Mariana Soares Fogo

Estudo de expressão gênica em células neurais
derivadas de células-tronco pluripotentes induzidas
de indivíduos com Transtorno do Espectro Autista

Gene expression study in neural cells derived from induced
pluripotent stem cells of individuals with Autism Spectrum
Disorder

São Paulo

2018

Mariana Soares Fogo

Estudo de expressão gênica em células neurais
derivadas de células-tronco pluripotentes induzidas
de indivíduos com Transtorno do Espectro Autista

Gene expression study in neural cells derived from induced
pluripotent stem cells of individuals with Autism Spectrum
Disorder

Dissertação apresentada ao
Instituto de Biociências da
Universidade de São Paulo, para
a obtenção de Título de Mestre
em Biologia, na Área de
Genética.

Orientadora: Dra. Andréa Laurato
Sertié

Coorientadora: Dra. Karina Griesi
Oliveira

São Paulo

2018

Fogo, Mariana Soares

Estudo de expressão gênica em células neurais derivadas de células-tronco pluripotentes induzidas de indivíduos com Transtorno do Espectro Autista

Número de páginas: 87

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

1. Transtorno do Espectro Autista
2. Transcriptoma 3. RNAseq 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).

Orientador(a)

Y que la técnica es un arma, y que quien sienta que el mundo no es perfecto como debiera ser, debe luchar porque el arma de la técnica sea puesta al servicio de la sociedad, y por eso rescatar antes a la sociedad para que toda la técnica sirva a la mayor cantidad posible de seres humanos, y para que podamos construir la sociedad del futuro (...).

Ernesto Che Guevara
(discurso feito no Encontro Internacional de Estudantes de Arquitetura, 1963)

AGRADECIMENTOS

Concluir esse trabalho não foi tarefa simples. Nele coloquei muito esforço e dedicação e com ele precisei aprender a aceitar minhas próprias limitações, lidar com as frustrações cotidianas e entender as limitações e contradições que o fazer científico carrega. Fazer essa caminhada sozinha não teria sido fácil, que bom que pude contar com pessoas que, de perto ou de longe e cada uma a sua maneira, me alegraram nos momentos de desânimo, festejaram comigo os momentos de alegria e me ajudaram direta ou indiretamente a avançar no trabalho. Meus mais sinceros agradecimentos vão a cada uma dessas pessoas.

A meus pais, Neide e Haroldo, agradeço por acreditarem em mim mesmo quando eu mesma tinha dúvidas e pelas palavras (nem sempre) doces, mas que sempre me fizeram seguir em frente. Ao meu irmão Daniel, agradeço pela parceria, pelas conversas e por compartilhar comigo sonhos e luta.

À Tati e à Ju, companheiras de lab que se tornaram grandes amigas, agradeço pelas eternas conversas sobre Ciência, projetos e experimentos, pelo companheirismo de sempre e por aguentarem minhas intermináveis reclamações e lamentações. Muito do que eu sei hoje, aprendi e continuo aprendendo com vocês desde a época em que a gente era o time da cultura no lab da Irene. Também gostaria de agradecer a todos do lab onde fiz minha IC, em especial à Irene, por me ensinar a dar os primeiros passos como pesquisadora, por toda a paciência e pelo acolhimento (inclusive hoje em dia quando eu resolvo aparecer no lab para conversar ou comer alguma coisa).

Agradeço também aos amigos do lab 200, com quem dividi pequenas angústias e alegrias fazendo experimentos nos horários mais bizarros, por toda ajuda dentro e fora do laboratório. Aos amigos do IIEP, agradeço pela companhia. Gostaria de agradecer especialmente à Sandra e à Juliana pelas discussões, ensinamentos e pela amizade. Fez toda a diferença ter vocês comigo todos os dias... A Ju sempre tão otimista me dizendo que tudo ia dar certo e sempre tendo toda a paciência do mundo para me explicar as coisas e ouvir minhas ideias, e a Sandrita sempre pronta pra me ajudar em tudo que eu precisasse, me ouvir reclamando (e reclamar junto) e fazer companhia no lab e nos rolês da vida.

Agradeço aos amigos da bio pela companhia, pela força e pelas eternas discussões sobre Ciência que me fizeram encontrar respostas para uma série de questionamentos que me acompanharam ao longo do mestrado. Agradeço a meus amigos do colégio, que caminham comigo há pelo menos 15 anos e nunca me deixaram desanimar. Agradeço também aos amigos do Levante, que me ensinaram que eu nunca estou sozinha e que a vida é mais fácil quando a gente sonha junto.

À Rita, à Karina e à Andrea, agradeço pela oportunidade de fazer parte do grupo de pesquisa e estudar essa condição tão intrigante que é o TEA. À Karina, que me acompanhou mais de perto, agradeço pela orientação, pelas discussões e por me fazer enxergar coisas que sozinha eu não conseguia. À Andrea, agradeço pela paciência e toda a ajuda com os experimentos do estudo de caso. Gostaria de agradecer também a todos os colaboradores do Genoma, do IIEP e da Química que contribuíram com meu trabalho, em especial à Suzana e ao Julio, que tiveram a paciência de me ensinar um pouco sobre bioinformática e programação.

Esse trabalho carrega nele um pouquinho de cada um de vocês.

Por fim, agradeço à FAPESP, Capes e CNPq pelo auxílio financeiro, fundamental para o desenvolvimento e conclusão desse trabalho no preocupante cenário de cortes orçamentários que vivenciamos atualmente.

ÍNDICE

RESUMO	1
ABSTRACT	2
INTRODUÇÃO.....	3
OBJETIVOS	15
CONCLUSÕES.....	16
ESTUDO DE CASO	17
BIBLIOGRAFIA.....	20

RESUMO

O Transtorno do Espectro Autista (TEA) é um transtorno neuropsiquiátrico grave caracterizado por comprometimento da capacidade de interação social e comunicação e pela presença de interesses, atividades e comportamentos repetitivos e restritos. O TEA apresenta alta heterogeneidade genética e pode ser enquadrado em diferentes modelos de herança, o que torna difícil apontar os fatores genéticos associados ao transtorno e compreender de que forma eles interagem e ocasionam, em última instância, sua manifestação clínica. Ainda, fenômenos como penetrância incompleta e discordância entre gêmeos monozigóticos são frequentemente observados em famílias com indivíduos afetados, tornando ainda mais complexo o entendimento do envolvimento de cada um dos fatores genéticos na etiologia do TEA. Uma forma de contornar esses problemas é a utilização de abordagens transcriptômicas no estudo desse transtorno. Nesse contexto, a busca por variantes patogênicas deixa de ser o principal foco do estudo e a atenção volta-se para a investigação de vias de sinalização e processos biológicos que estejam desregulados durante o processo de diferenciação neuronal e que sejam compartilhados pelos indivíduos afetados. Usando como principal ferramenta o sequenciamento do transcriptoma de células progenitoras neurais (NPC) e neurônios derivados de células-tronco pluripotentes induzidas (iPSC), identificamos uma série de alterações transcricionais importantes entre pacientes e controles, sugerindo que, a despeito da grande heterogeneidade genética associada ao transtorno, essas alterações se refletem em processos biológicos comuns. Nossos resultados são consistentes com os obtidos em trabalhos anteriores realizados com outros modelos de estudo e as linhagens celulares geradas a partir de nosso protocolo parecem refletir de forma fidedigna o perfil de expressão do cérebro fetal. Ainda, em um caso particular incluído neste estudo, investigamos as consequências transcricionais e funcionais de uma duplicação na região 17p13.3 - alteração previamente implicada em outros casos de autismo - nas células neuronais deste paciente, contribuindo para a compreensão do papel desta duplicação na patofisiologia do TEA. Dessa forma, nosso trabalho reforça a validade do uso de células neurais derivadas de iPSC e de abordagens transcricionais para o estudo de transtornos do neurodesenvolvimento.

ABSTRACT

Autism Spectrum Disorder (ASD) is a severe neuropsychiatry disorder characterized by impaired social interaction and communication, restricted interests and repetitive behaviors. ASD is highly genetically heterogeneous and can be caused by different inheritance models, which hampers the identification of the genetic factors associated to the disorder. Incomplete penetrance and discordance between monozygotic twins are often observed in families of affected individuals, making it even more complex to understand the involvement of genetic factors in ASD. Applying transcriptomic approaches to study this disorder is an interesting way to overcome these problems. In this context, searching for pathogenic variants is no longer the study's main focus. Instead, one aims to investigate signaling pathways and biological process that are deregulated during neuronal differentiation and shared by affected individuals. Based on the transcriptome sequencing of neural progenitor cells (NPC) and neurons derived from induced pluripotent stem cells (iPSC), we were able to identify a series of important transcriptional alterations between patients and controls, which suggest that despite the great genetic heterogeneity associated to ASD, those alterations are reflected in common biological processes. Our results are consistent with those obtained in previous studies performed with other models and the lineages generated by our protocol seem to reliably reflect fetal brain expression profile. Moreover, in a particular case included in this study, we investigated transcriptional and functional consequences of a duplication located at 17p13.3 - an alteration previously found in other autism cases – in the neuronal cells of this patient, contributing towards a better comprehension of the role of this duplication in ASD pathophysiology. Hence, our work reinforces the validity of the use of neural cell derived from iPSC and transcriptional approaches in the studies of neurodevelopmental disorders.

