers for Disease Control and Prevention, 2014; Miles, 2015; Werlin & Geschwind, 2013).

Patients ascertained at the *Hospital de Base* were 12.1 ± 7.8 years old at ascertainment, and only 3% of them were syndromic. All 71 patients were sporadic cases, with a male–female ratio of 4.5–1 and an ID rate of 55%.

The Ethics Committee of Bioscience Institute, University of São Paulo approved this research. Individuals were included in this study only after written informed consent signature (by themselves or legal caregivers).

2.2. Array CGH design

A custom oligonucleotide array with approximately 60,000 features (Custom 8×60 K Microarray, Agilent Technologies, Santa Clara, CA, USA) was designed with the eArray web software using the filters: similarity score, probe score above 0.8 and probe $\geq$59 bp, in order to select the best-performing oligonucleotides from Agilent's library. The array, which was designed before 2010, comprises probes targeted to all exons, including those at 5′ and 3′ UTR regions, of 269 selected genes based on the following criteria: (i) genes previously reported as candidates for ASD and/or other neuropsychiatric disorders in at least two different studies; (ii) genes involved in the ASD-associated pathways and with similar functions to known candidate ASD genes; (iii) genes mapped in ASD known/candidate regions and encoding proteins with neurological functions. The array included over 33,000 tiling probes targeted to 4229 exons of these genes, with an average of eight probes per exon and at least three probes covering the overwhelming majority of exons. This design should provide increased sensitivity for

detecting small CNVs. We also designed around 22,000 probes covering the genome (backbone probes), with an average spacing of 130 Kb, and a higher probe density (spaced from 10 to 30 Kb) in the regions 15q11-13, 16p11.2 and 22q13. The genomic coordinates are based on the February 2009 assembly of the human reference genome (GRCh37/hg19).

2.3. Array CGH hybridization and analysis

The procedures for DNA digestion, labeling with Cy3- and Cy5-dCTPs by random priming followed by purification, hybridization and washing were performed according to the manufacturer's protocol (Agilent Technologies, Santa Clara, CA, USA). Scanned images of the arrays were processed using the Feature Extraction software (Agilent Technologies, Santa Clara, CA, USA), and the analysis was carried out using the Agilent Genomic Workbench 7.0 (Agilent Technologies, Santa Clara, CA, USA). Identification of aberrant copy number segments in data that passed the Quality Control (QC) metrics was based on the Aberration Detection Method 2 (ADM-2) algorithm with default settings. Threshold was set to 5. We considered at least three consecutive probes spanning a minimum of two exons with $\log_2$ (Cy5/Cy3) ratios lower than -0.3 or greater than 0.3 for calling a segment. All samples were tested in duplicate, using a dye-swap strategy with a single sex-matched reference, and we only accepted concordant results.

Only rare CNVs that overlapped coding regions and at least two exons were considered potentially pathogenic. This last criterion was adopted based on the results of a initial validation experiment with deletions confined to one exon, all smaller than 1 Kb and involving regions with an extremely high degree of probe density/overlapping. None of the 10 events (deletions $<$ 1 kb) selected for validation by PCR were confirmed.

In order to be deemed rare, a CNV could not have been detected in any of 200 healthy Brazilian individuals tested with the same array CGH platform, should present a frequency $\leq$5% in our total sample of ASD patients, and should not have been reported in the Database of Genomic Variants (DGV) or, if reported, its frequency should be $\leq$1% (http://dgv.tcag.ca/dgv/app/home; Iafrate et al., 2004; MacDonald, Ziman, Yuen, Feuk, & Scherer, 2014). Rare CNVs detected in non-syndromic cases, which were the first to be tested with the customized array (before the syndromic patients), were confirmed using at least one of the following methods: another array CGH platform (SurePrint G3Human CGH Microarrays 180 K; Agilent Technologies, Santa Clara, CA, USA); a SNP-array platform (GeneChip Human Mapping 500 K Array Set, Affymetrix, Inc., Santa Clara, CA, USA); multiplex ligation dependent probe amplification (MLPA). Genetic origin of rare CNVs was tested in cases with available parental DNA.

