T.C. DOKUZ EYLÜL UNIVERSITY IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

COMPREHENSIVE TRANSCRIPTOME ANALYSIS OF DIZYGOTIC TWINS AND THEIR PARENTS WITH AUTISM SPECTRUM DISORDERS

KAAN OKAY

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

İZMİR-2020

Thesis code: DEU.HSI.MSc/2017850050

T.C. DOKUZ EYLÜL UNIVERSITY IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

COMPREHENSIVE TRANSCRIPTOME ANALYSIS OF DIZYGOTIC TWINS AND THEIR PARENTS WITH AUTISM SPECTRUM DISORDERS

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

KAAN OKAY

Supervisor: Asst. Prof. Dr. YAVUZ OKTAY

Co-supervisor: Assoc. Prof. Dr. GÖKHAN KARAKÜLAH

(This research project was supported by Dokuz Eylül University Scientific Research Projects Coordination Unit with the project no: 310-SBKAEK)

Thesis code: DEU.HSI.MSc/2017850050

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Moleküler Biyoloji ve Genetik Yüksek Lisans Programı 2017850050 numaralı öğrencisi Kaan Okay, "**COMPREHENSIVE TRANSCRIPTOME ANALYSIS OF DIZYGOTIC TWINS AND THEIR PARENTS WITH AUTISM SPECTRUM DISORDERS**" konulu Yüksek Lisans tezini 27/10/2020 tarihinde başarılı olarak tamamlamıştır.

Dr. Öğr. Üyesi Yavuz Oktay

Dokuz Eylül Üniversitesi

BAŞKAN

Dr. Öğr. Üyesi Cihangir Yandım

İzmir Ekonomi Üniversitesi

ÜYE

Prof. Dr. Şermin Genç

Dokuz Eylül Üniversitesi

ÜYE

Dr. Öğr. Üyesi Zerrin Işık

Dokuz Eylül Üniversitesi

YEDEK ÜYE

Dr. Öğr. Üyesi Hani Alotaibi

Dokuz Eylül Üniversitesi

YEDEK ÜYE

Dokuz Eylül University İzmir International Biomedicine and Genome Institute Department Of Genomics and Molecular Biotechnology, Molecular Biology and Genetics graduate program Master of Science student Kaan Okay has succesfully completed his Master of Science thesis titled "**COMPREHENSIVE TRANSCRIPTOME ANALYSIS OF DIZYGOTIC TWINS AND THEIR PARENTS WITH AUTISM SPECTRUM DISORDERS**" on the date of 27/10/2020.

Asst. Prof. Dr. Yavuz Oktay

Dokuz Eylül University

CHAIR

Asst. Prof. Dr. Cihangir Yandım

İzmir University of Economics

MEMBER

Prof. Dr. Şermin Genç

Dokuz Eylül University

MEMBER

Asst. Prof. Dr. Zerrin Işık

Dokuz Eylül University

SUBSTITUTE MEMBER

Asst. Prof. Dr. Hani Alotaibi

Dokuz Eylül University

SUBSTITUTE MEMBER

TABLE OF CONTENTS **Page Number:**

INDEX OF TABLES

INDEX OF FIGURES

LIST OF ABBREVIATIONS

5’ UTR: 5’ Untranslated Region

ALA: Alpha-Linolenic Acid

AS: Alternative Splicing

ASD: Autism Spectrum Disorder

BAM: Binary Alignment Map

CARS: Childhood Autism Rating Scale

CE: Cassette Exon

DAS: Differentially Alternatively Spliced

DDI: Domain-Domain Interaction

DDI-net: Domain-Domain Interaction Network

DE: Differentially Expressed

DNE: Dominant-Negative Effect

DSM–5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition

DZ: Dizygotic

FPKM: Fragments Per Kilobase Per Million Mapped Reads

GFF: General Feature Format

GMT: Gene Matrix Transposed

GO: Gene Ontology

GTF: Gene Transfer Format

K-SADS-PL: The Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version

KSJ: Known Splice Junction

LCR: Low-Complexity Region

MXE: Mutually Exclusive Exon

MZ: Monozygotic

NSJ: Novel Splice Junction

PBMC: Peripheral Blood Mononuclear Cells

PCA: Principal Component Analysis

PDH: Pyruvate Dehydrogenase

PPI-net: Protein-Protein Interaction Network

RNA-seq: RNA-sequencing

RPKM: Reads Per Kilo Base Per Million Mapped Reads

SAM: Sequence Alignment Map

SCID-I: Structured Clinical Interview for DSM-IV Axis I Disorders

SFARI: The Simons Foundation Autism Research Initiative

TPM: Transcripts Per Million Reads

WGCNA: Weighted Gene Co-expression Network Analysis

ACKNOWLEDGMENTS

For their academic contributions to the completion of this thesis, I would like to extend my sincerest gratitude to the following individuals and organizations:

Asst. Prof. Dr. Yavuz Oktay, for his continued guidance and patience throughout the planning and conducting of this thesis, as well as for paving the way on the knowledge-base of medical genetics;

Assoc. Prof. Dr. Gökhan Karakülah, for his continued guidance and patience throughout the planning and conducting of this thesis, as well as his mentorship regarding the field of computational biology;

Dokuz Eylül University Scientific Research Projects Coordination Unit, for funding the studies that produced the initial data this thesis is based on (Project #s: 310-SBKAEK);

Oktay laboratory members; Ece Sönmezler Adalı, Tutku Yaraş, Aykut Kuruoğlu, Burcu Ekinci, past and present, for their invaluable support and friendship; Burcu Ekinci and Tutku Yaraş, for producing the RNA-seq samples and qPCR experiments;

Prof. Dr. Süha Miral and Dr. Pelin Ünal Varış, for data providing, collection, and project funding;

My family, for their unique belief and support of 25+ years; And all members of the İzmir International Biomedicine and Genome Institute , İzmir Biomedicine and Genome Center, and, past and present, for giving me the opportunity and the environment to conduct my studies.

COMPREHENSIVE TRANSCRIPTOME ANALYSIS OF DIZYGOTIC TWINS AND THEIR PARENTS WITH AUTISM SPECTRUM DISORDERS

Kaan Okay, İzmir International Biomedicine and Genome Institute, Dokuz Eylül University, Balcova, İzmir/Turkey

ABSTRACT

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social skills, repetitive behaviors, abnormality of speech, and nonverbal communication. ASD has high heritability, however, understanding the complexity of the underlying genetic basis has proven to be a challenging task. We hypothesized that dissecting the aberrations in alternative splicing (AS) and their effects on expression networks might provide insight. Therefore, we performed combinatorial transcriptome analysis, by AS and co-expression analysis of total RNA isolated from Peripheral Blood Mononuclear Cells (PBMCs). In order to minimize effects of genetic background, we employed a twin-based family design, where two pairs of non-syndromic autistic dizygotic (DZ) twins and their parents were analyzed. We identified 183 differential AS events in 146 genes, seven of them being Simons Foundation Autism Research Initiative (SFARI) Category 1-3 genes, namely CUL3, APH1A, PPFIA1, PTPRC, ARID2, PTK7, and DMXL2, strikingly the first three being previously reported to be alternatively-spliced in ASD post-mortem brains. Gene co-expression analysis identified 7 modules with 513 genes, 21 of which were SFARI genes and 5 were differentially alternatively spliced (DAS) genes. Five of these 21 SFARI genes were either Category 1 (ADNP and NR4A2) or Category 2 (DIP2A, PYHIN1, and RAB43). Among five DAS genes within the modules, ZNF322 and NR4A1 could be potentially interesting targets for further investigations, as they display differential AS, have a central role in their networks, and were previously implicated in ASD or brain development. Briefly, our results may be a complementary approach to the current omic-based approaches with practical benefits of the combined analysis of blood transcriptome.

Keywords: Autism Spectrum Disorder, RNA-sequencing, Differential alternative splicing, Co-expression networks.

OTİZM SPEKTRUM BOZUKLUĞUNA SAHİP ÇİFT YUMURTA İKİZLERİ VE AİLELERİNİN KAPSAMLI TRANSKRİPTOM ANALİZİ

Kaan Okay, İzmir Uluslararası Biyotıp ve Genom Enstitüsü, Dokuz Eylül Üniversitesi, Balçova, İzmir/Türkiye

ÖZET

Otizm spektrum bozukluğu (OSB) sosyal becerilerdeki eksiklikler, tekrarlayan davranışlar, konuşma anormalliği ve sözsüz iletişim ile karakterize edilen nörogelişimsel bir bozukluktur. OSB yüksek kalıtıma sahiptir, ancak altta yatan genetik temelin karmaşıklığını anlamak zor bir iştir. Alternatif uçbirleştirme (AU) 'deki anormalliklerin ve onların eş-ifade gen ağları üzerindeki etkilerinin bir bakış açısı sağlayabileceğini ve bu tür değişikliklerin bazılarının kanda tespit edilebileceğini varsaydık. Bu nedenle, Periferik Kan Mononükleer Hücrelerin (PKMH)' den izole edilen toplam RNA'nın hem AU hem de eş-ifade analizini gerçekleştirerek birleştirici transkriptom analizi gerçekleştirdik. Genetik arka planın etkilerini en aza indirgemek için, iki çift sendromik olmayan otistik dizygotik ikiz ve ebeveynlerinin analiz edildiği ikiz tabanlı bir aile tasarımı kullandık. 146 gende 183 tane AU olayı tanımladık ve bu genlerden yedisi (CUL3, APH1A, PPFIA1, PTPRC, ARID2, PTK7 ve DMXL2) Simons Foundation Autism Research Initiative (SFARI) kategori 1-3 genleriydi. CUL3, APH1A ve PPFIA1 genleri ölüm sonrası beyinlerde AU olaylarına sahipti. Üstelik 513 eş-ifadeli gene sahip yedi tane eş-ifade modulü tanımladık. Bu genlerden 21 tanesi SFARI ve 5 tanesi ayrımsal AU genleriydi. 21 SFARI geninden 5 tanesi kategori 1 (ADNP ve NR4A2) ve kategori 2 (DIP2A, PYHIN1 ve RAB43) genleriydi. Modullerdeki 5 adet ayrımsal AU geni arasında olan ZNF322 ve NR4A1 genleri ileriki araştırmalar için potansiyel ilginç hedefler olabilirler çünkü; bu genler AU olayları sergiledikleri için kendi ağlarında merkezi bir role sahip olabilirler, ayrıca önceden OSB veya beyin gelişimi ile de ilişkilendirilmişlerdir. Kısaca, sonuçlarımız kan transkriptomunun birleştirilmiş analizinin, pratik faydaları olan mevcut omik tabanlı yaklaşımlara tamamlayıcı bir yaklaşım olabileceğini göstermektedir.

Anahtar kelimeler: Otizm spekturum bozukluğu, RNA dizileme, Ayrımsal alternatif uçbirleştirme, Eş-ifade ağları.

1. INTRODUCTION AND AIM

1.1. Statement and importance of the problem

Autism spectrum disorders (ASD) is a genetically complex and heterogeneous group of early-onset neurodevelopmental disorders. Diagnostic features of ASD include aberrations in reciprocal social behavior, language development, and repetitive behaviors. More than a thousand genes are thought to play a role in its etiology, some of which are non-coding as well as coding. Despite large-scale efforts aimed towards understanding the molecular basis of ASD, utilizing genomic approaches including genome-wide association, whole-exome and whole genome sequencing of families, transcriptomics approaches including co-expression, total expression, alternative splicing, allel-specific expression by using RNA-sequencing and microarray techniques, its etiology has deeply been not understood.

Even though there are many family-based genomics studies related to ASD, including whole genome and whole exome analyses, family-based RNA-seq studies are far less common. These RNA-seq studies are generally based on one-way approaches (e.g., only total expression), and do not have combinatorial interpretations for different biological processes such as AS and co-expression. Moreover, findings of RNA-seq studies where co-expression and splicing alterations in ASD patients are analyzed are limited by the genetic background effects, as cases are compared to unrelated healthy controls. These studies also provided little information on the relationship between gene co-expression and AS events in ASD patient transcriptomes. Besides, twin studies of ASD generally focused on monozygotic (MZ) twins rather than DZ twins, and one of the reasons for this is the low-concordance of ASD between DZ twins (~0-20%), compared to that of MZ twins (~60-90%). Briefly, the combinatorial dizygotic twin-based AS and co-expression analyses may shed light on ASD etiology.

1.2. Aim of study

We carried out RNA-seq, followed by co-expression and AS analyses from PBMCs of autistic DZ twins and their parents to examine relationship between AS and co-expression in ASD, thus presented a combinatorial perspective on transcriptome of ASD.

1.3. Hypothesis of study

The hypothesis of this study is ASD-susceptibility and co-expression genes underwent AS events and those events are associated with co-expression phenomenon and ASD etiology. AS events and their relationship with co-expression may shed light on ASD etiology and these combinatorial transcriptome signatures might be used as blood-based biomarkers.

2. GENERAL INFORMATION

2.1. Autism spectrum disorder

Autism Spectrum Disorder (ASD) is a genetically complex and heterogeneous group of early-onset neurodevelopmental disorders. Diagnostic features of ASD include aberrations in reciprocal social behavior, language development, and repetitive behaviors (Lord, Rutter et al. 1994). ASD has a strong heritability, estimated to be between 50-90% (Sandin et al. 2014, Le Couteur et al. 1995, Ronald, Larsson et al. 2011). Despite large-scale efforts aimed towards understanding the molecular basis of ASD, genomic approaches including genome-wide association (Weiss et al. 2009, An et al. 2018, Grove et al. 2019), whole-exome, and whole-genome sequencing of families (Iossifov et al. 2012, O ' Roak et al. 2012, Sanders et al. 2012, Yu, Chahrour et al. 2013, Yuen, Merico et al. 2017, Callaghan, Rogic et al. 2019), transcriptome analyses by RNA-sequencing (RNA-seq) and microarray techniques (Lin, Chang et al. 2018, Balestrieri, Emanuela et al. 2019, Nord, Roeb et al. 2011, Ch'ng, Kwok et al. 2015), yielded limited progress over the years.

2.1.1. Epidemiology

In United States, prevelance of ASD is estimated 1/88 (Baio et al. 2012), but global prevelance of it is estimated 62/10000 and prevelance of it increases day by day (Elsabbagh, Divan et al. 2012). In Turkey, there is no precise information about prevelance of ASD. ASD four times more appears in males than females, but it is more severe in females than males (Werling, Daniel et al. 2013), this difference between two sex is explained with female protective model that put forwad to more etiologic load (genetic and environmental) is required for appearance of ASD in females than males (Robinson, Lichtenstein et al. 2013, Wing, Gould et al. 1979). Genetic studies have revealed that female sex features have protective factors againts hereditary and de novo mutations causing the risk of ASD and gonozomal choromosomes are important in appearance of sexual differences in ASD (Levy et al. 2011,

Baron-Cohen, Auyeung et al. 2015). Besides, several studies have shown that affected females have more mutation load than affected males, which consolidated female protective effect hypothesis (Jacquemont, Coe et al. 2014). In addition to, more severe intellectual disability appears in affected females rather than affected males.

2.1.2. Diagnostic evaluation

The diagnosis of ASD is entirely made with observing behavioral-finding by clinicians from the age of two because there is no any available biomarker, molecular diagnosis test, or brain imaging technic due to heterogeneity of ASD. Besides, the determination of other psychiatric disorders comorbid to ASD is important during the diagnosis.

The Autism Diagnostic Interview-Revised (ADI-R) (Baker et al. 2010) and The Autism Diagnostic Observation Schedule (ADOS) (Bender et al. 1947) are the most common diagnostic tools for ASD and considered the gold standard. ADI-R and ADOS are based on Diagnostic and Statistical Manual of Mental Disorders, fourth Edition (DSM-IV) (Bell et al. 1994) which is publication for classification of mental disorders and published by American Psychiatric Association, used by mental health professionals to define the properties of mental disorders. Besides, Childhood Autism Rating Scale (CARS) is one of the most used diagnostic and rating scales for ASD (Schopler, Reichler et al. 1980). Almost 70% of individuals with ASD have other psychiatric disorders including intellectual disability (~45%), attention deficit hyperactivity disorder (28-44%), anxiety, depression, and obsessive compulsive disorder. In addition to, medical conditions coming along with ASD are epilepsy (~30%), sleeping disorders (%50-80), and fragile X syndrome (Mukaddes et al. 2013, Postorino, Kerns et al. 2017). The evaluation of autistic features in animal models is difficult task because of complexity and heterogeneity nature of ASD (de la Torre-Ubieta, Won et al. 2016).

2.1.3. Treatment

The most effective approach in treatment is educational approaches and one of them is Applied Behaviour Analysis (ABA) is based on the science of learning and behavior, increase language, communication, attention, focus, social skills, and memory, decrease problematic behaviors (Autism Speaks, 2020, Fernandes, Amato et al. 2013).

There is no direct drug therapy for ASD and therapies are usually utilized for other psychiatric conditions comorbid to ASD. A study in England has shown that drugs including

sleeping pills (9.7%), psychostimulants (7.9%), and antipsychotics (7.3%) are prescribed to 29% of individuals with ASD (Murray, Hsia et al. 2014). A study in the United States has shown that at least one psychotropic drug including stimulants (13%), selective serotonin reuptake inhibitors (8%), or atypical antipsychotics (8%) is prescribed to 27% of children with ASD (Coury, Anagnostou et al. 2012).

2.1.4. Etiology

ASD is a complex, heterogenous, and polygenic disorder (King, Bearman et al. 2009). The understanding genetic background is needed for elucidating pathophysiology of ASD (Miles et al. 2011). The elucidating gene-environment interactions, as well as pure genetic approach, is important for understanding the etiopathology of ASD (Chaste, Leboyer et al. 2012). The genes associating with immune system and epigenetic regulation have misregulated expression profile in individual with ASD (Quesnel-Vallières, Weatheritt et al. 2019, Siu, Weksberg et al. 2017). Despite there are many studies associating with genetics, neurochemistry, and brain imaging of ASD, etiologic, biological, or behavioral markers specific to it has not been described yet (Tordjman, Somogyi et al. 2014).

2.1.4.1. Genetic factors

ASD is genetically heterogeneous and heritable (between 50-90%) and 80% of its genetic background is unclear, however, it's all variants are heritable, de novo variants contribute to its etiology (Sandin et al. 2014, Colvert, Tick et al. 2014, Hallmayer, Cleveland et al. 2011). The genetic background of ASD has been shown with both direct and indirect evidence. One of indirect evidence is twin studies in which even MZ twins have different ASD severity with each other, indicates associations between gene-environment and ASD heritability (Hallmayer, Cleveland et al. 2011), however, the indirect evidence has been revealed with GWAS, WGS, and WES studies (Grove et al. 2019, Iossifov et al. 2012, Callaghan, Rogic et al. 2019).

The concordance rate in MZ twins is 60-90%, however, this rate in DZ twins is 0-20% (Hallmayer, Cleveland et al. 2011). A population-based with 14516 children with ASD, the relative recurrence risk for ASD among MZ twins was estimated to be 153.0; for DZ twins, 8.2; for full-siblings, 10.3; for maternal half-siblings, 3.3; for paternal half siblings, 2.9; and for cousins, 2.02. Besides, a study has shown that males (47,5%) have a higher ASD recurrence

rate than females (44,3%) (Werling, Geschwind et al. 2015). Another study has revealed that ASD recurrence rate is between 10-15% when the existing only an affected sibling, however, when the existing two affected siblings, this rate is 50% and %12 in males and females, respectively (Wood et al. 2015). The recurrence rate in full siblings is two times than of half-siblings has been reported in another study (Constantino, Todorov et al. 2013).

2.1.4.2. Common and rare variants

The common and rare variants contribute to aetiology of ASD, those variants are single nucleotide polymorphisms. The common variants are estimated to have a higher frequency than 5% in the population; however, rare variants are estimated to have less than 1% frequency in the population (Pierre, Génin et al. 2014). Even though in comparison to the rare variants, the common variants have higher an effect on ASD liability, common individual variants are less effective than rare individual variants, which means that ASD liability depends on the polygenic burden level of many common variants (Grove et al. 2019, Stein, Neelroop et al. 2013). Both common and rare variants enrich in genomic coordinations in which there are genes expressed during corticogenesis (Grove et al. 2019). The rare variants contribute to deletions in the gene-rich regions, aberrant splicing, or truncated protein products, may responsible for 15%-20% of ASD cases and inherited two rare variants effecting on two copies of a protein-coding gene may increase the ASD risk (Stein, Neelroop et al. 2013).

2.1.4.3. Perinatal and environmental factors

Environmental factors are very important development of ASD via distruption of DNA-repair mechanism, epigenetics, or generation of de novo mutations (Kinney et al. 2010). The several environmental risk contributing to ASD have been reported in a meta-analysis study (Gardener, Spiegelman et al. 2009). One of these risks is parent age, mothers with over 35 years old have a 1.5 times higher risk of having children with ASD than younger mothers; however, this risk is higher for father age due to de novo mutation occurring in sperm DNA (Hultman, Sandin et al. 2011). The exposure of chemical components such as valproic acid, psychoactive drugs, and pesticides, as well as prenatal viral infections and high-level maternal stress in gestation have been associated with ASD (Grabrucker et al. 2013, Beversdorf, Manning et al. 2005). A meta-analysis study has revealed that there is no relationship between ASD and vaccination, as well as thimerosal exposure (Taylor, Swerdfeger et al. 2014). The using folic acid at before and early pregnancy is thought to be protective (Surén, Roth et al. 2013).

Additionally, maternal influenza infection was related with twofold increased prevalence of having a child with ASD, prolonged period of fever during pregnancy was associated with a threefold increased prevalence of autism in children, and use of various antibiotics are risk factors for ASD (Atladóttir, Hjördis et al. 2012).

2.1.4.4. Gen-environment relationships

The association between prenatal air pollution exposure and MET gene has been shown for increasing ASD risk in a study (Volk, Heather et al. 2014). The maternal smoking and some genetic variations in gestation have been associated with increasing ASD-related behaviours in children with ADHD (Nijmeijer, Judith et al. 2010). In another study, Cntnap2 gene and prenatal valproic acid exposure have been related to the social deficiency in Cntnap2 knock-out mice (Kim, Park et al. 2019).

2.1.4.5. Epigenetic factors

Epigenetic is an inheritable mechanism regulating the chromatin conformation, gene expression, and developmental processes without any alterations in the DNA sequence. The authors emphasized that epigenetic alterations in the individual with ASD may be due to the aging of the gametes or may occur in early embryonic life (Berko et al. 2014). Besides, a study suggested that epigenetic mechanisms can activate immune responses in gestation and also raise susceptibility to ASD, briefly maternal immune responses can cause epigenetic changes and contribute to ASD development (Atladóttir, Hjördis et al. 2012).

Mutations in genes involving in the epigenetic mechanism were reported in individual with ASD in previous studies. For instance, a de novo mutation in the HIST1H1E gene synthesizing H1 histone linker protein involving in the regulation of chromatin structures and gene transcription in a boy with ASD (Duffney et al. 2018). A review has revealed that 42 out of 215 genes associating with ASD are implicated in epigenetic regulation of gene expression. Those genes encode proteins that modify DNA/histones and regulate chromatin remodeling (Duffney et al. 2018). Several genome-scale studies revealed aberrations in DNA methylation in the brains of individuals with ASD (Ladd-Acosta, Hansen et al. 2014, Ellis, Gupta et al. 2017). A meta-analysis study in peripheral blood samples of ~800 autistic patients showed that 55 CpG sites were associated with ASD (Andrews, Sheppard et al. 2018). Another study

reported the spreading of H3K4 trimethylation marks in the brains of ASD individuals (Shulha, Cheung et al. 2012).

A study reported the common DNA acetylation patterns in the brain of ASD individuals (Sun et al. 2016). Additionally, rare genetic variants related to ASD implicate in chromatin remodeling (Lyall et al. 2017).