INTRODUÇÃO

1.1 Aspectos clínicos e epidemiológicos do Transtorno do Espectro Autista

O Transtorno do Espectro Autista (TEA) é um transtorno neuropsiquiátrico grave caracterizado por comprometimento da capacidade de interação social e comunicação e pela presença de interesses, atividades e comportamentos repetitivos e restritos (American Psychiatric Association, 2013). Essas manifestações clínicas, quando precocemente apresentadas, são essenciais para que o diagnóstico do transtorno seja realizado (Jones, Gliga, Bedford, Charman, & Johnson, 2014). O grau de comprometimento que os indivíduos apresentam para cada um desses sintomas é bastante variado, indo desde formas mais leves, em que os pacientes têm uma vida independente, até aquelas que são bastante graves e estão, em muitos casos, associadas a quadros de deficiência intelectual (Griesi-Oliveira & Sertié, 2017). Além disso, outras manifestações clínicas não diagnósticas podem ainda ser observadas, tendo impacto negativo no cotidiano dos pacientes e de suas famílias, tais como comportamento agressivo, distúrbios do sono e da alimentação, problemas gastrointestinais, hiperatividade e ansiedade (Gorrindo et al., 2012; Krakowiak, Goodlin-Jones, Hertz-Picciotto, Croen, & Hansen, 2008; American Psychiatric Association, 1994).

Ainda que já tenham sido descritas estruturas do sistema nervoso central anatomicamente alteradas em indivíduos afetados pelo TEA, como o cerebelo, a amígdala e o córtex frontal (Donovan & Basson, 2017), essas alterações não são observadas em todos os pacientes. Também não foram ainda identificados quaisquer biomarcadores moleculares que possam ser usados no diagnóstico do TEA. Em função disso, existe uma grande dificuldade para se desenvolver ferramentas diagnósticas para este transtorno, e assim o diagnóstico pode ser definido apenas por meio de avaliações comportamentais, testes cognitivos e histórico clínico do paciente. Da mesma forma, não existem medicamentos usados para tratar os sintomas principais do transtorno; as terapias medicamentosas disponíveis tratam apenas as manifestações clínicas não diagnósticas citadas.

Estudos recentes apontam que o TEA tem prevalência de aproximadamente 1/68 indivíduos na população e é mais comum em meninos do que em meninas, apresentando um desvio sexual de 4:1 (Christensen et al., 2016). Em comparação com análises anteriores, esses dados indicam um aumento na prevalência do transtorno ao longo dos anos, o que pode ser reflexo de um aumento real no número de indivíduos afetados na população devido, por exemplo, à maior exposição a fatores ambientais de risco. Entretanto, parte desse aumento provavelmente se deve a mudanças em critérios

diagnósticos, melhorias em metodologias de análises epidemiológicas e aumento do conhecimento sobre o TEA tanto pelo público leigo quanto por profissionais da saúde (Baxter et al., 2015; Fombonne, 2009; Wing & Potter, 2002).

1.2 Aspectos genéticos do Transtorno do Espectro Autista

Sabe-se que tanto fatores ambientais, tais como infecções e exposição a agentes tóxicos (Bilbo, Block, Bolton, Hanamsagar, & Tran, 2018), quanto fatores genéticos estão envolvidos na manifestação do TEA. No entanto, diversas evidências nos permitem afirmar com segurança a grande relevância dos fatores genéticos na etiologia do transtorno, tais como: a maior concordância observada em gêmeos monozigóticos em comparação à observada em gêmeos dizigóticos, levando a estimativas de herdabilidade que variam entre 38 e 90%; o alto risco de recorrência do transtorno dentro de uma mesma família; a maior propensão que irmãos e pais de indivíduos afetados têm de apresentar características cognitivas e comportamentais qualitativamente similares às aquelas observadas em pacientes; a associação com doenças monogênicas e o aumento do risco de manifestação do transtorno associado a mutações e alterações cromossômicas em diversos *loci* (Sandin, Lichtenstein, Larsson, Hultman, & Reichenberg, 2014; Ronald & Hoekstra, 2011; Hallmayer et al., 2011; Abrahams & Geschwind, 2008; Trottier, Srivastava, & Walker, 1999).

Por meio do uso tanto de abordagens mais tradicionais, como a cariotipagem e o estudo de genes candidatos, quanto de abordagens mais modernas, como o *array* e o sequenciamento de nova geração, uma série de genes e regiões cromossômicas de risco para o TEA já foram identificadas. No banco de dados SFARI (Simons Foundation Autism Research Initiative – <https://gene.sfari.org/>), que agrupa todas as informações sobre *loci* de risco para o TEA já publicadas, é possível notar que, até o momento, quase 500 genes apresentam evidências de associação com o TEA. É importante ressaltar que nenhum desses genes sozinho consegue explicar mais do que 1% do total de casos de TEA e a causa genética ainda é desconhecida em aproximadamente 75% dos casos (Bourgeron, 2015), evidenciando uma alta heterogeneidade genética associada ao transtorno.

A identificação de genes e regiões cromossômicas por meio dessas abordagens e sua associação com a manifestação clínica do transtorno só é possível se as alterações ou variantes (rearranjos cromossômicos, CNVs (*copy number variants*), SNVs (*single nucleotide variants*) e pequenas deleções ou duplicações) contidas nesses *loci* forem raras de efeito médio ou alto, ou seja, se o efeito provocado por elas for forte o suficiente para que uma variante sozinha (ou poucas variantes em conjunto) consiga desestabilizar o desenvolvimento neuronal e levar à manifestação do fenótipo. Entretanto, considerando que o TEA é um transtorno comum na população, que estudos familiares indicam que o

risco de recorrência é diretamente proporcional à porcentagem do genoma compartilhado com um indivíduo afetado (Risch et al., 2014; Sandin et al., 2014) e que os indivíduos afetados apresentam adaptabilidade reprodutiva reduzida, evidenciada pelo comprometimento da capacidade de interação social inerente ao transtorno, é possível especular que variantes comuns de baixo efeito também tenham papel importante em sua etiologia. A ideia subjacente a essa hipótese é a de que, como variantes de baixo efeito sofrem pressão seletiva menor, elas acabam se tornando comuns na população e se acumulando nos indivíduos. O fenótipo se manifestaria então quando um indivíduo acumulasse uma carga suficientemente grande dessas variantes, capaz de provocar efeitos danosos no desenvolvimento neuronal. De fato, variantes raras de efeito médio e alto em conjunto são responsáveis por causar o transtorno em 10-26% dos casos, enquanto estima-se que SNVs comuns estejam envolvidas na manifestação do transtorno em 17-60% dos casos (Bourgeron, 2015; Gaugler et al., 2014; Klei et al., 2012), evidenciando a grande importância dessas variantes e reforçando a alta heterogeneidade genética associada ao TEA. Contudo, nenhuma dessas variantes de baixo efeito foi identificada até o momento, uma vez que cada uma está associada a um risco muito pequeno e nenhum estudo de associação global já realizado teve poder estatístico suficiente para apontar alguma que fosse indiscutivelmente associada ao TEA (Huguet, Benabou, & Bourgeron, 2016; Bourgeron, 2015).

Assim, atualmente acredita-se que tanto variantes raras quanto variantes comuns expliquem a arquitetura genética do TEA e que ele é, portanto, um transtorno complexo que pode ser enquadrado em diferentes modelos de herança (figura 1.1). Nesse contexto, parte dos casos deve ser causada por uma única variante *de novo* de efeito alto (representando um modelo de herança monogênico) ou por algumas variantes raras de efeito médio (modelo de herança oligogênico). Em um segundo conjunto de casos, os indivíduos podem estar carregando uma carga média de variantes comuns de risco e manifestar o fenótipo apenas com o acúmulo de poucas variantes de efeito médio. Ainda, é possível que alguns dos indivíduos apresentem uma carga muito alta de variantes comuns de risco e apenas isso seja suficiente para que eles sejam afetados. Essas duas últimas situações se aproximariam mais de um modelo de herança poligênico. Dessa forma, cada indivíduo deve carregar um conjunto próprio de variantes de risco para o TEA e o que determina, em última instância, se ele será afetado ou não é a quantidade de variantes raras e comuns presentes em seu genoma e como se dá a interação entre elas.

Figura 1.1: Interação entre variantes raras e comuns na manifestação do TEA. O TEA é um transtorno complexo que pode se encaixar em diferentes modelos de herança. Em alguns casos, uma única variante de novo de alto efeito ou algumas variantes raras de efeito médio são suficientes para sua manifestação, representando modelos de herança mono e oligogênicos, respectivamente (esquerda e centro). Ainda, de acordo com um modelo poligênico, uma carga média de variantes comuns de risco somada a poucas variantes de efeito médio ou uma carga muito alta de variantes comuns de risco podem levar à manifestação do transtorno (direita). Retirado de Griesi-Oliveira & Sertié, 2017.