2.4. Single nucleotide polymorphism array

The GeneChip Human Mapping 500 K Array Set (Affymetrix, Inc., Santa Clara, CA, USA) is comprised of two arrays, each capable of genotyping on average 500,000 SNPs (approximately 262,000 for *Nsp* arrays and 238,000 for *Sty* arrays). The assay was performed according to the standard protocol. The quality of hybridization samples was evaluated by QC Call Rate, using a threshold of 95.0. For the analysis we used the Genotyping Console software, (Affymetrix, Inc., Santa Clara, CA, USA), PennCNV (Wang et al., 2007) and dChip (Zhao et al., 2004) with default settings. We considered only CNVs detected by at least two different softwares.

2.5. Multiplex ligation dependent probe amplification (MLPA)

Probes were designed using the human MLPA probe design algorithm (H-MAPD) (Zhi, 2010) and synthesized by IDT Technologies. As for *SLC17A6* (NM_020346.2), it was possible to design probes for all exons but the first one. MLPA reactions were performed according to the manufacturer's instructions (MRC-Holland) and as previously described (Ahn et al., 2007). The reference commercial kit P200 was used in combination with custom probes for data normalization and quantity/quality control fragments. The amplified products were separated in an ABI-3730 DNA Analyzer (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA, USA). Results analyses were performed with the help of the softwares GeneMapper version 4.0 (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA, USA) and GeneMarker® version 1.97 (SoftGenetics, State College, PA, USA). Ratios below 0.75 and above 1.30 were considered as deletions and duplications, respectively.

2.6. Screening of SLC17A6 CNVs by whole exome sequencing (WES)

Screening of CNVs in the *SLC17A6* gene in 200 individuals of the control population was performed through whole exome sequencing (WES). CNV detection by WES was validated by the analysis of two patients who harbor *SLC17A6* intragenic duplications previously detected with both aCGH and MLPA.

Exome sequencing of genomic DNA was performed with the Nextera Rapid Capture Expanded kits (Illumina, San Diego, CA, USA), for library preparation and exome capture, and the HiSeq 2500 System (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. Reads were processed by the NextGENe software® (SoftGenetics, State College, PA, USA) for alignment and CNV calling (Shouyong & Liu, 2013). Briefly, the coverage (reads per kilobase per million, or RPKM) ratio (sample divided by sample plus control) is used as the basis for CNV detection. A beta-binomial model is fit to the coverage ratio in order to model the amount of dispersion (noise), and likelihood values are calculated based on the dispersion measurements and coverage ratios. The probabilities are then entered into a Hidden Markov Model (HMM) to make CNV

classifications for each region, namely "Insertion" (increased copy number), "Normal" (little evidence of a CNV), "Deletion", or "Uncalled" (due to low coverage). Additionally, each region receives three Phred-scaled probability scores (for insertion, normal and deletion status).

3. Results

A preliminary validation of the array was done using 10 ASD samples previously assigned by MLPA to harbor CNVs at the regions 15q11-q13, 16p11.2 and 22q13. All CNVs were detected by the customized array.

Next, we screened 98 patients with idiopathic ASD (77 non-syndromic and 21 syndromic) and 200 non-affected subjects for genomic imbalances using the custom CGH array. Based on this screening, ten CNVs detected in nine cases, ranging from 9 to 21,697 Kb, fulfilled our above-mentioned criteria for potential pathogenicity (Table 1). The two smallest CNVs, having less than 100 Kb, were detected in non-syndromic patients. No other alterations were detected in the non-syndromic patients using a whole genome SNP-array.

The 10 CNVs described in Table 1 include three single-gene CNVs. Among those genes, only *SLC17A6* (Fig. 1) did not present substantial documentation in the literature of involvement with neurodevelopmental disorders. This fact prompted us to screen other 407 individuals with idiopathic ASD and 350 non-affected controls for CNVs either involving this gene either by MLPA or WES. We found that another proband but also five controls harbor duplications which include two to four exons of *SLC17A6*, from exon 9 to exon 12.

4. Discussion

In this pilot study, we have found that 20% of the CNVs detected with our customized CGH array were less than 100 Kb (or 40%, considering only non-syndromic cases) in a small Brazilian cohort with idiopathic ASD, thus suggesting that small CNVs play a role in the etiology of ASD.

