UNIVERSIDADE SÃO PAULO

CINTIA MARQUES RIBEIRO

Estudo de genes candidatos aos Transtornos do Espectro Autista

Study of candidate genes to Autism Spectrum Disorders

São Paulo
2013

CINTIA MARQUES RIBEIRO

Estudo de genes candidatos aos Transtornos do Espectro Autista

Study of candidate genes to Autism Spectrum Disorders

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo para obtenção do título de Doutor em Ciências, Área de Concentração: Biologia/Genética

Orientador: Prof. Dr. Maria Rita Passos Bueno

São Paulo
2013

Ficha Catalográfica

Ribeiro, Cintia
Estudo de genes candidatos aos Transtornos do Espectro Autista.
100 páginas

Dissertação (Doutorado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Genética e Biologia Evolutiva.

1. Transtornos do Espectro Autista
2. Variações no número de cópias
3. Microarray-CGH
I. Universidade de São Paulo. Instituto de Biociências. Departamento de Genética e Biologia Evolutiva.

Comissão Julgadora:

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Orientador(a)

Dedicatórias

Ao meu marido, Alfredo Tavares, pela amizade, amor, companheirismo e apoio incondicional durante os anos de realização deste trabalho.

A minha amada família pelo carinho, apoio e dedicação

" Eu acredito demais na sorte.
E tenho constatado que, quanto mais duro eu trabalho,
mais sorte eu tenho ".

Thomas Jefferson

Agradecimentos

Agradeço ao meu marido e grande amigo Alfredo, quem me incentiva a batalhar por meus sonhos sem nunca desistir, pela compreensão, dedicação, amor e cumplicidade.

Agradeço a minha mãe Rosely, quem me ensinou a gostar de estudar, ser esforçada e determinada, ao meu pai Sérgio e a minha avó Anita pela imensa dedicação, amor, amizade e apoio.

Agradeço meus irmãos Alexandre e Marcelo pelos exemplos de perseverança, determinação e força de vontade.

Agradeço a Sílvia, Danielle, Patrícia, Érika, Maurício, Jefferson e as crianças mais lindas desse mundo: Ana Vitória, Eduardo, Marcelo Júnior, Pedro Henrique, Emilly e Joaquim.

Agradeço aos meus sogros pelo carinho.

Família, amo vocês mais do que tudo!

Agradeço a professora Maria Rita pela oportunidade, orientação, ensinamentos, amizade e por despertar em mim o senso crítico, tão importante para um pesquisador.

Agradeço a todos meus amigos queridos do laboratório por todas as risadas que demos juntos. As lições foram muitas mas talvez a mais importante tenha sido aprender a sorrir sempre, o sorriso deixa tudo mais leve. Ana, Andressa, Melina, Francine, Simone, Larissa e Joana, para vocês obrigada é pouco. São mais do que amigas, são irmãzinhas especiais, companheiras nas horas de diversão e trabalho. Cada almoço, cafezinho, *happy hours*, jogos de poker e festinhas foram momentos dos quais me lembrarei com carinho sempre. Ana, você acrescentou muito a este estudo, Si e Andressa, vocês também trabalharam duro e me ajudaram demais.

Agradeço de coração por tudo, inclusive pelos apelidos criativos que inventou para mim Ana. Mel obrigada por ser tão companheira e me ajudar quando eu estava totalmente perdida.

Agradeço a todos do grupo de Autismo, em especial Eloísa e Karina.

As minhas amigas de hoje e sempre, por deixarem minha vida tão mais alegre e divertida: Meire, Francine Ribeiro, Flávia, Mariane e Juliana. Apesar da distância carrego vocês no coração.

Agradeço toda a equipe do Centro de Estudos do Genoma Humano, aos psiquiatras e médicos, em especial a Dra Débora Bertola, que colaboraram com este estudo.

Agradeço com muito carinho a Sílvia Costa por me ensinar a técnica e a análise de microarray e a Dra Carla Rosenberg por permitir o uso dos equipamentos de seu laboratório.

Agradeço a Deus, quem me amparou nos momentos difíceis, me deu a força e a sabedoria necessária para conquistar mais este objetivo.

Agradeço o apoio financeiro da Fundação de Amparo a Pesquisa do estado de São Paulo – **FAPESP**.

Índice

Capítulo 1 – **Introdução Geral, Justificativa e Objetivos** ---------- 01
1_1 Características Clínicas, Prevalência dos TEA e Aspectos Genéticos ---------- 01
1_2 Variações no número de cópias (CNVs) e TEA ---------- 06
1_3 Sequenciamento de última geração e TEA ---------- 08
1_4 Justificativa ---------- 09
1_5 Objetivos ---------- 10
Capítulo 2 - **Materiais e métodos** ----------11
2_1 Fluxograma do projeto ----------11
2_2 Pacientes e Controles ----------12
2_3 Customização da lâmina de aCGH ----------13
2_4 aCGH ----------16
2_5 Análises in silico ----------19
2_6 SNP-array ----------20
2_7 PCR quantitativo ----------21
2_8 MLPA e PCR ----------23
2_9 Sequenciamento de última geração ----------25
Capítulo 3 - **Use of a custom-designed array-CGH to identify small CNVs in nonsyndromic Autism Spectrum Disorders.** ---------- 35
Abstract ---------- 35
Resumo ---------- 36
Introduction ---------- 37
Materials and methods ---------- 38
Results ---------- 43
Discussion ---------- 48
Reference ---------- 51
Capítulo 4 - **Characterization of microchromosomal rearrangements in syndromic autistic patients** ---------- 59
Abstract ---------- 59
Resumo ---------- 59
Introduction ---------- 60
Materials and methods ---------- 61
Results ---------- 63
Discussion ---------- 67
Reference ---------- 71
Capítulo 5 - **Search for genetic factors that modulate the severity of Autism Spectrum Disorders** ---------- 75
Abstract ---------- 75
Resumo ---------- 75
Introduction ---------- 76
Materials and methods ---------- 77