1.3 Desafios e possibilidades no estudo do Transtorno do Espectro Autista

Um grande obstáculo no estudo de doenças que acometem o sistema nervoso central sempre foi a obtenção de tecido para condução de experimentos celulares e moleculares. No caso do TEA, no qual o cérebro em desenvolvimento é o órgão-chave para a compreensão da doença, a impossibilidade de retirada de tecido por meio de biopsias e a dificuldade de se obter amostras *post-mortem* foram por algum tempo um grande entrave no avanço de estudos nessa área. Entretanto, o desenvolvimento das células-tronco pluripotentes induzidas (iPSC, do inglês *induced pluripotent stem cells*), resultado de um processo em que células somáticas são reprogramadas a um estado pluripotente a partir da superexpressão de genes específicos (Takahashi et al., 2007; Yu et

al., 2007), tornou possível a superação desse obstáculo. Após a reprogramação, as iPSC podem ser rediferenciadas em tipos celulares provenientes de qualquer um dos três folhetos embrionários, incluindo neurônios. A grande vantagem dessa técnica no estudo do TEA é a obtenção de modelos neuronais já predispostos à doença, uma vez que os neurônios obtidos apresentam o genoma do indivíduo doador. Dessa forma, é possível usar as iPSC como modelo na identificação de alterações na expressão gênica, morfologia e função neuronal de pacientes e no estudo da neurogênese, particularmente importante em doenças de neurodesenvolvimento como o TEA (Lin et al., 2012).

Trabalhos realizados tendo células progenitoras neurais (NPC, do inglês *neural progenitor cells*) e neurônios derivados de iPSC como modelo mostram que essa é de fato uma abordagem promissora para o estudo de transtornos que acometem o sistema nervoso. Griesi-Oliveira e colaboradores (2014) identificaram uma translocação equilibrada que causou o rompimento do gene TRPC6 (que codifica para um canal de cátion independente de voltagem e permeável a Ca^{2+}) em um paciente com TEA e reprogramaram células de polpa de dente a um estado pluripotente a fim de entender as consequências funcionais do rompimento desse gene durante o processo de diferenciação neuronal. Usando essa estratégia, o grupo foi capaz de mostrar que o influxo de Ca^{2+} estava alterado nas NPC do paciente e que seus neurônios apresentavam neuritos mais curtos, menor arborização, densidade reduzida de espinhos dendríticos, menor número de sinapses e correntes de Na^+ comprometidas em comparação às células derivadas dos controles. Nesse trabalho, também foi verificado o efeito de duas drogas diferentes sobre o fenótipo dos neurônios desse paciente, a hiperforina (a qual age ativando o canal TRPC6) e o IGF, e foram obtidos resultados positivos de resgate do fenótipo nas duas situações. Mais recentemente, Marchetto e colaboradores (2016), usando NPC e neurônios diferenciados a partir de iPSC derivadas de indivíduos afetados pelo TEA apresentando macrocefalia, mostraram que as células progenitoras dos pacientes apresentavam proliferação acelerada, enquanto alterações na neurogênese e redução na sinaptogênese foram observadas nos neurônios desses indivíduos. O grupo verificou ainda que as alterações observadas nos neurônios levaram a problemas na formação de redes neuronais, fenótipo que foi resgatado por meio do uso de IGF-1. Dessa forma, esses estudos mostram que a obtenção de células neurais a partir de reprogramação celular é de fato uma abordagem viável tanto para o estudo dos mecanismos celulares e moleculares subjacentes à doença quanto para a triagem de drogas que apontem para potenciais terapias.

Um segundo fator que representa um grande desafio no estudo do TEA é a própria heterogeneidade genética associada ao transtorno. Como já explicitado, o desenvolvimento de técnicas mais sensíveis e que possibilitam uma análise ampla do

genoma (ou do exoma) proporcionou um grande aumento no número de genes de risco identificados, porém, dada a complexidade da arquitetura genética do TEA, ainda há grande dificuldade de se identificar a causa genética em uma proporção importante dos casos. Nesse contexto, estudos de expressão gênica global surgem como uma estratégia interessante para contornar esse problema, conforme descrito a seguir.

1.4 Estudos de expressão gênica global

Os estudos de expressão gênica global surgem como uma alternativa para contornar o problema que a alta heterogeneidade genética e a dificuldade de identificar variantes de risco de baixo efeito trazem para o estudo do TEA, uma vez que o foco da análise migra do nível genômico, no qual os pacientes são altamente heterogêneos, para o nível transcriptômico, no qual é possível que se encontre uma maior homogeneidade entre os pacientes em termos de vias de sinalização e processos biológicos desregulados em função de alterações na expressão gênica (Han et al., 2014). Nesse contexto, os estudos de expressão gênica global representam também uma estratégia relevante para a identificação de marcadores moleculares que possam ser empregados no desenvolvimento de ferramentas diagnósticas, contribuindo para tornar o diagnóstico do TEA menos subjetivo do que ele é atualmente.

Para a realização desses estudos, são utilizadas técnicas que permitem a análise global da expressão gênica de um tecido em um dado momento, como o *microarray* e o RNAseq (estudo de expressão baseado no sequenciamento de nova geração) (Han et al., 2014). Enquanto o *microarray* quantifica a hibridação de cDNA a oligos específicos (pequenas sequências de DNA representando genes pré-determinados) para definir o nível de expressão dos genes que estão representados na lâmina, o RNAseq se baseia no sequenciamento do cDNA e alinhamento dos *reads* sequenciados a um genoma (previamente sequenciado ou montado a partir dos dados do RNAseq), sendo a expressão de determinado gene obtida a partir do número de *reads* que se alinham a ele (Zhao, Fung-Leung, Bittner, Ngo, & Liu, 2014; Govindarajan, Duraiyan, Kaliyappan, & Palanisamy, 2012); ambas as abordagens podem ser usadas para comparar as diferenças de expressão gênica entre duas amostras (figuras 1.2 e 1.3). As diferenças entre essas duas técnicas fazem com que o *microarray* venha sendo progressivamente substituído pelo RNAseq, uma vez que problemas inerentes à primeira, como o ruído de fundo das hibridações e a limitação da análise aos genes representados pelos oligos da lâmina, são eliminados na última. Além disso, o RNAseq torna possível a identificação de novos genes e isoformas, a quantificação de diferentes isoformas de um gene isoladamente e a avaliação de expressão alelo-específica, abrindo novas possibilidades para os estudos de

expressão (Zhao, Fung-Leung, Bittner, Ngo, & Liu, 2014; Malone & Oliver, 2011; Wang, Gerstein, & Snyder, 2009).

Figura 1.2: Princípios do *microarray*. O *microarray* é baseado na comparação entre a hibridação de cDNA com marcação fluorescente, proveniente de amostras distintas, a pequenas sequências de DNA que representam genes específicos. Nessa técnica, a diferença de expressão gênica entre as amostras é obtida por meio da quantidade relativa de fluorescência emitida em cada um dos poços onde os oligos se encontram.

Figura 1.3: Princípios do RNAseq. Nessa técnica, o mRNA é isolado, fragmentado e usado como molde para a síntese de cDNA. Às moléculas de cDNA, são adicionados adaptadores, necessários para a realização do sequenciamento de nova geração. Após o sequenciamento, os *reads* gerados são alinhados contra um genoma referência e a diferença de expressão gênica entre duas amostras pode ser obtida pela diferença no número de *reads* que se alinham aos genes.

Até o momento, uma série de trabalhos utilizando uma dessas técnicas e diversos modelos de estudos para o TEA, que variam desde diferentes tipos celulares até amostras de cérebro *post-mortem*, já foram publicados e sugerem o comprometimento de vias biológicas específicas nos pacientes (Ansel, Rosenzweig, Zisman, Melamed & Gesundheit, 2017).

Por meio do uso de *microarrays*, Kong e colaboradores (2013) usaram células de sangue periférico para analisar o perfil de expressão gênica de 20 indivíduos afetados pelo TEA pareados com seus irmãos não afetados. Controlando-se os possíveis efeitos de idade, sexo e família, foram identificados 189 genes diferencialmente expressos entre os dois grupos. Dentre eles, genes envolvidos em vias ribossômicas, spliceossômicas e mitocondriais foram encontrados com expressão aumentada entre os pacientes, enquanto genes envolvidos em vias neuroreceptor-ligante, resposta imune e sinalização de cálcio tiveram expressão reduzida nesse grupo.

Griesi-Oliveira e colaboradores (2013) realizaram estudos de *microarray* a partir do mRNA extraído de células-tronco de dentes humanos decíduos esfoliados (SHED, do inglês *stem cells from human exfoliated deciduous teeth*), um tipo celular proveniente do mesmo folheto embrionário que o cérebro, e identificaram 683 genes diferencialmente expressos entre pacientes e controles. Desse total, aproximadamente metade dos genes tem expressão no cérebro e estão envolvidos em mecanismo previamente associados ao TEA, como síntese proteica, processamento de mRNA, neurogênese, adesão celular e montagem do citoesqueleto. Os genes identificados como sendo diferencialmente expressos foram distribuídos em 68 vias de sinalização canônicas diferentes, sendo as vias relacionadas à síntese proteica, organização do citoesqueleto e formação de neuritos destacadas e apontadas pelos autores como especialmente relevantes para o TEA.

