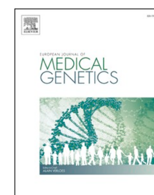
European Journal of Medical Genetics 66 (2023) 104752

Contents lists available at ScienceDirect

European Journal of Medical Genetics

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



Review

The non-coding genome in Autism Spectrum Disorders

S. Dominguez-Alonso [a], A. Carracedo [a,b], C. Rodriguez-Fontenla [a,*]

[a] *Grupo de Medicina Xenómica, Center for Research in Molecular Medicine and Chronic Diseases (CiMUS), Universidad de Santiago de Compostela, Santiago de Compostela, Spain*

[b] *Grupo de Medicina Xenómica, Fundación Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS), Center for Research in Molecular Medicine and Chronic Diseases (CiMUS), Universidad de Santiago de Compostela, Santiago de Compostela, Spain*

ARTICLE INFO

Handling Editor: A. Verloes

Keywords:
Autism spectrum disorders (ASD)
Neurodevelopmental disorders
Non-coding genome
Non-coding variation
Gene regulation

ABSTRACT

Autism Spectrum Disorders (ASD) are a group of neurodevelopmental disorders (NDDs) characterized by difficulties in social interaction and communication, repetitive behavior, and restricted interests. While ASD have been proven to have a strong genetic component, current research largely focuses on coding regions of the genome. However, non-coding DNA, which makes up for ~99% of the human genome, has recently been recognized as an important contributor to the high heritability of ASD, and novel sequencing technologies have been a milestone in opening up new directions for the study of the gene regulatory networks embedded within the non-coding regions. Here, we summarize current progress on the contribution of non-coding alterations to the pathogenesis of ASD and provide an overview of existing methods allowing for the study of their functional relevance, discussing potential ways of unraveling ASD's "missing heritability".

1. Introduction

Autism Spectrum Disorders (ASD) are a group of phenotypically and genetically heterogeneous neurodevelopmental disorders (NDDs) characterized by difficulties in social interaction and communication, repetitive behavior, and restricted interests (MIM209850). The population prevalence of ASD shows regional differences, but has been constantly increasing over the last decades (Ornoy et al., 2016). In Europe, there is an average estimated prevalence of 12.2 per 1000 children aged 7–9 years, ranging from 4.76/1000 in South-Eastern France to March 31, 1000 in Iceland ("Autism Spectrum Disorders in Europe (ASDEU)") (Delobel-Ayoub et al., 2015). ASD are 4 times more common in males than females and have an early onset, with an average age of diagnosis of 4 years (Hicks and Middleton, 2016).

On the basis of a large number of twin studies, it has been shown that ASD have a strong genetic component, with a high heritability estimated around 80% (Sandin et al., 2017). Still, the etiology of ASD has proven complex, and studies suggest that ASD may be considered as a multifactorial disease in which both genetic and environmental factors are involved. Despite ASD's remarkable heterogeneity, the vast majority of the variants proven to be associated with the disease colocalize within specific functional networks, showing a clear convergence in the biological processes affected, including synaptic function and neuronal activity, neuronal cell adhesion, Wnt signaling and chromatin remodeling during neurogenesis (Geschwind and State, 2015).

The first studies, in the 1990s, were mainly focused on the strategy of candidate genes that were selected a priori on the basis of their potential involvement in the etiology of ASD. Since 2005, the emergence of new technologies, such as microarrays, has allowed for the identification of different copy number variations (CNVs) and single nucleotide variations (SNVs). Although promising, it is estimated that CNVs only contribute at about 15% and SNVs at 7% to the disease, so that the cause of the other 75% ASD cases remains difficult to understand (Drivas et al., 2021). Next-generation sequencing (NGS) technologies (WES: Whole Exome Sequencing and WGS: Whole Genome Sequencing) allowed a paradigm shift, enabling simultaneous research of candidate genes, protein-coding genes, and even the entire genome. In fact, the latest WGS analysis has yielded numerous ASD-associated variants that provide new insight into the genomic architecture of ASD (Trost et al., 2022). These findings, especially some structural variants (SVs), would have been difficult to detect without WGS, as they are not detectable by other methods. Moreover, researchers are leveraging published WGS data from parent–child trio data from autism families (C Yuen et al., 2017) or are in the process of specifically analyzing these same WGS data from quad families with a focus on assessing noncoding mutations

* Corresponding author. CiMUS (Center for Research in Molecular Medicine and Chronic Diseases), Av Barcelona 31, 15706, Santiago de Compostela, A Coruña, Spain.