	Results	---------- 79
	Discussion	---------- 84
	Reference	---------- 86
Capítulo 6 -	**Discussão Geral e Conclusões**	---------- 91
	6_1 Discussão Geral	---------- 91
	6_2 Conclusões	---------- 94
Referências Bibliográficas		---------- 95

Lista de Tabelas e Figuras

Capítulo 2 -	Figura 2.1	---------- 11
	Figura 2.2	---------- 15
	Figura 2.3	---------- 18
	Figura 2.4	---------- 19
	Figura 2.5	---------- 24
	Tabela 2.1	---------- 27
	Tabela 2. 2	---------- 32
	Tabela 2.3	----------33
Capítulo 3 -	Tabela 3.1	---------- 45
Capítulo 4 -	Tabela 4.1	---------- 65
Capítulo 5 -	Figura 5.1	---------- 80
	Tabela 5.1	---------- 82

Os transtornos do espectro autista (TEA) são condições neuropsiquiátricas caracterizadas por padrões comportamentais estereotipados, ausência ou limitação de comunicação verbal e de interação social recíproca. Diversos estudos têm mostrado que esses transtornos possuem etiologia genética complexa e heterogênea, o que dificulta a identificação dos fatores causais. Estima-se que cerca de 70% dos casos de TEA são idiopáticos. Portanto, com o objetivo de identificar mecanismos etiológicos associados aos TEA, utilizamos as seguintes estratégias: customização de uma lâmina de microarray CGH que possibilite a detecção não só de grandes CNVs, mas também de alterações menores do que 10 kbp, em *exons* e regiões UTR de genes potencialmente candidatos; a comparação entre os tipos de rearranjos detectados em pacientes sindrômicos e em não sindrômicos e, ainda, a investigação mais detalhada de uma família com indivíduos portadores de transtorno autista e síndrome de Asperger. Foram avaliados 103 portadores de TEA não sindrômicos e 18 sindrômicos, sendo as taxas de detecção de alterações potencialmente patogênicas, respectivamente, de 11,6% e 38,9%. Dentre as alterações detectadas 44,4% são menores do que 10 Kbp. Portanto, a estratégia de usar uma lâmina customizada, com alta densidade de sondas complementares aos *exons* e regiões não codificantes de genes potencialmente envolvidos na etiologia dos TEA, capaz de detectar tanto alterações grandes quanto pequenas, parece ser relevante na tentativa de elucidar o maior número de casos possíveis e melhor compreender esses transtornos. Além disso, essa lâmina também pode ser utilizada como uma ferramenta para auxiliar o diagnóstico clínico e o aconselhamento genético com um custo mais acessível em comparação a outras comerciais ou ao sequenciamento de última geração. Cerca de 33,3% das CNVs observadas afetam região UTR, sugerindo que mutações nessas regiões podem explicar uma proporção significativa dos casos nos quais não são detectadas alterações através de outros testes genômicos, visto que a maioria desses ainda não exploram adequadamente regiões não codificantes. Entre os pacientes autistas não sindrômicos verificou-se que a maioria dos genes afetados por CNVs estão envolvidos em duas principais funções biológicas - sinapses glutamatérgicas e orientação axonal, sugerindo que TEA não sindrômico pode ser causado por disfunção em genes diferentes envolvidos em processos fisiológicos comuns.

Diferente do que observamos entre pacientes não sindrômicos, detectamos mais de uma alteração em um mesmo indivíduo ou alterações que englobam mais de um gene entre os pacientes sindrômicos, reforçando o modelo oligogênico para alguns casos de TEA. Por fim, os dados obtidos no estudo da família com portadores de síndrome de Asperger e transtorno autista sugere que a gravidade do quadro clínico possa estar relacionada ao número de mutações e possivelmente por duas mutações diferentes em ambos os alelos de um mesmo gene. Nossos resultados, além de apoiar o envolvimento dos genes *MDGA2*, *FHIT*, *HTR2A*, *SHANK2*, *GRIA3*, *ZNF778, PRKCα*, *CDH15, DIAPH3, GCH1, GRM5, MARK1*, *SLC17A6*, *IMMP2L, BZRAP1*, *SYNGAP1*, *ANK3*, *MAP1A*, *GABRR2* e *LAMC3* nos TEA também sugere novos genes candidatos: *LRRC7, LRRIQ3, CADPS1, NUFIP, SEMA3A, SNAP29, MBD2*, *GAD2, DGKH* e *PARD3*.

The autism spectrum disorders (ASD) are neuropsychiatric conditions typically characterized by social deficits, communication difficulties, stereotyped or repetitive behaviors and interests. Several studies have shown that these disorders have a complex and heterogeneous genetic etiology, which makes difficult to identify the causal factors. Approximately 70% of cases are idiopathic. In order to identify etiological mechanisms associated with ASD, we have used the following strategies: customized a microarray CGH platform that allows detection not only of large CNVs, but also alterations smaller than 10 kbp in exons and UTR regions of potential candidate genes, the comparison between the types of rearrangements detected in syndromic and non-syndromic patients and further, more detailed investigation of a family segregating both autistic disorder and Asperger syndrome. We evaluated 103 nonsyndromic and 18 syndromic patients by the custom-designed array and the detection rate of possibly pathogenic alterations were, respectively, 11.6% and 38.9%. Among these CNVs, 44.4% are smaller than 10 kbp. Therefore, the strategy of using a custom-designed array, enriched with probes targeted to genes potentially involved in the ASD etiology and able to detect both large and small CNVs, seems to be relevant in an attempt to elucidate the largest number of cases and to better understand these disorders. Furthermore, this platform can also be used as a tool to support the clinical diagnosis and genetic counseling with a more affordable cost compared to conventional other or next-generation sequencing. Approximately 33.3% of the observed CNVs affect UTR region, suggesting that mutations in non-coding regions might explain a significant proportion of ASD cases negative for most genomic screenings, which still do not explore adequately these regions. Among nonsyndromic autistic patients we found that most of the genes affected by CNVs are involved in two main biological functions – glutamatergic synapses and axonal guidance, suggesting that nonsyndromic ASD can be caused by dysfunction in different genes of a few common physiological processes. In contrast to our findings in nonsyndromic patients, we detected more than one alteration in a single individual or alterations that involve more than one gene among the syndromic patients, reinforcing the oligogenic model for some cases of ASD. Finally, the data obtained in the study of the family segregating both Asperger syndrome and autistic disorder suggests that the