Voineagu e colaboradores (2011), por sua vez, usando amostras de tecidos *post-mortem* de duas áreas corticais distintas do cérebro de 19 pacientes e 17 controles, realizaram análises de expressão gênica diferencial e de módulos de coexpressão gênica (isto é, grupos de genes cuja expressão é altamente correlacionada) entre esses grupos. Dentre os genes diferencialmente expressos entre pacientes e controles, aqueles com expressão diminuída nos pacientes estavam enriquecidos para categorias relacionadas à função sináptica, enquanto os com expressão aumentada mostraram enriquecimento para categorias relacionadas à resposta imune e inflamatória. A análise de módulos de coexpressão, por sua vez, apontou a existência de dois módulos correlacionados ao status da doença. Consistente com o observado na análise de expressão gênica diferencial, um desses módulos foi caracterizado pelo enriquecimento para genes envolvidos em função sináptica, transporte vesicular e projeção neural, enquanto o outro apresentou uma grande quantidade de genes relacionados a resposta imunológica e marcadores gliais. Análises adicionais mostraram ainda que o módulo neuronal estava enriquecido com genes de risco para o TEA e o módulo imune não, sugerindo que as alterações na expressão gênica observadas no primeiro têm base genética, enquanto as identificadas no segundo provavelmente se devem a fenômenos secundários ou a fatores ambientais.

A partir desse mesmo modelo, mas usando um número maior de amostras de tecido cortical (57 provenientes de controles e 47 de pacientes), Gupta e colaboradores (2014) identificaram três módulos de coexpressão diferencialmente regulados entre os dois grupos. Dentre esses módulos, um estava enriquecido para genes envolvidos em resposta imunológica e ativação M2 da micróglia enquanto os dois restantes estavam envolvidos em transmissão sináptica, estando um deles enriquecido para genes de atividade de canais iônicos relacionados a GABA e o outro para genes relacionados à sinalização por peptídeos e hormônios. Os autores acreditam que, ao aumentarem o número amostral em relação ao estudo de Voineagu e colaboradores (2011), foram capazes de criar módulos

mais específicos e apontar com maior precisão os genes diferencialmente expressos entre os grupos.

Mais recentemente, foram publicados estudos de expressão utilizando NPC e neurônios derivados de iPSC como modelo. Mariani e colaboradores (2015) ao realizarem análises de redes de coexpressão gênica em organóides cerebrais derivados de iPSC de pacientes com TEA apresentando macrocefalia e de pais não afetados, mostraram que módulos enriquecidos para genes envolvidos em regulação transcricional de proliferação e destino celular, diferenciação neuronal e transmissão sináptica estavam superexpressos nos indivíduos afetados. Buscando validar funcionalmente essas análises, o grupo observou a ocorrência de alterações nas células dos pacientes que sugeriam aumento na maturação neuronal, supercrescimento sináptico e superprodução de NPC e neurônios inibitórios, gerando um desbalanço entre o número de células excitatórias e inibitórias. Ainda, o grupo conseguiu demonstrar que a superprodução de neurônios inibitórios é decorrente da superexpressão de um fator de transcrição específico, denominado FOXP1. Marchetto e colaboradores (2016), por sua vez, usando NPC e neurônios derivados de iPSC de um grupo de pacientes com TEA apresentando macrocefalia (variando entre suave e extrema) e de controles, mostraram que as NPC dos indivíduos afetados apresentam aumento na proliferação e alterações na diferenciação em relação às NPC dos controles, enquanto seus neurônios apresentam neurogênese alterada e sinaptogênese reduzida, levando a problemas nas redes neuronais formadas. Os dados do RNAseq obtidos nesse trabalho corroboram o fenótipo celular observado, uma vez que, dentre os genes diferencialmente expressos identificados nas NPC, muitos dos que apresentam expressão aumentada entre os pacientes estão envolvidos no desenvolvimento do cérebro. Já nos neurônios, os genes com expressão diminuída estão enriquecidos para as categorias cílios e axonema, consistente com as alterações sinápticas observadas nessas células. Além disso, o grupo identificou 35 genes com expressão alterada entre pacientes e controles ao longo do processo de diferenciação de NPC para neurônios, muitos dos quais estão envolvidos na formação de canais de cálcio dependentes de voltagem.

Apesar das pequenas diferenças existentes entre os estudos, é possível notar uma grande consistência nos resultados apresentados por eles, mostrando que as alterações genéticas observadas nos indivíduos acometidos pelo transtorno, ainda que distintas, devem convergir em vias de sinalização e processos biológicos comuns e reforçando a ideia de que a análise global do transcriptoma é de fato uma abordagem bastante promissora para o estudo do TEA. Além disso, estudos que fazem esse tipo de análise usando células neurais derivadas de iPSC, ainda bastante escassos, podem acrescentar dados valiosos ao corpo de conhecimento já construído, uma vez que esse modelo permite não só que se estude mais profundamente as vias moleculares já previamente

identificadas e as refine, mas também que se monitore a célula desde o estágio em que está indiferenciada até se tornar um neurônio, tornando possível a análise dessas e de outras vias que possam estar alteradas durante toda a diferenciação neuronal.

1.5 Estudos de desvio de expressão alélica

Considerando que variantes que têm alto impacto no sucesso reprodutivo dos indivíduos tendem a ser raras na população, uma abordagem bastante popular no estudo do TEA é a busca por variantes *de novo* que possam estar associadas à manifestação de seus sintomas clínicos. Esse tipo de abordagem, apesar de importante na identificação de genes e *loci* relevantes, acaba por excluir quaisquer informações sobre variantes herdadas, as quais também devem ter papel na etiologia da doença, considerando-se, por exemplo, a agregação familiar observada em alguns casos. Nesse sentido, a expressão preferencial de um dos alelos de um gene em detrimento do outro (isto é, o desvio de expressão alélica) surge como um mecanismo que poderia explicar alguns dos fenômenos observados no TEA e em outras doenças neuropsiquiátricas, tais como penetrância incompleta e expressividade variável.

Nesse contexto, os dados de expressão gênica de um dado tipo celular provenientes de diferentes indivíduos, quando obtidos por meio de RNAseq, além de nos fornecer informações sobre as diferenças de expressão gênica associadas a uma determinada condição, podem ainda indicar a ocorrência de expressão preferencial de um dos alelos de um gene. O sequenciamento do mRNA, resultante do emprego dessa técnica, torna possível apontar a contribuição de cada alelo, quando em heterozigose, para o total de transcritos de um gene e, conseqüentemente, verificar se existe desvio de expressão para um deles.

Até pouco tempo atrás, apesar da expressão monoalélica em decorrência de *imprinting*, inativação do cromossomo X e eventos aleatórios específicos já ter sido bem documentada para determinados genes, ainda não estava claro para a maior parte do genoma se os dois alelos de um gene eram simultaneamente utilizados como molde para a transcrição e, em caso positivo, se ambos contribuíam de forma similar para o total de transcritos. A fim de investigar essa questão, Deng e colaboradores (2014) realizaram o RNAseq de células embrionárias individuais de camundongos e mostraram que entre 12 e 24% dos genes tem expressão monoalélica em um dado momento e que o alelo ativo em uma célula pode não ser o mesmo que está ativo nas células vizinhas, apontando a estocasticidade do processo. Os genes que apresentam expressão bialélica são, em geral, constitutivos. Ainda, o grupo obteve resultados similares em fibroblastos e hepatócitos murinos, demonstrando que o mesmo processo ocorre em células maduras. Em 2015, Borel e colaboradores, fazendo uso de RNAseq em fibroblastos humanos individuais,

obtiveram resultados semelhantes, sugerindo que o mesmo fenômeno deve estar ocorrendo também em células humanas.

O trabalho de Lin e colaboradores (2012) também demonstrou a ocorrência de desvio de expressão alélica em neurônios humanos. Usando um *pool* de células neuronais derivadas de iPSC, o grupo identificou 801 genes apresentando desvio de expressão alélica. O grupo verificou ainda que o desvio observado não se devia a origem parental do alelo, também sugerindo que eventos estocásticos devam estar envolvidos nesse processo. Além disso, dentre os genes identificados, havia um pequeno enriquecimento para genes previamente associados à esquizofrenia e ao TEA.

Ainda, Gregg e colaboradores (2010) verificaram, no cérebro de camundongos adultos, a ocorrência de desvio de expressão alélica de acordo com o sexo do indivíduo e a origem parental do alelo. Alguns dos genes identificados pelo grupo são candidatos para transtornos neurológicos ou tem função relevante no comportamento.