2.1.4.6. Neurobiological factors

Electrophysiological, functional, and structural neuroimaging studies proven that individuals with ASD have atypical neuronal connections (i.e., synaptic connections) rather than atypical brain regions. Decreased fronto-posterior and increased parieto-occipital connections, as well as dropped long connections and increased short connections, are proof of the atypical connections (Ha, Sohn et al. 2015). One of the common neuroanatomical features of ASD is brain volume growth between 6th and 24th months; however this growth tends to decreases with further ages (Courchesne, Campbell et al. 2011). Amygdala volume is larger in autistic children than typically developed children, but such a difference was not detected in adolescence. According to meta-analysis outcomes, life-span permanent neuroanatomical differences have been found in both the gray and white matter of autistic individuals compared to healthy controls (Via, Radua et al. 2011, Radua, Via et al. 2011). Decreasing the corpus callosum volume was reported in a study (Frazier, Antonio et al. 2009). Decreasing the cell amount and permanent nöroinflammation signatures were described in postmortem brain samples of autistic adolescences and adults. A study containing younger autistic individuals reported that there was an increasing number of neuron cells in the prefrontal cortex rather than decreasing (Courchesne, Mouton et al. 2011). Moreover, the differences in gastrointestinal microbiota between autistic and typically developed children have been shown in a study (Tomova et al. 2015).

2.1.4.7. The diagnosis methods for autism spectrum disorder

There are several methods used for the diagnosis of ASD. One of them is fluorescence in situ hybridization (FISH) in which fluorescent probes that bind to parts of a nucleic acid sequence with sequence complementarity are used. The second method is chromosomal microarray analysis (CMA) provides simultaneous genomic testing and detect chromosomal

anomalous with higher confidence compared to FISH (Miller et al. 2010). Copy number variants (CNVs) can be detected by using CMA (Ho et al. 2016).

NGS technologies such as WGS and WES can detect both CNVs related to other neurodevelopmental disorders sharing common genetic backgrounds with ASD, as well as ASD-specific CNVs; however, CMA provides the only detection of CNVs related to other neurodevelopmental disorders sharing common genetic backgrounds with ASD. The novel variants containing indels (i.e., insertion and deletion) as well as CNVs were found using WGS and WES technologies, and they are important for diagnosing and treating ASD (Ungar et al. 2015).

2.2. Alternative splicing

Alternative splicing (AS) of primary transcripts is a regulatory mechanism contributing to the increased complexity of higher eukaryotic organisms resulting in structurally and functionally distinct transcript isoforms and protein variants (Braunschweig, Gueroussov et al 2013, Chen, Manley et al. 2009, Kalsotra, Cooper et al. 2011). Over 95% of human multi-exonic genes undergo AS, and the resulting distinct splice isoforms are expressed at different levels in various cells and tissues (Pan, Shai et al. 2008, Wang, Sandberg et al. 2008). These splice isoforms are created through processes such as usage of alternative promoters or polyadenylation sites, intron retention, cassette exon (CE) inclusion/exclusion, usage of alternative 5' or 3' splice sites, and mutually exclusive exon (MXE) usage in which one of two exon groups is spliced- in, while the other group is spliced-out (Blencowe et al 2006, Pohl, Bortfeldt et al. 2013). These processes are regulated by the interactions of cis-acting motifs and trans-acting factors that control the assembly of spliceosomes, which is regulated by RNA-binding proteins (Chen, Manley et al. 2009, Wahl, Will et al. 2009). Furthermore, AS is known to be most prominent in the brain compared to other organs and thought to play important roles in nervous system development, neuronal differentiation, and maturation (Black, Grabowski et al. 2003, Mazin, Xiong et al. 2013, Weyn-Vanhentenryck, Feng et al. 2018).

There are several mechanisms causing aberrations in AS including misrecognition of 5' or 3' splice sites, non-preferential selection of splice site during the embryonic stage (the proportion of usage of a normally preferred splice site is altered), masking of small exonic sequences (~80% <200 bp) by large intronic sequences and the improper functioning of multiple splicing machinery in splicing process (Scotti, Swanson et al. 2016). Aberrations in

AS were associated with neurodevelopmental disorders (Morel, Thomas et al. 2006, Otomo, Kure et al. 2005, Arnold, Ling et al. 2013, Tian, Liao et al. 2011) including ASD (Voineagu, Wang et al. 2011, Gupta, Ellis et al. 2014, Stamova, Tian et al. 2013, Irimia, Weatheritt et al. 2014).

AS events causing functional domain alterations may also lead to disturbances in protein-protein interaction networks (PPI-nets) if the affected domains are involved in PPIs or domain-domain interactions (DDIs) (Mosca, Céol et al. 2014). Although 21306 differentially expressed (DE) exons and 211 differentially alternatively spliced (DAS) genes were described in autistic brains in the previous studies (Voineagu, Wang et al. 2011, Gupta, Ellis et al. 2014), DAS genes in ASD-associated modules and their relationship with PPI-nets are largely unknown.

2.3. Gene co-expression networks

In gene co-expression networks, each network has genes with similar expression patterns with each other. This similarity is determined by correlation or distance-based methods. The genes with similar expression patterns (i.e. correlated) can be assumed to have similar functions, involved in the same pathways, and regulated by common gene regulatory networks (i.e. transcription factors). Each node and edge in networks, represents genes and interactions between genes, respectively. The gene with highest connection (eg: edges with other genes) is named as the hub gene and it has highest degree within interest of network according to number of edges. There are two classes of co-expression networks called unweighted networks and weighted networks. In unweighted networks, the predicted connections with using correlation-based methods between the genes are identified in a straightforward-manner. The correlation scores range from -1 to 1 and those scores diverge from 0 (i.e. with a higher absolute value) indicates stronger negative or positive relationships between the genes, genes with absolute correlation scores higher than a cut-off value are considered as the connected. In comparison to unweighted correlation networks, weighted correlation networks do not implement a sticky correlation threshold to determine potential gene relationships. Such weighted correlation networks are based on methods similar to those used in hierarchical clustering (Fuller et al. 2011).

One of the software tools used for weighted correlation network construction and analysis is weighted gene co-expression network analysis (WGCNA), which allows users to

group genes into modules based on their co-expression. WGCNA, also known as weighted correlation network analysis, is a systems biology method using gene expression data, used to construct co-expression networks. WGCNA generates the correlation patterns among genes across samples/conditions with using Pearson or Spearman correlations. WGCNA uses hierarchical clustering to identify modules from the co-expression network using different tree-cutting thresholds by summarizing each module by the module eigengene or an intramodular hub gene. Eigengene helps associate modules by calculating module memberships and building their relationships with external sample traits (i.e. clinical phenotypes with numerical measurement), thus associations based on correlations between co-expressed genes and sample traits can be carried out in this way (Fuller et al. 2011).

After construction of co-expressed networks, Gene Ontology (GO) enrichment analysis, pathway enrichment analysis (eg: KEGG or Reactome), hub gene identification, and gene regulatory network analysis can be performed for more in-depth biological information (Fuller et al. 2011).

Detection of condition-specific alterations in expression patterns are generally based on differential expression (DE) analysis, but DE does not consider potential interacitons among genes and only considers individual gene behaviors across different conditions. To look with broader perspective to gene expression alterations, an approach called differential co-expression analysis have recently been appeared in molecular biology, which provides insights into the altered mechanisms between different conditions by analyzing the gene co-expression patterns. The main viewpoint of differential co-expression analysis is that co-expressed genes have different expression patterns between conditions because of many the reasons such as mutations and epigenetic alterations, which may resulted in aberrations in their interactions with other genes. Differential co-expression analysis can identify co-expressed genes, which undergo distortion during transition from control to disease conditions and can be used to identify candidate biomarkers, which exhibit distinct properties during progression of the diseases (Kakati, Tulika et al. 2019).

Co-expression analysis requires for normalized gene expression data because the aligment of transcripts to a relevant gene can be potentially misleading due to diverse technical and biological factors. One of these factors is sequencing depth of data (Liu et al. 2013, Kim et al. 2015), a uniform sequencing depth is preferable if the sequence data is produced during a single research, but it is not always possible, especially when working with public data. Besides, at even the single research, different samples can have different sequencing depth (i.e. read depth). In addition to sequencing depth, other factor is the length of the transcript, which affects the alignment of reads, with longer transcripts having a higher number of mapped reads compared to shorter transcripts, at even similar expression levels, causes differences between read counts of different transcripts across samples. Therefore, normalizasyon methods are used in minimizing of such techical/biological biases stemming from sequencing depths and/or differences in length of transcipts. Variables with different distribution across samples (i.e. genes) are scaled to the same range to reduce technical biases of RNA-seq and thus, samples are become more comparable according to samples with non-normalized data. Briefly, instead of using integer raw counts directly, using normalized expression units such as Reads Per Kilobase of transcript per Million reads mapped (RPKM), Fragments Per Kilobase of transcript per Million fragments mapped (FPKM), and Transcripts Per Million (TPM) should be utilized in RNA-seq studies as well as co-expression studies (Zhao, Zhan et al. 2020).

RPKM normalization method is calculated like below:

$$\mathrm{RPKM} = 10^9 * \frac{Reads\ mapped\ to\ the\ transcript}{Total\ reads * Transcript\ length}$$

RPKM normalization is not recommended for RNA-seq data anymore, instead, TPM normalization can be used for RNA-seq data (Wagner, Kin et al. 2012).

TPM normalizasyon is a slight modification of RPKM, is calculated like below:

$$\mathrm{TPM} = 10^6 * \frac{reads\ mapped\ to\ transcript\ /\ transcript\ length}{Sum\ (reads\ mapped\ to\ transcript\ /\ transcript\ length)}$$

TPM and RPKM are closely related, they can be converted to each other using the following formula:

$$\mathrm{TPM} = 10^6 * \frac{RPKM}{Sum\ (RPKM)}$$

After normalization, normalized gene expression data can be filtered to eliminate genes with low-expression across samples, which may cause statistical biases during construction of co-expression networks. Eventually, normalized and filtered expression data can be used to construct co-expression networks.

2.4. Dominant-negative effect

The dominant-negative effect (DNE) can be created through aberrations in AS events, where short an isoform of a gene with truncated or extended functional domain prevents the function of other isoforms of the same gene and such AS aberrations causing truncated domains and intron retention event were associated with DNE (Kelemen et al. 2012, Donaldson, Lucy et al. 2016). Moreover, DNE was associated with a neurodevelopment disorder (Sollis, Graham et al. 2016).

2.5. Next generation sequencing

In sequencing technologies, biomolecular components such as DNA and RNA are sequenced to elucidate their structure as well as interactions within cells. With this aim, various sequencing technologies were appeared in molecular biology, first of them is chain termination method known as Sanger sequencing which was developed by Frederick Sanger (Sanger Steven et al. 1977) and second is the chemical degradation method known as Maxam-Gilbert sequencing which was developed by Allan Maxam and Walter Gilbert (Maxam, Gilbert et al. 1980). Sanger sequencing is more preferable than Maxam-Gilbert sequencing because Maxam-Gilbert sequencing process requires various deleterious chemical components. Sanger sequencing is based on the DNA polymerase inhibitory activity of dideoxyribonucleotides (ddNTP) due to their chemical structures with lack of two OH- groups. When a ddNTP is added to a DNA chain instead of deoxyribonucleotides (dNTP), DNA polymerase does not continue to synthesize chain of interest and then DNA elongation process is halted. Each ddNTPs are labelled by radioactively or fluorescently for automated sequencing machines. After sequencing process, sequenced nucleotide chain having a ddNTP at the end of its sequence are examined

by using of gel electrophoresis to detect the position of sites where the ddNTP substitution took place.

Next Generation Sequencing (NGS) were developed in the early 21st century (Barba, Henryk et al. 2014). NGS provides high-throughtput sequencing and produces a parallel sequencing of multiple samples in one run and has four main steps, library preparation, amplification with PCR, sequencing, and data analysis. Despite NGS has similar working principle with Sanger sequencing, it is more sensitivity, faster, and cheaper to get reliable sequences results (Raza, Ahmad et al. 2016).

Sanger and Maxam-Gilbert sequencing technologies were named as the first generation of sequencing, but NGS technology was named as the second generation of sequencing. In today, due to similar reasons in the passed from Sanger sequencing to NGS, the third generation technologies are being developed. The third-generation sequencing compared to second generation sequencing can produce more reliable and accurate sequences results because of its long-read ability (>10,000 bp) (van Dijk et al. 2018, Lee et al. 2016).

2.5.1. NGS technologies and their applications

Depending on the biological question, data type, and chemistries, NGS technologies provide various methods such as whole genome, exome, targeted sequencing, metagenomics, RNA-seq, CHIP-seq, and methylome. Different NGS technologies and platforms have various advantages and disadvantages, which can be considered according to aim of study. There are various companies improving NGS technologies which use different sequencing machines such as Illumina and PacBio. For examples, Illumina sequencing machines provide varying levels of throughput, including the MiniSeq, MiSeq, NextSeq, NovaSeq and HiSeq models. The MiniSeq provides 7.5 Gb of information with 25 million reads/run in segments of 2×150 bp reads, the MiSeq can perform 2×300 bp reads, 25 million reads for an output of 15 Gb, the NextSeq can provide 120Gb with 400 million reads at 2×150bp read length (Slatko, Andrew et al. 2018).

Various NGS approaches available based on chemicals and methods, one of them is the sequencing-by-synthesis (SBS) method used by Illumina sequencing technologies. It similar to Sanger sequencing, it is based on use of ddNTPs, however, in contrast to Sanger sequencing, the ddNTPs used in SBS are not irreversible and bound to fluorescent labels unique for each base (Illumina, 2015).

There are other approaches called ion semiconductor sequencing and pyrosequencing, both of them based on the detection of the release of radicals from the added base during DNA synthesis. In ion semiconductor sequencing, added base to DNA strand releases hydrogen ions altering the pH of the solution which is recorded as a voltage change by an ion sensor. If no nucleotide is incorporated in DNA strand, no voltage spike occurs. In pyrosequencing, added base to DNA strand releases pyrophosphate ions (Slatko, Andrew et al. 2018).

The use of NGS in genome sequencing is variant identification in bulk tissues and individual cells of patients or populations and it can also be used to identify epigenetic modifications to a genome, namely sites of cytosine methylation and hemimethylation (Metzker et al. 2010).

One of the more common uses of NGS in recent years is RNA-seq that has the major advantage in transcriptome research, differential expression and splicing, co-expression network analysis, identification of transcription start sites, gene fusion detection. In splicing research, long-read lengths have a potential to identify known exon-boundaries as well as novel exon-boundaries. Known disease-causing mutations and their associations with aberrant splicing patterns have recently been discovered and therefore, understading aberrations in splicing phenomenon are crucial in explaining the underlying mechanism of complex diseases as well as mendelian diseases. In co-expression network analysis, genes existing in the same network can have similar functions and their expressions can be regulated the same transcription factors. The different co-expression patterns between two conditions can be used in elucidating disease mechanism. Additionally, RNA-seq is used in variant-calling analyses to corroborate variants detected in genome sequencing as well as discovering novel variants that are not detected in genome sequencing (Voineagu, Wang et al. 2011, Han et al. 2015, Ozsolak, Milos et al. 2011, Gonorazky et al. 2019, Stark, Grzelak et al. 2019).

2.6. Transcriptome research with Next Generation Sequencing

2.6.1. Raw data pre-processing and quality control

The raw read data is immediately not appropriate for alignment process. The presence of low-quality reads, base call error causing insertions and deletions, and adapter contamination are drawbacks for confidence alignment, thus raw data with such drawbacks should be underwent pre-processing step before the aligment. There are many quality-control checking

tools for raw reads such as NGS QC (Patel, Jain et al. 2012). The reads or bases with poor quality, as well as adapter sequence content introduced during library prepration steps, can be removed with using several tools such as Trimmomatic (Bolger, Lohse et al. 2014) and FASTX-Toolkit (Gordon, Hannon et al. 2010).

2.6.2. Alignment of sequencing reads

The raw RNA-seq reads can be aligned to a reference sequence, after the pre-processing steps. There are two reference sequences for alignment step: the first one is genomic sequences, the second one is transcriptome sequences. The selection of appropriate reference depends on the kind of data and aim of the study. The alignment of RNA reads to genome is carried out with splice-aware and gapped aligners, such as HISAT (Kim, Langmead et al. 2015), or STAR (Dobin et al. 2013). Different aligners have distinct accuracy rates to determine splice sites, polymorphisms, and these should be utilized in the research requires variant analysis/discovery or alternative splicing.

Reference transcriptome alignment is mostly faster and more accurate than reference genome alignment, but is only appropriate for known transcript analyses. Such alignments are not proposed for discovery of novel transcripts or splice junctions, and have little variant analysis power (Garber et al. 2011). When aligning to the transcriptome, ungapped aligners such as Bowtie are utilized (Langmead et al. 2009). Furthermore, multiple mapping reads are more common in transcriptome alignment than genome alignment.

2.6.3. Transcriptome assembly and novel splice junction discovery

The identifying transcripts with novel exons requires for transcriptome assembly. There are two main techniques for transcriptome assembly named “reference-free” (de novo) and “genome-guided”.

In both techniques, the reads obtained from RNA-seq, are assembled into transcripts via overlapping end regions. Reference-free assembly (i.e., de novo assembly) generally constructs the transcriptome based merely on the RNA reads by using “de Brujin graphs” to identify isoforms of genes and splicing sites (Robertson et al. 2010). This tecnique is calculating power density and has a higher error rate compared to genome-guided assemblies. The producing misleading initial results containing chimera genes, missing alleles, gene fragments by this tecnique require a deeper quality control (Cahais et al. 2012). It is often utilized for working

with organisms having unannotated genome or transcriptome. When working with annotated organisms, such as Drosophila melanogaster, it is preferable to utilize former sequence assemblies, both genome and transcriptome.

Genome-guided transcriptome assembly requires alignment files, with information on the mapping of reads to the genome. When assembling a transcriptome in this manner, the selection of aligner algorithm is critical. Mature transcripts existing in the cell's transcriptome do not have intronic sequences settled on the genome, and the 5' end and the 3' end of a single read can align to two bordering exons of a single intron. For accurate and confidence mapping of such reads to the genome, gapped aligners (i.e., splice-aware or gap-aware) recognizing potential splice sites to prevent large intronic sequences from creating false negatives when calculating the alignment scores (Garber et al. 2011) must be used.

Once the transcriptome is assembled, it can be compared with previously annotated assemblies to find known transcripts with annotated exons; however, it can also be compared with previously unannotated assemblies to find novel transcripts with unannotated exons (i.e., novel exons). There are several tools providing splicing analysis such as DEXSeq (Anders, Reyes et al. 2012) and JunctionSeq (Hartley, Mullikin et al. 2016), first is based on annotated transcriptome assemblies, and cannot directly identify novel exons and splice junctions, latter is particularly working with an incomplete or partly annotated transcript annotation in which all alternative isoforms for genes of interest were not included, based on generalized linear models, do not requires for an additional isoform assembly step, can directly identify novel exons, as well as novel splice junctions.

2.7. Transcriptome studies in ASD

There is extensive variation in transcriptome of ASD due to studies including different tissues (e.g., postmortem brain and blood) and its heterogeneity. Despite the hundreds, ASD-susceptibility genes have already been found; mutations account for less than 8% of ASD cases; therefore, transcriptomic studies are thought of as the complementary of genetic studies to elucidate the etiology of ASD. Moreover, transcriptome analysis is more efficient than genomic studies in identifying differences between ASD and healthy controls postulated in a study (Ansel et al. 2017). Transcriptomic studies related to ASD focus mainly on total gene expression, alternative splicing, and co-expression analyses. In total gene expression analysis, all transcripts of genes across samples are calculated to find total expression alterations in ASD

and control groups. In alternative splicing analysis, each exon count for all genes across samples is calculated to find exon expression alterations in ASD and control groups. In co-expression analysis, differences in gene modules in which genes are simultaneously expressed between ASD and control groups can be calculated to find aberrations in gene co-expression harmony.

2.7.1. Differential expression studies

There are many studies reporting differential gene expression analysis in various tissues of ASD. For examples, 593 downregulated and 145 up-regulated genes in peripheral blood lymphocytes of both children with autism and children of with older fathers; misexpressed two genes were reported in the lymphoblast cell line of 35 autistic individuals compared to controls; 31 dysregulated genes in whole blood between 37 autistic children and 15 controls; 2017 dysregulated genes in the dorsolateral prefrontal cortex of autistic individuals; misexpressed six genes in ileum and cecum of autistic children; one dysregulated gene in scalp hair follicles of autistic individuals; 156 genes in adult nasal olfactory stem cells that were DE in at least one ASD patient, of which 31 were dysregulated in more than a third of the cohort. There are many genes DE in the brain, intestinal, and blood samples of autistic individuals. For examples, ABHD3, CMKOR1, CX3CR1, CXCL10, CXCR4, GPR56, SLC9A9 gene were up-regulated in both blood and brain samples; ACTG2, ALAD, LAMP2 genes were down-regulated in both brain and intestinal tissues; UBD, IL2RA, CCL17 genes were up-regulated in both blood and intestinal tissues; ANKRD22, DRD4, FAM46C, HCK, IGHG1, NEURL3, PXDN, RELN, RPS21 gene were reported as the either up-regulated or down-regulated in the same tissues in different studies (Ansel et al. 2017).

The most enriched pathways of DE genes in different studies are cell cycle, cell death, gastrointestinal disease, immune function, neurogenesis, alternative splicing, arrhythmogenic right ventricular cardiomyopathy, cellular assembly and organization, cell-to-cell signaling and interaction, gap junction, inflammation, small molecule biochemistry, ubiquitin-mediated proteolysis, and Wnt and calcium pathways. For example, Wnt and calcium pathways are essential in the development of the nervous system and need to search for further investigation of ASD (Ansel et al. 2017).

There are many genes related to the etiology of ASD and can be roughly divided into a brain and immune system genes. For example of brain-related genes, DE genes in autistic-brain, protein ASS is up-regulated in autistic individuals, which controls nitric oxide (NO) production

through the regeneration of arginine from citrulline byproduct of the nitric oxide synthase (NOS) reaction. NO is a significant signaling molecule in the brain; thus it can play a role in ASD phenotype (Hu, Frank et al. 2006); A2A receptor-signaling pathway was determined as the top dysregulated pathology in the autistic brain, which is vital for neuronal stem cell proliferation, synaptic plasticity, motor function, cognition and emotion-related behaviors (Chow et al. 2012); decreased expression level of CNTNAP2 gene was reported in a study (Maekawa et al. 2015) and encoded the contactin associated protein-like 2, a neurexin protein that acts like neuronal adhesion molecule and receptor. Moreover, it is one of the neuronal targets of FOXP2 protein, and mutations of FOXP2 and CNTNAP2 were associated with language and speech disorders in ASD (Maekawa et al. 2015); down-regulation of MOCOS was found in ASD individuals, and its altered expression was associated with neurotransmission and synaptic defects. Besides, its expression plays a role in increased oxidative stress sensitivity (Féron, Gepner et al. 2016).

For example of immune-related genes; SH2D1B/EAT2 gene has been reported as a DE gene in ASD patients, is expressed in natural killer cells, B cells, and macrophages, suppress natural killer cell activity. Another DE gene is RUNX3, function in leukocytes has been associated with gastric mucosal hyperplasia, might be related to ASD because children with ASD appear to have gastrointestinal abnormalities (Gregg et al. 2008). Increasing the circulation of natural killer cells has already been observed in ASD patients instead of quiescent (Enstrom et al. 2009). Reducing expression level of NLGN3 and SHANK3 genes in lymphoblast cells has been observed in ASD patients (Yasuda et al. 2011).