severity of ASD seems to be modulated by the number of hits and possibly by hits in both alleles of the same gene. Our results support the involvement of genes *MDGA2, FHIT, HTR2A, SHANK2, GRIA3, ZNF778, PRKCα, CDH15, DIAPH3, GCH1, GRM5, MARK1, SLC17A6, IMMP2L, BZRAP1, SYNGAP1, ANK3, MAP1A, GABRR2* and *LAMC3* in ASD etiology and also suggests new candidates: *LRRC7, LRRIQ3, CADPS1, NUFIP, SEMA3A, SNAP29, MBD2, GAD2, DGKH* and *PARD3*.

1_1 Características Clínicas, Prevalência dos TEA e Aspectos Genéticos

De acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais – DSM-IV – os Transtornos Globais do Desenvolvimento (TGD) compreendem: transtorno autista, transtorno desintegrativo da infância, síndrome de Asperger, transtorno global do desenvolvimento sem outra especificação (*Pervasive Developmental Disorders Not Otherwise Specified* – PDD-NOS) e a síndrome de Rett.

O transtorno autista se caracteriza por comprometimento do uso de múltiplos comportamentos não-verbais como por exemplo: contato visual direto, expressão facial, postura e linguagem corporal. Pode haver fracasso em desenvolver relacionamentos com seus pares e esse fracasso assume diferentes formas em diferentes idades. Os indivíduos mais jovens podem demonstrar pouco ou nenhum interesse em estabelecer amizades; os mais velhos podem ter interesse por amizades, mas não compreendem as convenções da interação social. Pode ocorrer uma ausência da busca espontânea pelo prazer compartilhado, interesses ou realizações com outras pessoas e falta de reciprocidade social ou emocional (DSM-IV).

Pode ocorrer atraso ou ausência total de desenvolvimento da linguagem falada. Em indivíduos que desenvolvem a fala, pode existir um acentuado comprometimento da capacidade de iniciar ou manter uma conversação, uso estereotipado e repetitivo da linguagem ou uma linguagem idiossincrática (linguagem que só pode ser entendida claramente por pessoas familiarizadas com o estilo de comunicação do indivíduo). Há um prejuízo acentuado no uso pragmático da linguagem evidenciado pela incapacidade de integrar palavras e gestos ou compreender os aspectos figurado da fala, tais como a ironia, duplo sentindo e o humor (DSM-IV).

Quanto ao comportamento repetitivo, crianças com transtorno autista geralmente apresentam: preocupação com um ou mais padrões estereotipados e restritos de interesse, anormais em intensidade ou foco como, por exemplo, decorar número de telefones, datas, prefixos, estações de rádio, saber o nome de todas as espécies de dinossauros, saber itinerários de ônibus e trens, entre outros; adesão aparentemente inflexível a rotinas ou rituais específicos e não funcionais, por exemplo, percorrer exatamente o mesmo caminho para a escola todos os dias, enfileirar um número exato de brinquedos da mesma maneira repetidas vezes ou imitar a ação de um personagem de desenho inúmeras vezes, apresentar uma reação extremamente negativa a uma mudança mínima ocorrida no ambiente, tal como a mudança de uma cortina ou a mudança na posição de um móvel; maneirismos motores estereotipados e repetitivos (bater palmas, estalar os dedos, balançar-se); ou uma preocupação persistente com partes de objetos (rodinha dos carrinhos) (DSM-IV).

Outras condições médicas, tais como: anomalias no sono e alimentação, hiperatividade, ansiedade, impulsividade, comportamento hétero e auto agressivo, aumento do perímetro cefálico, respostas incomuns a estímulos sensoriais, epilepsia, hipotonia e distúrbios gastrointestinais, podem também estar presentes (AMERICAN PSYCHIATRIC ASSOCIATION, 2000).

Se a criança apresentar um período de pelo menos dois anos de desenvolvimento normal, refletido pela presença de comunicação verbal e não-verbal, relacionamentos sociais, jogos e comportamento adaptativo próprios da idade, seguido de regressão, ou seja, perda de habilidades já adquiridas, o diagnóstico adequado, segundo o DSM-IV, é de transtorno desintegrativo da infância. Essa condição geralmente está associada a grave retardo mental.

A Síndrome de Asperger difere do transtorno autista devido à ausência de atraso na comunicação verbal e de prejuízo significativo de desenvolvimento cognitivo. Já portadores de PDD-NOS são chamados de “autistas atípicos” porque a manifestação clínica desses pacientes não satisfaz por completo os critérios para o diagnóstico de transtorno autista (VEENSTRA-VANDERWEELE; COOK, 2004).

A síndrome de Rett é geralmente causada por mutações *de novo* no gene *MeCP2,* localizado na região Xq28 e o padrão de herança dessa condição é dominante ligado ao X. Estas

mutações são geralmente letais em homens, o que explica o aparecimento quase exclusivo da síndrome no sexo feminino. A maioria das pacientes são aparentemente saudáveis ao nascimento e se desenvolvem normalmente até os 6-18 meses de vida, quando então ocorre um período de estagnação do desenvolvimento, seguido de rápida regressão da linguagem e controle motor. A característica mais marcante da síndrome é a perda de controle de movimentos voluntários das mãos e a sua substituição por movimentos automáticos repetitivos e estereotipados. Episódios de gritos e choro incontrolável são comuns no final do segundo ano de vida. Características adicionais incluem ataques de pânico, bruxismo, apnéia e ou hiperpnéia episódicas, ataxia da marcha e apraxia, tremores, convulsões (50% dos casos) e microcefalia adquirida (PENA, 2005). Durante os anos pré escolares, as crianças com Rett podem exibir dificuldades na interação social similares as observadas no transtorno autista mas essas tendem a ser temporárias (DSM-IV).