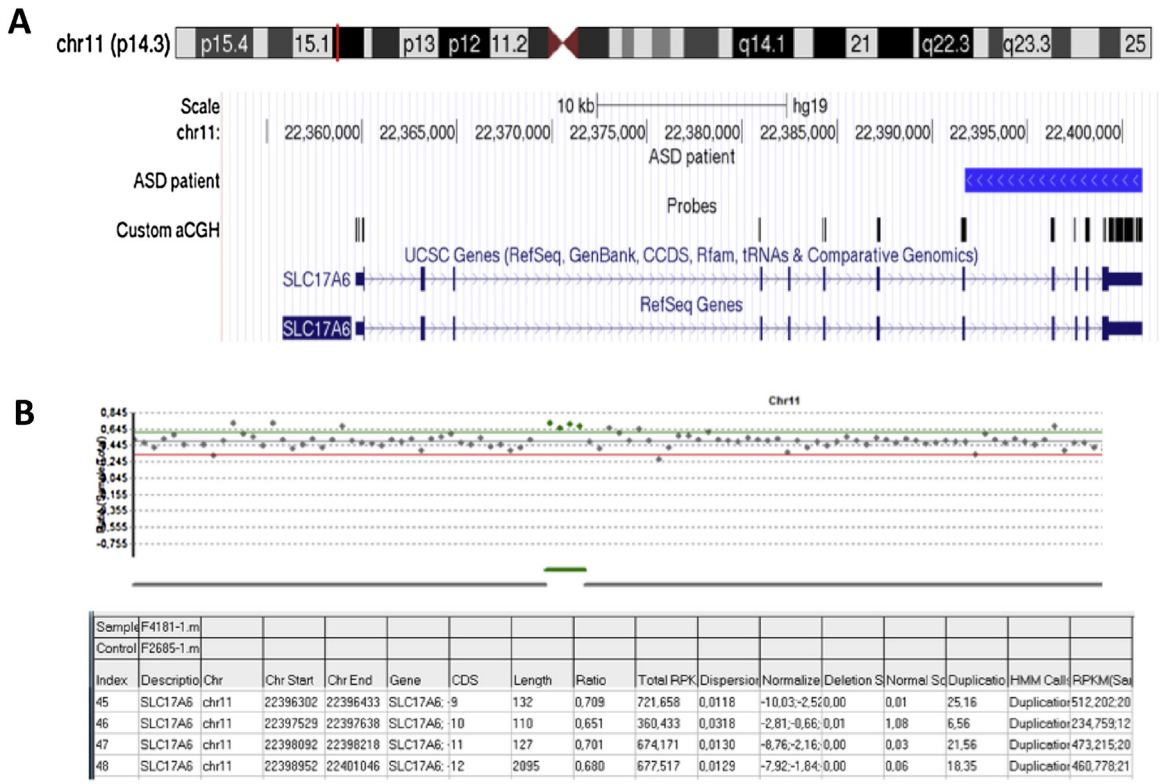
The prevalence of all rare, potentially pathogenic CNVs in this study was 9%. At first glance, one might think that our detection rate is similar to the 10% frequency commonly described in studies with conventional CGH or SNP-array in ASD. Nevertheless, our sample was submitted to a prescreening in order to exclude patients with clinically recognizable syndromes and those harboring the most prevalent CNVs in ASD, at proximal 16p11.2, 15q11-q13, and 22q13. It is of note that the combined prevalence of CNVs in these regions among Brazilian ASD patients has been estimated at 2.1% (Moreira et al., 2014). Also, our array covers only 269 candidate genes, contrasting to whole-genome coverage in most studies and to the current number of ASD candidate genes (740) compiled by the Simons Foundation Autism Research Initiative (SFARI) (http://sfari.org/; Basu, Kollu, & Banerjee-Basu, 2009).

We could determine the origin of four out of the ten detected CNVs, with only one of them being de novo. Nevertheless, both de novo and inherited genetic variations are implicated in the current models of ASD risk (De Rubeis & Buxbaum, 2015), with recent reports suggesting a specific role for small transmitted rare CNVs (Krumm et al., 2013, 2015).

Among the ten alterations, seven large CNVs (>100 Kb) that encompass several genes (detected in the syndromic patients and in one non-syndromic patient) have already been described as associated with ASD at Decipher (https://decipher.sanger.ac.uk/; Firth et al., 2009), SFARI (Basu, Kollu, & Banerjee-Basu, 2009) and/or in other studies (Prasad et al., 2012; Vorstman

Table 1
Results of customized array comparative genomic analysis of 98 patients affected by autism spectrum disorder.