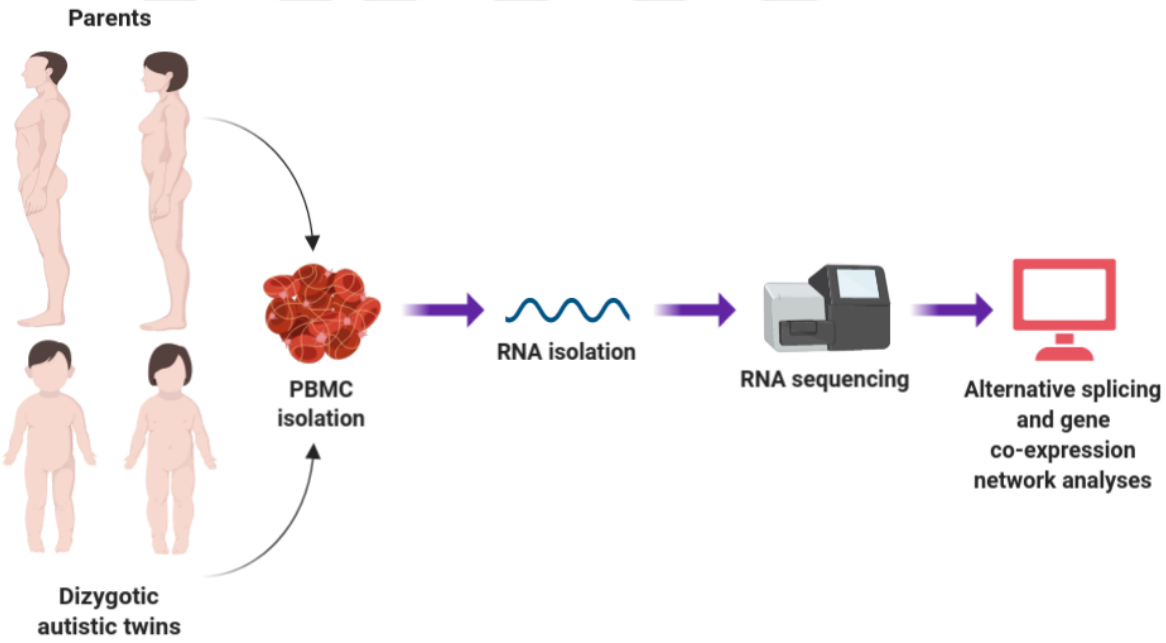
2.7.2. Alternative splicing studies

AS mechanism is abnormal in ASD patients. For instance, A2BP1/FOX1, a neural and muscle specific alternative splicing regulator was repoted as a down-regulated gene in ASD patients (Voineagu, Eapen et al. 2013). Many differentially alternatively spliced genes enriched for Gene Ontology terms such as “synaptogenesis”, “cell-junction”, and “actin-binding” had been identified in autistic brains (Voineagu, Wang et al. 2011, Gupta, Ellis et al. 2014).

2.7.3. Co-expression studies

Several ASD studies revealed that co-expression harmony of genes is disrupted in ASD patients, that is, when some genes co-express in healthy controls together, the same genes do not co-express in the ASD patients together and this may result in dysregulated biological pathways. Co-expression modules have been associated with several Gene Ontologies including Type-I interferon signalling pathway, defense response to virus, cytokine-mediated signalling pathway, regulation of cell proliferation, and inflammatory response (Voineagu, Wang et al. 2011, Gupta, Ellis et al. 2014). Overall, neuroimmunological mechanism may play a role in pathophysiology of ASD.

3. MATERIALS AND METHODS



The summary representation of materials and methods. (PBMC = Peripheral blood mononuclear cells).

3.1. Type of study

This study was conducted by using computational approaches.

3.2. Time and place of study

Computational analyses, literature screening, interpretation of results, and writing of this study was conducted at İzmir International Biomedicine and Genome Institute, between February 2019 and September 2020.

3.3. Materials of study

3.3.1. Participants

Ethics approval for this study was obtained from Dokuz Eylül University, Ethics Committee for Non-invasive Clinical Research (approval no. 310-SBKAEK 2016/03-24). The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. In two families (arbitrarily named F1 and F2), two pairs of DZ twins with ASD, aged 7-11 years, and their parents, followed at the Dokuz Eylül University Hospital Child and Adolescent Psychiatry Department outpatient clinic, participated in the study.

ASD diagnosis was determined based on the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM–5) criteria (American Psychiatric Association, 2013) and the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL) (Kaufman, Birmaher et al. 1997) was performed to detect other diagnoses comorbid with ASD. The Childhood Autism Rating Scale (CARS) (Schopler, Reichler et al. 1980) was applied to assess the diagnosis and severity of ASD. The presence of other psychiatric disorders in the parents was determined using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (First, Spitzer et al. 2002).

3.3.2. Sample collection and PBMC isolation

Blood samples were collected from twins and parents in lavender-cap (EDTA containing) blood collection tubes. To isolate PBMCs, blood samples were diluted 1:1 with 1X PBS in 15 ml falcon tubes. Diluted blood samples were spread out onto Lymphocyte Separation Medium (Lonza, #17-829E) at 1:1:1 (blood: PBS: Lymphocyte Separation Medium) ratio. Samples were centrifuged at 2200 rpm for 25 minutes by using acceleration and brake setting "0". Four layers were formed after the centrifugation and PBMCs were settled between plasma and Lymphocyte Separation Medium and those PBMCs were collected into a new falcon tube and washed twice with 10 mL of 1X PBS subsequently samples were centrifuged at 1500 rpm

for 10 minutes by using acceleration and break setting “9”. The supernatant was discarded, and the pellet was dissolved in TRI reagent (Sigma, #93289). Samples were stored at -80 oC until RNA isolation.

3.3.3. RNA isolation and preparation of strand-specific total RNA-seq library

PBMC samples in TRI reagent were mixed with chloroform (1:5 ratio). These samples were vortexed and incubated at RT for 2 minutes, subsequently centrifuged at 12000 rpm for 15 minutes at 4 oC. The top liquid part containing RNA was transferred into a clean microcentrifuge tube and mixed with 1.5 X (vol:vol) isopropanol. Then it was loaded on to Qiagen RNeasy Mini Kit columns (#74104) and followed by the manufacturer’s protocol. Eluted RNAs were immediately stored at -80 o C.

The RNA quality was measured with the Bioanalyzer 2100 system, and all samples were confirmed to have RIN > 9. For whole transcriptome sequencing, strand-specific libraries were generated by using Illumina’s TruSeq Stranded Total RNA LT Sample Prep Kit. Libraries were sequenced with Illumina HiSeq 2500 to obtain 2x150bp paired-end 50 million reads per sample.

3.4. Variables of study

The total transcript expression and exon expression values of genes are dependent on blood of DZ twins and their parents. Novel exons identified as the previously unannotated exons are also dependent variables.

3.5. Tools for data collection

All data collection and processing operations were carried out in Intel® Core™ i7-9750H CPU @ 2.60GHz × 12 Processor, 16 GB RAM, 2 TB hard disk, and 512 GB SSD computer with Ubuntu release 16.04 LTS operating system installed.

3.6. Study plan and calendar

3.7. Data evaluation

3.7.1. Quality control of RNA sequencing data and adapter sequence removal

The quality controls of the raw sequencing files in .FASTQ format were performed with FastQC (v0.11.8) (Andrews et al. 2010) to examine their base qualities and adaptor sequence contents.

3.7.2. Alignment of sequnce reads to human reference genome

The current human reference genome (GRCh38) and gene annotation file in the gene transfer format (GTF) were collected from the GENCODE website (https://www.gencodegenes.org/, release 30), and genome indexes were generated by using

splice-aware aligner tool HISAT2 (v2.1.0) (Kim, Langmead et al. 2015, Kim et al. 2019) with default parameters. With the use of these indexes, FASTQ files were aligned to reference genome through HISAT2 with default parameters to obtain Sequence Alignment Map (SAM) files; subsequently, those SAM files were converted to Binary Alignment Map (BAM) files (sorted by coordinate) through SAMtools (v1.9) (Li, Handsaker et al. 2009). Total and aligned reads of each sample are shown in (Figure 1).

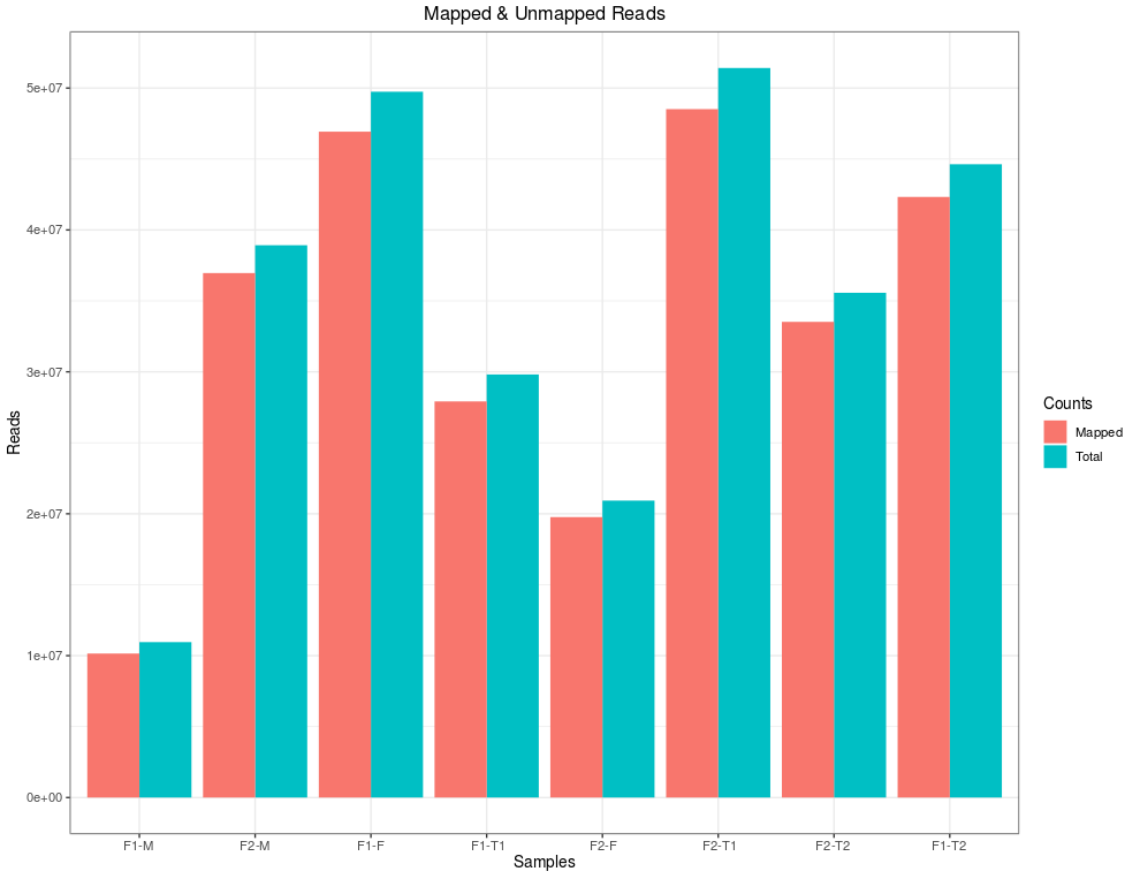


Figure 1. The distribution of total and genome-mapped reads across samples.

3.7.3. Exon-level quantification for differential alternative splicing analysis

Hartley's QoRTs/JunctionSeq pipeline (Hartley, Mullikin et al. 2016) was employed to perform differential AS analysis between autistic twins and parents. BAM files were processed by QoRTs (v1.3.6) with parameters including "--runFunctions", "--stranded", "--prefilterImproperPairs", and "--minMAPQ 30". We utilized the GENCODE annotation file in GTF format to generate a flat annotation file in the general feature format (.GFF) containing known and novel splice junctions through QoRTs's "makeFlatGff" and "--stranded"

parameters, then this flat annotation file and other files containing exon and splice junction counts across samples were transferred to R statistical computation environment (v3.6.2) for detection of DAS genes between twins and parents. Sashimi plot was created with ggsashimi command-line tool by using the BAM files (Garrido-Martín et al. 2018). The Gene Ontology (GO) and pathway enrichment analyses of DAS genes were performed with the R package clusterProfiler (v3.12) (Yu, Wang et al. 2012).

3.7.4. Gene-level quantification for weighted gene co-expression network analysis

Aligned sequencing reads were processed by using StringTie (v1.3.5) (Perteas et al. 2015) with parameters including "-e", "-B", "--rf", and "-A" to obtain expression values as the Transcripts Per Million (TPM) at the gene-level. We obtained 20787 genes (> 1 TPM), of which 13159 were protein-coding genes. In this step, the GENCODE annotation file in GTF format was used for the annotations of genes.

Non-coding genes were filtered out from downstream analysis steps, and only the protein-coding genes having expression > 1 TPM across samples were used for construction of co-expression modules through R package CEMiTool (Russo, Ferreira et al. 2018). The cemitool() function with "filter=T", "filter_pval=0.05", "cor_method=spearman", "cor_function=bicor", "network_type=signed", "tom_type=signed", "min_ngen=30", and "merge_similar=T" arguments were utilized to construct co-expression gene networks and the β value was set in 12. The files in the gene matrix transposed file format (GMT) were collected from the GSEA database (https://www.gsea-msigdb.org/gsea/index.jsp) for GO and pathway analyses of each module. The read_gmt() and mod_ora() functions were used for reading files in GMT format and running over-representation analysis (ORA), respectively.

3.7.5. Construction of protein-protein and domain-domain interaction networks

We downloaded only high confidence interactions (i.e., 325036 interactions of 19336 genes) from HitPredict database (http://www.hitpredict.org/, update 01 Aug 2019) to construct PPI-nets; however, all low confidence interactions were excluded. The Pfam domain-domain interaction networks (DDI-nets) information for DAS SFARI and module genes was obtained from the 3did database (https://3did.irbbarcelona.org/, accessed 20 Nov 2019) and visualization of those DDI-nets was created with Cytoscape (Michael, Keiichiro et al. 2011).

3.7.6. Hierarchical clustering and principal component analysis (PCA)

The log2 transformed values of DE exons were used to generate heatmap and PCA plots. Autistic twins and parents were separated with the PCA plot based on the expression values of these exons. Heatmaps were created using the R package pheatmap (Kolde et al. 2015) with the Euclidean distance and average linkage parameters.

3.8. Limitations of study

The number of participants is small; the findings are limited to the transcriptome of only two families, the findings are dependent on the algorithm of used software, and outcomes are needed to experimental validation to corroborate them and remove false-positive suspicion.

3.9. Ethics committee approval

Our analyses are a continuation of the previously approved autism study, where the whole genome and transcriptome analysis of parents and their autistic twins were performed. However, part of transcriptome analysis was performed only to a limited extent, without advanced analysis. We used the same data with more different transcriptome analysis tools and methods to perform a comprehensive analysis of the transcriptomic changes. Therefore we have taken a permission to use the existing data from the Dokuz Eylul University Clinical Studies Ethical Committee, without any new sample or data request. The pending approval states that the existing data was used in our proposed study titled “Comprehensive transcriptome analysis of dizygotic twins and their parents with autism spectrum disorders”.

Previous ethical committee approval date = 24.03.2016

Ethical committee approval number = 310-SBKAEK

Decision number= 2016/06-01

4. RESULTS

4.1. Identification of differential AS events in twins with ASD

We isolated PBMCs from DZ twins with non-syndromic ASD and their parents and carried out RNA-seq to identify AS alterations at both gene and network-level (Figure 2). Our analysis revealed 146 genes having 183 differential AS events and 103, 74, and 6 of these events

corresponded to CEs, KSJs, and NSJs, respectively (Table 1). Some exons of KSJs and NSJs that were used in a mutually exclusive manner (i.e., both exons in an exon cluster do not co-exist in a mature transcript, as one of two exons is spliced in, while other is spliced out) between twins and parents, are identified as the "MXEs" (Pohl, Bortfeldt et al. 2013, Hatje, Rahman et al. 2017). Unsupervised hierarchical clustering and PCA analyses based on log2 transformed expression levels of differential AS events were performed, and good separation of twins from parents was observed (Figure 3A-B). Moreover, differential AS events were highly correlated among autistic twins (Figure 3C). GO and pathway enrichment analyses of 146 DAS genes revealed that there were statistically significant enrichments for "cellular component" terms including focal adhesion (GO:0005925; p-adjusted = 0.034), cell-substrate adherens junction (GO:0005924; p-adjusted = 0.034), and cell-substrate junction (GO:0030055; p-adjusted = 0.034) (Table 2, Figure 4). We did other enrichment analyses including biological process, molecular function, and KEGG pathway using the Enrichr tool (https://maayanlab.cloud/Enrichr/) owing to clusterProfiler found no significant enrichments for them, and found that 146 DAS genes enriched for; KEGG pathways including hedgehog signaling pathway (CUL3, CUL1, and CSNK1D), hippo signaling pathway (PPP2R2D, CSNK1D, DCHS1, and ACTB), dopaminergic synapse (PPP3CA, PPP2R2D, and PPP2R5C), and synaptic vesicle cycle (ATP6V1G1 and ATP6V1H); biological processes including protein retention in endoplasmic reticulum (ER) lumen (OS9), morphogenesis of a polarized epithelium (PTK7 and CELSR1), actin crosslink formation (FLNA), and amyloid precursor protein catabolic process (APH1A); molecular functions including coreceptor activity involved in Wnt signaling pathway, planar cell polarity pathway (PTK7) and cadherin binding (FLNA, RPS26, and PFN1). Besides, four of these 146 genes (i.e., OS9, CAST, MYO15B, and SCML4) overlapped with 725 genes that were identified as the DAS genes between the blood of ASD and typically developed children (Table 1) (Stamova, Tian et al. 2013).

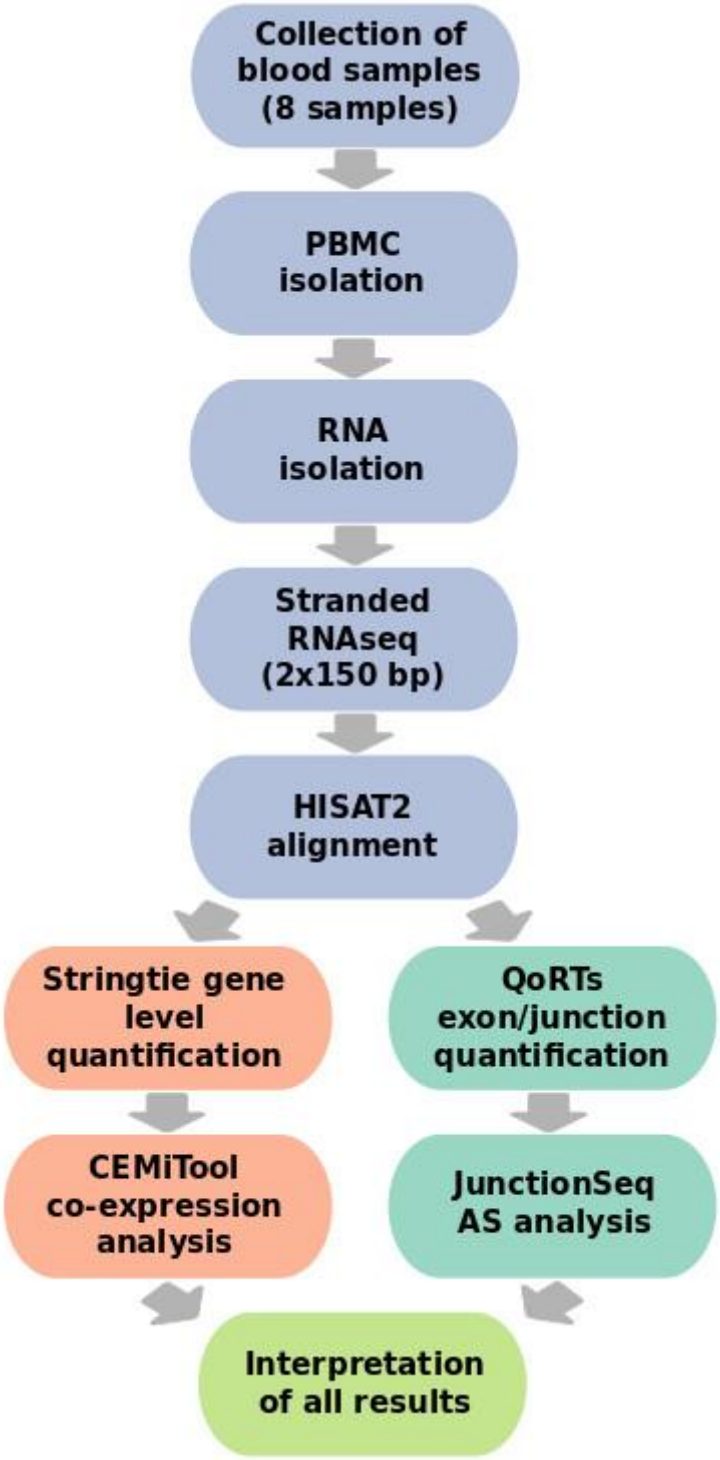


Figure 2. Flowchart showing detail of data preparation, processing, analysis, and interpretation.

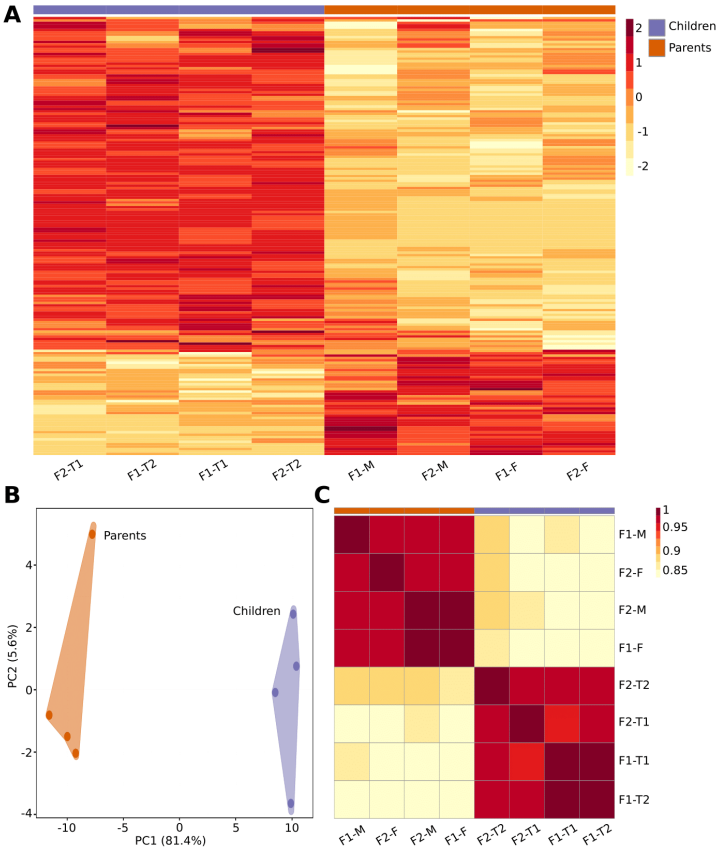


Figure 3. Differentially expressed exons/junctions of 146 genes. (A) Unsupervised hierarchical clustering of DE 183 exons/junctions. Each row represents a single exon/junction and each column represent an individual (F1: Family-1; F1-M= Mother in Family-1, F1-F= Father in Family-1, F1-T1= Twin-1 in Family-1, F1-T2= Twin 2 in Family-1; F2: Family-2; F2-M= Mother in Family-2, F2-F= Father in Family-2, F2-T1= Twin-1 in Family-2, F2-T2= Twin-2 in Family-2). High expression is shown in red color and low expression is shown in yellow color (FDR <0.05). (B) Principal component analysis (PCA) based on DE exons/junctions. PCA plot shows the separation of the twins from the parents. The two principal components accounted for 87% of the total variance. (C) The correlation heatmap showing Pearson correlations of DE exons/junctions across individuals. High-correlation is shown in red color and low-correlation is shown in yellow color. Each row and column represents individuals.

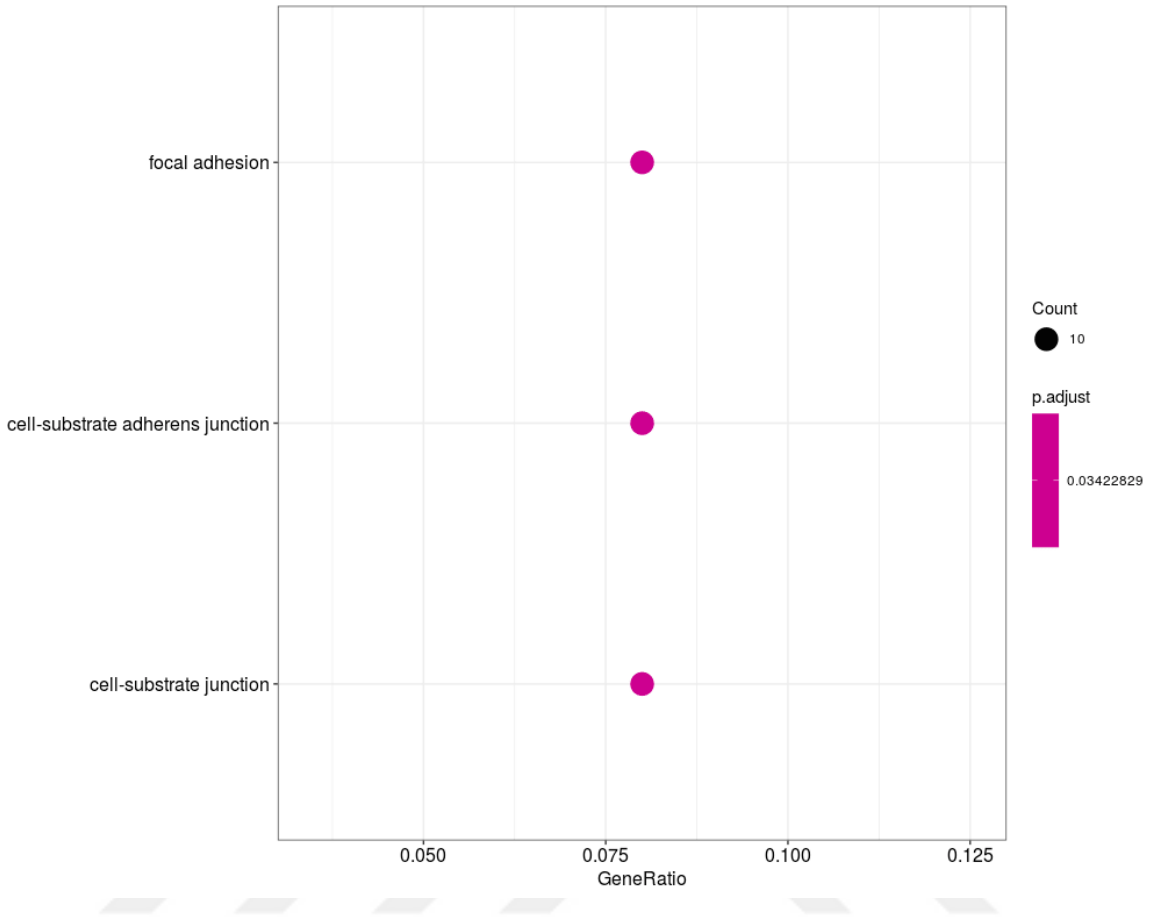


Figure 4. ClusterProfiler enrichment results of differentially spliced genes.