E-mail address: mariacristina.rodriguez.fontenla@usc.es (C. Rodriguez-Fontenla).

https://doi.org/10.1016/j.ejmg.2023.104752
Received 24 October 2022; Received in revised form 10 March 2023; Accepted 19 March 2023
Available online 5 April 2023
1769-7212/© 2023 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

List of abbreviations

ADHD	Attention deficit hyperactivity disorder
APA	Alternative cleavage and polyadenylation
AS	Alternative splicing
ASD	Autism Spectrum Disorders
BELD	Broad enhancer-like domain
CNV	Copy number variation
CRE	Cis-regulatory element
DNM	*De novo* mutation
GWAS	Genome-wide association studies
IRES	Internal ribosome entry site
lncRNA	Long non-coding RNA
miRNA	MicroRNA
mRNA	Messenger RNA
ncRNA	Non-coding RNA
NDD	Neurodevelopmental disorder
NGS	Next-generation sequencing
RBP	RNA-binding proteins
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variations
sQTL	Splicing quantitative trait locus
STR	Short tandem repeat
SZP	Schizophrenia
TAD	Topologically-associated domains
TF	Transcription factor
TRS	Tandem repeat sequence
TSS	Transcription start site
uORF	Upstream open reading frame
UTR	Untranslated region
WES	Whole exome sequencing
WGS	Whole genome sequencing

(Werling et al., 2019) or restricting to a more detailed analysis of CNV mutations (Brandler et al., 2018). With the birth of genome-wide association studies (GWAS), which provides access to a broader spectrum of genetic variation, further single nucleotide polymorphisms (SNPs) and indels were identified (Wang et al., 2009).

However, while hundreds of ASD loci have been identified, the genetic basis of ASD remains elusive, and only a small fraction of ASD patients have been associated with specific genetic variants. Therefore, the significant genetic component to ASD suggested by family and twin studies has not been consistently supported by genetic analyses. This has resulted in a discrepancy, referred to as "missing heritability", in our understanding of the genetic basis of ASD (Masini et al., 2020; Grove et al., 2019; Manolio et al., 2009) (*i.e.* the incomplete explanation of the genetic factors that contribute to the development of ASD). Other factors, such as rare genetic variations, epigenetic changes, gene-gene interactions, or environmental factors, may also play a role in the development of ASD, highlighting the need for continued research to better understand the genetic and non-genetic factors that contribute to the disease.

In contrast, much less is known about the contribution of non-coding elements to ASD risk. Nevertheless, progress is being achieved by virtue of projects aiming to characterize all the functional elements present in the human genome such as ENCODE, Roadmap Epigenomics and FANTOM5. Also, the growing popularity of WGS studies remarks the importance of surveying the non-coding DNA, which makes up for ~99% of the human genome (Devanna et al., 2018). Recent reports have shown that non-coding regulatory DNA variation play an important role in NDDs, finely tuning biological pathways relevant for appropriate brain development (Zhang et al., 2019a; D'haene and Vergult, 2021; Williams et al., 2019). Moreover, latest analyses combining WGS data and homozygosity mapping have been able to identify candidate biallelic variants within human brain-specific regulatory regions for known ASD and neurodevelopmental disease genes as well as new candidate ASD genes (Tuncay et al., 2022). It is known that rare biallelic events are estimated to contribute to 5% of the ASD burden of disease, with that percentage increasing to 10% in females. This has shown that WGS and homozygosity mapping is an effective method for identifying non-coding genomic regions that likely harbor recessive disease-causing mutations, contributing to the understanding of the underlying biology of the disorder and promoting further investigation.

The characteristics of genetic variants lying within non-coding regions (Fig. 1) are mainly the following: i) they would affect gene