Esses estudos demonstram que o desvio de expressão alélica é um mecanismo comum (ocorrendo inclusive em neurônios) que pode ter relação com o desvio sexual ou a penetrância incompleta observados em transtornos que afetam o sistema nervoso, reforçando a hipótese de que ele pode ter um papel relevante na etiologia do TEA e que, portanto, investigações mais profundas neste sentido podem ser uma abordagem promissora para o estudo dessa doença.

OBJETIVOS

Os estudos de expressão gênica global conduzidos até o momento demonstram a existência de alterações em processos relevantes para o neurodesenvolvimento que são compartilhadas pelos indivíduos afetados pelo TEA. Trabalhos que fazem esse tipo de análise usando células neurais derivadas de iPSC ainda são bastante escassos, mas são importantes para a compreensão do transtorno uma vez que possibilitam o monitoramento da expressão gênica em células provenientes de indivíduos afetados durante diversos estágios da diferenciação neuronal. Os dois estudos realizados até o momento a partir desse modelo são baseados em coortes de indivíduos afetados pelo TEA que apresentam macrocefalia, o que indica a necessidade de condução desses experimentos em outros grupos de pacientes. Além disso, o estudo do transcriptoma pela técnica de RNAseq permite a análise de parâmetros ainda pouco explorados no TEA, como por exemplo, o possível papel do mecanismo de desvio de expressão alélica na etiologia do transtorno.

Posto isso, a partir da realização do RNAseq em NPC e neurônios derivados de iPSC de seis indivíduos afetados pelo TEA e seis controles, os objetivos desse trabalho são: (1) identificar vias de sinalização e processos biológicos que estejam alterados entre pacientes e controles por meio de análises de expressão gênica diferencial e de redes de coexpressão gênica em células neurais e (2) verificar se genes candidatos ao TEA agrupados no banco de dados SFARI apresentam desvio de expressão alélica nesses mesmos tipos celulares. Ainda, durante a caracterização genética de nossos pacientes, identificamos uma duplicação na região 17p13.3 afetando seis genes em um dos indivíduos. Os impactos transcricionais e funcionais da duplicação foram avaliados em NPC e neurônios desse paciente e os resultados desse estudo de caso são descritos no capítulo 7 da presente dissertação.

CONCLUSÕES

Nossas análises apontam a ocorrência de desregulações transcricionais em um conjunto de genes importantes no neurodesenvolvimento, alterações também observadas em um contexto sistêmico quando são analisados módulos de coexpressão gênica. Os neurônios diferenciados a partir de nosso protocolo manifestam de forma consistente as alterações observadas nos cérebros de indivíduos afetados pelo TEA e em modelos celulares semelhantes aos empregados nesse estudo, sugerindo que a grande heterogeneidade genética associada ao transtorno se reflete em alterações de processos biológicos específicos, as quais são observadas mesmo em modelos de estudo e/ou subgrupos de pacientes diferentes. Verificamos também que nossas amostras neuronais parecem representar de forma mais próxima o perfil de expressão do cérebro fetal e a comparação entre nossos resultados e aqueles obtidos em estudos anteriores sugere que esse conjunto de genes associados à função sináptica apresenta desregulações opostas no período fetal em relação ao período pós-natal, apontando a relevância do uso de iPSC no estudo de transtornos do neurodesenvolvimento. Ainda, ao demonstrar a ocorrência de desregulações transcricionais associadas ao *status* da doença que não se mantem em estágios posteriores do neurodesenvolvimento, fomos capazes de apontar a importância do uso das NPC para entender de forma mais ampla os processos envolvidos na patofisiologia do TEA. Nesse último caso, a necessidade da utilização de reprogramação celular se torna ainda mais evidente devido à impossibilidade de obtenção de amostras de cérebro de indivíduos afetados que correspondam a estágios tão precoces do desenvolvimento.

Ao migrar o foco das análises do nível genômico, no qual os pacientes são altamente heterogêneos, para o transcriptômico, no qual é possível observar desregulações comuns entre esses indivíduos, obtivemos resultados que abrem caminho para uma melhor compreensão sobre os mecanismos envolvidos na manifestação do TEA e, eventualmente, das comorbidades associadas a ele. Em particular, a observação consistente da desregulação do módulo M1-turquoise em indivíduos afetados pelo TEA sugere sua importância para o desenvolvimento de ferramentas diagnósticas e de terapias medicamentosas voltadas para o tratamento dos principais sintomas do transtorno, focando em drogas que tenham como alvo genes desse módulo.

ESTUDO DE CASO

Introdução

O *locus* 17p13.3 apresenta grande instabilidade genômica e é caracterizado pela presença de diversas CNV (*copy number variants*) não recorrentes (Gu et al., 2015). Ainda que duplicações de diferentes tamanhos identificadas nessa região apresentem penetrância incompleta, elas têm sido associadas a transtornos do neurodesenvolvimento, como TEA e deficiência intelectual (Curry et al., 2013; Prasad et al., 2012; Capra et al., 2012; Bruno et al., 2010). Com base em análises de sobreposição dessas duplicações, Bruno e colaboradores (2010) identificaram uma região mínima dentro desse *locus* que parece estar fortemente associada à manifestação do TEA e contém apenas um gene, o *YWHAE*. Ainda, discute-se a possibilidade do *CRK*, gene adjacente ao *YWHAE*, também estar contribuindo com a manifestação do transtorno (Curry et al., 2013).

O gene *YWHAE* codifica a proteína 14-3-3epsilon, pertencente a uma família de proteínas cujos membros interagem com mais de 200 parceiros por meio de ligações a resíduos de serina e treonina fosforilados e estão envolvidos em uma série de eventos celulares, tais como regulação do ciclo celular e transdução de sinais (Cornell, Wachi, Zhukarev, & Toyo-oka, 2016; Berg, Holzmann, & Riess, 2003). Ainda, as proteínas dessa família são altamente expressas no cérebro, sugerindo sua importância na função neuronal (Berg, Holzmann & Riess, 2003). A relevância da proteína 14-3-3epsilon na proliferação, migração e diferenciação de células progenitoras neurais já foi demonstrada (Toyo-oka et al., 2014; Toyo-Oka et al., 2003). Mais recentemente, Cornell e colaboradores (2016) mostraram também o papel dessa proteína na morfogênese neuronal por meio da interação e estabilização de Dcx, uma proteína que se liga a microtúbulos. Além disso, Wachi e colaboradores (2017) observaram alterações em uma série de comportamentos neuropsiquiátricos (como atividade locomotora, memória e sociabilidade) em camundongos nocaute para o gene *YWHAE*, sugerindo o envolvimento de 14-3-3epsilon na patofisiologia de transtornos neurológicos.

O gene *CRK*, por sua vez, codifica a proteína Crk, pertencente a uma família de proteínas adaptadoras que são responsáveis por mediar a formação de complexos proteicos envolvidos na transdução de sinais em resposta a uma série de estímulos extracelulares (Feller, 2001). As proteínas da família Crk são expressas em muitos tecidos e participam de diversos processos biológicos, incluindo regulação da morfologia, migração, proliferação e diferenciação celular (Feller, 2001). No desenvolvimento do cérebro, as proteínas Crk estão envolvidas no controle da migração neuronal e formação das camadas do córtex por meio da ativação da via C3G/Rap1 e na promoção de tradução

proteica e, possivelmente, crescimento de dendritos e desenvolvimento de espinhos dendríticos por meio da ativação da via mTOR em resposta ao estímulo por Relina, proteína codificada por um gene cuja associação com o TEA já está bem estabelecida (figura 7.1A e B; Lee & D’Arcangelo, 2016; Park & Curran, 2008; Ballif et al., 2004).

Figura 7.1: A proteína Crk participa da ativação de vias relevantes para o neurodesenvolvimento, tais como a C3G/Rap1, relacionada a migração neuronal, e Akt/mTOR, envolvida no controle de tradução.

A associação de ambos os genes ao TEA e deficiência intelectual é baseada em análises de sobreposição de duplicações e estudos funcionais conduzidos em linhagens celulares e organismos modelo que apontam sua importância no neurodesenvolvimento, mas até o momento nenhum estudo usando células de pacientes foi realizado. Nosso grupo identificou uma duplicação *de novo* de 345kb no *locus* 17p13.3 envolvendo seis genes, dentre eles *YWHAE* e *CRK*, em um indivíduo do sexo masculino diagnosticado com TEA. Assim, nesse estudo, analisamos as consequências funcionais dessa duplicação avaliando a contribuição dos genes *YWHAE* e *CRK* nas possíveis alterações moleculares em células progenitoras neurais (NPC) e neurônios derivados de células-tronco pluripotentes induzidas (iPSC). Considerando ainda a penetrância incompleta do fenótipo em indivíduos apresentando duplicações nessa região, verificamos a existência de outras variantes que pudessem estar contribuindo para a manifestação do quadro nesse paciente.