4.2. Differential AS events that affect SFARI genes

To explore whether SFARI genes had any differential AS event between autistic twins and parents, we inquired 831 genes from SFARI database (https://www.sfari.org/: release 01 Nov 2019) (genes in categories: S = "syndromic"; 1 = "high confidence"; 2 = "strong candidate"; 3 = "suggestive evidence"; 3S = "suggestive evidence, syndromic" were included) and found out that APH1A (Entrez ID: 51107), PTPRC (Entrez ID: 5788), CUL3 (Entrez ID: 8452), ARID2 (Entrez ID: 196528), PTK7 (Entrez ID: 5754), DMXL2 (Entrez ID: 23312), and PPFIA1 (Entrez ID: 8500) genes had differential AS events (Figure 5A), all of which occurred in coding-exons. Some of the enriched GO "cellular component" terms included presynapse (APH1A, DMXL2, and PPFIA1), synaptic vesicle (APH1A and DMXL2), and focal adhesion (PTPRC, PTK7, and PPFIA1) (Table 2). Aberrations in these cellular components are well-known to be associated with ASD (Waites, Garner et al. 2011, Fassio, Patry et al. 2011, Stewart et al. 2015).

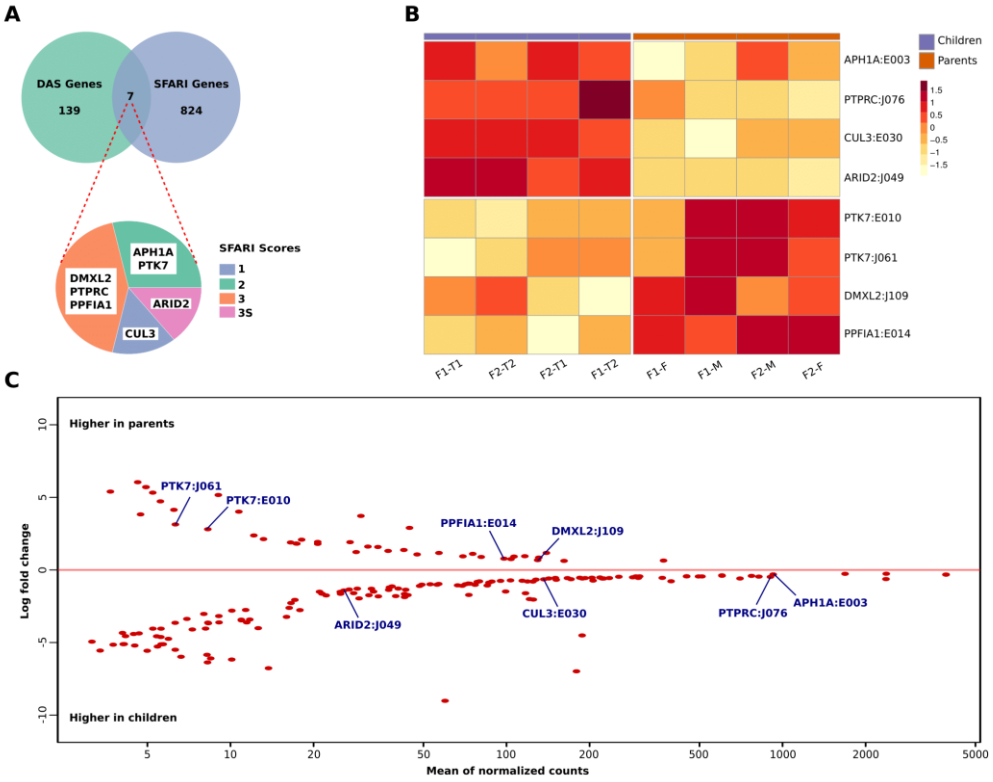


Figure 5. Differentially alternatively spliced (DAS) SFARI genes. (A) Venn diagram showing 7 DAS SFARI genes and their scores based on scoring-criteria of SFARI database (https://gene.sfari.org/). (B) Heatmap plot showing the expression level of each exon/junction of SFARI genes. Each row represents a single exon/junction and each column represents an individual. High expression level is shown in red color and low expression level is shown in yellow color. (C) MA plot showing exons/junctions of each SFARI gene. Each dot represents a single exon/junction of the genes.

Comparing the expression of these genes at exon-level among twins and parents revealed that PTK7, DMXL2, and PPFIA1 had lower exon expression levels in twins. In contrast, APH1A, PTPRC, CUL3, ARID2 had higher exon expression levels. Four of the differential splicing events affecting these SFARI genes were CE usage, while the other four were KSJ usage (Figure 5B-C). Although all eight splicing events affect coding exons, only exons of CUL3 and PTK7 encode functional Pfam domains, namely Cullin (PF00888.22) and Ig_3 (PF13927.6), respectively. Furthermore, APH1A, CUL3, and PPFIA1 genes were previously identified among the DAS genes between post-mortem autistic and healthy brain samples in a study (Gupta, Ellis et al. 2014) and the exon 7 (791 bp) of APH1A encoding non-

functional protein and 3' UTR sequences had differential usage in both respective and our study (Table 1).

4.3. Differential AS events creating novel isoforms and their associations with ASD

A closer examination of the 146 DAS genes showed that TRAF3IP3, CSGALNACT2, PFN1, COL6A2, AC139495.3, and FLNA genes had NSJs (Figure 6), that is, they had novel isoforms expressed. The NSJs of these genes had higher expression levels in twins compared to the parents, except for the NSJ of the AC139495.3 gene. Moreover, FLNA, PFN1, and TRAF3IP3 exons that were present in novel isoforms were utilized in a mutually exclusive manner in parents, but not in twins.

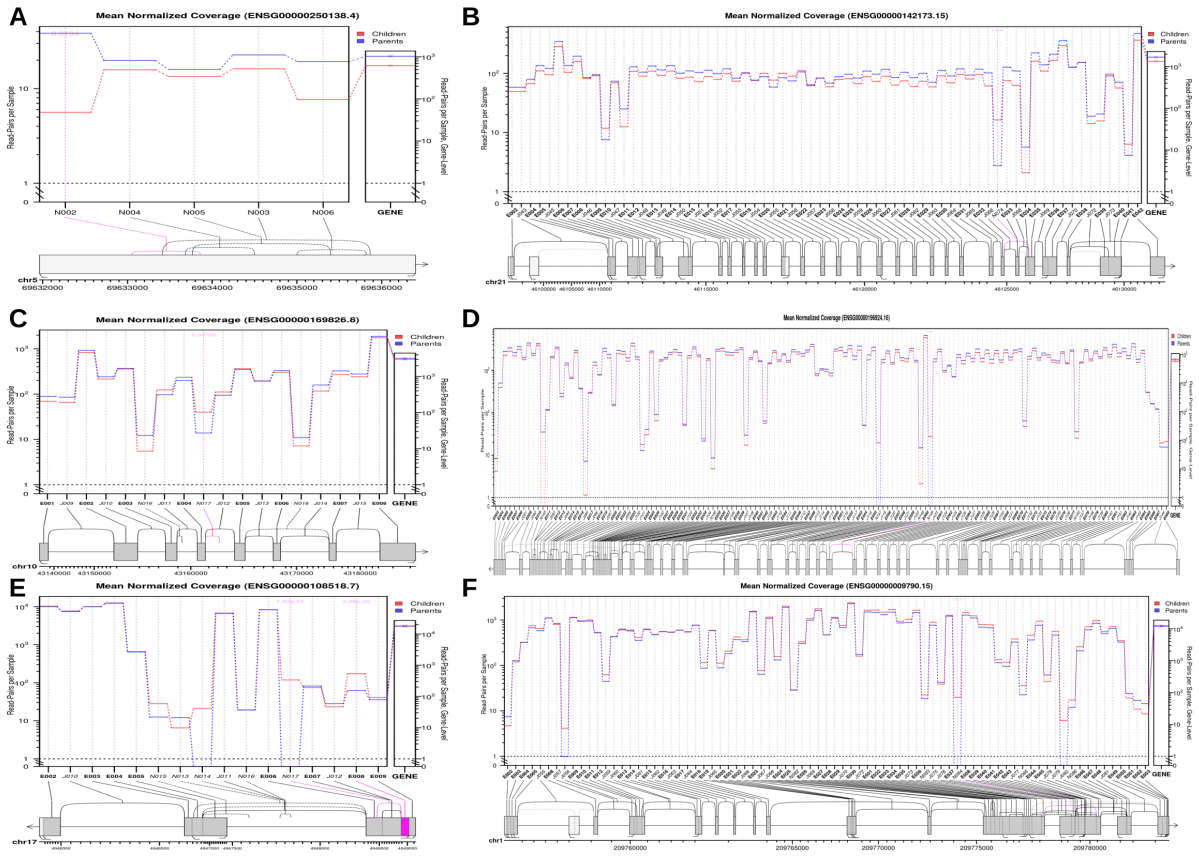


Figure 6. JunctionSeq plots showing differentially expressed novel isoforms. (A) AC139495.2 gene. (B) COL6A2 gene. (C) CSGALNACT2 gene. (D) FLNA gene. (E) PFN1 gene. (F) TRAF3IP3 gene.

4.4. Gene co-expression network analysis

In addition to the gene-centric approach, we also employed a network-based approach to analyzing the AS events. Only protein-coding genes were included, and all non-coding genes

were excluded from further analysis to construct the co-expression networks. Based on the gene expression levels across samples, we obtained eight co-expression network modules in total. Seven of the eight modules had a total of statistically correlated 513 co-expressed genes (Table 3) and the size of those modules ranged in size from 37 to 141 genes (Figure 7A). Enrichment analysis was performed to identify Reactome pathways that were overrepresented in each module (Figure 8A-H) and top enrichments are shown in (Figure 7B). M3 and M6 modules were enriched with terms that are associated with ASD. M3 module was enriched for interferon signaling, cytokine signaling in the immune system, and alpha-linolenic acid (ALA) metabolism.

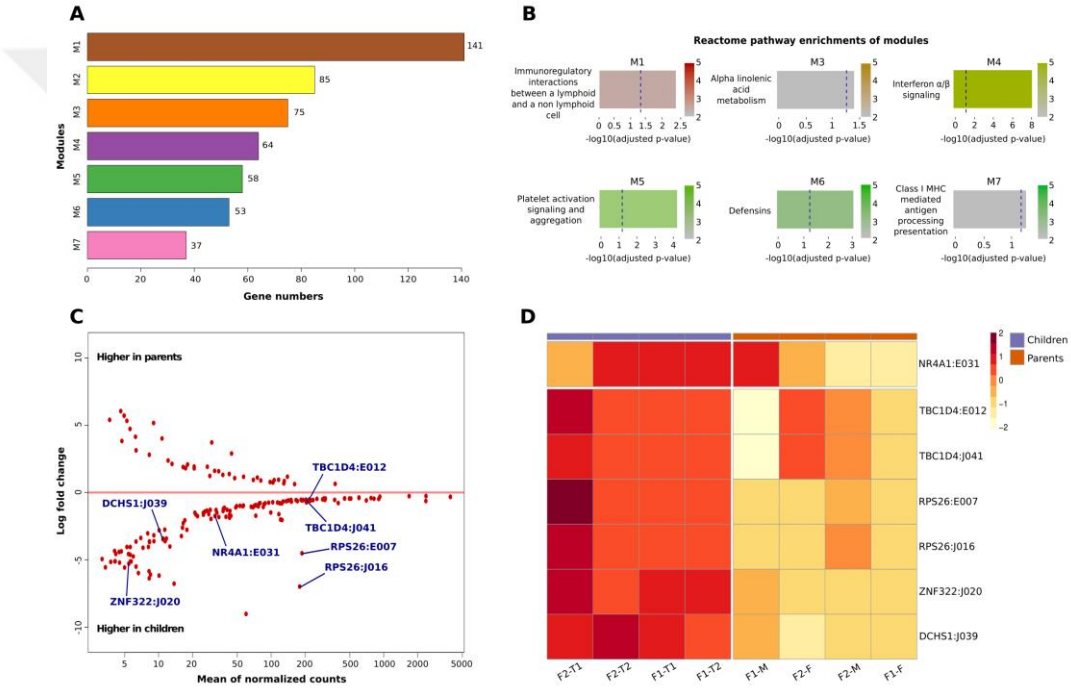


Figure 7. Co-expression network analysis and DAS module genes. (A) The horizontal color bar plot showing gene numbers in each module. (B) ORAs of six modules. Each horizontal color bar represents the most statistically significant Reactome pathways. X-axis represents the False Discovery Rate (FDR). Each vertical color bar on the y-axis represents degree of FDR value for module enrichments. Each vertical blue dashed line corresponds to FDR cut-offs (≤ 0.05). (C) MA plot showing exons/junctions of module genes. Each dot represents a single exon/junction. (D) Heatmap plot showing expression levels of exon/junction of module genes. Each row represents a single exon/junction and columns represent individuals. High expression level is shown in red color and low expression level is shown in yellow color.

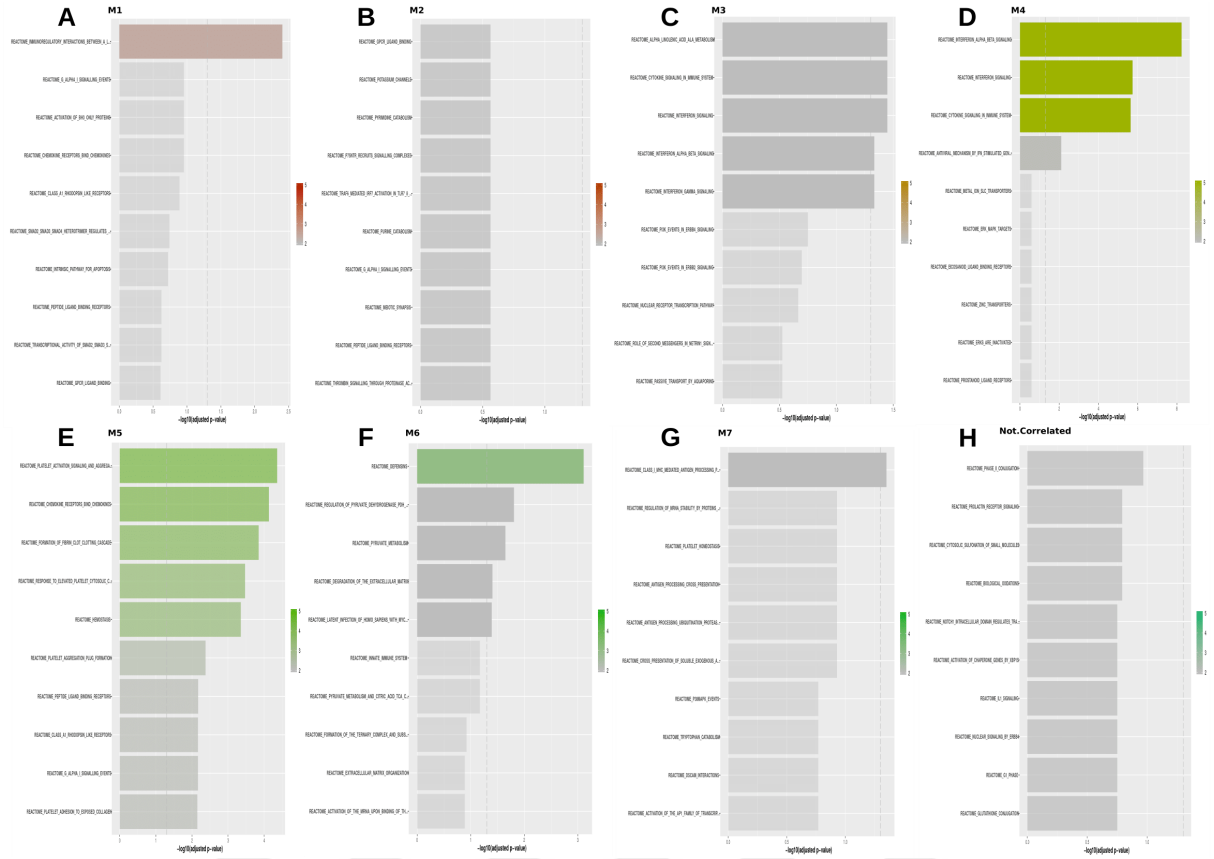


Figure 8. Over-representation analyses of eight modules. Each horizontal colorful bar represents Reactome pathways. Each x-axis represents the False Discovery Rate (FDR). Each colorful vertical bar on the y-axis represents FDR degrees of enrichments of modules. Each vertical gray dashed line corresponds to FDR-cut-off ($\leq$ 0.05).

On the other hand, the M6 module was enriched for defensins, regulation of pyruvate dehydrogenase (PDH) complex, and pyruvate metabolism. M6 module activity was higher in twins compared to the parents (Figure 9).

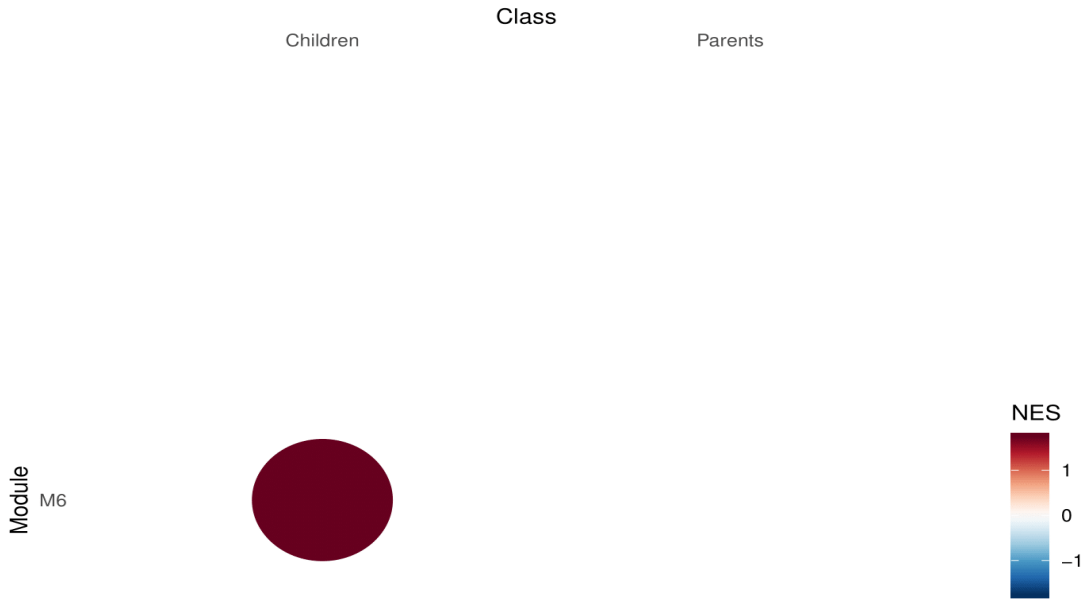


Figure 9. Gene set enrichment analysis showing the M6 module activity at children and parents groups. Symbol size and color reflect normalized enrichment score (NES).

Moreover, 21 SFARI genes were distributed among all modules, except the M7 module, and most of them (16 genes) belong to the SFARI Category 3 (Table 3). Of note, two of them are Category 1 genes (ADNP and NR4A2), and three are Category 2 genes (DIP2A, PYHIN1, and RAB43). Also, ADNP in the M2 module and USP9Y (SFARI Category 3) in M6 module genes were hub genes in their respective modules (Figure 10).

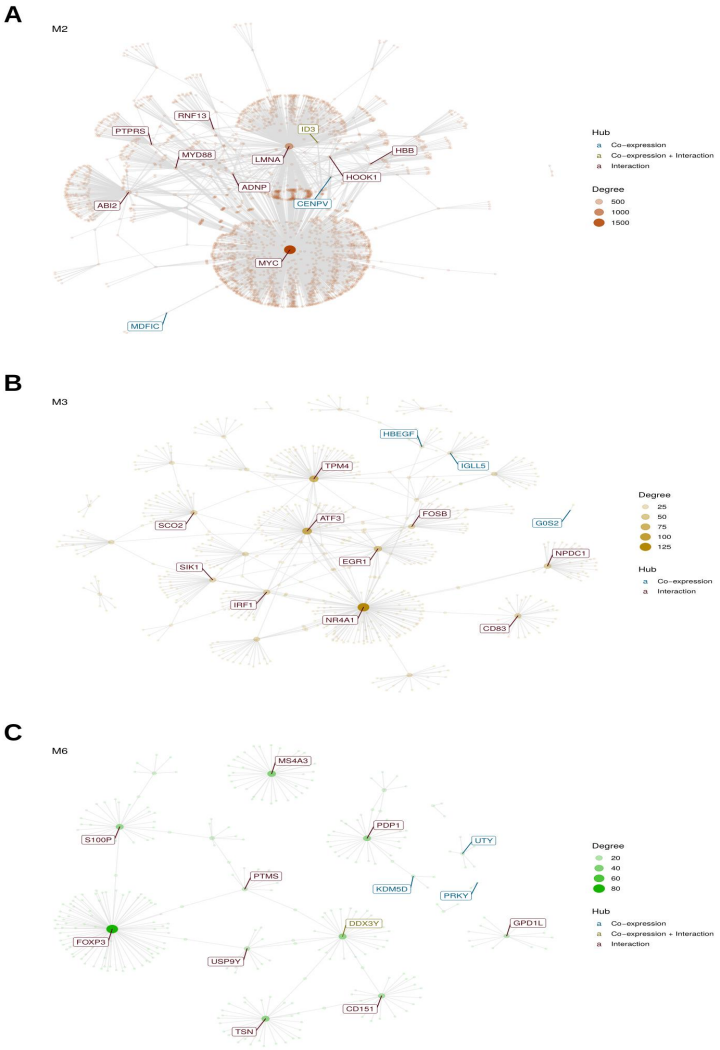


Figure 10. Interaction plots for modules, with hub genes highlighted. (A) M2 module. (B) M3 module. (C) M6 module.

4.5. Differential AS events that affect co-expressed genes

AS analysis of co-expression module genes revealed that ZNF322 (Entrez ID: 79692), DCHS1 (Entrez ID: 8642), NR4A1 (Entrez ID: 3164), TBC1D4 (Entrez ID: 9882), and RPS26 (Entrez ID: 6231) had differential AS events between autistic twins and parents. Moreover, the exon/junction expression levels of all these genes were higher in twins compared to the parents (Figure 7C).

ZNF322 and DCHS1 in the M1 module had differential KSJ usage, while NR4A1 in the M3 module had differential CE usage; however, TBC1D4 and RPS26 in M6 module had both CE and KSJ usage (Figure 7D). Furthermore, the differentially used exons of DCHS1, RPS26,

and TBC1D4 encode Cadherin (PF00028), S26e (PF01283), and DUF3350 (PF11830) Pfam domains, respectively.