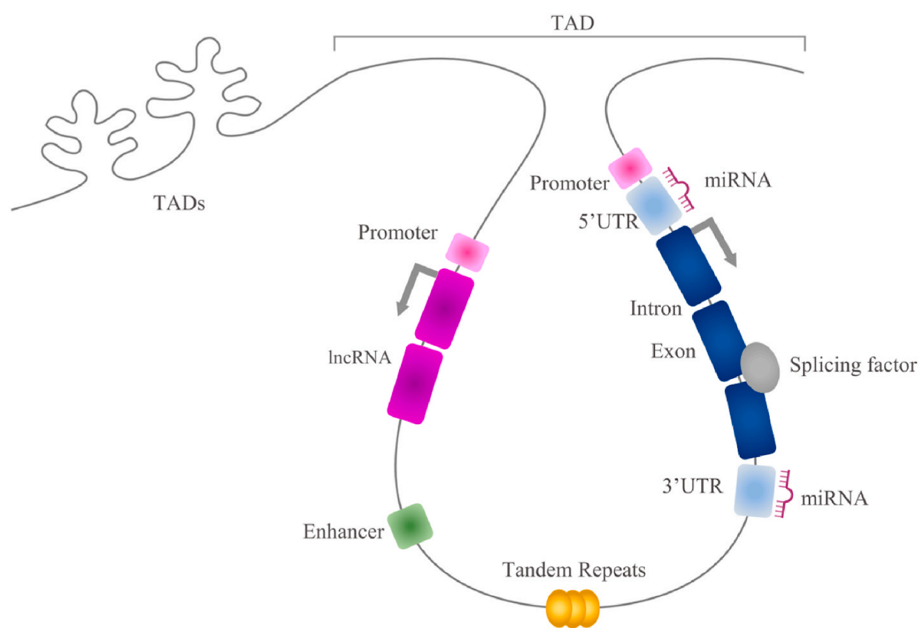


Fig. 1. Non-coding regulatory elements affecting gene expression. Three-dimensional (3D) conformation of the genome plays a fundamental role in gene regulation: DNA is organized into topologically associating domains (TADs), self-interacting genomic regions. Within TADs, enhancers may form chromatin loops with proximal or distal promoters through the action of transcription factors. Transcription factors also mediate tandem-repeat's gene expression regulation. Within a specific gene, variants altering 5′ and 3′ untranslated regions (UTRs), promoters and enhancers also play a role in gene regulation, as well as long non-coding RNAs (lncRNAs) and microRNAs (miRNAs), via different mechanisms. Lastly, splicing factors determine different gene-isoform levels through alternative splicing, a process further discussed in this review.

regulation in some way if they are located within 5′ and 3′UTR (5′ and 3′ Untranslated Regions) and cis-regulatory elements (promoters, enhancers, insulators) and/or ncRNAs (non-coding RNAs), ii) they can only be detected using WGS and specific algorithms, iii) they have been dismissed in previous studies (WES) that do not cover the non-coding genome, iv) they are part of the "missing heritability" in ASD that remains to be discovered.

Post-transcriptional mechanisms, such as ncRNAs, that widely influence gene expression levels without altering the underlying DNA sequence, constitute one means of altering entire gene networks (Hicks and Middleton, 2016). Moreover, the untranslated gene regions, including the 5′ and 3′ untranslated regions and introns, are the major regulators of gene expression and account for up to ~35% of the human genome. Tandem repeats (minisatellite, microsatellite and VNTR) further comprise a large proportion of the human genome; there are over 1 million distinct tandem repeats that can affect the structure and function of DNA, RNA and proteins, affecting a wide range of molecular and cellular processes (Hannan, 2018). Cis-regulatory elements (CREs), such as promoters, enhancers, and insulators, are also accountable for gene transcription regulation, mainly through the recruitment of transcription factors (TFs).

However, the extent by which variants in non-coding elements contribute to ASD risk is poorly understood (Marti et al., 2020). This is partly because only a few WGS studies have been conducted, and techniques for annotating the function of variants in non-coding regions need to be improved. We herein provide an overview of the current progress and challenges in unraveling the non-coding architecture of ASD (see Supplementary Fig. 1 and Supplementary Table 1 for more details).

2. Genetic variants in 5′ and 3′ untranslated regions (UTRs)

The 5′ and 3′ untranslated regions are non-coding regions of DNA that are located at the beginning (5′) and end (3′) of a protein-coding gene. These regions are transcribed into mRNA but are not translated into protein. The 5′ and 3′ UTRs harbor several regulatory elements. These include upstream open reading frames (uORFs), which can act as translational repressors by binding to ribosomes and initiating translation before the main coding sequence, and internal ribosome entry sites (IRESs), which can allow for cap-independent translation initiation. The secondary structure, sequence composition, and length of the UTRs may also act as regulatory elements and can affect the stability and localization of the mRNA. These elements are recognized by RBPs (RNA-binding proteins) or ncRNAS (mainly microRNAs (miRNAs)) in a sequence- or structure-dependent manner, playing a major role in post-transcriptional gene regulation. This includes regulation of mRNA stability, localization, and secondary structure, as well as translational initiation (Steri et al., 2018; Mayr, 2017).

Because of their relevance on these regulatory mechanisms, variants in UTRs can modify molecular pathways and cellular processes, potentially affecting disease onset. In fact, approximately 3.7% of disease-associated variants in the NHGRI-EBI GWAS catalog localize to the