Devido a sobreposição clínica existente entre o transtorno autista, a síndrome de Asperger e o PDD-NOS, tornou-se comum o uso do termo Transtornos do Espectro Autista (TEA) para se referir a estas três categorias de TGD.

A nova edição do Manual de Classificação de Doenças Mentais da Associação Americana de Psiquiatria, DSM-V, a qual está prevista para ser lançada em maio de 2013, trará modificações no que diz respeito aos TEA. A principal mudança será a eliminação das categorias: transtorno autista, síndrome de Asperger e PDD-NOS. Existirá apenas uma denominação e essa também mudará, não será "transtornos do espectro autista" e sim "transtorno autista" (American Psychiatric Association DSM-5 Development, disponível *on-line*). O autismo era anteriormente definido por alterações na capacidade de comunicação, na capacidade de interagir socialmente e comportamento esteriotipado e repetitivo, já no DSM-V os dois primeiros sintomas citados se tornam apenas um. Essa mudança é justificada com o argumento de que a interação social depende da comunicação verbal. Ainda, segundo o DSM-IV, um desses sintomas deveria ocorrer pela primeira vez antes dos 3 anos. No DSM-V, esse limite com relação a idade desaparece e o que se coloca é que os sintomas devem surgir na infância (American Psychiatric Association DSM 5 Development, disponível *on-line*).

Visto que o DSM-V entrará em vigor apenas em maio de 2013, no presente trabalho seguimos as definições, classificações e categorias definidas no DSM-IV.

Fombonne, em 2009, relatou que a prevalência estimada dos TEA é de 6,37/1000 (equivalente a uma em cada 157 crianças). Quando o transtorno autista é considerado isoladamente, o estudo aponta uma prevalência de cerca de 2 casos para 1000 indivíduos, com uma razão sexual entre homens e mulheres de 4,2/1. A prevalência de casos de PDD-NOS é a mais alta dentre os TGD, estimada em 3,71/1000, e a de síndrome de Asperger é a mais baixa, calculada em 0,6/1000.

A prevalência dos TEA no Brasil foi recentemente estimada em 2,72/1000 (uma em cada 368 crianças) por um grupo de pesquisadores que avaliaram um bairro da cidade de Atibaia, São Paulo. Os autores discutem que a frequência mais baixa encontrada nesse estudo, em comparação a de outros trabalhos, pode ser explicada pela falta de conscientização da população local sobre esses transtornos, diagnósticos incorretos e falta de sistema para registro dos casos (PAULA *et al*., 2011).

Existem fortes evidências de que fatores genéticos são determinantes para a etiologia dos TEA: estudos de famílias e de gêmeos demonstram uma taxa de concordância entre gêmeos monozigóticos de 60% a 90% em contraste à taxa de concordância de 1% a 10% vista em gêmeos dizigóticos. O risco de recorrência entre irmãos é de 2% a 8% e a herdabilidade é estimada em 90% (BAILEY *et al*., 1995; FOLSTEIN; ROSEN-SHEDLEY, 2001; TIDMARSH; VOLKMAR, 2003). Além disso, cerca de 10% a 15% dos casos estão associados a doenças de herança monogênica, tais como: síndrome do X Frágil, esclerose tuberosa, neurofibromatose e em aproximadamente 3% dos casos, os indivíduos possuem alguma alteração cromossômica detectável por cariótipo (VORSTMAN *et al*., 2006). Porém, em cerca de 70% dos casos de TE, o mecanismo etiológico é ainda desconhecido (SCHAAF *et al*., 2011).

Até o início da década passada, o modelo de herança mais aceito para TEA era o multifatorial ou oligogênico, no qual o fenótipo resultaria de combinações de variantes alélicas comuns de pequeno efeito, somadas ou não com fatores ambientais. Assim, inicialmente muitos estudos de ligação e de associação foram realizados com o objetivo de identificar regiões e alelos de susceptibilidade para TEA. Yang e Gill (2007) revisaram 14 desses estudos e relataram um total de 94 *loci* descritos como significantes ou sugestivos de ligação. Entretanto, apenas os resultados obtidos para as regiões 7q21.2–q36.2, 16p12.1–p13.3, 1q21–q44, 2q24.1–q33.1, 3q21.3–q29, 5p12–p15.33, 6q14.3–q23.2, 10q23.3–q26.3, 17q11.1–q21.2, 19p13.11–p13.3,

1p12–p21.1, 4q21.3–q35.1, 10p12.1–p14, 13q12.13–q33.1, 15q13.1–q26.1 e 19q12–q13.12 foram replicados por pesquisas independentes. Quanto aos estudos de associação, também observamos baixa replicação de resultados, isto é, enquanto alguns estudos reportaram associação entre a variante de um gene e TEA, outros rejeitaram tal associação.

O surgimento de tecnologias mais avançadas, como por exemplo microarranjos de polimorfismos de base única (Single Nucleotide Polymorphism array; *SNP-array*), possibilitaram a identificação de polimorfismos em todo o genoma humano e não apenas em genes alvos. Isto possibilitou estudos de associação genômica ampla (frequentemente chamados pela sigla *GWAS* do nome em inglês *Genome-wide Association Studies*), onde todo o genoma é analisado em vez de apenas regiões ou genes candidatos. Os dados obtidos a partir desses estudos apontaram novas variações em um único nucleotídeo (SNVs - *single nucleotide variant*), porém, novamente observamos baixa replicação entre diferentes estudos e, além disso, o risco atribuído a cada SNV é baixo e não explica a alta herdabilidade dos TEA (EL-FISHAWAY; STATE, 2010).