Patient	Sex	Phenotype	CNV position	Type	Origin	Min (Max) length (Kb)	Position (Hg19)	Gene(s) involved
1	Male	Non-syndromic	11p14.3	Gain	Not maternal; father's DNA not available	9 (31)	22,391,688–22,401,068	*SLC17A6*
2	Male	Non-syndromic	7q31.1	Loss	Parents not avaible	480 (533)	112,644,855–113,125,338	*GPR85, LOC401397*
3	Male	Non-syndromic	13q21.2	Loss	Paternal	306 (376)	60,420,512–60,727,008	*DIAPH3*
4	Male	Non-syndromic	1q21.1–21.2	Gain	Maternal	1322 (3722)	146,506,310–147,828,089	>15
5	Male	Non-syndromic	18q21.2	Gain	Maternal	12 (85)	51,680,552–51,692,525	*MBD2*
6	Male	Syndromic	Xq24–25	Loss	*De novo*	142 (2779)	122,318,291–122,460,062	*GRIA3* (>15)
7	Male	Syndromic	11q13.3–13.4	Loss	Not paternal; mother's DNA not avaible	1260 (2677)	69,482,657–70,742,731	>15
8	Female	Syndromic	1p31.3–22.3	Gain	Parents not avaible	21,697 (25,415)	64,708,662–86,405,527	>15
			14q21.2–21.3	Loss		1620 (3458)	46,375,799–47,995,925	*LINC0087, RPL10L, MDGA2* (>15)
9	Female	Syndromic	3p14.3–14.1	Loss	Parents not available	10,225 (14,469)	58,196,379–68,421,182	>15



Sample	F4181-1.m															
Control	F2685-1.m															
Index	Descriptio	Chr	Chr Start	Chr End	Gene	CDS	Length	Ratio	Total RPK	Dispersion	Normalize	Deletion S	Normal Sc	Duplicatio	HMM Call	RPKM(Sa
45	SLC17A6	chr11	22396302	22396433	SLC17A6;	9	132	0,709	721,658	0,0118	-10,03;-2,52	0,00	0,01	25,16	Duplicatior	512,202;20
46	SLC17A6	chr11	22397529	22397638	SLC17A6;	10	110	0,651	360,433	0,0318	-2,81;-0,66	0,01	1,08	6,56	Duplicatior	234,759;12
47	SLC17A6	chr11	22398092	22398218	SLC17A6;	11	127	0,701	674,171	0,0130	-8,76;-2,16	0,00	0,03	21,56	Duplicatior	473,215;20
48	SLC17A6	chr11	22398952	22401046	SLC17A6;	12	2095	0,680	677,517	0,0129	-7,92;-1,84	0,00	0,06	18,35	Duplicatior	460,778;21

Fig. 1. Schematic representations and tabular results of the duplication at *SLC17A6* detected in patient 1 using (A) custom array comparative genomic hybridization (light blue bar) and (B) exome sequencing (green dots and green line on the scheme). The exome sequencing results obtained with the software NextGENe® (SoftGenetics) include: (i) a graphical representation of the coverage (reads per kilobase per million, or RPKM) ratio (sample divided by sample plus control) along a region of chromosome 11 (top); with dot colors corresponding to the predicted copy number status (gray = normal, green = gain); (ii) a CNV report showing chromosome number, position, gene, exon number (CDS column), length and Phred-scaled probability score (15th column), among other parameters, for the duplication (bottom). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

et al., 2011). Among these CNVs, the deletion presented by proband 2 is worth noting, since one of the two genes lost (*GPR85*, or *SREB2*) encodes a G protein-coupled receptor that belongs to the SREB (Super Conserved Receptor Expressed in Brain) family. Although only one deletion involving *GPR85* has been reported so far in ASD (Prasad et al., 2012; SFARI database), two SNPs at this gene were associated with risk for schizophrenia (Matsumoto et al., 2008). Our data thus reinforce the relevance of *GPR85* to ASD etiology.