In our analysis, ZNF322 displayed differential splice junction usage with the second (89 bp) and the third (70 bp) exons, encoding the 5' Untranslated Region (5'UTR). Twins did not utilize these 5'UTR-coding exons in a mutually exclusive manner, in contrast to parents (Figure 11A-B).

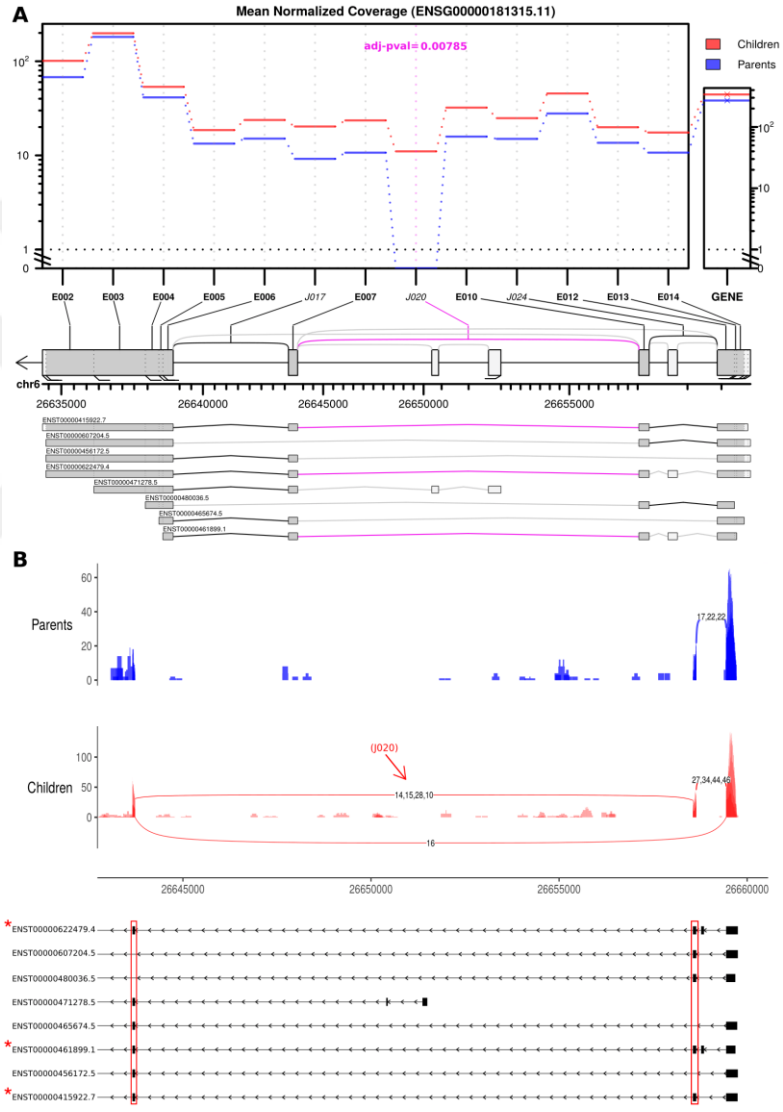


Figure 11. Differentially expressed isoforms of the ZNF322 gene. (A) JunctionSeq plot showing differentially used splice junction J020 (pink line). (E=exon, J=junction). (B) Sashimi plot summarizing differential usage of splice junction J020. The exons of splice junction J020 is highlighted with red boxes. Each red asterisk represents transcripts having differentially used splice junction J020. (Mean coverage cut-off set in 10).

Other interesting DAS module genes were NR4A1 and DCHS1. NR4A1, which was one of the hub genes of the M3 module (Figure 10). Moreover, NR4A1 is a paralog of NR4A2, an SFARI Category 1 gene, also present in the M3 module.

4.6. Differential AS events and their possible effects at domain-level

Aberrations in AS mechanisms may contribute to alterations in PPI-nets and DNEs via truncated or extended functional domains (Kelemen et al. 2013, Donaldson, Lucy et al. 2016, Yang, Coulombe-Huntington et al. 2016). Especially, PPI-nets rely on interactions between functional domains, and AS events could modulate them through the alternative inclusion or exclusion of domains participating in DDIs (Yang, Coulombe-Huntington et al. 2016); therefore, we focused on the possible effect of DAS SFARI and co-expressed genes on the PPI-nets. To this end, DDI information of the DE exons encoding functional domains of these genes was obtained from 3did database, except for the 14th exon of TBC1D4 encoding DUF3350 domain due to its unknown function. Cullin, Ig_3, Cadherin, and Ribosomal S26e Pfam domains had DDIs (Figure 12), involving in ubiquitin-dependent protein catabolic process, immune response, homophilic cell adhesion via plasma membrane adhesion molecules, and translation, respectively.

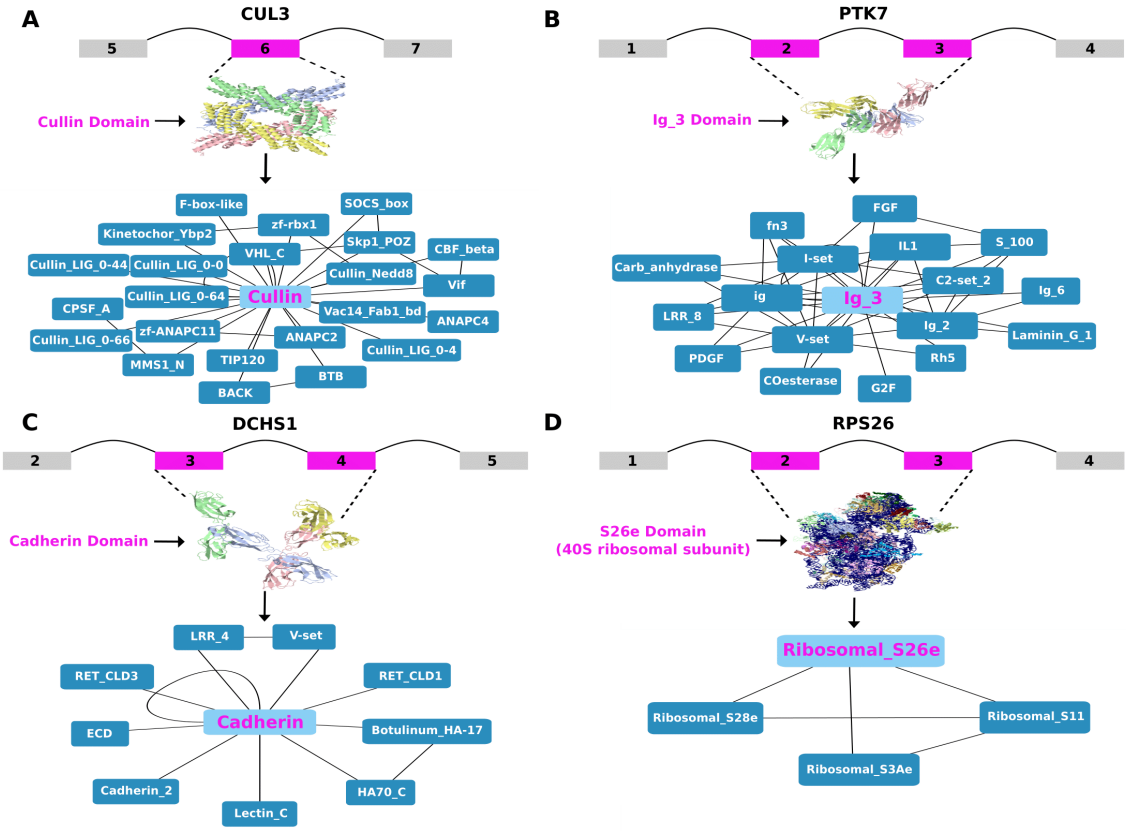


Figure 12. Differentially expressed exons/junctions of SFARI and module genes and their domain domain interaction networks. (A-B) The representation of DDIs of SFARI genes. (C-D) The representation of DDIs of module genes. Each pink box represents statistically significant differentially used exons (FDR<0.05) and each grey box represents neighbor exons (FDR>0.05). The canonical transcript of each gene was used to obtain domain information of exons/junctions. Information of domain-domain interaction networks and 3D protein structure images obtained from 3did database.

5. DISCUSSION

By performing AS analysis of blood transcriptomes of autistic DZ twins and their parents, we found 183 differential AS events in 146 genes, some of which had already been strongly linked to ASD. The enrichment results of these genes for ASD-related GO terms (e.g., presynapse, synaptic vesicle, and focal adhesion) and pathways (e.g., hippo and hedgehog) suggest that the splicing mechanisms might be important for differential synaptic transmission, and neuronal development between twins and parents. Nevertheless, when we did closer examination for some genes, realized that OS9, CAST, MYO15B, and SCML4 DAS genes were among the previously identified DAS genes in ASD blood samples. When we used Enrichr

tool to find out enrichment of GO terms among these four genes, only OS9 enriched for "protein retention in endoplasmic reticulum (ER) lumen" GO biological process term, "ER lumen" GO cellular component term (Table 2), and "protein processing in ER" Reactome pathway. Aberrations in ER-related protein processing had been reported in ASD (Comoletti et al. 2004). The protein traffic in the ER is critical for initiating the formation and maturation of synaptic junctions. Given that ASD is a synaptic disease, differential splicing pattern of ER-related genes may result in differentiation of synaptic junctions between children and parents. In addition to ER-related genes, seven DAS genes were SFARI Category 1-3 genes, namely CUL3, APH1A, PPFIA1, PTPRC, ARID2, PTK7, and DMXL2; the first three had previously been reported being alternatively-spliced in ASD post-mortem brains. These genes were enriched for ASD-related synaptic cellular component terms. Besides, we observed a splicing event of the APH1A gene, which was reported in another ASD study using brain samples. Taken together, synapse and focal adhesion-related genes may have a role in the diversification of synaptic transmission and plasticity through their differential splicing events causing changes in both non-coding sequences (i.e., UTRs) and coding sequences (i.e., functional protein domains), and several of these events might be detected from both blood and brain samples of autistic children.

According to their predicted effects on protein domains, the novel isoforms of PFN1, COL6A2, CSGALNACT2, and FLNA may exert their effects through encoded functional domains. Because splice junctions of novel isoforms have the potential to result in altered functional domains (truncated or extended) or low-complexity regions (LCRs), proteins with such alterations may show dominant-negative effects (DNEs) on interacting proteins and PPI-nets. Therefore, these ASD-associated genes may contribute to the etiology of ASD through DNEs and altered PPIs. In addition to the exons encoding functional domains and LCRs, exons encoding UTR sequences of novel isoforms may also potentially affect PPI-nets through post-transcriptional gene regulation mechanisms (Mignone, Gissi et al. 2002). Among those genes with novel isoforms, TRAF3IP3, a transcription factor that may mediate TRAF3-induced JNK activation, was previously reported to be DE between ASD and healthy control blood samples (Kunkel, Kohane et al. 2011). Alterations in splicing and expression of CSGALNACT2, whose protein product is an enzyme required for chondroitin sulfate synthesis, were previously reported in ASD and Tourette syndrome (Kunkel, Kohane et al. 2011, Liao, Catoire et al. 2019). FLNA encodes an actin-binding protein that interacts with actin filaments to remodel the cytoskeleton, regulate the cell shape and migration. It may assist in neuroblast migration and

play a role in cell-cell contacts and adherens junctions during the development of blood vessels, heart, and brain. Various alterations in FLNA, including differential expression, de novo duplications, and interactions with SFARI genes (i.e., SHANK3, NLGN3), were previously associated with ASD (Sakai, Shaw et al. 2011, Shen, Huo et al 2015, Wall, Esteban et al. 2009, Gilman, Iossifov et al. 2011). Finally, PFN1 (Entrez ID: 5216) and FLNA (Entrez ID: 2316) are associated with "glutamatergic synapse" (Table 2), a cellular component well-known to be associated with ASD (Volk, Chiu et al. 2015, El-Ansary, Al-Ayadhi et al. 2014). Briefly, these genes involved in brain development, neuronal migration, and synaptic transmission may exert their diversified functionalities between twins and parents through differential splicing resulting in novel isoforms.

In addition to AS analysis, we employed co-expression analysis and identified seven modules. M3 and M6 modules were enriched for ALA and pyruvate metabolism, respectively. ALA is an essential short-chain omega-3 fatty acid and significant for brain development and function (Bourre, Dumont et al. 1991, Stark, Crawford et al. 2008), and its differential amounts between ASD and healthy controls have been reported in several studies (James, Montgomery et al. 2011, Jory, McGinnis et al. 2008, El-Ansary, Ben et al. 2011, Gordon, Miller et al. 2010, Bu, Ashwood et al. 2006, Esparham, Smith et al. 2015).

Pyruvate metabolism is a component of the cellular energy metabolism pathways. Mitochondrial key enzyme PDH converts pyruvate to acetyl-CoA and alterations in this enzyme may have a substantial effect on cellular energy metabolism. Aberrations in pyruvate levels were reported in patients with ASD compared to healthy controls (Anney, Kenny et al. 2011, Haas et al. 2010, Palmieri, Persico et al. 2010, Clark-Taylors et al. 2004), and the mitochondrial dysfunctionality was proposed as the underlying cause (Rossignol, Frye et al. 2012). The co-occurrence of ASD and mitochondrial abnormalities in children were estimated between 4-5% (Haas et al. 2010) and dysfunction of the electron transport chain was associated with increased blood pyruvate levels observed in patients with ASD (Weissman, Kelley et al. 2008). Moreover, higher activation of the M6 module in twins rather than their parents might be associated with increased ASD-related blood pyruvate levels and mitochondrial activity. Taken together, co-expression modules that are related to pyruvate and ALA metabolism may reflect the role of mitochondrial dysfunction and metabolic alterations in ASD. They may also be interpreted as that such defects can be detected by transcriptomic analysis of PBMCs from ASD patients and their parents.

In addition to co-expression module enrichments, we looked at gene distribution across co-expression modules and found that the seven co-expression modules have 513 genes, 21 of which were SFARI genes and 5 of which were DAS genes. Five of these 21 SFARI genes were either Category 1 (ADNP and NR4A2) or Category 2 (DIP2A, PYHIN1, and RAB43). As an example for co-expressed DAS genes, considering the well-known roles of 5'UTRs in gene-regulation, the differential mutually exclusive usage of exons of co-expressed gene ZNF322 may alter its post-transcriptional regulation during development and/or adulthood, which may eventually result in the altered expression of its target genes. ZNF322 is a zinc-finger transcription factor, expressed ubiquitously during development and adulthood, believed to be a significant gene during human brain development and function (Li, Wang et al. 2004). Its human-specific duplications occurred after the divergence of human and chimpanzee lineages (Sassa et al. 2013). Expression and methylation analyses pointed to ZNF322 as an ASD-related gene (Tian, Liao et al 2011, Kunkel, Kohane et al. 2011, Mordaunt, Jianu et al. 2019). Another co-expressed gene NR4A1 had previously been reported as one of the top genes in ASD-related PPI-nets and as a DE gene in different ASD studies (Kunkel, Kohane et al. 2011, Correia, Oliveira et al. 2014) and it may have a potential for alterations in PPI-nets due to both centrality role within its own module and differential splicing pattern. Taken together, ZNF322 and NR4A1 could be potentially interesting targets for further investigations, as they display differential AS, have a central role in their networks, and were implicated previously in ASD or brain development.

When we looked at DE exons encoding functional domains, realized that SFARI genes (CUL3 and PTK7), co-expressed genes (DCHS1 and RPS26), and some other genes (FLNA and PFN1) had such exons. Of note, DCHS1 and CUL3 genes involve in ASD-related pathways called hippo and hedgehog signaling pathways, implicating neural development and neuronal maintenance, respectively. DCHS1 was reported to be a DAS gene in the post-mortem autistic brains compared to healthy control brains (Gupta, Ellis et al. 2014). FLNA, RPS26, and PFN1 genes involved in cadherin binding molecular function related to ASD. Overall, according to the known roles of functional domains on PPI-nets (Yang, Coulombe-Huntington et al. 2016), PPI-nets related to neurodevelopmental genes might be diversified by differential splicing events causing alterations in the expression of domains of interest between autistic twins and their parents.

6. CONCLUSION

The closer examination of DE exons within SFARI genes, co-expressed genes, and the novel isoforms between twins with ASD and their healthy parents revealed that those exons could potentially code UTR or functional domain sequences. Thus, the differential AS events in coding and non-coding exons of genes associated with synaptic transmission, brain development, and neuronal migration-related might exert their effects on PPI-nets through altered protein structure or protein amount, because of altered functional domains (truncated or extended) or post-transcriptional regulation, respectively. Briefly, differential AS events might be source of diversification of PPI-nets between autistic twins and parents via altering post-transcriptional regulation and/or functional domains.

Results of co-expression analysis suggest that further investigation of mitochondrial and ALA pathways may help elucidate their roles in ASD etiology, and co-expressed genes involved in those pathways might be potential targets.

In addition, our observation that some of, the same splicing events of SFARI genes reported for post-mortem brains from ASD patients can be detected from family and blood-based ASD samples suggest that they might be used as the biomarkers for the diagnosis and possibly treatment of ASD.

Overall, a combinatorial approach that integrates AS and co-expression analysis of PBMC transcriptomes could provide an alternative and possibly complementary tool to other methods employed in ASD studies. Testing and validation of our findings in larger cohorts may result in the development of blood-based biomarkers and could also potentially shed light on ASD etiology.

7. REFERENCES

Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J. Autism Dev. Disord. 24, 659–685 (1994).

Sandin, S. et al. The familial risk of autism. JAMA - J. Am. Med. Assoc. 311, 1770–1777 (2014).

Le Couteur, A. L. et al. Autism as a strongly genetic disorder evidence from a british twin Study. Psychol. Med. 25, 63–77 (1995).

Ronald, A., Larsson, H., Anckarsäter, H. & Lichtenstein, P. A twin study of autism symptoms in Sweden. Mol. Psychiatry 16, 1039–1047 (2011).

Weiss, L. A. et al. A genome-wide linkage and association scan reveals novel loci for autism. Nature 461, 802–808 (2009).

An, J. Y. et al. Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. Science (80-.). 362, (2018).

Grove, J. et al. Identification of common genetic risk variants for autism spectrum disorder. Nat. Genet. 51, 431–444 (2019).

Iossifov, I. et al. De Novo Gene Disruptions in Children on the Autistic Spectrum. Neuron 74, 285–299 (2012).

O ' Roak, B. J. et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485, 246–250 (2012).

Sanders, S. J. et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485, 237–241 (2012).

Yu TW, Chahrour MH, Coulter ME, Jiralerspong S, Okamura-Ikeda K, Ataman B, et al. Using Whole-Exome Sequencing to Identify Inherited Causes of Autism. Neuron. 2013 Jan 23;77(2):259–73.

Yuen RKC, Merico D, Bookman M, Howe JL, Thiruvahindrapuram B, Patel R V., et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci. 2017 Mar 29;20(4):602–11.

Callaghan DB, Rogic S, Tan PPC, Calli K, Qiao Y, Baldwin R, et al. Whole genome sequencing and variant discovery in the ASPIRE autism spectrum disorder cohort. Clin Genet. 2019;96(3):199–206.

Lin CY, Chang KW, Lin CY, Wu JY, Coon H, Huang PH, et al. Allele-specific expression in a family quartet with autism reveals mono-to-biallelic switch and novel transcriptional processes of autism susceptibility genes. Sci Rep. 2018 Dec 1;8(1):1–15.

Balestrieri, Emanuela, et al. "Children with Autism Spectrum Disorder and their mothers share abnormal expression of selected endogenous retroviruses families and cytokines." Frontiers in Immunology 10 (2019): 2244.

Nord AS, Roeb W, Dickel DE, Walsh T, Kusenda M, O'Connor KL, et al. Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. Eur J Hum Genet. 2011 Jun;19(6):727–31.

Ch'ng C, Kwok W, Rogic S, Pavlidis P. Meta-analysis of gene expression in Autism spectrum disorder. Autism Res. 2015 Oct 1;8(5):593–608.

Baio, Jon. "Prevalence of Autism Spectrum Disorders: Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. Morbidity and Mortality Weekly Report. Surveillance Summaries. Volume 61, Number 3." Centers for Disease Control and Prevention (2012).

Elsabbagh M, Divan G, Koh YJ, Kim YS, Kauchali S, Marcín C, et al. Global Prevalence of Autism and Other Pervasive Developmental Disorders. Autism Res. 2012;5(3):160–79.

Werling, Donna M., and Daniel H. Geschwind. "Sex differences in autism spectrum disorders." Current opinion in neurology 26.2 (2013): 146.

Robinson EB, Lichtenstein P, Anckarsäter H, Happé F, Ronald A. Examining and interpreting the female protective effect against autistic behavior. Proc Natl Acad Sci U S A. 2013;110(13):5258–62.

Wing L, Gould J. Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. J Autism Dev Disord. 1979;9(1):11–29.

Levy, Dan, et al. "Rare de novo and transmitted copy-number variation in autistic spectrum disorders." Neuron 70.5 (2011): 886-897.

Baron-Cohen S, Auyeung B, Nørgaard-Pedersen B, Hougaard DM, Abdallah MW, Melgaard L, et al. Elevated fetal steroidogenic activity in autism. Mol Psychiatry. 2015;20(3):369–76.

Jacquemont S, Coe BP, Hersch M, Duyzend MH, Krumm N, Bergmann S, et al. A higher mutational burden in females supports a “female protective model” in neurodevelopmental disorders. Am J Hum Genet. 2014;94(3):415–25.

Baker, J. P. (2010). Autism in 1959: Joey the mechanical boy. Pediatrics, 125(6), 1101-1103.

Bender, L. (1947). Clinical study of one hundred schizophrenic children. American Journal of Orthopsychiatry, 17, 40-56.

Bell CC. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders. JAMA J Am Med Assoc [Internet]. 1994 Sep 14 [cited 2020 Jul 13];272(10):828. Available from: https://jamanetwork.com/journals/jama/fullarticle/379036.

Schopler E, Reichler RJ, DeVellis RF, Daly K. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). J Autism Dev Disord. 1980;10(1):91–103.

Mukaddes, N. M. "Otizm spektrum bozuklukları tanı ve takip." İstanbul: Nobel Tıp Kitapevleri (2013).

Postorino V, Kerns CM, Vivanti G, Bradshaw J, Siracusano M, Mazzone L. Anxiety Disorders and Obsessive-Compulsive Disorder in Individuals with Autism Spectrum Disorder. Curr Psychiatry Rep. 2017;19(12).

de la Torre-Ubieta, L., Won, H., Stein, J. et al. Advancing the understanding of autism disease mechanisms through genetics. Nat Med 22, 345–361 (2016). https://doi.org/10.1038/nm.4071.

Applied Behavior Analysis (ABA) | Autism Speaks [Internet]. [cited 2020 Jul 13]. Available from: https://www.autismspeaks.org/applied-behavior-analysis-aba-0.

Fernandes FDM, Amato CA de la H. Applied Behavior Analysis and Autism Spectrum Disorders: Literature review. Codas. 2013;25(3):289–96.

Murray ML, Hsia Y, Glaser K, Simonoff E, Murphy DGM, Asherson PJ, et al. Pharmacological treatments prescribed to people with autism spectrum disorder (ASD) in primary health care. Psychopharmacology (Berl). 2014;231(6):1011–21.

Coury DL, Anagnostou E, Manning-Courtney P, Reynolds A, Cole L, McCoy R, et al. Use of psychotropic medication in children and adolescents with autism spectrum disorders. Pediatrics. 2012;130(SUPPL. 2).

King M, Bearman P. Diagnostic change and the increased prevalence of autism. Int J Epidemiol. 2009;38(5):1224–34.

Miles JH. Autism spectrum disorders-A genetics review. Genet Med. 2011;13(4):278–94.

Chaste P, Leboyer M. Autism risk factors: genes, environment, and gene-environment interactions. Dialogues Clin Neurosci. 2012 Sep;14(3):281-92. PMID: 23226953; PMCID: PMC3513682.

Quesnel-Vallières, M., Weatheritt, R.J., Cordes, S.P. et al. Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. Nat Rev Genet 20, 51–63 (2019). https://doi.org/10.1038/s41576-018-0066-2.

Siu M.T., Weksberg R. (2017) Epigenetics of Autism Spectrum Disorder. In: Delgado-Morales R. (eds) Neuroepigenomics in Aging and Disease. Advances in Experimental Medicine and Biology, vol 978. Springer, Cham.