Experimentos com microarranjos de DNA [SNP-array e Hibridização Genômica Comparativa por microarranjos (aCGH)] também permitiram a identificação de variações no número de cópias (*Copy Number Variation* - CNVs - duplicações ou deleções submicroscópicas) e, recentemente, tem sido sugerido que essas alterações possuem um importante papel na etiologia dos TEA, podendo estar associados ao mecanismo etiológico em cerca de 20% dos casos (MILES *et al*., 2010). O grande número de estudos reportando CNVs exclusivas de TEA (não observadas em indivíduos aparentemente saudáveis) que compreendem um único gene, sugere que ao contrário do que se pensava anteriormente, uma parte dos casos de TEA pode ser explicada por modelos monogênicos, seja por herança de um alelo causador ou por mutações *de novo*. Nesse cenário as variantes seriam de alto efeito e, por isso, mais raras na população. O que contribuiria para alta frequência dos TEA seria, neste caso, a existência de diferentes alterações individualmente raras em um mesmo gene ou em genes diferentes (EL-FISHAWAY; STATE, 2010).

Recentemente, crescente esforço tem sido investido em técnicas de sequenciamento de última geração (*NGS – Next Generation Sequencing*) que permitem analisar toda a região codificante (exoma) ou todo o genoma humano em busca de pequenas alterações do DNA, como a troca de um único nucleotídeo e InDels (pequenas inserções ou deleções), as quais, por sua vez,

possam ser causativas do TEA. A taxa de indivíduos afetados nos quais são identificadas, através de NGS, mutações possivelmente causativas têm sido estimada em cerca de 15% (IOSSIFOV *et al*., 2012; NEALE *et al*., 2012; O'ROAK *et al*., 2012; SANDERS *et al.,* 2012).

A seguir faremos uma revisão mais detalhada sobre os estudos de CNVs e sequenciamento de última geração.

1_2 Variações no número de cópias (CNVs) e TEA

A origem das CNVs é atribuída a mecanismos como NAHR (*Nonalelic Homologous Recombination*), NHEJ (*Nonhomologus End-Joining*), FoSTeS (*Fork Stalling and Template Switching*) e retrotransposição (para uma revisão sobre o assunto consultar Stankiewicz e Lupski, 2010). Essas alterações podem englobar parte ou todo um gene, ou podem englobar vários genes ou, ainda, elementos regulatórios. Assim, muitas dessas variações podem estar envolvidas na etiologia de doenças através do mecanismo de dosagem gênica ou da formação de produtos gênicos errôneos devido a ruptura de sequências codificantes (MERIKANGAS *et al*., 2009; STANKIEWICZ; LUPSKI, 2010). Entretanto, estima-se que 12% a 15% do genoma humano varie significantemente em número de cópias e muitas dessas variações, presentes em indivíduos aparentemente saudáveis, abrangem um ou mais genes (CARTER *et al*., 2007). Portanto, não é ainda conhecida a razão para algumas alterações apresentarem efeito fenotípicos e outras não, sendo a interpretação do significado clínico de uma CNV um grande desafio.

Lee *et al*. (2007) propôs que alterações *de novo* ou herdadas de um genitor afetado são mais prováveis de serem patogênicas, entretanto, não se pode afirmar que alterações herdadas de genitores saudáveis não estejam contribuindo para a manifestação de um fenótipo, uma vez que essas podem ter penetrância incompleta. De fato, como veremos adiante, uma parcela considerável das CNVs observadas em pacientes com TEA são herdadas de pais não afetados

(DEVLIN; SCHERER, 2012). Esses autores também ressaltaram a importância de observar se a CNV encontrada é similar a outras já catalogadas no banco de dados que compilam variações presentes em indivíduos saudáveis (DGV- *Database of Genomic Variants;* IAFRATE *et al*., 2004; ZHANG *et al*., 2006) ou em bancos de dados que copilam CNVs clinicamente relevantes (DECIPHER *Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources*) e também observar se a região deletada ou duplicada contem genes já associados a doenças (LEE *et al.*, 2007).

A frequência comumente reportada de CNVs detectadas em pacientes autistas é de 8 a 11,6%. Considerando-se apenas alterações *de novo* a frequência varia entre famílias simplex (apenas um indivíduo afetado - 5,6% a 10%) e famílias multiplex (mais de um afetado - 2% para 5,5%), sugerindo que nos casos familiais, alelos de susceptibilidade podem ser transmitidos por pais não afetados e atuam em conjunto para modular o risco para TEA (SEBAT *et al*., 2007; MARSHALL *et al*., 2008; CHRISTIAN *et al*., 2008; QIAO *et al.,* 2009; PINTO *et al*., 2010; ITSARA *et al*., 2010; GIRIRAJAN *et al*., 2010; BREMER *et al*., 2011).

As CNVs são em geral extremamente raras e apenas algumas alterações recorrentes foram descritas, tais como as duplicações 15q11-q13, as microdeleções/microduplicações 16p11.2, as deleções 22q11-13 e alterações nos genes *SHANK3, NLGN3, NLGN4 e NRXN1* (JAMAIN *et al*., 2003; LAUMONNIER *et al*., 2004; JACQUEMONT *et al.,* 2006; MOESSNER *et al*., 2007). Embora as CNVs sejam particulares de cada paciente, há evidências de que essas abrangem genes envolvidos em funções biológicas comuns, mais significativamente implicadas na etiologia do TEA, como por exemplo sinapse e orientação axonal. Estes dados mostram que danos em diferentes genes envolvidos na mesma via biológica ou envolvidos no mesmo processo fisiológico podem ter um mesmo efeito final, levando ao fenótipo autista (SBACCHI *et al*., 2010).

Cerca de 40% das CNVs detectadas em indivíduos autistas são herdadas de pais saudáveis, evidenciando a penetrância incompleta (DEVLIN; SCHERER, 2012). Em alguns casos, o autismo se comporta como um padrão de herança semelhante ao autossômico dominante, onde alterações em apenas um alelo estão relacionadas com a determinação do transtorno. Já em outros, o autismo está relacionado a um padrão de herança autossômico

recessivo, onde se observam mutações em ambos os alelos, como foi descrito em famílias com deleções em homozigose dos genes *PCDH10*, *DIA1* e *NHE92* (MORROW *et al*., 2008).

Todas essas informações, derivadas dos estudos de CNVs, mostram o quanto os TEA são geneticamente heterogêneos, o que dificulta o entendimento de sua etiologia.