Other two CNVs detected in two non-syndromic cases involve only one gene each: a deletion involving most of *DIAPH3* and a duplication involving exons 4–7 of *MBD2*. MBD2 is a methyl-CpG-binding protein, and a number of deletions and duplications disrupting this gene in ASD patients are listed in the SFARI database. DIAPH3 acts in the assembly of actin filaments essential for neuron migration during brain development. The importance of DIAPH3 in ASD is further supported by the description of four ASD patients with intragenic deletions in the SFARI database. The *DIAPH3* deletion in our patient was inherited from his non-affected father, suggesting the need for at least a second hit to disorder initiation. Indeed, Vorstman et al. (2011) reported an ASD proband who inherited a large deletion at 13q21.2 from his mother and a missense, most likely loss of function mutation at *DIAPH3* from his father. These data reinforce that inherited CNVs should also be considered in the molecular dissection of ASDs.

The *SLC17A6* gene encodes VGLUT2, one of the three known isoforms of vesicular glutamate transporters (Berry, Sceniak, Zhou, & Sabo, 2012), which is responsible for glutamate uptake into synaptic vesicles in presynaptic neurons (Chiocchetti, Bour, & Freitag, 2014; Moechars et al., 2006). We also found *SLC17A6* duplications at a frequency of 0.9% in Brazilian non-affected controls. Considering our relatively small sample size and the potential of this gene as functional candidate to ASD, it will be important to validate this finding in larger cohorts of both ASD and matched controls. These results demonstrate the importance of screening control samples for small CNVs using not only ethnically matched controls but also appropriate method(s), since small CNVs may be underrepresented in the current databases due to the resolution of most arrays.

Our study suggests the search of small CNVs implicated in the etiology of ASD is worth to be pursued. It also seems that our approach has a significant potential in terms of improving the sensitivity for detecting small disease-related CNVs. Indeed, the 500 K SNP-array did not detect (using the recommended criteria of greater than 20 consecutive probes) the smallest gains in the gene *SLC17A6* and the gain in *MBD2*, detected by the customized array (data not shown). We are aware

that the number of genes included in our array (40% of the current number of ASD candidate genes compiled by SFARI) represents a drawback in our study; nevertheless it serves the purpose to test our hypothesis.

Regarding the design of our array, we would like to point two other limitations: (i) an exaggerated degree of overlapping/probe density in some exons, resulting in noise and false positive indels/small CNVs in those regions, as proven by our validation strategy; as a consequence, we were conservative in our analysis and discarded all alterations involving only one exon, a strategy that may be associated to an underestimation of the number of small CNVs in our sample (see Supplementary figure); (ii) the average backbone probe spacing (130 Kb) provided a genome-wide suboptimal resolution (roughly 390 Kb).

Based on these findings and on current Agilent guidelines to customize CGH arrays, we have designed a new version of the customized array (4 × 180 K) aimed to clinical applications, with probes covering 1527 known/candidate ASD genes (based on SFARI Dec/14 and on recent studies refs) and a genomic resolution of around 109 Kb. Although SNP arrays with millions of probes are currently available at a similar cost, they often fail to detect small CNVs due to the limited, uneven distribution of SNPs in the genome (Celestino-Soper et al., 2011; Wiszniewska et al., 2014). A study by Haraksingh, Abyzov, Gerstein, Urban, and Snyder (2011) comparing twelve different platforms based on four main designs, whole genome SNP/CGH/SNP + CGH arrays and "CNV focused arrays" (CGH customized arrays), concluded that the latter was the best to detect known CNVs in a well characterized European sample. Indeed, dozens of studies on a range of human diseases have been published in the last few years based on results obtained with an exon targeted CGH array with 180 K probes covering 1700 clinically relevant genes, designed in the Medical Genetics Laboratories at Baylor College of Medicine. Strikingly, in their original report the researchers showed that eight out of 15 cases, the CNVs would have been missed by the standard 180 K CGH array (Boone et al., 2010). WES has also low sensitivity to detect shorter exonic CNVs, and can present other problems such as a high false positivity rate and duplication bias (Guo et al., 2013; Samarakoon et al., 2014).

While the cost of whole genome sequencing remains prohibitive, especially in less developed countries, the combination of WES and customized aCGH seems to represent the best strategy to detect the bulk of the genetic alterations, including small CNVs, associated with ASD.

Conflict of interest

Authors declare no conflict of interest.