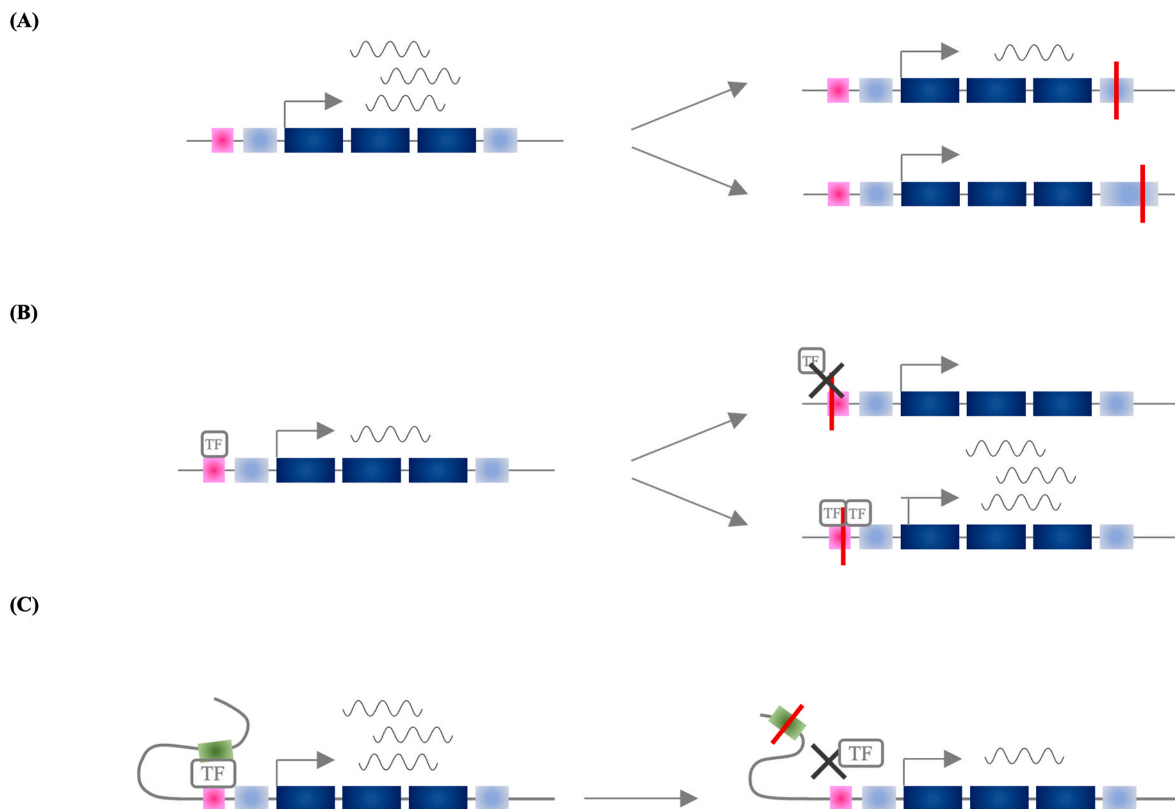


Fig. 2. Examples of noncoding ASD-associated variants affecting proper functioning of untranslated regions (UTRs), promoters and enhancers. Representation of a gene divided into exons (dark blue rectangles). The untranslated regions (UTRs) are represented as light blue rectangles, the promoters are represented in pink and the enhancers are represented in green. mRNA levels are represented as undulating lines. Risk variants are represented as red lines. (A) UTRs. (Top) Variants in the 3′UTR may give rise to a decrease in mRNA stability, with lower mRNA levels and thereby protein abundance (Harrison et al., 2015; Campbell et al., 2011). (Bottom) Changes in 3′UTR length induced by a variant that results in alternative polyadenylation and cleavage. This modification of the 3′UTR length involves changes in gene expression because of the presence of new regulatory elements binding sites or the absence of existing ones (Göpferich et al., 2020). (B) Promoters. Risk variants involving loss (top) or gain (bottom) of transcription factor's binding sites provoke alterations in gene expression. (C) Cis-regulatory elements. Variants harbored in gene enhancers may reduce/ablate their interaction with target gene promoters, reducing mRNA expression (Brandler et al., 2018; Padhi et al., 2020; Turner et al., 2017; Zhao et al., 2018). TF, transcription factor. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

UTRs of protein-coding genes (Steri et al., 2018). 5′ and/or 3′ UTR dysregulation has been linked to ASD in several recent studies (LoParo and Waldman, 2015; Harrison et al., 2015; Campbell et al., 2011; Göpferich et al., 2020) (Fig. 2). For example, SNPs overlapping the 3′ UTR of the oxytocin receptor gene (*OXTR*) may affect the stability of the mRNA due to changes in transcription factor binding (LoParo and Waldman, 2015; Harrison et al., 2015; Campbell et al., 2011). Multiple lines of evidence suggest that *OXTR* variants are associated with ASD owing to the role of oxytocin in modulating social disruptions characteristic of ASD (*i.e.,* regulation of affiliative behavior and social bonding) through the regulation of the underlying neural circuitry.

Moreover, alternative cleavage and polyadenylation (APA) affects 3′UTR length, influencing the presence or absence of regulatory elements, which can be used to control gene expression during neuronal differentiation. Through single cell 3′end sequencing in neural stem cells (NSCs) Göpferich et al., 2020 (Göpferich et al., 2020), detected altered APA in ASD patients, with longer 3′UTRs resulting in a decreased number of neural stem cells.