Conclusão

Nesse estudo, demonstramos pela primeira vez o impacto funcional da duplicação na região 17p13.3 em células neurais derivadas de um indivíduo afetado pelo TEA. Ainda que sejam necessários experimentos adicionais para responder algumas questões em aberto, nossos resultados mostram a ocorrência de alterações importantes em vias e processos relevantes para o neurodesenvolvimento nesse paciente, as quais são condizentes com o aumento de expressão de *CRK* e *YWHAE*, reforçando o protagonismo desses genes no fenótipo decorrente de duplicações em 17p13.3.

BIBLIOGRAFIA

- Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nature Reviews. Genetics*, 9(5), 341–355. <https://doi.org/10.1038/nrg2346>
- Amano, M., Tsumura, Y., Taki, K., Harada, H., Mori, K., Nishioka, T., ... Kaibuchi, K. (2010). A proteomic approach for comprehensively screening substrates of protein kinases such as rho-kinase. *PLoS ONE*, 5(1), 1–9. <https://doi.org/10.1371/journal.pone.0008704>
- American Psychiatric Association. (1994). Diagnostic and Statistical Manual of Mental Disorders. 4thed. Washington (DC): American Psychiatric Association.
- American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. 5thed. Washington (DC): American Psychiatric Association.
- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166–169. <https://doi.org/10.1093/bioinformatics/btu638>
- Ansel, A., Rosenzweig, J. P., Zisman, P. D., Melamed, M., & Gesundheit, B. (2017). Variation in gene expression in autism spectrum disorders: An extensive review of transcriptomic studies. *Frontiers in Neuroscience*, 10(JAN). <https://doi.org/10.3389/fnins.2016.00601>
- Ballif, B.A., Arnaud, L., Arthur, W.T., Guris, D., Imamoto, A. & Cooper, J.A. (2004) Activation of a Dab1/CrkL/C3G/Rap1 pathway in Reelin-stimulated neurons. *Curr Biol*, 14(7):606-10.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, 45(3), 601–613. <https://doi.org/10.1017/S003329171400172X>
- Berg, D., Holzmann, C., & Riess, O. (2003). 14-3-3 Proteins in the Nervous System. *Nature Reviews Neuroscience*, 4(9), 752–762. <https://doi.org/10.1038/nrn1197>
- Bilbo, S. D., Block, C. L., Bolton, J. L., Hanamsagar, R., & Tran, P. K. (2018). Beyond infection - Maternal immune activation by environmental factors, microglial development, and relevance for autism spectrum disorders. *Experimental Neurology*, 299, 241–251. <https://doi.org/10.1016/j.expneurol.2017.07.002>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Borel, C., Ferreira, P. G., Santoni, F., Delaneau, O., Fort, A., Popadin, K. Y., ... Antonarakis, S. E. (2015). Biased allelic expression in human primary fibroblast single cells. *The American Journal of Human Genetics*, 96, 70–80. <https://doi.org/10.1016/j.ajhg.2014.12.001>
- Bourgeron, T. (2015). From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nature Reviews Neuroscience*, 16(9), 551–563. <https://doi.org/10.1038/nrn3992>
- Bruno, D. L., Anderlid, B. M., Lindstrand, A., Van Ravenswaaij-Arts, C., Ganesamoorthy, D., Lundin, J., ... Schoumans, J. (2010). Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *Journal of Medical Genetics*, 47(5), 299–311. <https://doi.org/10.1136/jmg.2009.069906>
- Capra, V., Mirabelli-Badenier, M., Stagnaro, M., Rossi, A., Tassano, E., Gimelli, S., & Gimelli, G. (2012). Identification of a rare 17p13.3 duplication including the BHLHA9 and YWHA9 genes in a family with developmental delay and behavioural problems. *BMC Medical Genetics*, 13, 7–9. <https://doi.org/10.1186/1471-2350-13-93>
- Christensen, D. L., Baio, J., Braun, K. V. N., Bilder, D., Charles, J., Constantino, J. N., ...

- Yeargin-Allsopp, M. (2016). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *Morbidity and Mortality Weekly Report. Surveillance Summaries*, 65(3), 1–23. <https://doi.org/10.15585/mmwr.ss6503a1>
- Cornell, B., & Toyo-oka, K. (2017). 14-3-3 Proteins in Brain Development: Neurogenesis, Neuronal Migration and Neuromorphogenesis. *Frontiers in Molecular Neuroscience*, 10(October), 1–17. <https://doi.org/10.3389/fnmol.2017.00318>
- Cornell, B., Wachi, T., Zhukarev, V., & Toyo-oka, K. (2016). Regulation of neuronal morphogenesis by 14-3-3epsilon (Ywhae) via the microtubule binding protein, doublecortin. *Human Molecular Genetics*, 25(20), ddw270. <https://doi.org/10.1093/hmg/ddw270>
- Costa-Mattioli, M., & Monteggia, L. M. (2013). mTOR complexes in neurodevelopmental and neuropsychiatric disorders. *Nature Neuroscience*, 16(11), 1537–1543. <https://doi.org/10.1038/nn.3546>
- Costa, A., Bueno, D. F., Martins, M. T., Kerkis, I., Kerkis, A., Fanganiello, R. D., ... Passos-Bueno, M. R. (2008). Reconstruction of Large Cranial Defects in Nonimmunosuppressed Experimental Design With Human Dental Pulp Stem Cells. *The Journal of Craniofacial Surgery*, 19, 204–210.
- Curry, C. J., Rosenfeld, J. A., Grant, E., Gripp, K. W., Anderson, C., Aylsworth, A. S., ... Dobyns, W. B. (2013). The duplication 17p13.3 phenotype: Analysis of 21 families delineates developmental, behavioral and brain abnormalities, and rare variant phenotypes. *American Journal of Medical Genetics, Part A*, 161(8), 1833–1852. <https://doi.org/10.1002/ajmg.a.35996>
- Deng, Q., Ramsköld, D., Reinius, B. & Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*, 343, 193–196. <https://doi.org/10.1126/science.1245316>
- Doers, M. E., Musser, M. T., Nichol, R., Berndt, E. R., Baker, M., Gomez, T. M., ... Bhattacharyya, A. (2014). iPSC-Derived Forebrain Neurons from FXS Individuals Show Defects in Initial Neurite Outgrowth. *Stem Cells and Development*, 23(15), 1777–1787. <https://doi.org/10.1089/scd.2014.0030>
- Donovan, A. P. A., & Basson, M. A. (2017). The neuroanatomy of autism - a developmental perspective. *Journal of Anatomy*, 230(1), 4–15. <https://doi.org/10.1111/joa.12542>
- Dotti, M. T., Manneschi, L., Malandrini, A., De Stefano, N., Caznerale, F., & Federico, A. (1993). Mitochondrial dysfunction in Rett syndrome. An ultrastructural and biochemical study. *Brain & Development*, 15, 103–106. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8214327>
- Eisenberg, E., & Levanon, E. Y. (2013). Human housekeeping genes, revisited. *Trends in Genetics*, 29(10), 569–574. <https://doi.org/10.1016/j.tig.2013.05.010>
- Faus-Garriga, J., Novoa, I., & Ozaita, A. (2017). mTOR signaling in proteostasis and its relevance to autism spectrum disorders. *AIMS Biophysics*, 4(1), 63–89. <https://doi.org/10.3934/biophy.2017.1.63>
- Feller, S. M. (2001). Crk family adaptors - signalling complex formation and biological roles. *Oncogene*, 20(44), 6348–6371. <https://doi.org/10.1038/sj.onc.1204779>
- Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, 65(6), 591–598. <https://doi.org/10.1203/PDR.0b013e31819e7203>
- Frye, R., & Rossignol, D. (2011). Mitochondrial Dysfunction Can Connect the Diverse Medical Symptoms Associated With Autism Spectrum Disorders. *Pediatr Res*, 69(5), 41R–47R. <https://doi.org/10.1203/PDR.0b013e318212f16b>
- Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., ... Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, 46(8), 881–885. <https://doi.org/10.1038/ng.3039>
- Gorrindo, P., Williams, K. C., Lee, E. B., Walker, L. S., McGrew, S. G., & Levitt, P. (2012). Gastrointestinal dysfunction in autism: Parental report, clinical evaluation, and associated factors. *Autism Research*, 5(2), 101–108. <https://doi.org/10.1002/aur.237>
- Govindarajan, R., Duraiyan, J., Kaliyappan, K., & Palanisamy, M. (2012). Microarray and its