1_3 Sequenciamento de última geração e TEA

Foram publicados recentemente quatro estudos que analisaram o exoma de trios (pais e probando) e/ou quartetos (pais, probando e irmão não afetado). A porcentagem de pacientes com alterações *de novo*, consideradas patogênicas por introduzir um códon de parada prematuro, mudar a matriz de leitura do mRNA ou um sítio de excisão ("*splice"*) do pré-mRNA*,* variou entre os estudos de 10,3% a 17,2%. Os resultados obtidos foram concordantes quanto a maior ocorrência de mutações *de novo* em crianças cujos pais possuem idade avançada (35,9 a 58 anos) em comparação a crianças com pais mais jovens (30,9 a 35,9 anos). Também foi visto que cerca de 60% a 75% das alterações *de novo* detectadas estão em alelos de origem paterna, sugerindo que essas podem ter ocorrido durante a formação do gameta masculino. Outra descoberta foi maior taxa de alterações *de novo* em mulheres afetadas em comparação aos homens afetados (IOSSIFOV *et al*., 2012; NEALE *et al*., 2012; O'ROAK *et al*., 2012; SANDERS *et al.,* 2012).

Os quatro estudos citados identificaram um total de 127 mutações possivelmente patogênicas porém, novamente nos deparamos com a grande heterogeneidade de *locus*, visto que apenas os genes *CHD8, KATNAL2, DYRK1A* e *POGZ* foram apontados como candidatos aos TEA por pelo menos dois desses estudos (IOSSIFOV *et al*., 2012; NEALE *et al*., 2012; O'ROAK *et al*., 2012; SANDERS *et al.,* 2012). Avaliando o conjunto de genes alterados (excluindo alterações sinônimas e mudanças de aminoácido), Ben-David e Shifman (2012) relataram que a maior parte desses genes codificam proteínas envolvidas na regulação da transcrição durante o

desenvolvimento do cérebro, com enriquecimento significativo de proteínas reguladoras de cromatina (p=0,021 (corrigido); *Enrichment Score* = 2,04).

Assim como elucidar o significado clínico de uma CNV, a interpretação do grande volume de dados gerados nas análises de sequenciamento de última geração também é um grande desafio. As perguntas sobre o quanto cada uma das variações identificadas contribuem para o fenótipo ou, ainda, quantas dessas podem ser causativas independente da existência ou não de outras alterações, permanecem sem respostas. Desse modo, mesmo após o surgimento do sequenciamento de última geração, a maioria dos casos de TEA continuam sem causa conhecida.

1_4 Justificativa

Os TEA causam prejuízos funcionais e impactos graves durante toda a vida, tanto para os indivíduos afetados quanto para os familiares. Estes transtornos possuem alta prevalência sendo atualmente considerados um problema de saúde pública. O diagnóstico nos primeiros anos de vida possibilita que abordagens terapêuticas comportamental, educacional e farmacológica sejam aplicadas precocemente. Isto traz benefícios significativos na aquisição de habilidades funcionais, na redução de comportamentos mal-adaptativos e, consequentemente, aumento no sucesso de integração da criança à sociedade. A investigação dos mecanismos etiológicos de origem genética associados ao TEA pode trazer novas perspectivas para a compreensão desses transtornos bem como para o desenvolvimento de ferramentas de diagnóstico precoce. Visto que 70% dos casos são idiopáticos, hipotetizamos que uma parte desses pode ser causada por pequenas CNVs (>10 kbp), geralmente não detectáveis pelas plataformas comerciais de microarray disponíveis atualmente. Na tentativa de identificar novos genes candidatos e testar essa hipótese, customizamos uma lâmina de aCGH direcionada a detecção de pequenas CNVs em genes potencialmente envolvidos na fisiopatologia dos TEA. Essa lâmina também poderia ser utilizada como uma ferramenta para auxiliar o diagnóstico clínico e o aconselhamento genético, com um

Os nossos resultados, além de corroborar com o envolvimento dos genes *MDGA2*, *FHIT*, *HTR2A*, *SHANK2*, *GRIA3*, *ZNF778, PRKCα*, *CDH15, DIAPH3, GCH1, GRM5, MARK1*, *SLC17A6*, *IMMP2L, BZRAP1*, *SYNGAP1*, *ANK3*, *MAP1A*, *GABRR2* e *LAMC3* nos TEA também sugere novos genes candidatos: *LRRC7, LRRIQ3, CADPS1, NUFIP, SEMA3A, SNAP29, MBD2*, *GAD2, DGKH* e *PARD3*.

Por fim, nossos dados são consistentes com os outros estudos que reportam os TEA como distúrbios complexos, com grande heterogeneidade alélica e heterogeneidade de *locus.* Quanto a perspectivas futuras, para uma melhor compreensão dos mecanismos pelos quais as CNVs identificadas podem levar ao fenótipo pretendemos realizar estudos *in vitro* e em *in vivo.*

6.2 Conclusões e Considerações finais

- A triagem de 269 genes candidatos revelou que a taxa de detecção de alterações em pacientes não sindrômicos e sindrômicos foi, respectivamente, 11,6% e 38,9%. A taxa de alterações *de novo* foi maior em sindrômicos (25%) em comparação a taxa observada em não sindrômicos (7,89%);
- A lâmina de aCGH customizada parece ser eficaz na identificação de deleções/duplicações em autistas e poderia ser utilizada como uma ferramenta para auxiliar o diagnóstico;
- Pequenas CNVs (<10 Kbp), não detectáveis pelas plataformas comerciais de *microarray* disponíveis atualmente, em *exons* e regiões UTR, parecem contribuir para o fenótipo de TEA;
- Os genes *LRRC7, LRRIQ3, CADPS1, NUFIP, SEMA3A, SNAP29, MBD2, GAD2, DGKH e PARD3* podem ser novos candidatos aos TEA;
- A gravidade do quadro clínico pode ser modulada pelo número de mutações em genes envolvidos com neurogênese.

ADZHUBEI, I.; JORDAN, D. M.; SUNYAEV, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. **Curr Protoc Hum Genet**. Chapter 7:Unit7.20, 2013.