Altogether, these data provide evidence that variants in UTRs can impact mRNA and/or protein levels through various mechanisms, thus becoming an important risk factor. However, regulation of protein abundance is not the only function of UTR regulation. Future work should focus on how UTR regulation may lead to changes in mRNA's subcellular localization, probably affecting protein function and not just abundance. Another limitation with the studies aforementioned is the broad ASD phenotype, so that aiming to specific autism-related behaviors may help to add clarity to this body of research.

3. Promoters

Core promoters are short sequences that can extend ±50 bp upstream and downstream of the transcription start site (TSS), serving as a binding platform for the RNA Pol II and general transcription factors to initiate transcription (Haberle and Stark, 2018). Numerous complex diseases and disorders have been associated with variants in core promoters, such as Polycystic Kidney disease, Cancer, Alzheimer or Parkinson (Cabezas et al., 2017; Maraganore, de Andrade, Elbaz, Farrer, Ioannidis; Theuns et al., 2006). Regarding autism, several studies revealed association with promoter variation (An et al., 2018; Tansey et al., 2011; Nagarajan et al., 2006), although further investigation is needed in order to explain the specific mechanisms by which gene expression is being altered (Fig. 2).

Shorter repeats (microsatellites) and de novo mutations (DNMs) located in promoter regions have been described as associated with changes in the expression of certain genes that are believed to play a role in the development of the autistic phenotype. For example, Tansey et al.,2011 (Tansey et al., 2011), among many others, analyzed the implications of variations in the promoter of the arginine vasopressin receptor 1A gene (*AVPR1A*) for autism. *AVPR1A* is a key receptor involved in the regulation of social behavior. The researchers found that shorter repeat alleles of the RS1 microsatellite decrease *AVPR1A* transcription, suggesting an increased susceptibility to the autistic phenotype. Also, a WGS study with the largest ASD cohort analyzed for variants in promoters (An et al., 2018) demonstrated an association between DNMs in regulatory regions and ASD. The strongest signal was found in conserved transcription factor binding sites that were distal to the transcription start site (TSS) and correlated with lower nonverbal IQ scores.

Moreover, epigenetic changes of state of promoters have also been associated with ASD. Nagarajan et al., 2006 (Nagarajan et al., 2006), found significantly increased methylation of the *MECP2* (methyl CpG binding protein 2) promoter in the frontal cortex of male individuals with autism compared to controls, leading to reduced MeCP2 protein expression. Furthermore, a rare *MECP2* promoter variant found in an ASD patient was also correlated with reduced expression of *MeCP2*. Because expression defects of MeCP2 were previously found in the autistic brain, these data suggest that both genetic and epigenetic mechanisms affecting promoters may be important in ASD pathology (Göpferich et al., 2020).

So far, studies implicating promoter variants in ASD have analyzed small cohorts, so that larger populations (and different ethnic groups) must be taken into account in order to gain statistical power.

4. Cis-regulatory elements (CREs)

CREs are non-coding elements, such as enhancers, silencers and insulators, that modulate expression of target distal genes and ensure a tight spatio-temporal regulation by the interplay with core promoters (Haberle and Stark, 2018).

Although the core promoter is sufficient to direct transcriptional initiation by the RNA polymerase II, alone it only produces basal levels of mRNA. Enhancers augment the rate of transcription of a cognate gene, while silencers result in the opposite outcome. Insulators act to either block enhancers from targeting genes or to provide a barrier between a protein and associated silencer (Haberle and Stark, 2018).

An increasing number of studies have shown that a high fraction of causative variants in neurodevelopmental disorders are linked to variants in CREs (Mathelier et al., 2015). Moreover, rare inherited and DNMs on CREs have been proven to be causative of ASD (Brandler et al., 2018; Padhi et al., 2020; Turner et al., 2017; Perenthaler et al., 2019; Chen et al., 2021) (Fig. 2). It has been reported that rare inherited structural variants disrupting CREs in genes that are highly dosage sensitive, particularly those with a paternal-origin, put children at risk for ASD (Brandler et al., 2018). An enhancer that targets *EBF3* (a genome-wide associated gene in ASD) has been found to accumulate an excess of DNMs that lead to shared phenotypic characteristics among carrier individuals, including being male, not having intellectual disability, and having hypotonia or motor delay (Padhi et al., 2020). WGS data also reveals a higher burden of missense variants mapping to fetal brain promoters and embryonic stem cell enhancers with an enriched expression in striatal neurons in probands (Turner et al., 2017). These data suggest that variation in CREs contributes substantially to ASD pathogenesis.

A set of long genes containing broad enhancer-like domains (BELDs) have been reported as involved in synapse function and development and highly related to NDDs and ASD. Their expression is mostly downregulated in autistic individuals, thus challenging RNA polymerase activity and resulting in the transcriptional impairment of these BELD genes, contributing to cellular and functional abnormalities found in ASD patients (Zhao et al., 2018).