Acknowledgments

We would like to express our gratitude to the patients, parents and family members who participated in this study. This study was supported by FAPESP/CEPID and CNPq/INCT grants. MRPB is a CNPq1A researcher.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.rasd.2015.12.012.

References

Ahn, J. W., Ogilvie, C. M., Welch, A., Thomas, H., Madula, R., Hills, A., et al. (2007). Detection of subtelomere imbalance using MLPA: validation, development of an analysis protocol, and application in a diagnostic centre. *BMC Medical Genetics*, *8*, 9.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*, 5th ed. Arlington: American Psychiatric Association.

Basu, S. N., Kollu, R., & Banerjee-Basu, S. (2009). AutDB: a gene reference resource for autism research. *Nucleics Acids Research*, *37*(Database issue), 832–836.

Berry, C. T., Sceniak, M. P., Zhou, L., & Sabo, S. L. (2012). Developmental upregulation of vesicular glutamate transporter-1 promotes neocortical presynaptic terminal development. *PLoS One*, *7*, e50911.

Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Research*, *1380*, 42–77.

Boone, P. M., Bacino, C. A., Shaw, C. A., Eng, P. A., Hizson, P. M., Pursley, A. N., et al. (2010). Detection of clinically relevant exon copy-number changes by array CGH. *Human Mutation*, *31*, 1326–1342.

Buxbaum, J. D., Bolshakova, N., Brownfeld, J. M., Anney, R. J. L., Bender, P., Bernier, R., et al. (2014). The autism simplex collection: an international, expertly phenotyped autism sample for genetic and phenotypic analyses. *Molecular Autism*, *5*, 34.

Campos, F. (2004). *Teste das matrizes progressivas: escala geral series A, B, C, D e E. Manual*. Rio de Janeiro: CEPA.

Celestino-Soper, P. B., Shaw, C. A., Sanders, S. J., Li, J., Murtha, M. T., Ercan-Sencicek, A. G., et al. (2011). Use of array CGH to detect exonic copy number variants throughout the genome in autism families detects a novel deletion in TMLHE. *Human Molecular Genetics*, *20*, 4360–4370.

Centers for Disease Control and Prevention (2014). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveillance Summaries*, *63*, 1–21.

Chiocchetti, A. G., Bour, H. S., & Freitag, C. M. (2014). Glutamatergic candidate genes in autismo spectrum disorder: an overview. *Journal of Neural Transmission*, *121*, 1081–1106.

De Rubeis, S., & Buxbaum, J. D. (2015). Genetics and genomics of autism spectrum disorder: embracing complexity. *Human Molecular Genetics*, *24*, R24–R31.

Devlin, B., & Scherer, S. W. (2012). Genetic architecture in autism spectrum disorder. *Current Opinion in Genetics & Development*, *22*, 229–237.

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., et al. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *American Journal of Human Genetics*, *84*, 524–533.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*, 591–598.

Gillberg, C., Cederlund, M., Lamberg, K., & Zeijlon, L. (2006). Brief report: "the autism epidemic": the registered prevalence of autism in a Swedish urban area. *Journal of Autism and Developmental Disorders*, *36*, 429–435.