Tordjman S, Somogyi E, Coulon N, Kermarrec S, Cohen D, Bronsard G, et al. Gene × environment interactions in autism spectrum disorders: Role of epigenetic mechanisms. Front Psychiatry. 2014;5(AUG):1–17.

Colvert E, Tick B, McEwen F, et al. Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. JAMA Psychiatry. 2015;72(5):415–423. doi:10.1001/jamapsychiatry.2014.3028.

Hallmayer J, Cleveland S, Torres A, et al. Genetic Heritability and Shared Environmental Factors Among Twin Pairs With Autism. Arch Gen Psychiatry. 2011;68(11):1095–1102. doi:10.1001/archgenpsychiatry.2011.76.

Werling, D.M., Geschwind, D.H. Recurrence rates provide evidence for sex-differential, familial genetic liability for autism spectrum disorders in multiplex families and twins. Molecular Autism 6, 27 (2015). https://doi.org/10.1186/s13229-015-0004-5.

Wood, Claire L., et al. "Evidence for ASD recurrence rates and reproductive stoppage from large UK ASD research family databases." Autism Research 8.1 (2015): 73-81.

Constantino, J., Todorov, A., Hilton, C. et al. Autism recurrence in half siblings: strong support for genetic mechanisms of transmission in ASD. Mol Psychiatry 18, 137–138 (2013). https://doi.org/10.1038/mp.2012.9.

Pierre A Saint, Génin E. How important are rare variants in common disease? Brief Funct Genomics. 2014;13(5):353–61.

Stein, Jason L., Neelroop N. Parikshak, and Daniel H. Geschwind. "Rare inherited variation in autism: beginning to see the forest and a few trees." Neuron 77.2 (2013): 209-211.

Kinney, Dennis K., et al. "Environmental risk factors for autism: do they help cause de novo genetic mutations that contribute to the disorder?." Medical hypotheses 74.1 (2010): 102-106.

Gardener H, Spiegelman D, Buka SL. Prenatal risk factors for autism: comprehensive meta-analysis. British Journal of Psychiatry. Cambridge University Press; 2009;195(1):7–14.

Hultman, C., Sandin, S., Levine, S. et al. Advancing paternal age and risk of autism: new evidence from a population-based study and a meta-analysis of epidemiological studies. Mol Psychiatry 16, 1203–1212 (2011). https://doi.org/10.1038/mp.2010.121.

Grabrucker AM. Environmental factors in autism. Front Psychiatry. 2013;3(JAN):1–13.

Beversdorf, D.Q., Manning, S. ., Hillier, A. et al. Timing of Prenatal Stressors and Autism. J Autism Dev Disord 35, 471–478 (2005). https://doi.org/10.1007/s10803-005-5037-8.

Taylor, L. E., Swerdfeger, A. L., & Eslick, G. D. (2014). Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies. Vaccine, 32(29), 3623-3629.

Surén P, Roth C, Bresnahan M, et al. Association Between Maternal Use of Folic Acid Supplements and Risk of Autism Spectrum Disorders in Children. JAMA. 2013;309(6):570–577. doi:10.1001/jama.2012.155925.

Atladóttir, Hjördis Ósk, et al. "Autism after infection, febrile episodes, and antibiotic use during pregnancy: an exploratory study." Pediatrics 130.6 (2012): e1447-e1454.

Volk, Heather E., et al. "Brief Report: Autism Spectrum Disorder: Interaction of Air Pollution with the MET Receptor Tyrosine Kinase Gene." Epidemiology (2014): 44-47.

Nijmeijer, Judith S., et al. "Perinatal risk factors interacting with catechol O-methyltransferase and the serotonin transporter gene predict ASD symptoms in children with ADHD." Journal of Child Psychology and Psychiatry 51.11 (2010): 1242-1250.

Kim, J., Park, K., Kang, R.J. et al. Gene-environment interaction counterbalances social impairment in mouse models of autism. Sci Rep 9, 11490 (2019). https://doi.org/10.1038/s41598-019-47680-w.

Berko, Esther R., et al. "Mosaic epigenetic dysregulation of ectodermal cells in autism spectrum disorder." PLoS Genet 10.5 (2014): e1004402.

Duffney, Lara J., et al. "Epigenetics and autism spectrum disorder: A report of an autism case with mutation in H1 linker histone HIST1H1E and literature review." American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 177.4 (2018): 426-433.

Ladd-Acosta, C., Hansen, K., Briem, E. et al. Common DNA methylation alterations in multiple brain regions in autism. Mol Psychiatry 19, 862–871 (2014). https://doi.org/10.1038/mp.2013.114.

Ellis, S.E., Gupta, S., Moes, A. et al. Exaggerated CpH methylation in the autism-affected brain. Molecular Autism 8, 6 (2017). https://doi.org/10.1186/s13229-017-0119-y.

Andrews, S.V., Sheppard, B., Windham, G.C. et al. Case-control meta-analysis of blood DNA methylation and autism spectrum disorder. Molecular Autism 9, 40 (2018). https://doi.org/10.1186/s13229-018-0224-6.

Shulha HP, Cheung I, Whittle C, et al. Epigenetic Signatures of Autism: Trimethylated H3K4 Landscapes in Prefrontal Neurons. Arch Gen Psychiatry. 2012;69(3):314–324. doi:10.1001/archgenpsychiatry.2011.151.

Sun, Wenjie, et al. "Histone acetylome-wide association study of autism spectrum disorder." Cell 167.5 (2016): 1385-1397.

Lyall, Kristen et al. “The Changing Epidemiology of Autism Spectrum Disorders.” Annual review of public health vol. 38 (2017): 81-102. doi:10.1146/annurev-publhealth-031816-044318.

Ha S, Sohn I-J, Kim N, Sim HJ, Cheon K-A. Characteristics of Brains in Autism Spectrum Disorder: Structure, Function and Connectivity across the Lifespan. Exp Neurobiol. 2015;24(4):273–84.

Courchesne E, Campbell K, Solso S. Brain growth across the life span in autism: Age-specific changes in anatomical pathology. Brain Res. 2011 Mar 22;1380:138–45.

Via E, Radua J, Cardoner N, Happé F, Mataix-Cols D. Meta-analysis of Gray Matter Abnormalities in Autism Spectrum Disorder: Should Asperger Disorder Be Subsumed Under a Broader Umbrella of Autistic Spectrum Disorder? Arch Gen Psychiatry. 2011;68(4):409–418. doi:10.1001/archgenpsychiatry.2011.27.

Radua J, Via E, Catani M, Mataix-Cols D. Voxel-based meta-analysis of regional white-matter volume differences in autism spectrum disorder versus healthy controls. Psychological Medicine. Cambridge University Press; 2011;41(7):1539–1550.

Frazier, Thomas W., and Antonio Y. Hardan. "A meta-analysis of the corpus callosum in autism." Biological psychiatry 66.10 (2009): 935-941.

Courchesne E, Mouton PR, Calhoun ME, et al. Neuron Number and Size in Prefrontal Cortex of Children With Autism. JAMA. 2011;306(18):2001–2010. doi:10.1001/jama.2011.1638.

Tomova, Aleksandra, et al. "Gastrointestinal microbiota in children with autism in Slovakia." Physiology & behavior 138 (2015): 179-187.

Miller, David T., et al. "Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies." The American Journal of Human Genetics 86.5 (2010): 749-764.

Ho, Karen S., et al. "Chromosomal microarray analysis of consecutive individuals with autism spectrum disorders using an ultra-high resolution chromosomal microarray optimized for neurodevelopmental disorders." International journal of molecular sciences 17.12 (2016): 2070.

Ungar, Wendy J. “Next Generation Sequencing and Health Technology Assessment in Autism Spectrum Disorder.” Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent vol. 24,2 (2015): 123-7.

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ. Dynamic integration of splicing within gene regulatory pathways. Vol. 152, Cell. Cell; 2013. p. 1252–69.

Chen M, Manley JL. Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. Vol. 10, Nature Reviews Molecular Cell Biology. Nat Rev Mol Cell Biol; 2009. p. 741–54.

Kalsotra A, Cooper TA. Functional consequences of developmentally regulated alternative splicing. Vol. 12, Nature Reviews Genetics. NIH Public Access; 2011. p. 715–29.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 2008 Dec;40(12):1413–5.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008 Nov 27;456(7221):470–6.

Blencowe BJ. Alternative Splicing: New Insights from Global Analyses. Vol. 126, Cell. Cell; 2006. p. 37–47.

Pohl M, Bortfeldt RH, Grützmann K, Schuster S. Alternative splicing of mutually exclusive exons-A review. Vol. 114, BioSystems. Elsevier Ireland Ltd; 2013. p. 31–8.

Wahl MC, Will CL, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. Vol. 136, Cell. Cell; 2009. p. 701–18.

Black DL, Grabowski PJ. Alternative pre-mRNA splicing and neuronal function. Vol. 31, Progress in molecular and subcellular biology. Prog Mol Subcell Biol; 2003. p. 187–216.

Mazin P, Xiong J, Liu X, Yan Z, Zhang X, Li M, et al. Widespread splicing changes in human brain development and aging. Mol Syst Biol. 2013;9:633.

Weyn-Vanhentenryck SM, Feng H, Ustianenko D, Duffié R, Yan Q, Jacko M, et al. Precise temporal regulation of alternative splicing during neural development. Nat Commun. 2018 Dec 1;9(1):1–17.

Scotti MM, Swanson MS. RNA mis-splicing in disease. Vol. 17, Nature Reviews Genetics. Nature Publishing Group; 2016. p. 19–32.

Morel CF, Thomas MA, Cao H, O'Neil CH, Pickering JG, Foulkes WD, et al. A LMNA splicing mutation in two sisters with severe Dunnigan-type familial partial lipodystrophy type 2. J Clin Endocrinol Metab. 2006;91(7):2689–95.

Otomo J, Kure S, Shiba T, Karibe A, Shinozaki T, Yagi T, et al. Electrophysiological and histopathological characteristics of progressive atrioventricular block accompanied by familial dilated cardiomyopathy caused by a novel mutation of lamin A/C gene. J Cardiovasc Electrophysiol. 2005 Feb;16(2):137–45.

Arnold ES, Ling SC, Huelga SC, Lagier-Tourenne C, Polymenidou M, Ditsworth D, et al. ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. Proc Natl Acad Sci U S A. 2013 Feb19;110(8).

Tian Y, Liao IH, Zhan X, Gunther JR, Ander BP, Liu D, et al. Exon expression and alternatively spliced genes in tourette syndrome. Am J Med Genet Part B Neuropsychiatr Genet. 2011 Jan;156(1):72–8.

Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature. 2011 Jun 16;474(7351):380–6.

Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, et al. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. Nat Commun. 2014 Dec 10;5(1):1–8.

Stamova BS, Tian Y, Nordahl CW, Shen MD, Rogers S, Amaral DG, et al. Evidence for differential alternative splicing in blood of young boys with autism spectrum disorders. Mol Autism. 2013;4(1).

Irimia M, Weatheritt RJ, Ellis JD, Parikshak NN, Gonatopoulos-Pournatzis T, Babor M, et al. A highly conserved program of neuronal microexons is misregulated in autistic brains. Cell. 2014 Dec 18;159(7):1511–23.

Mosca R, Céol A, Stein A, Olivella R, Aloy P. 3did: A catalog of domain-based interactions of known three-dimensional structure. Nucleic Acids Res. 2014;42(D1):374–9.

Fuller, Tova, et al. "Review of weighted gene coexpression network analysis." Handbook of Statistical Bioinformatics. Springer, Berlin, Heidelberg, 2011. 369-388.

Kakati, Tulika, et al. "Comparison of Methods for Differential Co-expression Analysis for Disease Biomarker Prediction." Computers in Biology and Medicine 113 (2019): 103380.

Liu, Yichuan, et al. "Evaluating the impact of sequencing depth on transcriptome profiling in human adipose." PLoS One 8.6 (2013): e66883.

Kim, Kyung, et al. "Effect of next-generation exome sequencing depth for discovery of diagnostic variants." Genomics & informatics 13.2 (2015): 31.

Zhao, Shanrong, Zhan Ye, and Robert Stanton. "Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols." RNA (2020): rna-074922.

Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 2012;131(4):281-285. doi:10.1007/s12064-012-0162-3.

Kelemen, Olga et al. “Function of alternative splicing.” Gene vol. 514,1 (2013): 1-30. doi:10.1016/j.gene.2012.07.083.

Donaldson, Lucy F, and Nicholas Beazley-Long. “Alternative RNA splicing: contribution to pain and potential therapeutic strategy.” Drug discovery today vol. 21,11 (2016): 1787-1798. doi:10.1016/j.drudis.2016.06.017.

Sollis E, Graham SA, Vino A, et al. Identification and functional characterization of de novo FOXP1 variants provides novel insights into the etiology of neurodevelopmental disorder. Hum Mol Genet. 2016;25(3):546-557. doi:10.1093/hmg/ddv495.

Ansel, Ashley, et al. "Variation in gene expression in autism spectrum disorders: an extensive review of transcriptomic studies." Frontiers in neuroscience 10 (2017): 601.

Hu, V.W., Frank, B.C., Heine, S. et al. Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes. BMC Genomics 7, 118 (2006). https://doi.org/10.1186/1471-2164-7-118.

Chow, Maggie L., et al. "Age-dependent brain gene expression and copy number anomalies in autism suggest distinct pathological processes at young versus mature ages." PLoS Genet 8.3 (2012): e1002592.

Maekawa, Motoko, et al. "Utility of scalp hair follicles as a novel source of biomarker genes for psychiatric illnesses." Biological psychiatry 78.2 (2015): 116-125.

Féron, F., Gepner, B., Lacassagne, E. et al. Olfactory stem cells reveal MOCOS as a new player in autism spectrum disorders. Mol Psychiatry 21, 1215–1224 (2016). https://doi.org/10.1038/mp.2015.106.

Gregg, Jeffrey P., et al. "Gene expression changes in children with autism." Genomics 91.1 (2008): 22-29.

Enstrom, Amanda M., et al. "Altered gene expression and function of peripheral blood natural killer cells in children with autism." Brain, behavior, and immunity 23.1 (2009): 124-133.

Yasuda, Yuka, et al. "Gene expression analysis in lymphoblasts derived from patients with autism spectrum disorder." Molecular Autism 2.1 (2011): 9.

Voineagu, Irina, and Valsamma Eapen. "Converging pathways in autism spectrum disorders: interplay between synaptic dysfunction and immune responses." Frontiers in human neuroscience 7 (2013): 738.

Sanger, Frederick, Steven Nicklen, and Alan R. Coulson. "DNA sequencing with chain-terminating inhibitors." Proceedings of the national academy of sciences 74.12 (1977): 5463-5467.

Maxam, Allan M., and Walter Gilbert. "[57] Sequencing end-labeled DNA with base-specific chemical cleavages." Methods in enzymology. Vol. 65. Academic press, 1980. 499-560.

Barba, Marina, Henryk Czosnek, and Ahmed Hadidi. "Historical perspective, development and applications of next-generation sequencing in plant virology." Viruses 6.1 (2014): 106-136.

Raza, Khalid, and Sabahuddin Ahmad. "Principle, analysis, application and challenges of next-generation sequencing: a review." arXiv preprint arXiv:1606.05254 (2016).

van Dijk, Erwin L., et al. "The third revolution in sequencing technology." Trends in Genetics 34.9 (2018): 666-681.

Lee, Hayan, et al. "Third-generation sequencing and the future of genomics." BioRxiv (2016): 048603.

Slatko, Barton E., Andrew F. Gardner, and Frederick M. Ausubel. "Overview of next-generation sequencing technologies." Current protocols in molecular biology 122.1 (2018): e59.

Illumina, I. "An introduction to next-generation sequencing technology." (2015).

Metzker, Michael L. "Sequencing technologies—the next generation." Nature reviews genetics 11.1 (2010): 31-46.

Han, Yixing, et al. "Advanced applications of RNA sequencing and challenges." Bioinformatics and biology insights 9 (2015): BBI-S28991.

Ozsolak, Fatih, and Patrice M. Milos. "RNA sequencing: advances, challenges and opportunities." Nature reviews genetics 12.2 (2011): 87-98.

Gonorazky, Hernan D., et al. "Expanding the boundaries of RNA sequencing as a diagnostic tool for rare mendelian disease." The American Journal of Human Genetics 104.3 (2019): 466-483.

Stark, Rory, Marta Grzelak, and James Hadfield. "RNA sequencing: the teenage years." Nature Reviews Genetics 20.11 (2019): 631-656.

Patel, Ravi K., and Mukesh Jain. "NGS QC Toolkit: a toolkit for quality control of next generation sequencing data." PloS one 7.2 (2012): e30619.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics 30.15 (2014): 2114-2120.

Gordon, Assaf, and G. J. Hannon. "Fastx-toolkit." FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab. cshl. edu/fastx_toolkit 5 (2010).

Kim, Daehwan, Ben Langmead, and Steven L. Salzberg. "HISAT: a fast spliced aligner with low memory requirements." Nature methods 12.4 (2015): 357-360.

Dobin, Alexander, et al. "STAR: ultrafast universal RNA-seq aligner." Bioinformatics 29.1 (2013): 15-21.

Garber, Manuel, et al. "Computational methods for transcriptome annotation and quantification using RNA-seq." Nature methods 8.6 (2011): 469-477.

Langmead, Ben, et al. "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome biology 10.3 (2009): R25.

Robertson, Gordon, et al. "De novo assembly and analysis of RNA-seq data." Nature methods 7.11 (2010): 909-912.

Cahais, V., et al. "Reference-free transcriptome assembly in non-model animals from next-generation sequencing data." Molecular ecology resources 12.5 (2012): 834-845.

Anders, Simon, Alejandro Reyes, and Wolfgang Huber. "Detecting differential usage of exons from RNA-seq data." Nature Precedings (2012): 1-1.

Hartley, Stephen W., and James C. Mullikin. "Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq." Nucleic acids research 44.15 (2016): e127-e127.

American Psychiatric Association. DSM-5 Diagnostic Classification. In: Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association; 2013.

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. J Am Acad Child Adolesc Psychiatry. 1997;36(7):980–8.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. (SCID-I/P) New York: Biometrics Research, New York State Psychiatric Institute.

Andrews, Simon. "FastQC: a quality control tool for high throughput sequence data." (2010).

Kim, Daehwan, et al. "Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype." Nature biotechnology 37.8 (2019): 907-915.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.

Garrido-Martín, Diego, et al. "ggsashimi: Sashimi plot revised for browser-and annotation-independent splicing visualization." PLoS computational biology 14.8 (2018): e1006360.

Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi A J Integr Biol. 2012 May 1;16(5):284–7.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015 Feb 18;33(3):290–5.

Russo, P.S.T., Ferreira, G.R., Cardozo, L.E. et al. CEMiTool: a Bioconductor package for performing comprehensive modular co-expression analyses. BMC Bioinformatics 19, 56 (2018). https://doi.org/10.1186/s12859-018-2053-1.

Michael E. Smoot, Keiichiro Ono, Johannes Ruscheinski, Peng-Liang Wang, Trey Ideker, Cytoscape 2.8: new features for data integration and network visualization, Bioinformatics, Volume 27, Issue 3, 1 February 2011, Pages 431–432, https://doi.org/10.1093/bioinformatics/btq675.

Kolde, Raivo, and Maintainer Raivo Kolde. "Package 'pheatmap'." R Package 1.7 (2015).

Hatje K, Rahman R, Vidal RO, Simm D, Hammesfahr B, Bansal V, et al. The landscape of human mutually exclusive splicing. Mol Syst Biol. 2017 Dec;13(12):959.

Waites CL, Garner CC. Presynaptic function in health and disease. Vol. 34, Trends in Neurosciences. Elsevier; 2011. p. 326–37.

Fassio A, Patry L, Congia S, Onofri F, Piton A, Gauthier J, et al. SYN1 loss-of-function mutations in autism and partial epilepsy cause impaired synaptic function. Hum Mol Genet. 2011;20(12):2297–307.

Stewart LT. Cell adhesion proteins and the pathogenesis of autism spectrum disorders. J Neurophysiol 2015; 113: 1283–6.

Kunkel LM, Kohane IS, Kong SW, Collins CD. Methods and compositions for characterizing autism spectrum disorder based on gene expression patterns. 2011 Sep 15. https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2011112961&tab=PCTBIBLIO.

Liao C, Catoire H, Bourassa C, Dion PA, Rouleau GA. Increased expression of genetically-regulated FLT3 implicated in Tourette Syndrome. bioRxiv Genet. 2019;812420.

Sakai Y, Shaw CA, Dawson BC, Dugas D V., Al-Mohtaseb Z, Hill DE, et al. Protein interactome reveals converging molecular pathways among autism disorders. Sci Transl Med. 2011;3(86).

Shen C, Huo L rong, Zhao X liang, Wang P rong, Zhong N. Novel Interactive Partners of Neuroligin 3: New Aspects for Pathogenesis of Autism. J Mol Neurosci. 2015 May 1;56(1):89–101.

Wall DP, Esteban FJ, DeLuca TF, Huyck M, Monaghan T, Velez de Mendizabal N, et al. Comparative analysis of neurological disorders focuses genome-wide search for autism genes. Genomics. 2009 Feb;93(2):120–9.

Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D. Rare De Novo Variants Associated with Autism Implicate a Large Functional Network of Genes Involved in Formation and Function of Synapses. Neuron. 2011 Jun 9;70(5):898–907.

Volk L, Chiu S-L, Sharma K, Huganir RL. Glutamate Synapses in Human Cognitive Disorders. Annu Rev Neurosci. 2015 Jul 8;38(1):127–49.

El-Ansary A, Al-Ayadhi L. GABAergic/glutamatergic imbalance relative to excessive neuroinflammation in autism spectrum disorders. J Neuroinflammation. 2014;11(1):1–9.

Li Y, Wang Y, Zhang C, Yuan W, Wang J, Zhu C, et al. ZNF322, a novel human C 2H 2 Krüppel-like zinc-finger protein, regulates transcriptional activation in MAPK signaling pathways. Biochem Biophys Res Commun. 2004 Dec 24;325(4):1383–92.

Sassa T. The Role of human-specific gene duplications during brain development and evolution. Vol. 27, Journal of Neurogenetics. J Neurogenet; 2013. p. 86–96.

Mordaunt CE, Jianu JM, Laufer B, et al. Cord blood DNA methylome in newborns later diagnosed with autism spectrum disorder reflects early dysregulation of neurodevelopmental and X-linked genes. bioRxiv; 2019. DOI: 10.1101/850529.

Correia C, Oliveira G, Vicente AM. Protein interaction networks reveal novel autism risk genes within GWAS statistical noise. PLoS One. 2014;9(11):1–11.

Yang X, Coulombe-Huntington J, Kang S, Sheynkman GM, Hao T, Richardson A, et al. Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing. Cell. 2016 Feb 11;164(4):805–17.

Comoletti, Davide, et al. "The Arg451Cys-neuroligin-3 mutation associated with autism reveals a defect in protein processing." Journal of Neuroscience 24.20 (2004): 4889-4893.

Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. Genome Biol. 2002;3(3):1–10.

Bourre JM, Dumont O, Piciotti M, Clément M, Chaudière J, Bonneil M, et al. Essentiality of omega 3 fatty acids for brain structure and function. World Rev Nutr Diet. 1991;66:103–17.

Stark AH, Crawford MA, Reifen R. Update on alpha-linolenic acid. Vol. 66, Nutrition Reviews. Nutr Rev; 2008. p. 326–32.

James S, Montgomery P, Williams K. Omega-3 fatty acids supplementation for autism spectrum disorders (ASD). Cochrane Database Syst Rev. 2011 Nov 9;(11).

Jory J, McGinnis WR. Red-cell trace minerals in children with autism. Am J Biochem Biotechnol. 2008 Jun 30;4(2):101–4.

El-Ansary AK, Ben Bacha AG, Al- Ayahdi LY. Impaired plasma phospholipids and relative amounts of essential polyunsaturated fatty acids in autistic patients from Saudi Arabia. Lipids Health Dis. 2011;10:63.

Gordon Bell J, Miller D, MacDonald DJ, MacKinlay EE, Dick JR, Cheseldine S, et al. The fatty acid compositions of erythrocyte and plasma polar lipids in children with autism, developmental delay or typically developing controls and the effect of fish oil intake. Br J Nutr. 2010 Apr;103(8):1160–7.

Bu B, Ashwood P, Harvey D, King IB, Water J Van de, Jin LW. Fatty acid compositions of red blood cell phospholipids in children with autism. Prostaglandins Leukot Essent Fat Acids. 2006 Apr;74(4):215–21.