Undisputedly, it has become evident that investigating the role of the regulatory elements is crucial for understanding common biological pathways concerning ASD as well as functional mechanisms affecting gene function.

5. Non-coding RNAs

Non-coding RNAs (ncRNAs), molecules that do not harbor canonical open reading frames and are therefore not normally translated into proteins, are emerging as important regulators in various neurodevelopmental diseases due to their indispensable roles in brain development. Important types of regulatory non-coding RNAs mainly include microRNA (miRNA), PIWI-interacting RNA (piRNA), small nucleolar RNA (snoRNA), small interfering RNAs (siRNAs), long non-coding RNA (lncRNA), enhancer RNAs (eRNAs) and Circular RNAs (CircRNAs) (Wang et al., 2009; Lekka and Hall, 2018). This review will summarize current knowledge of miRNAs and lncRNAs, both of which have been extensively studied in ASD (Fig. 3).

5.1. Long non-coding RNAs (lncRNA)

LncRNA are a family of long-chain non-coding RNA that are usually

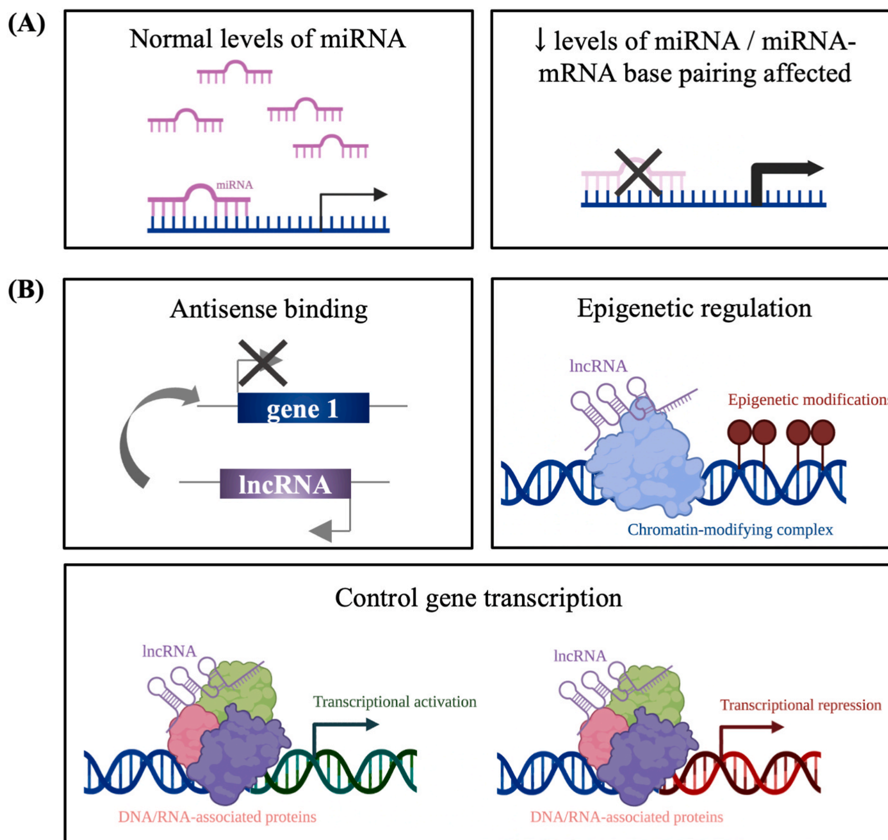


Fig. 3. Examples of the impact of microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) on ASD. (A) Pairing of miRNAs (purple) to their target mRNA (blue). (Left) Normal levels of miRNA prompt decreased levels of target gene through: i) inhibition of target mRNA transcription, ii) silencing of target mRNA translation (D'haene and Vergult, 2021). (Right) Numerous factors (i.e. reduction in miRNA levels or variants affecting miRNA-mRNA base pairing) may impair miRNA-mRNA pairing, thus increasing the levels of target gene expression. (B) (Top left) LncRNAs (purple rectangle) may function in several ways (Zhang et al., 2019a; Lekka and Hall, 2018). One possible role, already reported in ASD, include antisense binding to target gene (blue rectangle), inhibiting its expression (Wang et al., 2009; Luo et al., 2019). (Top right) LncRNAs may also recruit chromatin-modifying to the target gene locus, leading to activation or repression of local genes. (Bottom) They can also act as scaffolds for DNA/RNA-associated pr[58]oteins to regulate transcriptional expression. For example, they can stimulate transcription by bringing key proteins to target gene promoters or inhibit gene transcription by binding to gene repressors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