Griesi-Oliveira, K., Moreira, D. P., Davis-Wright, N., Sanders, S., Mason, C., Orabona, G. M., et al. (2012). A complex chromosomal rearrangement involving chromosomes 2, 5, and X in autism spectrum disorder. *American Journal of Medical Genetics Part B, Neuropsychiatric Genetics*, *159B*, 529–536.
Guo, Y., Sheng, Q., Samuels, D. C., Lehmann, B., Bauer, J. A., Pietenpol, J., et al. (2013). Comparative study of exon copy number variation estimation tools using array comparative genomic hybridization as control. *BioMed Research International*915636.
Haddad, L. A., Mingroni-Netto, R. C., Vianna-Morgante, A. M., & Pena, S. D. (1996). A PCR-based test suitable for screening for fragile X syndrome among mentally retarded males. *Human Genetics*, *97*, 808–812.
Haraksingh, R. R., Abyzov, A. A., Gerstein, M., Urban, A. E., & Snyder, M. (2011). Genome-wide mapping of copy number variations in humans: comparative analysis of high resolution array platforms. *PLoS One*, *6*, e27859.
Hoischen, A., Krumm, N., & Eichler, E. E. (2014). Prioritization of neurodevelopmental disease genes by discovery of new mutations. *Nature Neuroscience*, *17*, 764–772.
Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., et al. (2004). Detection of large-scale variation in the human genome. *Nature Genetics*, *36*, 949–951.
Krumm, N., O'Roak, B. J., Karakoc, E., Mohajeri, K., Nelson, B., Vives, L., et al. (2013). Transmission disequilibrium of small CNVs in simplex autism. *American Journal of Human Genetics*, *93*, 595–606.
Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Whiterspoon, K., et al. (2015). Excess of rare, inherited truncating mutations in autism. *Nature Genetics*, *47*, 582–588.
MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., & Scherer, S. W. (2014). The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Research*, *42*(Database issue), D986–D992.
Matsumoto, M., Straub, R. E., Marenco, S., Nicodemus, K. K., Matsumoto, S., Fujikawa, A., et al. (2008). The evolutionarily conserved G protein-coupled receptor SREB2/GPR85 influences brain size, behavior, and vulnerability to schizophrenia. *Proceedings of the National Academy of Sciences of the USA*, *105*, 6133–6138.
Miles, J. H. (2015). Complex autism spectrum disorders and cutting-edge molecular diagnostic tests. *The Journal of The American Medical Association*, *314*, 879–880.
Moechars, D., Weston, M. C., Leo, S., Callaerts-Vegh, Z., Goris, I., Daneels, G., et al. (2006). Vesicular glutamate transporter VGLUT2 expression levels control quantal size and neuropathic pain. *Journal of Neuroscience*, *26*, 12055–12066.
Moreira, D. P., Griesi-Oliveira, K., Bossolani-Martins, A. L., Lourenço, N. C., Takahashi, V. N., Rocha, K. M., et al. (2014). Investigation of 15q11-q13, 16p11.2 and 22q13CNVs in autism spectrum disorder Brazilian individuals with and without epilepsy. *PLoS One*, *9*, e107705.
Nascimento, E., & Figueiredo, V. L. M. (2002). WISC-III e WAIS-III: alterações nas versões originais americanas decorrentes das adaptações para uso no Brasil. *Psicologia: Reflexão e Crítica*, *15*, 603–612.
Orabona, G. M., Griesi-Oliveira, K., Vadasz, E., Bulcão, V. L., Takahashi, V. N., Moreira, E. S., et al. (2009). HTR1B and HTR2C in autism spectrum disorders in Brazilian families. *Brain Research*, *1250*, 14–19.
Prasad, A., Merico, D., Thiruvahindrapuram, B., Wei, J., Lionel, A. C., Sato, D., et al. (2012). A discovery resource of rare copy number variations in individuals with autism spectrum disorder. *G3*, *2*, 1665–1685.
Samarakoon, P. S., Sorte, H. S., Kristiansen, B. E., Skodje, T., Sheng, Y., Tjønnfjord, G. E., et al. (2014). Identification of copy number variants from exome sequence data. *BMC Genomics*, *15*, 661.
Shouyong, N., & Liu, C. S. (2013). Copy number variation detection through the utilization of unique algorithms in the CNV tool in NextGENe Software®. *Journal of Biomolecular Techniques*, *24*(Suppl), S72.
Vorstman, J. A., van Daalen, E., Jalali, G. R., Schmidt, E. R., Pasterkamp, R. J., de Jonge, M., et al. (2011). A double hit implicates DIAPH3 as an autism risk gene. *Molecular Psychiatry*, *16*, 442–451.
Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F., et al. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research*, *17*, 1665–1674.
Werlin, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology*, *26*, 146–153.
Wiszniewska, J., Bi, W., Shaw, C., Stankiewicz, P., Kang, S. H., Pursley, A. N., et al. (2014). Combined array CGH plus SNP genome analyses in a single assay for optimized clinical testing. *European Journal of Human Genetics*, *22*, 79–87.
Zhao, X., Li, C., Paez, J. G., Chin, K., Janne, P. A., Chen, T. H., et al. (2004). An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Research*, *64*, 3060–3071.
Zhi, J. (2010). MA PD: a probe design suite for multiplex ligation-dependent probe amplification assays. *BMC Research Notes*, *3*, 137.