AMERICAN PSYCHIATRIC ASSOCIATION. **Diagnostic and statistical manual of mental disorders: DSM-IV**. 4th ed. Washington (DC): American Psychiatric Association. 943 p., 2000.

BAILEY, A. et al. Autism as a strongly genetic disorder: evidence from a British twin study. **Psychol Med**., 25(1):63-77, 1995.

BEN-DAVID, E.; SHIFMAN. S. Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism. Mol Psychiatry. 2012.

BERUMENT, S. et al. Autism Screening Questionnaire: Diagnostic Validity. **British Journal of Psychiatry**, v. 175, p.444-451, 1999.

BLANCHETTE, M. et al. Aligning multiple genomic sequences with the threaded blockset aligner. **Genome Res**. 14(4):708-15, 2004.

BLATT, G. J. et al. Density and distribution of hippocampal neurotransmitter receptors in autism: an autoradiographic study. **J. Autism Dev. Disord**., 31:537–543, 2001.

BLATT, G. J. et al. GABAergic cerebellar system in autism: a neuropathological and developmental perspective. **Int. Rev. Neurobiol**., 71:167–178, 2005.

BREMER, A. et al. Copy number variation characteristics in subpopulations of patients with autism spectrum disorders. **Am J Med Genet B Neuropsychiatr Genet**.;156(2):115-24, 2011

BREMER, A. et al. Screening for copy number alterations in loci associated with autism spectrum disorders by two-color multiplex ligation-dependent probe amplification. **Am J Med Genet B Neuropsychiatr Genet**. 5;153B(1):280-5, 2010.

BUCAN, M. et al. Genome-Wide Analyses of Exonic Copy Number Variants in a Family-Based Study Point to Novel Autism Susceptibility Genes. **PloS Genetics**, 5: e1000536, 2009.

CARTER, N. P. Methods and strategies for analyzing copy number variation using DNA microarrays. **Nat Genet**. 39(7 Suppl):S16-21, 2007.

CHENG, Y.; MIURA, R. M.; TIAN, B. Prediction of mRNA polyadenylation sites by support vector machine. **Bioinformatics**. 1;22(19):2320-5, 2006.

CHIAROMONTE, F.; YAP, V. B.; MILLER, W. Scoring pairwise genomic sequence alignments. **Pac Symp Biocomput**. 115-26, 2002.

CHRISTIAN, S. L. et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. **Biol Psychiatry**, 63:1111–1117, 2008.

CUSCO, I. et al. Autism-specific copy number variants further implicate the phosphatidylinositol signaling pathway and the glutamatergic synapse in the etiology of the disorder. **Hum. Mol. Genet**. 18, 1795-1804, 2009.

DEVLIN, B.; SCHERER, S. W. Genetic architecture in autism spectrum disorder. **Curr Opin Genet Dev**. 22(3):229-37, 2012.

EL-FISHAWY, P.; STATE, M. W. The genetics of autism: key issues, recent findings, and clinical implications. **Psychiatr Clin North Am**. 33(1):83-105, 2010.

FATEMI, S. H.; FOLSOM, T. D. Dysregulation of fragile × mental retardation protein and metabotropic glutamate receptor 5 in superior frontal cortex of individuals with autism: a postmortem brain study. **Mol Autism**. 6;2:6, 2011.

FOLSTEIN, S. E.; ROSEN-SHEIDLEY, B. Genetics of autism: complex a etiology for a heterogeneous disorder. **Nat. Rev. Genet**., 2:943–955, 2001.

FRIEDMAN, R. C. et al. Most Mammalian mRNAs Are Conserved Targets of MicroRNAs. **Genome Research**. 19:92-105, 2009.

GIRIRAJAN, S., EICHLER, E. E. Phenotypic variability and genetic susceptibility to genomic disorders. **Hum Mol Genet**. 15;19(R2):R176-87, 2010.

GLESSNER, J. T. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. **Nature**. 459, 569-573, 2009.

GOIZET, C. et al. Case with autistic syndrome and chromosome 22q13.3 deletion detected by FISH. **Am. J. Med. Genet**., 96: 839–844, 2000.

GRIMSON, A. et al. MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. **Molecular Cell**. 6;27(1):91-105, 2007.

HADDAD, L. A. et al. A PCR-based test suitable for screening for fragile X syndrome among mentally retarded males. **Hum Genet**. 97(6):808-12, 1996.

HUSSMAN, J. P. et al. A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism. **Mol Autism**. 19;2(1):1, 2011.

IAFRATE, A. J. et al. Detection of large-scale variation in the human genome. **Nat Genet**. 36(9):949-51, 2004.

IOSSIFOV, I. et al. De novo gene disruptions in children on the autistic spectrum. **Neuron**. 26;74(2):285-99, 2012.

ITSARA, A. De novo rates and selection of large copy number variation. **Genome Res**. 20(11):1469-81, 2010.

JACQUEMONT, M. L. et al. Array-based comparative genomic hybridization identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. **J. Med. Genet**., 43:843–849, 2006.

JAMAIN, S. et al. Linkage and association of the glutamate receptor 6 gene with autism. **Mol. Psychiatry** 7:302–310, 2002.

KENT, W. J. et al. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. **Proc Natl Acad Sci U S A**.; 100(20):11484-9, 2003.

KEN, W. J. et al. The human genome browser at UCSC. **Genome Res**. 12(6):996-1006, 2002.

KUMAR, R. A. et al. SHANK2 redemption: another synaptic protein for mental retardation and autism. **Clin Genet**. 78(6):519-21, 2010.

LAUMONNIER, F. et al. X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. **Am J Hum Genet.** 74(3):552-7, 2004.

LEE, C. et al. Copy number variation and clinical cytogenetic diagnosis of constitutional disordes. **Nat. Genet**., 39:S 48-54, 2007.

LEPAGNOL-BESTEL, A. M. SLC25A12 expression is associated with neurite outgrowth and is upregulated in the prefrontal cortex of autistic subjects. **Mol Psychiatry**. 13(4):385-97, 2008.

LEWIS, B. P.; BURGE, C. B.; BARTEL, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. **Cell**. 14;120(1):15-20, 2005.