Esparham AE, Smith T, Belmont JM, Haden M, Wagner LE, Evans RG, et al. Nutritional and metabolic biomarkers in autism spectrum disorders: An exploratory study. Integr Med. 2015;14(2):40–53.

Anney RJL, Kenny EM, O'Dushlaine C, Yaspan BL, Parkhomenka E, Buxbaum JD, et al. Gene-ontology enrichment analysis in two independent family-based samples highlights biologically plausible processes for autism spectrum disorders. Eur J Hum Genet. 2011 Oct 1;19(10):1082–9.

Haas RH. Autism and mitochondrial disease. Vol. 16, Developmental Disabilities Research Reviews. John Wiley and Sons Inc.; 2010. p. 144–53.

Palmieri L, Persico AM. Mitochondrial dysfunction in autism spectrum disorders: Cause or effect? Vol. 1797, Biochimica et Biophysica Acta - Bioenergetics. Biochim Biophys Acta; 2010. p. 1130–7.

Clark-Taylor T, Clark-Taylor BE. Is autism a disorder of fatty acid metabolism? Possible dysfunction of mitochondrial β-oxidation by long chain acyl-CoA dehydrogenase. Med Hypotheses. 2004;62(6):970–5.

Rossignol DA, Frye RE. Mitochondrial dysfunction in autism spectrum disorders: A systematic review and meta-analysis. Vol. 17, Molecular Psychiatry. Nature Publishing Group; 2012. p. 290–314.

Weissman JR, Kelley RI, Bauman ML, Cohen BH, Murray KF, Mitchell RL, et al. Mitochondrial disease in autism spectrum disorder patients: A cohort analysis. PLoS One. 2008 Nov 26;3(11).

Table 1. Differential alternative splicing events.

Exons and Junctions	P-value	P-adjust	Chr	Start	End	Strand	Log2FC Parents/Children
ENSG00000223972.5:E010	0.000004025	0.0185266927	1	13452	13670	+	4.2002403381
ENSG00000116685.16:E007	0.00000489556969907771	0.0203291366	1	11926302	11926403	-	-1.2973417783
ENSG00000243970.3:E051	0.00000261661498906852	0.0143366152	1	39559633	39559671	-	-3.671269055
ENSG00000243970.3:J089	0.000000235921005594104	0.0034470069	1	39559365	39559633	-	-4.1188817194
ENSG00000270231.4:E021	0.00000131118806845071	0.0095787873	1	120453908	120454114	+	-0.6000253199
ENSG00000117362.13:E003	0.00000972870886050375	0.0314582586	1	150265405	150266194	-	-0.3049430652
ENSG00000122218.15:E061	0.0000170143978213959	0.041368648	1	160339436	160339908	-	1.2970982192
ENSG00000185278.15:J013	0.00000447246126140924	0.0193885128	1	173868405	173868924	+	26.7879367469

ENSG00000135828.11:J011	0.0000110841018002493	0.0343719561	1	182580007	182581224	-	1.2031315364
ENSG00000116698.21:E030	0.0000160047891669019	0.040215142	1	183529397	183529517	+	-0.5423870182
ENSG00000116668.13:E006	0.000015016	0.0390095877	1	185168339	185168398	+	-3.1006734176
ENSG00000081237.20:J076	0.00000375376227756968	0.0178413607	1	198702530	198703297	+	-0.4820417399
ENSG00000143851.15:E010	0.0000149821961855553	0.0390095877	1	202148580	202148699	-	-0.6221910697
ENSG00000117335.19:J049	0.00000920269964731509	0.0302662212	1	207767195	207767606	+	2.3021125977
ENSG00000009790.15:N084	0.000000067078638500799	0.0013231027	1	209778173	209780469	+	-27.707685943
ENSG00000163041.10:E003	0.0000167952904675892	0.041368648	1	226062715	226062734	+	1.1234594774
ENSG00000163041.10:J012	0.0000154174314723252	0.0394512035	1	226062094	226062734	+	-4.8213317134
ENSG00000057608.17:J033	0.000018305751919481	0.0427307159	10	5766295	5768212	-	-4.9913970177

ENSG00000134453.16:J046	0.00000564959085113128	0.0222408338	10	6110127	6112209	+	-0.4596261602
ENSG00000169826.8:N017	0.000000897408875658267	0.0078458814	10	43160595	43162199	+	-1.6254356585
ENSG00000156650.14:J129	0.0000000033209356686572	0.000093892	10	74975454	74977315	+	-27.5737486881
ENSG00000151553.15:J029	0.0000150305767984582	0.0390095877	10	114836246	114842932	+	-1.3104280522
ENSG00000175470.20:E022	0.000023067	0.0497526594	10	131947529	131947791	+	-0.4439719299
ENSG00000132275.11:E007	0.00000408580701149059	0.0185266927	11	6600925	6601055	-	-1.0024941117
ENSG00000132275.11:E009	0.00000225351059117989	0.0138905337	11	6601148	6601253	-	-0.9398089378
ENSG00000166341.8:J039	0.00000676816431953959	0.0244953527	11	6634020	6634117	-	-3.7398903716
ENSG00000167986.13:E058	0.000000118347161080612	0.0020298751	11	61324237	61324266	-	28.5298362975
ENSG00000131626.18:E014	0.0000116537793214781	0.0347420796	11	70325499	70325574	+	0.7756607026

ENSG00000187240.15:E053	0.000000934760382934133	0.0078458814	11	103188496	103188648	+	-27.5662065078
ENSG00000075239.13:J044	0.0000142415321519556	0.0390095877	11	108133937	108134220	+	-1.1563293199
ENSG00000198851.9:E003	0.0000000518955891339997	0.0011373581	11	118304736	118304776	+	-2.125607439
ENSG00000198851.9:J020	0.000000066560385306 1188	0.0013231027	11	118304776	118304893	+	-2.0858195268
ENSG00000134954.14:E008	0.000004045	0.0185266927	11	128463627	128463950	-	0.9493466187
ENSG00000111669.15:E003	0.0000012180626625 5662	0.0092407153	12	6867530	6867681	+	-0.4342207947
ENSG00000111669.15:J029	0.000000897733872986025	0.0078458814	12	6867681	6868863	+	-0.434564978
ENSG00000172243.17:J041	0.0000023716813 5795643	0.0140465894	12	10125448	10127746	-	-0.6253296468
ENSG00000189079.16:J049	0.0000198904694244503	0.0453563639	12	45836991	45837320	+	-1.4347578659
ENSG00000123358.20:E031	0.000000526133937073772	0.0057861222	12	52055720	52055961	+	-2.0162166843

ENSG00000197728.11:E007	2.9029651803389E-052	5.725997E-047	12	56042139	56042169	+	-4.7742258255
ENSG00000197728.11:J016	0.000000000000171171287211	0.00000002250	12	56042169	56042424	+	-5.9931587513
ENSG00000135506.16:E051	0.0000170930526521035	0.041368648	12	57719182	57720098	+	-1.0943390245
ENSG00000120833.14:E013	0.0000173743960698731	0.0415460682	12	93572674	93572684	+	1.7654652997
ENSG00000136048.14:E013	0.000000071	0.0013328394	12	101914173	101914232	+	-1.9093771364
ENSG00000136048.14:E015	0.0000147099293818868	0.0390095877	12	101920108	101920201	+	-1.3175646125
ENSG00000136048.14:J028	0.000002139	0.0135017137	12	101908363	101914173	+	-1.7306521167
ENSG00000120837.8:J024	0.00000229447604633751	0.0139254575	12	104131838	104135447	-	-3.6350080038
ENSG00000089060.11:J059	0.0000128536981029934	0.0359623683	12	113306575	113307690	-	-0.7813905853
ENSG00000136111.13:E012	0.000008581	0.0289327073	13	75306465	75306471	-	-0.506583284

ENSG00000136111.13:J041	0.0000124228710820386	0.0350052549	13	75306471	75309941	-	-0.5154522036
ENSG00000126217.21:J128	0.00000000075126959140910	0.00002963705	13	113065085	113066045	+	-28.0732745937
ENSG00000092010.15:J041	0.000000165860880380308	0.0025966196	14	24137565	24137699	+	-0.4716747743
ENSG00000020577.13:E016	0.0000181401286851695	0.0425961535	14	54702061	54702396	+	2.1678590089
ENSG00000100650.15:E031	0.000000493257166478673	0.0057231323	14	69769251	69769383	+	-0.592767908
ENSG00000100650.15:E038	0.00000245036485793539	0.0140465894	14	69770447	69770466	+	-0.7686892082
ENSG00000078304.19:J096	0.00000675071561728583	0.0244953527	14	101810036	101856685	+	-0.5748509196
ENSG00000104093.13:J109	0.00000333733993820887	0.017098146	15	51536862	51537487	-	0.6802746959
ENSG00000185088.14:J019	0.00000724057777336778	0.0255938595	15	63154060	63155620	-	-0.7514191739
ENSG00000028528.15:E024	0.0000207143015663751	0.0464298123	15	64130227	64130321	+	-0.428604437

ENSG00000028528.15:J063	0.00000194273859028582	0.0132137375	15	64130321	64131686	+	-0.539238895
ENSG00000166938.13:J057	0.000005793	0.0222408338	15	66306952	66308708	+	-1.2863476575
ENSG00000225151.10:E020	0.00000456211446753545	0.0195621981	15	84204786	84204975	-	3.8051116085
ENSG00000225151.10:J037	0.0000230795899187232	0.0497526594	15	84204975	84205230	-	3.8451369211
ENSG00000140511.11:J021	0.0000154484809593399	0.0394512035	15	88880610	88881356	-	-3.6861191719
ENSG00000103365.15:E022	0.000016923	0.041368648	16	23480644	23480770	-	-0.4407768678
ENSG00000140853.15:J149	0.000000125585976320234	0.0020642829	16	57055519	57058064	+	-0.7712136594
ENSG00000102934.10:E008	0.00000130515411717932	0.0095787873	16	57258461	57258584	-	-27.1799475782
ENSG00000141098.13:E003	0.000000999579823909466	0.0082151509	16	67674535	67675899	-	-0.7756594138
ENSG00000141098.13:E015	0.0000206066589544592	0.0464298123	16	67685612	67685802	-	-0.9944647508

ENSG00000157350.13:J021	0.00000638982767318163	0.0237805876	16	70383235	70388366	-	0.7621673587
ENSG00000127774.7:E003	0.0000102372892914162	0.0325688626	17	3668899	3669012	+	-1.5061204734
ENSG00000108518.7:E008	0.00000000070917527483504	0.00002963705	17	4948943	4949014	-	-1.5111160361
ENSG00000108518.7:N017	4.56276483942161E-061	1.799978E-055	17	4948461	4948943	-	-31.3016866219
ENSG00000141696.13:E012	0.00000816290003459676	0.0280017993	17	41807858	41808004	-	1.7837025497
ENSG00000126561.16:E017	0.0000119437338389524	0.0347420796	17	42299750	42299881	+	-0.4450334018
ENSG00000131470.14:J027	0.00000852928598936827	0.0289327073	17	42573166	42573310	-	-5.051709904
ENSG00000167107.12:E047	0.00000187520444906512	0.0129781584	17	50471188	50471400	+	2.1025757479
ENSG00000172809.13:E007	0.00000380505667744055	0.0178698598	17	74203703	74203710	+	-0.2624000286
ENSG00000172809.13:E011	0.0000144929372537846	0.0390095877	17	74203917	74203958	+	-0.3222813483

ENSG00000172809.13:J026	0.00000537770863920532	0.0216476369	17	74203958	74204129	+	-0.3216360122
ENSG00000266714.9:E016	0.000000000001374182027479	0.00000108421	17	75592427	75592463	+	3.0581742216
ENSG00000266714.9:J163	0.000020368393102772	0.0461792443	17	75592301	75592427	+	4.0522166992
ENSG00000266714.9:J164	0.00000000124037742770311	0.00004077668	17	75592301	75592463	+	-1.5050781564
ENSG00000129646.14:J042	0.000017377	0.0415460682	17	76282115	76287191	-	-3.4652915021
ENSG00000173818.17:J094	0.00000505749746608906	0.0207827849	17	80425620	80428595	+	-1.9273708617
ENSG00000141551.14:E023	0.0000159638237392226	0.040215142	17	82249014	82249164	-	0.9023084094
ENSG00000141551.14:E024	0.00000321310403311191	0.0166782506	17	82249164	82249238	-	0.9961021264
ENSG00000235552:E001	0.000000019	0.0005008391	18	6462143	6462151	-	0.9407634473

ENSG00000198858.10:E009	0.000000702929517743127	0.0071102763	19	900828	900952	-	-0.3945795929
ENSG00000198858.10:J032	0.000001186	0.0092407153	19	900952	901421	-	-0.4250872819
ENSG00000186300.12:E009	0.000014507	0.0390095877	19	2852496	2854209	+	-1.0406740133
ENSG00000186300.12:E010	0.00000215620546907294	0.0135017137	19	2854209	2860471	+	1.1101894937
ENSG00000142459.8:J061	0.000000542687754247813	0.0057861222	19	7863280	7863423	+	-2.3288913171
ENSG00000127507.18:J055	0.00000000087422127025856	0.00003135219	19	14755844	14756237	-	-28.3669553426
ENSG00000124299.14:J049	0.000014375271739943 5	0.0390095877	19	33391479	33401720	-	-0.6703645436
ENSG00000123815.12:E012	0.000014656	0.0390095877	19	40703714	40703721	-	-0.7935615382
ENSG00000105053.11:E022	0.000018958137373893 3	0.0434817005	19	49992910	49992952	-	-0.5721490755
ENSG00000057935.13:E027	0.00000360727179495565	0.0177257233	2	42682400	42682589	+	-0.9590829349

ENSG00000116062.15:E022	0.00000445474037822066	0.0193885128	2	47790926	47791123	+	-0.8498875428
ENSG00000115355.16:E017	0.000000914233434991535	0.0078458814	2	55290508	55290654	-	-1.0831457443
ENSG00000035141.8:E011	0.000003072	0.0161586886	2	70301924	70301994	-	-1.0679516997
ENSG00000174501.14:E069	0.000005807	0.0222408338	2	95938986	95939015	-	1.1616443403
ENSG00000174501.14:J159	0.000000484274247101914	0.0057231323	2	95939015	95941159	-	1.3737289905
ENSG00000158411.11:E021	0.000002592	0.0143366152	2	99178289	99178335	-	4.7448863024
ENSG00000135953.11:J030	0.0000211538261565107	0.0468822266	2	102731068	102732023	-	3.8434556335
ENSG00000115827.14:E028	0.0000161538321499454	0.040332745	2	171476859	171476950	+	-1.175550342
ENSG00000115827.14:J056	0.00000786593342863552	0.0272197866	2	171473975	171476859	+	-1.4861932447
ENSG00000197121.15:J067	0.0000179691321615063	0.0424472865	2	196875821	196880075	-	-27.501893923

ENSG00000144579.7:E025	0.000011636	0.0347420796	2	218402348	218402405	+	-0.3723942771
ENSG00000036257.13:E030	0.00000782752314458969	0.0272197866	2	224511353	224511411	-	-0.6376111658
ENSG00000198612.11:E023	0.0000118326190010497	0.0347420796	2	237098618	237098957	+	1.0796179535
ENSG00000144504.15:E034	0.000001579	0.0113252339	2	240529040	240529509	-	0.8991767815
ENSG00000115694.15:J090	0.0000010495369521154	0.0084496935	2	241495701	241496397	-	-0.5457750627
ENSG00000101347.9:E029	0.000000918590475893927	0.0078458814	20	36904249	36905363	-	1.5734303507
ENSG00000101109.12:E023	0.0000217828497356478	0.0477398986	20	45053085	45053180	+	1.7800331208
ENSG00000087586.17:J032	0.00000604124737852761	0.0229156712	20	56373556	56381432	-	-26.9313072493
ENSG00000142173.15:N074	0.00000261641657059298	0.0143366152	21	46124920	46125464	+	-2.9159222154
ENSG00000100097.12:E014	0.00000245686151990444	0.0140465894	22	37678482	37678654	+	-0.6716798396

ENSG00000177096.9:E005	0.00000461372200311929	0.0195707638	22	42077618	42079438	+	-1.5586697337
ENSG00000075275.16:E042	0.0000061634101059256	0.0231564014	22	46534200	46535294	-	5.789233754
ENSG00000164054.15:J054	0.0000000306760802932913	0.0007563437	3	48468982	48469022	-	-26.9002969551
ENSG00000114423.21:J089	0.00000726631970463305	0.0255938595	3	105776542	105853413	-	-0.9520918691
ENSG00000151576.10:E016	0.00000469966954392007	0.0197232632	3	114065329	114065457	+	-1.1015075537
ENSG00000151576.10:E017	0.0000000314935098935416	0.000093892	3	114066227	114066244	+	-1.6708368603
ENSG00000151576.10:E018	0.000000000000877008163277	0.00000086493	3	114066244	114066283	+	-1.7354286619
ENSG00000151576.10:J051	0.000000000006887903725035	0.00000452871	3	114065457	114066227	+	-1.8119179197
ENSG00000151576.10:J052	0.000000000162777770151424	0.00000899677	3	114066283	114067986	+	-1.8129895327
ENSG00000163406.11:J057	0.0000120652708843615	0.0347420796	3	121939495	121940383	+	-1.2989539027

ENSG00000197763.16:E012	0.00000688376040960132	0.02468723	3	126615354	126615462	-	-4.2344749133
ENSG00000114770.17:E055	0.000000313125111251623	0.004260671	3	184017916	184017939	-	-27.6875196364
ENSG00000138814.17:J034	0.0000100398610482698	0.0322004464	4	101026061	101029165	-	-26.8321811028
ENSG00000145348.16:J104	0.0000111525689524257	0.0343719561	4	106252007	106260436	-	-1.4514785534
ENSG00000109586.12:E019	0.000000115850269386279	0.0020298751	4	173315840	173317633	+	1.6175666061
ENSG00000109771.15:E017	0.0000155007478897858	0.0394512035	4	185394778	185395899	-	-3.3787596301
ENSG00000113387.12:E027	0.000000884	0.0078458814	5	32601093	32601130	+	-0.4052183228
ENSG00000197603.14:E001	0.0000109721671935228	0.0343719561	5	37106227	37106977	-	-1.6644795704
ENSG00000250138.4:N002	0.0000044372303831037	0.0193885128	5	69633053	69633860	+	2.1260635862
ENSG00000153113.23:J139	0.0000000001824473400694	0.00000899677	5	96729725	96736171	+	-28.6116308998

ENSG00000176783.15:J072	0.00000946744299526029	0.0308664462	5	179607659	179608284	+	-2.799000073
ENSG00000272168.7:E044	0.0000110699341813596	0.0343719561	6	22147292	22151668	+	-3.2337965133
ENSG00000112312.10:J029	0.0000020892007159289	0.0135017137	6	24784554	24785637	+	-27.6889104202
ENSG00000181315.11:J020	0.00000088441743419799 2	0.0078458814	6	26643728	26658557	-	-27.6200788647
ENSG00000204576.11:J015	0.0000150093028330711	0.0390095877	6	30557450	30561833	+	-4.8779587742
ENSG00000112655.16:E010	0.00000039130918430964 4	0.0051456245	6	43129726	43129829	+	2.7303001695
ENSG00000112655.16:J061	0.0000057281410226567 9	0.0222408338	6	43129264	43129726	+	3.0322789631
ENSG00000146285.14:E017	0.000017726	0.0421252036	6	107744948	107745143	-	-0.5315937484
ENSG00000112282.18:E002	0.00000368449963072141	0.0177257233	6	131586731	131586737	-	27.6079744488
ENSG00000112282.18:E003	0.00000368449963072141	0.0177257233	6	131586737	131586748	-	27.6079744488

ENSG00000146425.11:E003	0.00000360977361986486	0.0177257233	6	158637127	158637205	-	-0.5444383547
ENSG00000146425.11:J011	0.00000164306274702747	0.0115745849	6	158636897	158637127	-	-0.5691211632
ENSG00000075624.15:E034	0.000003514	0.0177257233	7	5530590	5530595	-	-0.2493872169
ENSG00000122674.11:E013	0.000000313210780486228	0.004260671	7	5901704	5902297	+	1.9447814731
ENSG00000010270.13:J032	0.00000920661846984048	0.0302662212	7	38179145	38207446	+	-26.5479125424
ENSG00000233369.7:E015	0.00000445766831705541	0.0193885128	7	73195878	73196062	+	-1.6307652173
ENSG00000233369.7:E022	0.0000113565983970858	0.0347294463	7	73204562	73204604	+	0.6268011923
ENSG00000237513.1:J006	0.000000037355359506648	0.0008668487	7	104941670	104961983	+	-28.0063784186
ENSG00000135269.18:J027	0.0000119818774730064	0.0347420796	7	116210734	116249019	+	-3.4159682432
ENSG00000128534.8:E006	0.000000484605630763841	0.0057231323	7	118184280	118185646	+	0.6145951724

ENSG00000211728.2:E001	0.00000087420838659290 5	0.0078458814	7	142500027	142500117	+	-1.8568932337
ENSG00000055130.16:E023	0.0000214526213665141	0.047278821	7	148797811	148797859	+	-0.5596746309
ENSG00000243836.5:E002	0.0000087045974008008	0.0291008707	7	151409235	151409270	+	-4.0770557337
ENSG00000243836.5:J015	0.00000210518207608984	0.0135017137	7	151409382	151411669	+	-4.8211266935
ENSG00000164733.21:E056	0.0000023868152911179	0.0140465894	8	11864375	11864463	-	-0.8448806728
ENSG00000253616.5:E001	0.0000123485370556836	0.0350052549	8	23071376	23071807	-	26.7706898489
ENSG00000164808.16:E057	0.00000071136391052306	0.0025966196	8	47727252	47727293	+	-0.8055298542
ENSG00000164808.16:J117	0.0000116865486208547	0.0347420796	8	47727293	47728932	+	-0.725602161
ENSG00000047249.18:J059	0.0000122519375862599	0.0350052549	8	53756656	53769617	-	-0.8383445737
ENSG00000172167.8:E024	0.00000052928817666412 2	0.0057861222	8	120497392	120497554	+	-3.4278444177

ENSG00000147996.16:E054	0.000001988	0.0132926858	9	65708901	65708905	+	-2.6972629336
ENSG00000147996.16:E055	0.00000120594617478195	0.0092407153	9	65708905	65708958	+	-2.8553317894
ENSG00000147996.16:J115	0.000000558	0.005792555	9	65708958	65720040	+	-4.1055906646
ENSG00000135048.14:E005	0.0000186107995021443	0.043187236	9	71685055	71685393	-	-0.5117471555
ENSG00000136888.7:J008	0.00000287853449415309	0.0153454285	9	114588118	114592551	+	-5.1265830469
ENSG00000123453.18:E009	0.00000548763782054593	0.0218670172	9	133670583	133670752	-	-1.3150078187
ENSG00000188566.13:E031	0.000022524854261724	0.0490933554	9	137216989	137217009	+	-1.0946743121
ENSG00000073464.12:E016	0.00000284560469042697	0.0153454285	X	10212466	10212653	+	-2.7918946538
ENSG00000073464.12:J033	0.00000649229148593277	0.0239361079	X	10212653	10213680	+	-5.2543119621
ENSG00000102226.9:E036	0.00000523623021288049	0.0212954244	X	47245086	47245325	+	1.8871021261

ENSG00000172943.19:E043	0.000011893	0.0347420796	X	54022258	54022262	-	-1.0071542241
ENSG00000172943.19:J088	0.0000139089303684702	0.0386406737	X	54017821	54022258	-	-0.9678073222
ENSG00000188917.15:J042	0.0000169424820381237	0.041368648	X	101042312	101051643	-	-4.1752120773
ENSG00000101974.14:J051	0.0000210244551984693	0.0468587593	X	139731755	139737915	-	-3.6794149499
ENSG00000196924.16:N149	0.00000045936275660065	0.0057231323	X	154360341	154361307	-	-29.0300970832

Table 2. ClusterProfiler enrichment results of differentially spliced genes.