longer than 200 nt but lack open reading frames, having a low protein-coding potential and representing a large proportion of the transcriptional output of the cell. LncRNA function through heterogeneous mechanisms, conferring additional layers of regulation upon gene expression. They are highly expressed in the central nervous system, especially in the brain, where lncRNAs can have specific spatial and temporal patterns of expression. LncRNAs can regulate gene expression through a variety of mechanisms, including interacting with chromatin complexes to influence epigenetic regulation, functioning as modulators of proteins or protein complexes, binding to DNA/RNA-associated proteins to control transcriptional expression, maintaining DNA stability through R-loop and triple helix formation, and contributing to the formation of higher-order chromatin structure (Graf and Kretz, 2020). Thus, lncRNAs are implicated in a wide variety of developmental and physiological processes, such as neural development and brain function, organelle formation, competition for miRNA binding and regulation of protein localization (Willingham et al., 2005; Cesana et al., 2011). For example, Wang et al., 2009 (Wang et al., 2009), found an ASD genome-wide significant association of a specific genetic variant (rs4307059) on chromosome 5p14.1 ($p = 3.4x10^{-8}$, OR = 1.19). This variant was also associated with social communication in the general population but was not correlated with the expression of nearby protein-coding genes (*CDH9* and *CDH10*). The researchers identified a lncRNA transcribed at the ASD GWAS peak on chromosome 5p14.1, called *MSNP1AS*, which is encoded by the opposite strand of a pseudogene (*MSNP1*). *MSNP1AS* is expressed at higher levels in the temporal cortex of individuals with ASD and in those with the rs4307059 risk allele. It can bind to and decrease the expression of moesin, a protein that regulates neuronal architecture and immune response. This finding suggests that *MSNP1AS*, a functional lncRNA, may contribute to ASD risk. Also, it shows how the study of lncRNAs may explain mechanisms of action underlying previously reported ASD associated variants.

Moreover, among 168 human diseases that have been found to be associated with lncRNAs (recorded in the lncrnadisease database; http://cmbi.bjmu.edu.cn/lncrnadisease) and neurological diseases account for 8.3% of them (Theuns et al., 2006). Numerous studies have found altered lncRNA levels in ASD brains (Luo et al., 2019; Graf and Kretz, 2020; Willingham et al., 2005) (Fig. 3). Ziats and Rennert, 2013 (Ziats and Rennert, 2013), analyzed over 33,000 annotated lncRNAs from postmortem brain tissue of autistic and control prefrontal cortex and cerebellum, detecting 222 differentially expressed lncRNAs in ASD that were enriched for genomic regions containing genes related to neurodevelopment and psychiatric disease. In one study performing a co-expression network analysis on the developing brain transcriptome, a co-enrichment of lncRNA genes and ASD risk genes was found, comprising mostly transcriptional regulators and genes associated with synapse formation (Cogill et al., 2018). In a genome-wide differential expression study of lncRNAs, a total of 3929 lncRNAs were found to be differentially expressed in ASD peripheral leukocytes. Functional pathway analysis of those lncRNAs revealed that neurological pathways involved in synaptic vesicle cycling, long-term depression and long-term potentiation were primarily affected, once again, raising evidence to conclude that not only gene variants at ASD-causing gene loci, but also deregulation of its lncRNAs represents a new approach for exploring possible genetic mechanisms underlying ASD (Wang et al., 2015). Overall, the numerous studies implicating lncRNAs highlight how dysregulation of lncRNAs is an integral component of the transcriptomic signature of ASD.

5.2. miRNA

miRNAs are endogenous non-coding single-stranded RNAs, approximately 22 nucleotides in length, that are capable of regulating gene expression at post-transcriptional level by binding to the 3′UTR of target messenger RNAs (mRNAs). They repress protein production by translational repression or messenger RNA destabilization and decay

(Constantin, 2017). Each miRNA has the ability to bind many mRNAs, simultaneously influencing the expression of hundreds of target genes. Collectively, miRNAs are predicted to target >60% of the transcriptome, potentially modulating the corresponding cellular networks (Wu et al., 2016).

miRNAs are abundant in the brain, with approximately 70% of experimentally detectable miRNAs being expressed in this tissue, playing a vital role in many important brain development facets, such as neuronal survival, function and plasticity (Hicks and Middleton, 2016). Recent studies have involved miRNA in neurogenesis, synaptogenesis, axon outgrowth and guidance, and neuronal migration and integration (Cao et al., 2006, 2016). In addition, accumulating evidence supports the implication of disturbances in the miRNA system in the pathology of neurological diseases, such as fragile X syndrome, schizophrenia (SZP), spinal muscular atrophy and autism (Devanna and Vernes, 2015; Jin et al., 2004; Bădescu et al., 2016). This is because of miRNAs' participation in every step of neural development.