- applications. *Journal of Pharmacy and Bioallied Sciences*, 4(6), 310–312. <https://doi.org/10.4103/0975-7406.100283>
- Gregg, C., Zhang, J., Butler, J. E., Haig, D. and Dulac, C. (2010). Sex-specific parent-of-origin allelic expression in the mouse brain. *Science*, 329, 682–685. <https://doi.org/10.1126/science.1190831>
- Griesi-Oliveira K., Sunaga D. Y., Alvizi L., Vadasz E. & Passos-Bueno M. R. (2013). Stem cells as a good tool to investigate dysregulated biological systems in autism spectrum disorders. *Autism Res.*, 6(5), 354–361. <https://doi.org/10.1002/aur.1296>
- Griesi-Oliveira, K., Acab, A., Gupta, a R., Sunaga, D. Y., Chailangkarn, T., Nicol, X., ... Muotri, a R. (2014). Modeling non-syndromic autism and the impact of TRPC6 disruption in human neurons. *Molecular Psychiatry*, (March), 1–16. <https://doi.org/10.1038/mp.2014.141>
- Griesi-Oliveira, K., & Sertié, A. L. (2017). Autism spectrum disorders: an updated guide for genetic counseling. *Einstein (São Paulo)*, 15(2), 233–238. <https://doi.org/10.1590/s1679-45082017rb4020>
- Gu, S., Yuan, B., Campbell, I. M., Beck, C. R., Carvalho, C. M. B., Nagamani, S. C. S., ... Lupski, J. R. (2015). Alu-mediated diverse and complex pathogenic copy-number variants within human chromosome 17 at p13.3. *Human Molecular Genetics*, 24(14), 4061–4077. <https://doi.org/10.1093/hmg/ddv146>
- Gupta, S., Ellis, S. E., Ashar, F. N., Moes, A., Bader, J. S., Zhan, J., ... Arking, D. E. (2014). Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nature Communications*, 5, 1–8. <https://doi.org/10.1038/ncomms6748>
- Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., ... Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism, *Arch Gen Psychiatry*, 68(11), 1095–1102.
- Han, G., Sun, J., Wang, J., Bai, Z., Song, F., & Lei, H. (2014). Genomics in Neurological Disorders. *Genomics, Proteomics and Bioinformatics*, 12(4), 156–163. <https://doi.org/10.1016/j.gpb.2014.07.002>
- Hart, T., Komori, H. K., LaMere, S., Podshivalova, K., & Salomon, D. R. (2013). Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics*, 14(1). <https://doi.org/10.1186/1471-2164-14-778>
- Højsgaard, S., Halekoh, U., & Yan, J. (2006). The R Package geepack for generalized estimating equations. *Journal of Statistical Software*, 15(2), 1–11. <https://doi.org/10.18637/jss.v015.i02>
- Hsueh, Y. P. (2012). Neuron-specific regulation on F-actin cytoskeletons: The role of CTTNBP2 in dendritic spinogenesis and maintenance. *Communicative and Integrative Biology*, 5(4), 334–336. <https://doi.org/10.4161/cib.20364>
- Huguet, G., Benabou, M., & Bourgeron, T. (2016). The Genetics of Autism Spectrum Disorders. In P. Sassone-Corsi & Y. Christen (Eds.), *A Time for Metabolism and Hormones* (pp. 101–129). Springer. <https://doi.org/10.1007/978-3-319-27069-2>
- Jones, E. J. H., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience and Biobehavioral Reviews*, 39, 1–33. <https://doi.org/10.1016/j.neubiorev.2013.12.001>
- Jung, M.-Y., Lorenz, L., & Richter, J. D. (2006). Translational Control by Neuroguadin, a Eukaryotic Initiation Factor 4E and CPEB Binding Protein. *Molecular and Cellular Biology*, 26(11), 4277–4287. <https://doi.org/10.1128/MCB.02470-05>
- Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., ... Šestan, N. (2011). Spatio-temporal transcriptome of the human brain. *Nature*, 478, 483–489. <https://doi.org/10.1038/nature10523>
- Kircher, M., Witten, D. M., Jain, P., O’roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, 46(3), 310–315. <https://doi.org/10.1038/ng.2892>
- Klei, L., Sanders, S. J., Murtha, M. T., Hus, V., Lowe, J. K., Willsey, A. J., ... Devlin, B. (2012). Common genetic variants, acting additively, are a major source of risk for

- autism. *Molecular Autism*, 3(1), 1–13. <https://doi.org/10.1186/2040-2392-3-9>
- Kong, S. W., Shimizu-Motohashi, Y., Campbell, M. G., Lee, I. H., Collins, C. D., Brewster, S. J., ... Kunkel, L. M. (2013). Peripheral blood gene expression signature differentiates children with autism from unaffected siblings. *Neurogenetics* 14(2), 143–152.
- Konradi, C., Eaton, M., MacDonald, M. L., Walsh, J., Benes, F. M., & Heckers, S. (2004). Molecular evidence for mitochondrial dysfunction in bipolar disorder. *Arch. Gen. Psychiat.*, 61, 300–308. <https://doi.org/10.1001/archpsyc.61.3.300>
- Krakowiak, P., Goodlin-Jones, B., Hertz-Picciotto, I., Croen, L. A., & Hansen, R. L. (2008). Sleep problems in children with autism spectrum disorders, developmental delays, and typical development: A population-based study. *Journal of Sleep Research*, 17(2), 197–206. <https://doi.org/10.1111/j.1365-2869.2008.00650.x>
- Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, 9. <https://doi.org/10.1186/1471-2105-9-559>
- Langfelder, P., Luo, R., Oldham, M. C., & Horvath, S. (2011). Is my network module preserved and reproducible? *PLoS Computational Biology*, 7(1). <https://doi.org/10.1371/journal.pcbi.1001057>
- Lee, G. H., & D'Arcangelo, G. (2016). New Insights into Reelin-Mediated Signaling Pathways. *Frontiers in Cellular Neuroscience*, 10, 1–8. <https://doi.org/10.3389/fncel.2016.00122>
- Lee, Y., & McKinnon, P. J. (2007). Responding to DNA double strand breaks in the nervous system. *Neuroscience*, 145(4), 1365–1374. <https://doi.org/10.1016/j.neuroscience.2006.07.026>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, Z., Dong, X., Wang, Z., Liu, W., Deng, N., Ding, Y., ... Wu, D. (2005). Regulation of PTEN by Rho small GTPases. *Nature Cell Biology*, 7(4), 399–404. <https://doi.org/10.1038/ncb1236>
- Lin, M., Hrabovsky, A., Pedrosa, E., Wang, T., Zheng, D., & Lachman, H. M. (2012). Allele-Biased Expression in Differentiating Human Neurons: Implications for Neuropsychiatric Disorders. *PLoS ONE*, 7(8), 1–10. <https://doi.org/10.1371/journal.pone.0044017>
- Lin, Y.-C., Frei, J. A., Kilander, M. B. C., Shen, W., & Blatt, G. J. (2016). A Subset of Autism-Associated Genes Regulate the Structural Stability of Neurons. *Frontiers in Cellular Neuroscience*, 10, 1–35. <https://doi.org/10.3389/fncel.2016.00263>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Magdalon, J., Sánchez-Sánchez, S. M., Griesi-Oliveira, K., & Sertié, A. L. (2017). Dysfunctional mTORC1 signaling: A convergent mechanism between syndromic and nonsyndromic forms of autism spectrum disorder? *International Journal of Molecular Sciences*, 18(3), 1–27. <https://doi.org/10.3390/ijms18030659>
- Malone, J. H., & Oliver, B. (2011). Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biology*, 9. <https://doi.org/10.1186/1741-7007-9-34>
- Manji, H., Kato, T., Di Prospero, N. A., Ness, S., Beal, M. F., Krams, M., & Chen, G. (2012). Impaired mitochondrial function in psychiatric disorders. *Nature Reviews Neuroscience*, 13, 293–307. <https://doi.org/10.1038/nrn3229>
- Marchetto, M. C. N. M., Carromeu, C., Acab, A., Yu, D., Yeo, G. W., Mu, Y., ... Muotri, A. R. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell*, 143(4), 527–539. <https://doi.org/10.1016/j.cell.2010.10.016>
- Marchetto, M. C., Bellinson, H., Tian, Y., Freitas, B. C., Fu, C., Vadodaria, K. C., ... Muotri, A. R. (2016). Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Molecular Psychiatry* 00, 1–16. <https://doi.org/10.1038/mp.2016.95>
- Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., ... Vaccarino, F. M. (2015). FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron

- Differentiation in Autism Spectrum Disorders. *Cell*, 162(2), 375–390. <https://doi.org/10.1016/j.cell.2015.06.034>
- McKenna, A., Matthew, H., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., ... DePristo, M. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110.20>
- Miller, F. D., & Gauthier, A. S. (2007). Timing Is Everything: Making Neurons versus Glia in the Developing Cortex. *Neuron*, 54, 357–369. <https://doi.org/10.1016/j.neuron.2007.04.019>
- Parikshak, N. N., Luo, R., Zhang, A., Won, H., Lowe, J. K., Chandran, V., ... Geschwind, D. H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*, 155(5), 1008–1021. <https://doi.org/10.1016/j.cell.2013.10.031>
- Park, T., & Curran, T. (2008). Crk and CrkL play essential overlapping roles downstream of Dab1 in the Reelin pathway. *J Neurosci*, 28(50), 13551–13562. <https://doi.org/10.1523/JNEUROSCI.4323-08.2008.Crk>
- Prabakaran, S., Swatton, J., Ryan, M., Huffaker, S., Huang, J.-J., Griffin, J., ... Bahn, S. (2004). Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. *Molecular Psychiatry*, 9, 684–697. <https://doi.org/10.1038/sj.mp.4001511>
- Prasad, A., Merico, D., Thiruvahindrapuram, B., Wei, J., Lionel, A. C., Sato, D., ... Scherer, S. W. (2012). A Discovery Resource of Rare Copy Number Variations in Individuals with Autism Spectrum Disorder. *G3 Genes/Genomes/Genetics*, 2(12), 1665–1685. <https://doi.org/10.1534/g3.112.004689>
- Pucilowska, J., Puzerey, P. A., Karlo, J. C., Galan, R. F., & Landreth, G. E. (2012). Disrupted ERK Signaling during Cortical Development Leads to Abnormal Progenitor Proliferation, Neuronal and Network Excitability and Behavior, Modeling Human Neuro-Cardio-Facial-Cutaneous and Related Syndromes. *Journal of Neuroscience*, 32(25), 8663–8677. <https://doi.org/10.1523/JNEUROSCI.1107-12.2012>
- Pucilowska, J., Vithayathil, J., Tavares, E. J., Kelly, C., Karlo, J. C., & Landreth, G. E. (2015). The 16p11.2 Deletion Mouse Model of Autism Exhibits Altered Cortical Progenitor Proliferation and Brain Cytoarchitecture Linked to the ERK MAPK Pathway. *Journal of Neuroscience*, 35(7), 3190–3200. <https://doi.org/10.1523/JNEUROSCI.4864-13.2015>
- Risch, N., Hoffmann, T. J., Anderson, M., Croen, L. A., Grether, J. K., & Windham, G. C. (2014). Familial recurrence of autism spectrum disorder: Evaluating genetic and environmental contributions. *American Journal of Psychiatry*, 171(11), 1206–1213. <https://doi.org/10.1176/appi.ajp.2014.13101359>
- Risso, D., Ngai, J., Speed, T. P., & Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology*, 32(9), 896–902. <https://doi.org/10.1038/nbt.2931>
- Ronald, A. & Hoekstra, R.A. (2011). Autism Spectrum Disorders and Autistic Traits: A Decade of New Twin Studies. *Am J Med Genet Part B: Neuropsychiatric Genetics*, 156(3), 255–274. <https://doi.org/10.1002/ajmg.b.31159>
- Rossignol, D. A., & Frye, R. E. (2012). Mitochondrial dysfunction in autism spectrum disorders: A systematic review and meta-analysis. *Molecular Psychiatry*, 17(3), 290–314. <https://doi.org/10.1038/mp.2010.136>
- Russo, F. B., Freitas, B. C., Pignatari, G. C., Fernandes, I. R., Sebat, J., Muotri, A. R., & Beltrão-Braga, P. C. B. (2017). Modeling the Interplay Between Neurons and Astrocytes in Autism Using Human Induced Pluripotent Stem Cells. *Biological Psychiatry*, 83(7), 569–578. <https://doi.org/10.1016/j.biopsych.2017.09.021>
- Samuels, I. S., Karlo, J. C., Faruzzi, A. N., Pickering, K., Herrup, K., Sweatt, J. D., ... Landreth, G. E. (2008). Deletion of ERK2 Mitogen-Activated Protein Kinase Identifies Its Key Roles in Cortical Neurogenesis and Cognitive Function. *Journal of Neuroscience*, 28(27), 6983–6995. <https://doi.org/10.1523/JNEUROSCI.0679-08.2008>
- Sandin, S., Lichtenstein, P., Larsson, H., Hultman, C., & Reichenberg, A. (2014). The

- familial risk of autism. *JAMA*, 311(17), 1770–1777. <https://doi.org/10.1001/jama.2014.4144>.The
- Santini, E., & Klann, E. (2014). Reciprocal signaling between translational control pathways and synaptic proteins in autism spectrum disorders. *Science Signaling*, 7(349), 1–10. <https://doi.org/10.1126/scisignal.2005832>
- Sheridan, S. D., Theriault, K. M., Reis, S. A., Zhou, F., Madison, J. M., Daheron, L., ... Haggarty, S. J. (2011). Epigenetic characterization of the FMR1 gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile X syndrome. *PLoS ONE*, 6(10). <https://doi.org/10.1371/journal.pone.0026203>
- Shih, P.-Y., Lee, S.-P., Chen, Y.-K., & Hsueh, Y.-P. (2014). Cortactin-binding protein 2 increases microtubule stability and regulates dendritic arborization. *Journal of Cell Science*, 127(16), 3521–3534. <https://doi.org/10.1242/jcs.149476>
- Sroka, R., Van Lint, J., Katz, S.-F., Schneider, M. R., Kleger, A., Paschke, S., ... Eiseler, T. (2016). Cortactin is a scaffolding platform for the E-cadherin adhesion complex and is regulated by protein kinase D1 phosphorylation. *Journal of Cell Science*, 129(12), 2416–2429. <https://doi.org/10.1242/jcs.184721>
- Suzuki, A. M., Griesi-Oliveira, K., De Oliveira Freitas Machado, C., Vadasz, E., Zachi, E. C., Passos-Bueno, M. R., & Sertie, A. L. (2015). Altered mTORC1 signaling in multipotent stem cells from nearly 25% of patients with nonsyndromic autism spectrum disorders. *Molecular Psychiatry*, 20(5), 550–552. <https://doi.org/10.1038/mp.2014.175>
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, 131(5), 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>
- Toyo-Oka, K., Shionoya, A., Gambello, M. J., Cardoso, C., Leventer, R., Ward, H. L., ... Wynshaw-Boris, A. (2003). 14-3-3 ϵ is important for neuronal migration by binding to NUDEL: A molecular explanation for Miller-Dieker syndrome. *Nature Genetics*, 34(3), 274–285. <https://doi.org/10.1038/ng1169>
- Toyo-oka, K., Wachi, T., Hunt, R. F., Baraban, S. C., Taya, S., Ramshaw, H., ... Wynshaw-Boris, A. (2014). 14-3-3 ϵ and ζ Regulate Neurogenesis and Differentiation of Neuronal Progenitor Cells in the Developing Brain. *Journal of Neuroscience*, 34(36), 12168–12181. <https://doi.org/10.1523/JNEUROSCI.2513-13.2014>
- Trottier, G., Srivastava, L., & Walker, C. D. (1999). Etiology of infantile autism: A review of recent advances in genetic and neurobiological research. *Journal of Psychiatry and Neuroscience*, 24(2), 103–115.
- Trushina, E., & McMurray, C. T. (2007). Oxidative stress and mitochondrial dysfunction in neurodegenerative diseases. *Neuroscience*, 145, 1233–1248. <https://doi.org/10.1016/j.neuroscience.2006.10.056>
- Turner, T. N., Sharma, K., Oh, E. C., Liu, Y. P., Collins, R. L., Sosa, M. X., ... Chakravarti, A. (2015). Loss of δ -catenin function in severe autism. *Nature*, 520(7545), 51–56. <https://doi.org/10.1038/nature14186>
- Voineagu, I., Wang, X., Johnston, P., Lowe, J. K., Tian, Y., Horvath, S., ... Geschwind, D. H. (2011). Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, 474(7351), 380–386. <https://doi.org/10.1038/nature10110>
- Wachi, T., Cornell, B., & Toyo-oka, K. (2017). Complete ablation of the 14-3-3epsilon protein results in multiple defects in neuropsychiatric behaviors. *Behavioural Brain Research*, 319, 31–36. <https://doi.org/10.1016/j.bbr.2016.11.016>
- Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), 1–7. <https://doi.org/10.1093/nar/gkq603>
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, 57–63.
- Williams, E. C., Zhong, X., Mohamed, A., Li, R., Liu, Y., Dong, Q., ... Chang, Q. (2014). Mutant astrocytes differentiated from Rett syndrome patients-specific iPSCs have adverse effects on wildtype neurons. *Human Molecular Genetics*, 23(11), 2968–2980. <https://doi.org/10.1093/hmg/ddu008>

- Wing, L. & Potter, D. (2002). The epidemiology of autistic spectrum disorders: is the prevalence rising? *Mental Retardation and Developmental Disabilities Research Reviews*, 8(3), 151-61.
- Wu, J., Li, Y., & Jiang, R. (2014). Integrating Multiple Genomic Data to Predict Disease-Causing Nonsynonymous Single Nucleotide Variants in Exome Sequencing Studies. *PLoS Genetics*, 10(3). <https://doi.org/10.1371/journal.pgen.1004237>
- Yoon, K., & Gaiano, N. (2005). Notch signaling in the mammalian central nervous system: Insights from mouse mutants. *Nature Neuroscience*, 8(6), 709–715. <https://doi.org/10.1038/nn1475>
- Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., ... Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318(5858), 1917–1920. <https://doi.org/10.1126/science.1151526>
- Zhao, S., Fung-Leung, W. P., Bittner, A., Ngo, K., & Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS ONE*, 9(1), 1–13. <https://doi.org/10.1371/journal.pone.0078644>
- Zoncu, R., Sabatini, D., & Efeyan, A. (2011). mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol*, 12(1), 21–35. <https://doi.org/10.1038/nrm3025.mTOR>