ID	Description	GeneRatio	P-value	P-adjust	geneID
GO:0005925	focal adhesion	10/125	0.0003402	0.034228	5216/5754/2316/60/5788/6169/4179/8500/26136/2665
GO:0005924	cell-substrate adherens junction	10/125	0.0003540	0.034228	5216/5754/2316/60/5788/6169/4179/8500/26136/2665

GO:0030055	cell-substrate junction	10/125	0.0003904	0.034228	5216/5754/2316/60/5788/6169/4179/8500/26136/2665
GO:0005912	adherens junction	10/125	0.0018545	0.091194	5216/5754/2316/60/5788/6169/4179/8500/26136/2665
GO:0005876	spindle microtubule	3/125	0.0053718	0.165342	1453/6790/8452
GO:0031461	cullin-RING ubiquitin ligase	4/125	0.0096783	0.254539	1642/80067/8452/8454
GO:0032580	Golgi cisterna membrane	3/125	0.0124781	0.256160	55454/6483/51107
GO:0098562	cytoplasmic side of membrane	4/125	0.0138984	0.256160	6231/5788/5530/5778
GO:0042788	polysomal ribosome	2/125	0.0171020	0.256160	6231/6169
GO:0000151	ubiquitin ligase complex	5/125	0.0174446	0.256160	1642/80067/8452/10956/8454
GO:0031985	Golgi cisterna	3/125	0.0245499	0.323302	55454/6483/51107
GO:0005769	early endosome	6/125	0.0269086	0.323302	6642/1183/150368/80230/51107/23062

GO:0005874	microtubule	6/125	0.0330237	0.361885	79659/6993/1453/6790/8452/256364
GO:0009898	cytoplasmic side of plasma	3/125	0.0411643	0.416393	5788/5530/5778
GO:0005795	Golgi stack	3/125	0.0506339	0.446628	55454/6483/51107
GO:0005844	polysome	2/125	0.0611617	0.450531	6231/6169
GO:0005819	spindle	5/125	0.0681756	0.450531	6993/1453/6790/8452/93627
GO:0098552	side of membrane	5/125	0.0935567	0.450531	6231/916/5788/5530/5778
GO:0098554	cytoplasmic side of endoplasmic	1/125	0.0958507	0.450531	6231
GO:0030863	cortical cytoskeleton	2/125	0.1024165	0.450531	2316/60
GO:0097440	apical dendrite	1/125	0.1079219	0.450531	2316
GO:0031941	filamentous actin	1/125	0.1198333	0.463473	2316

GO:0022626	cytosolic ribosome	2/125	0.1330817	0.468233	6231/6169
GO:0030867	rough endoplasmic reticulum	1/125	0.1431848	0.468233	6231
GO:0005788	endoplasmic reticulum lumen	4/125	0.1495499	0.468233	51246/3956/1292/10956
GO:0030018	Z disc	2/125	0.1584353	0.478948	2316/5530
GO:0097431	mitotic spindle pole	1/125	0.1602942	0.479061	6790
GO:0031674	I band	2/125	0.1821909	0.504664	2316/5530
GO:0005938	cell cortex	3/125	0.1982181	0.504664	5216/2316/60
GO:0031463	Cul3-RING ubiquitin ligase	1/125	0.2042799	0.504664	8452
GO:0030133	transport vesicle	4/125	0.2061607	0.504664	55704/23312/51107/1314
GO:0098978	glutamatergic synapse	4/125	0.2103292	0.504664	5216/2316/60/5530

GO:0005813	centrosome	5/125	0.2108417	0.504664	84166/79692/1453/115752/6790
GO:0044448	cell cortex part	2/125	0.2113226	0.504664	2316/60
GO:0043198	dendritic shaft	1/125	0.2149143	0.504664	2316
GO:0022627	cytosolic small ribosomal subunit	1/125	0.2306022	0.522830	6231
GO:0072562	blood microparticle	2/125	0.2582166	0.547669	5216/60
GO:0000922	spindle pole	2/125	0.2706167	0.553753	6790/8452
GO:0008021	synaptic vesicle	2/125	0.2805348	0.554741	23312/51107
GO:0031965	nuclear membrane	3/125	0.2906583	0.562082	51246/3164/23133
GO:0005911	cell-cell junction	4/125	0.3007630	0.573193	916/5754/2316/5530
GO:0070382	exocytic vesicle	2/125	0.3052759	0.577608	23312/51107

GO:0044391	ribosomal subunit	2/125	0.3225089	0.595528	6231/6169
GO:0030017	sarcomere	2/125	0.3249628	0.595528	2316/5530
GO:0048786	presynaptic active zone	1/125	0.3320429	0.595528	8500
GO:0098793	presynapse	4/125	0.3371051	0.595528	23312/60/51107/8500
GO:0097223	sperm part	2/125	0.3371977	0.595528	8452/4179
GO:0015935	small ribosomal subunit	1/125	0.3628148	0.595528	6231
GO:0098589	membrane region	3/125	0.3638395	0.595528	51090/5788/4179
GO:0044449	contractile fiber part	2/125	0.3686812	0.595528	2316/5530
GO:0030016	myofibril	2/125	0.3806410	0.595528	2316/5530
GO:0009897	external side of plasma membrane	2/125	0.3853988	0.595528	916/5788

GO:0005791	rough endoplasmic reticulum	1/125	0.3921799	0.595528	6231
GO:0005884	actin filament	1/125	0.3921799	0.595528	2316
GO:0099056	integral component of presynaptic	1/125	0.3921799	0.595528	51107
GO:0044445	cytosolic part	2/125	0.4042700	0.600695	6231/6169
GO:0043292	contractile fiber	2/125	0.4089457	0.600853	2316/5530
GO:0098889	component of presynaptic	1/125	0.4279703	0.616725	51107
GO:0036126	sperm flagellum	1/125	0.4318156	0.616725	8452
GO:0030672	synaptic vesicle membrane	1/125	0.4431984	0.616725	23312
GO:0099501	exocytic vesicle membrane	1/125	0.4431984	0.616725	23312
GO:0097729	9+2 motile cilium	1/125	0.4506610	0.623809	8452

GO:0005840	ribosome	2/125	0.5047099	0.653499	6231/6169
GO:0099568	cytoplasmic region	3/125	0.5064064	0.653499	5216/2316/60
GO:0005635	nuclear envelope	3/125	0.5359439	0.674417	51246/3164/23133
GO:0030667	secretory granule membrane	2/125	0.5862765	0.707257	5788/4179
GO:0043025	neuronal cell body	3/125	0.5889327	0.707257	2316/6993/6790
GO:0042734	presynaptic membrane	1/125	0.6134033	0.724535	51107
GO:0045121	membrane raft	2/125	0.6153309	0.724535	51090/5788
GO:0098857	membrane microdomain	2/125	0.6170945	0.724535	51090/5788
GO:0099699	component of synaptic membrane	1/125	0.6411392	0.738170	51107
GO:0031514	motile cilium	1/125	0.6646359	0.742259	8452

GO:0099240	omponent of synaptic membrane	1/125	0.6691488	0.742259	51107
GO:0036464	cytoplasmic ribonucleoprotein	1/125	0.6949884	0.755297	60
GO:0045177	apical part of cell	2/125	0.7047840	0.762790	79659/8642
GO:0030658	transport vesicle membrane	1/125	0.7188247	0.771636	23312
GO:0015629	actin cytoskeleton	2/125	0.7885351	0.819702	2316/60
GO:0098797	membrane protein complex	2/125	0.8081767	0.836812	916/51107
GO:0097060	synaptic membrane	1/125	0.9314945	0.938632	51107

Table 3. Co-expressed genes in each module. SFARI genes are represented in the bold red colour.

Modules	Genes
M1	NKG7, GNLY, CCL5, FCGR3A, SELPLG, CST7, **DIP2A**, HLA.DRB1, SPON2, CD69, PPT1, FGFBP2, GZMA, TNFAIP3, JUN, YWHAB, ADGRG1, TNFAIP8, NAPRT, GZMH, PATL2, KLRC4.KLRK1, ANPEP ,DUSP2, CCL4, SLAMF7, **PYHIN1**, CD163, REPIN1, RAB5B, TBX21, PRNP, S1PR5, PCGF3, SULT1A1, MYBL1, **FCRL6**, SYTL2, IL18RAP, UPK3BL1, RNF11, MILR1, ATP6V0E2, FAS, VSTM1, KLRC4, CYP1B1, RHOH, FAM153A, HLA.DQA1, KLRG1, BCL2, TGFBR3, ARFRP1, GFI1, KLRD1, FOLR3, **VASH1**, TMEM164, ZNF683, MYCL, TIGIT, METTL7A, ITPRIP ,CCL3, **THBS1**, MPST, SMAD7, SFMBT2, TM9SF1, P2RY11, HLA.DQA2, LAX1, CYB561D1, NMUR1, ZNF780A, NQO2, NPIPB13, LRRN3, KLRC3, ADGRG5, APOL1, IL18R1, ADTRP ,RGS1, TENT5C, OLIG1, ZNF468, TBKBP1, PRSS23, S100B, MARCO, ROBO3, KLRC2, RABL3, SLC14A1, ANKRD55, FCGBP, TGIF2, EPHA4, AK5, CDC42EP2, PDXP, NLRP2, GNAZ, CLDND2, PNMA3, ALOX12, Cxorf65, FAM219A, LAG3, RBAK, METTL8, LY6G6F, **GSTM1**, AJM1, HPGD, SCART1, ZNF322, KIR3DL1, LIPN, FAM153B, ACOT7, DTHD1, KIR2DL1, ZNF564, ARMC2, **KANK1**, DCHS1, C19orf33, DAG1, GALNT4, ICAM4, SOX13, KIF21A, ITPRIPL1, TMED7.TICAM2, EOGT, APOBEC3H, RTL10, CATSPER1.
M2	MT.ATP6, HBB, HBA1, HBA2, TSC22D3, PRF1, TCF7, MYD88, GNG2, SPN, TSPAN14, TARDBP, LRPAP1, FCRL5, C3AR1, CD8B, CORO1C, MDFIC, PTGDS, UHMK1, ITM2C, ELF2, DMTN,RNF13, MFSD14A, MAF, FAM129A, VPREB3, **MEMO1**, CYB561, CHI3L2, TAX1BP3, ZDHHC12, MYC, NOG, PRDM1, **ADNP**, F2R, ID3, SLC25A24, LMNA, ZNF518A, FAM110A, CR2, CRIP2, PCCB, TTC24, **RAB43**, TTC16, NT5E, RPL17.C18orf32, ARMT1, TRAPPC5, ZNF808, MCOLN2, RNF185, ZNF7, CD59, ZNF664, LYPD2, YPEL2, UMAD1, LPP, PLEKHG4, LGR6, PPP1R12B, CENPV, CTLA4, ZNF23, P2RX5.TAX1BP3, BTBD9, KLHDC8B, CKB, TCEAL4, CD248, HOOK1, ALAS2, LIMS2, STAG3, SSTR3, ABI2, **CACNA1I**, PROK2, PTPRS, KCNG1.
M3	DUSP1, GPSM3, JCHAIN, TPM4, IRF1, NFKBIZ, GBP5, TUBA4A, IGLL5, XAF1, TXNDC5, GBP1, RASSF2, SELENOF, **SCP2**, TMEM230, SLC35E2A, PCMTD1, EGR1, SCO2, CPT1B, NR4A1, NPDC1, PITPNA, MZB1, TRIOBP, AQP9, RAB31, COL18A1, EPSTI1, **CD38**, CXCL8, LRRC8D, PMAIP1, HOPX, DUSP5, CD83, G0S2, CHKB.CP,T1B, FOSB, SIK1, FADS2, **PTGS2**, HLA.DQB1, HBEGF, SIK1B, OSM, GSKIP, DERL3, IDNK, CSF2RA, **NR4A2**, SGMS2, NCAPG2,CLEC11A, FECH, MRPL42, TNNT3, TNFRSF7, XRRA1, ARPIN, ATF3, HOXB3, LGALS9B,NUSAP1, MMAA, LGALS9C, SERPING1, MARC1, RRM2, CLEC5A, RNASE4, LAIR2, TRIM73, ZNF136.

M4	RPL29, AHNAK, KLF13, TMEM176, APOBEC3A, IER2, MX1, IFI44, IFI6, TMEM176A, OAS3, FAM118A, STX16.NPEPL1, TTC32, SAMD9L, HLA.DRB5, FLYWCH1, ZBTB38, NOL12, PCED1B, MDM4,BLOC1S6, CES1, PDZD4, SLC39A1, JUP, **FHIT**, IL1B, IFI44L, DUSP3, IFIT3, IFIT2, **ZNF559**, KCTD18, NPIPA7, **CMPK2**, PTGDR, SENP3.EIF4A1, AGAP9, KRT73, C2orf74, RSAD2, NT5C3B, SNURF, NUDT4, RETN,XCL2, SEPT5, ORAI3, GPR15, NSG1, SIGLEC1, SUSD4, AFDN, SERF1A, KCTD7, ZFP1, ZNF677, SPATA2L, PLCXD1, USP18, IFIT1, CLEC4F, FCGR3B.
M5	PPBP, IFITM3, PF4, SPARC, GP1BB, IDS, ARPC5, CLU, SIRPB2, ITGA2B, ERAP2, GAPT, GZMK, AP1M1, CCR2, HIGD1A, TLR2, GP9, TREML1, ZBED6, ZBTB14, GOPC,**ZC3H11A**, CLEC4E, YTHDF3, KCTD2, SH2D1A, **KLF16**, MPIG6B, HP, RAD50, NBPF14, GAS2L1, CXCR2, SORT1, ZNF44, CHAMP1, ASB1, CRCP, C5AR2, C1QA, TCN2, ZNF182, TUBA8, PTCRA, ENKUR, CCL3L1, ITGB5, PDZK1IP1, MED18, HIST3H2A, HIST1H2BJ, LDOC1, ZBED1, C7orf25, P2RY12, RTKN2, PF4V1.
M6	DEFA1, DEFA3, RPS26,DEFA1B, CGGBP1, LGALS2, ZNF600, LEPROT, TBC1D4, CD151, RPS4Y1, KCTD20, ANKRD49, MRPL49, HDC, LCN2, TSN, LTF, DDX3Y, PDK4, SERPINB8, MMP25, PDP1, KDM5D, DTD1, SLC10A3, USP53, MOCS2, BPI, CAMP, PTMS, MS4A3, PLB1, GPD1L, ZNF226, ZNF138, MPO, PRKY, NME1.NME2, **USP9Y**, ELANE, DEFA4, EIF1AY, FOXP3, C2orf76, PPP2R3B, AZU1, TCN1, CEACAM1, S100P, APOBEC3B, UTY, FAM217B.
M7	CLC, PSMA4, C4orf3,MAPK14, RNF4, H2AFJ, UBE2F, POLR2J3, CD99, CXXC5, TOR1AIP1, RASA4B, CD9, TRIM56, ANKRD52, PGRMC2, DDA1, CALHM6, ZNF672, GPR160, ZNF791, RBM12, POLR2J2, TMEM110.MUSTN1, ZNF595, NKRF,KYNU, PDE6B, ARVCF, MS4A2, PMF1.BGLAP, FCGR1B, SPATA2, JRK, FXYD7, CXCR1, LIMS4.
Not correlated	TAGAP, SKP1, ITPK1, TRAF1, TCTA, WASHC1, HOXB2, FAM162A, CDK2AP1, MRPL30, MRM2, TKTL1, SULT1A4, GSTM4, MRPL53, KYAT1, ARMCX5, ZNF227, MERTK.

APPENDIX 1

Research Ethics Committee Approval

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Sendromik Olmayan Otizm Spektrum Bozukluğu Tanılı İkisi de Etkilenmiş İkiz Çocuk ve Ergenlerin ve Anne Babalarının Genomlarının Tüm Genom Dizileme Yöntemi ile Analizi
VARSA ARAŞTIRMANIN PROTOKOL KODU	
ETİK KURUL PROTOKOL NUMARASI	310-SBKAEK

ETİK KURUL BİLGİLERİ	ETİK KURULUN ADI	Dokuz Eylül Üniversitesi Klinik Araştırmalar Etik Kurulu
	AÇIK ADRESİ:	Dokuz Eylül Üniversitesi Sağlık Yerleşkesi Dekanlık Binası Kat:2 İnciraltı 35340 İZMİR-TÜRKİYE
	TELEFON	0 232 4122254 - 0 232 4122258
	FAKS	0232 4122243
	E-POSTA	etikkurul@deu.edu.tr

BAŞVURU BİLGİLERİ	KOORDİNATÖR/SORUMLU ARAŞTIRMACI UNVANI/ADI/SOYADI	Prof.Dr.Süha MİRAL			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ UZMANLIK ALANI	Çocuk ve Ergen Ruh Sağlığı			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ BULUNDUĞU MERKEZ	Dokuz Eylül Üniversitesi Çocuk ve Ergen Ruh Sağlığı A.D			
	VARSA İDARİ SORUMLU UNVANI/ADI/SOYADI				
	DESTEKLEYİCİ				
	PROJE YÜRÜTÜCÜSÜ UNVANI/ADI/SOYADI (TÜBİTAK vb. gibi kaynaklardan destek alanlar için)				
	DESTEKLEYİCİNİN YASAL TEMSİLCİSİ				
	ARAŞTIRMANIN FAZI VE TÜRÜ	FAZ 1	☐		
		FAZ 2	☐		
		FAZ 3	☐		
		FAZ 4	☐		
		Gözlemsel ilaç çalışması	☐		
		Tıbbi cihaz klinik araştırması	☐		
		İn vitro tıbbi tanı cihazları ile yapılan performans değerlendirme çalışmaları	☐		
		İlaç dışı klinik araştırma	☒		
		Diğer ise belirtiniz			
	ARAŞTIRMAYA KATILAN MERKEZLER	TEK MERKEZ ☒	ÇOK MERKEZLİ ☐	ULUSAL☒	ULUSLARARASI ☐

	Belge Adı	Tarihi	Versiyon Numarası	Dili
DEĞERLENDİRİLEN BELGELER	ARAŞTIRMA PROTOKOLÜ	mevcut		Türkçe ☒ İngilizce ☐ Diğer ☐
	Bilgilendirilmiş Gönüllü Olur Formu	mevcut		Türkçe ☒ İngilizce ☐ Diğer ☐
	OLGU RAPOR FORMU	mevcut		Türkçe ☒ İngilizce ☐ Diğer ☐
	ARAŞTIRMA BROŞÜRÜ			Türkçe ☐ İngilizce ☐ Diğer ☐

	Belge Adı		Açıklama
DEĞERLENDİRİLEN DİĞER BELGELER	SİGORTA	☐	
	ARAŞTIRMA BÜTÇESİ	☐	
	BİYOLOJİK MATERYEL TRANSFER FORMU	☐	
	İLAN	☐	
	YILLIK BİLDİRİM	☐	
	SONUÇ RAPORU	☐	
	GÜVENLİLİK BİLDİRİMLERİ	☐	
	DİĞER:	☐	

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof.Dr.Aysegül Yıldız
İmza:

Not: Etik kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Sendromik Olmayan Otizm Spektrum Bozukluğu Tanılı İkisi de Etkilenmiş İkiz Çocuk ve Ergenlerin ve Anne Babalarının Genomlarının Tüm Genom Dizileme Yöntemi ile Analizi
VARSA ARAŞTIRMANIN PROTOKOL KODU	
ETİK KURUL PROTOKOL NUMARASI	310-SBKAEK

KARAR BİLGİLERİ		
	Karar No:2016/06-01	**Tarih:24.03.2016**
	Yukarıda bilgileri verilen başvuru dosyası ile ilgili belgeler araştırmanın/çalışmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş ve uygun bulunmuş olup araştırmanın/çalışmanın başvuru dosyasında belirtilen merkezlerde gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına toplantıya katılan etik kurul üye tam sayısının salt çoğunluğu ile karar verilmiştir.	

KLİNİK ARAŞTIRMALAR ETİK KURULU	
ETİK KURULUN ÇALIŞMA ESASI	İlaç ve Biyolojik Ürünlerin Klinik Araştırmaları Hakkında Yönetmelik, İyi Klinik Uygulamaları Kılavuzu
BAŞKANIN UNVANI / ADI / SOYADI:	Prof.Dr.Ayşegül Yıldız

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet		Araştırma ile ilişki		Katılım *		İmza
Prof.Dr.Ayşegül YILDIZ	Psikiyatri	DEU Tıp Fakültesi Psikiyatri Anabilim Dalı	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Prof.Dr.Hülya ELLİDOKUZ	Halk Sağlığı	DEU Onkoloji Enstitüsü Prevantif Onkoloji A.D.	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Prof.Dr.Nuray DUMAN	Çocuk Sağlığı ve Hastalıkları (Yeni Doğan)	DEU Tıp Fakültesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Prof.Dr.Hale ÖREN	Çocuk Sağlığı ve Hastalıkları (Çocuk Hematoloji)	DEU Tıp Fakültesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Prof.Dr.A.Necati GÖKMEN	Anesteziyoloji ve Reanimasyon	Anesteziyoloji ve Reanimasyon A.D	E ☒	K ☐	E ☐	H ☒	E ☐	H ☐	
Prof.Dr.Taner DAĞCI	Fizyoloji	Ege Üniversitesi Tıp Fakültesi	E ☒	K ☐	E ☐	H ☒	E ☐	H ☐	
Doç.Dr.Pembe KESKİNOĞLU	Biyoistatistik	DEU Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim A.D	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Doç.Dr.Erdem YAKA	Nöroloji	DEU Tıp Fakültesi Nöroloji A.D	E ☒	K ☐	E ☐	H ☒	E ☐	H ☐	
Doç.Dr.Uğur Önsel TÜRK	Kardiyoloji	Ege Üniversitesi İlaç ve Farmakokinetik Arş-Uyg.Merk	E ☒	K ☐	E ☐	H ☒	E ☐	H ☐	
Doç.Dr.Yasemin BASKIN	Temel Onkoloji	DEU Onkoloji Enstitüsü Temel Onkoloji A.D	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Yard.Doç.Dr.Yasemin ERAÇ	Farmakoloji	Ege Üniversitesi Eczacılık Fakültesi Farmakoloji Anabilim Dalı	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Av.Semra MARMARA	Hukuk	DEÜ Rektörlüğü	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Av.Nazan PEDÜKCOŞKUN	Hukuk	Alsancak Nevvar Salih İşgören Hastanesi	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	
Hayat ALBAYRAK	Emekli	Sağlık Mesleği Mensubu Olmayan Üye	E ☐	K ☒	E ☐	H ☒	E ☐	H ☐	

*:Toplantıda Bulunma

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof.Dr.Ayşegül Yıldız
İmza:

Not: Etik kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

APPENDIX 2

Thesis Approval

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
İZMİR ULUSLARARASI
BİYOTIP VE GENOM ENSTİTÜSÜ

Sayı : 96350690-302.14.01-/180 21 /02/2019
Konu : Tez Konusu

Sayın Kaan OKAY

"Comprehensive transcriptome analysis of dizygotic twins and their parents with autism spectrum disorders" başlıklı tez konunuz 13.02.2019 tarih ve 03/7 sayılı Yönetim Kurulu Kararı ile kabul edilmiştir.

Bilgilerinizi rica ederim.

Prof. Dr. Sefa KIZILDAĞ
Enstitü Müdürü

Ek: 1 YKK

İBG-İzmir
Dokuz Eylül Üniversitesi
İzmir Uluslararası Biyotıp ve Genom Enstitüsü
Sağlık Kampüsü
Balçova 35340 - İzmir
Telefon : 0 (232) 412 6770
Faks : 0 (232) 277 6353

https://www.ibg.deu.edu.tr
ibg@deu.edu.tr

T.C.

DOKUZ EYLÜL ÜNİVERSİTESİ

İZMİR ULUSLARARASI BİYOTIP VE GENOM ENSTİTÜSÜ

ENSTİTÜ YÖNETİM KURULU KARARI

Toplantı Sayısı	2019/03	Toplantı Günü	Çarşamba
Toplantı Tarihi	13.02.2019	Toplantı Saati	12:30

GÜNDEM 7 Enstitümüz Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Moleküler Biyoloji ve Genetik Yüksek Lisans programı 2017850050 numaralı öğrencisi Kaan OKAY'ın tez öneri başvurusuna ilişkin Anabilim Dalı teklifinin görüşülmesi.

Görüşmeler sonunda;

KARAR 7 Enstitümüz Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Moleküler Biyoloji ve Genetik Yüksek Lisans programı 2017850050 numaralı öğrencisi Kaan OKAY'ın tez konusunun "Comprehensive transcriptome analysis of dizygotic twins and their parents with autism spectrum disorders" olmasının uygunluğuna; kararın adı geçen öğrenciye ve danışmanına bildirilmesine oy birliği ile karar verildi.

ASLI GİBİDİR

RAPORTÖR

Meral Yakut Öztürk
Enstitü Sekreter V.