MA, D. Q. et al. Identification of significant association and gene-gene interaction of GABA receptor subunit genes in autism. **Am. J. Hum. Genet**., 77:377–388, 2005.

MARSHALL, C. R. et al. Structural variation of chromosomes in autism spectrum disorder. **Am. J. Hum. Genet**. 82:477–488, 2008.

MERIKANGAS, A. K.; CORVIN, A. P.; GALLAGHER, L. Copy-number variants in neurodevelopmental disorders: promises and challenges. **Trends Genet**.

Dec;25(12):536-44, 2009.

MEYER, L. R. et al. The UCSC Genome Browser database: extensions and updates. **Nucleic Acids Res**. 2012 Nov 15, 2013.

MOESSNER, R. et al. Contribution of SHANK3 mutations to autism spectrum disorder. **Am. J. Hum. Genet**. 81(6):1289-97, 2007.

MORROW, E. M. et al. Identifying autism loci and genes by tracing recent shared ancestry. **Science**. 11;321(5886):218-23, 2008.

MURPHY, W. J. et al. Resolution of the early placental mammal radiation using Bayesian phylogenetics. **Science**. 2001 Dec 14;294(5550):2348-51.

NEALE, B. M. et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. **Nature**. 4;485(7397):242-5, 2012.

NG, P. C.; HENIKOFF, S. SIFT: Predicting amino acid changes that affect protein function. **Nucleic Acids Res**. 1;31(13):3812-4, 2003.

O'ROAK, B. J. et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. **Science**. 21;338(6114):1619-22, 2012.

O'ROAK, B. J. et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. **Nature**. 4;485(7397):246-50, 2012.

PAULA, C. S. Autism in Brazil: perspectives from science and society. **Rev Assoc Med Bras**. 57(1):2-5, 2011.

PENA, S. D. J. **Enfoque genético da criança com deficiência mental e sua família**. Deficiencia Mental: Abrodagem Multidisciplinar. Fascículo 4. 2005.

PINTO, D. et al. Functional impact of global rare copy number variation in autism spectrum disorders. **Nature**. 15;466(7304):368-72, 2010.

POLLARD, K. S.; HUBISZ, M. J.; SIEPEL, A. Detection of non-neutral substitution rates on mammalian phylogenies. **Genome Res**. 2009.

PONCHEL, F. et al. Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. **BMC Biotechnology** 3:18, 2003.

QIAO, Y. et al. Phenomic determinants of genomic variation in autism spectrum disorders. **J Med Genet**. 46(10):680-8, 2009.

QIAO, Y. et al. Clinical application of 2.7M Cytogenetics array for CNV detection in subjects with idiopathic autism and/or intellectual disability. **Clin Genet**. 2013

PRUITT, K. D. et al. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. **Nucleic Acids Res**. 40(Database issue):D130-

5, 2012.

SANDERS, S. J. et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. **Nature** 4;485(7397):237-41, 2012.

SBACCHI, S. et al. Functional Annotation of Genes Overlapping Copy Number Variants in Autistic Patients: Focus on Axon Pathfinding. **Current Genomics**, 11, 136-145, 2010.

SCHAAF, C. P. et al. Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorder. **Human Molecular Genetics**, v.20, n.17, p.3366-3375, 2011.

SCHOPLER, E. et al. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). **Journal of Autism and Developmental Disorders**, v. 10, n. 1, p. 91-103, 1980.

SCHMITZ, C. et al. Autism: neuropathology, alterations of the GABAergic system, and animal models. **Int Rev Neurobiol.**, 71:1-26. 2005.

SCHOPLER, E. et al. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). **Journal of Autism and Developmental Disorders**, v. 10, n. 1, p. 91-103, 1980.

SCHWARTZ, S. et al. Human-mouse alignments with BLASTZ. **Genome Res.** 13(1):103-7, 2003.

SEBAT, J. et al. Strong association of de novo copy number mutations with autism. **Science**, 316:445–449, 2007.

SIEPEL, A. et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. **Genome Res**. 15(8):1034-50, 2005.

SORTE, H. S. et al. Copy number variation findings among 50 children and adolescents with autism spectrum disorder. **Psychiatr Genet**. 23(2):61-9, 2013.

STANKIEWICZ, P.; LUPSKI. J. R. Structural variation in the human genome and its role in disease. **Annu. Rev. Med**., 61:437-55, 2010.

TIDMARSH, M. D.; VOLKMAR. F. R. Diagnosis and epidemiology of autism spectrum disorders. **Can J Psychiatry**, 48:517–525, 2003.

VEENSTRA-VANDERWEELE J.; COOK, E. H. Jr. Molecular genetics of autism spectrum disorder. **Molecular Psychiatry**. 9:819-832, 2004.

VORSTMAN, J. A. et al. A double hit implicates DIAPH3 as an autism risk gene. **Mol Psychiatry**. 16(4):442-51, 2011.

WANG, K. et al. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. **Genome Res**. 17(11):1665-74, 2007.

WANG, K.; BUCAN, M. Copy Number Variation Detection via High-Density SNP Genotyping. **Cold Spring Harb**. Protoc.; DOI:10.1101/pdb.top46, 2008.

WEISS, L. A. et al. Association between microdeletion and microduplication at 16p11.2 and autism. **N. Engl. J. Med**. 358, 667–675, 2008.

WEKSBERG, R. et al. A method for accurate detection of genomic microdeletions using real-time quantitative PCR. **BMC Genomics**. 13;6:180, 2005.

ZHANG, H.; HU, J.; RECCE, M.; TIAN B. PolyA_DB: a database for mammalian mRNA polyadenylation. **Nucleic Acids Res**. 1;33(Database issue):D116-120, 2005.

ZHANG, J. et al. Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. **Cytogenet Genome**. 115(3-4):205-14, 2006.

ZHI, J.; HATCHWELL, E. Human MLPA Probe Design (H-MAPD): a probe design tool for both electrophoresis-based and bead-coupled human Multiplex ligation-dependent probe amplification assays**. BMC Genomics**, v.9, n. 407, 2008.