Numerous studies compared the expression of human miRNAs in individuals with ASD and healthy controls, as previously discussed in Hicks and Middleton, 2016) (Hicks and Middleton, 2016)(Fig. 3). Four studies examined postmortem brain tissue (Abu-Elneel et al., 2008; Ander et al., 2015; Mor et al., 2015; Wu et al., 2016), six involved peripheral blood (Huang et al., 2015; Kichukova et al., 2017; Mundalil Vasu et al., 2014; Nakata et al., 2019; Ozkul et al., 2020; Popov et al., 2012), one examined saliva (Hicks et al., 2016), one examined olfactory precursor cells (Nguyen et al., 2016), and three employed cultured lymphoblasts (Sarachana et al., 2010; Talebizadeh et al., 2008; Ghahramani Seno et al., 2011). Altogether, these 15 studies identified more than 200 miRNAs with potential implications in ASD. However, no specific miRNA was uniformly dysregulated across all sample sets. Vaishnavi et al., 2013 (Vaishnavi et al., 2013), focused on autism-associated CNV loci, as these genomic structural variants are a major contributor to the pathophysiology of autism, finding 71 CNV-associated miRNAs (5 of them were previously reported in ASD (Bădescu et al., 2016)).

Also, genetic variants in miRNA genes impact their regulatory function. Brum et al., 2020 (Brum et al., 2021), found that polymorphisms in miRNAs that are involved in neurodevelopment confer higher risk to psychiatric diseases, such as ADHD (Attention deficit hyperactivity disorder) and SZP. However, no difference was observed for ASD, probably due to the fact that studies relating to ASD are restricted by having lower power to detect significant associations. In contrast, Williams et al., 2019 (Williams et al., 2019), identified several rare inherited SNVs within the mature sequence of microRNAs predicted to affect the regulation of ASD-risk genes, such as *NRXN2* and *CNTNAP2*. Moreover, Toma et al., 2015 (Toma et al., 2015), explored the possible contribution of common and rare variants in miRNA genes in ASD. Their study found significant associations with two miRNA clusters: miR-133b/miR-206 (rs16882131, $P = 0.037$) and miR-17/miR-18a/miR-19a/miR-20a/miR19b-1/miR92a-1 (rs6492538, $P = 0.019$). Both miR-133b and miR-206 regulate the *MET* gene, which has been previously associated with ASD. In addition, rare variant analysis identified mutations in several miRNA genes, among them miR-541, a brain-specific miRNA that regulates *SYN1*, which was found to be mutated in individuals with ASD (Toma et al., 2015). Finally, studies in animal models (KO mice for some ASD candidate genes) have enabled a further characterization of 11 miRNAs in correlation with ASD (Schepici and Bramanti, 2019).

6. Repetitive DNA

It is estimated that over half of the human genome, which consists of approximately 3 billion base pairs of DNA, is made up of repetitive elements that make up the repeatome (Hannan, 2018). Repetitive DNA can be classified as tandem and interspersed repeats and copy number variants. Tandem repeat sequences (TRS) and CNVs may occur in the coding, non-coding and regulatory regions of the genes (Siwach, 2008), while interspersed repeats correspond to retrotransposons, discrete pieces of DNA that are reversed transcribed and inserted at new genomic locations (Cordaux and Batzer, 2009). In this section, we will focus on TRS in the non-coding and regulatory regions of the genome (Fig. 4).

6.1. Tandem repeats

TRS are consecutive and head-to-tail copies of a sequence, with a compositional complexity ranging from a few bases to more than a hundred, although the most commonly observed repeats in the genome are homonucleotide, dinucleotide and trinucleotide repeats (Hannan, 2018). Particularly, short tandem repeats (STRs), which consist of repeating units of 1–6 base pair motifs, represent a key component of the repeatome, involving 3% of the genome. Hundreds of STRs have been involved in gene expression regulation, replication, transcription, translation, recombination and chromatin organization (Usdin, 2008). On the other hand, tandem repeats have also been associated with molecular and cellular dysfunction, with implications in more than 50 monogenic diseases (López Castel et al., 2010). Polygenic diseases, such as ASD, are recently being connected with TRS (Mitra et al., 2021; Trost et al., 2020) (Fig. 4): two studies have identified a large number of complementary, non-overlapping tandem-repeat variants associated with fetal brain regulatory regions and genes related to the development of the nervous system (Mitra et al., 2021; Trost et al., 2020). Additionally, Trost et al., 2020 (Trost et al., 2020), aimed to assess STR's possible functional effects and detected specific associations between tandem repeat expansions and particular clinical features such as lower IQ and adaptive ability.

Overall, these results accentuate the importance of considering repeat variants in the study of genetic variants underlying ASD. However, this kind of research faces several limitations: i) the identification of TR variants remains a challenge and requires stringent filtering in order to achieve high validation rates; ii) there is a lack of specific datasets, so that validation heavily relies on simulated data; iii) TRs with complex structures are ambiguous to define and their boundaries highly depend on the choice of parameters used to create the reference, originating huge variance among different methodologies and studies.

Future method improvements, such as long read data, are likely to pinpoint the specific TR variants most relevant to ASD.

7. Splicing variants

90% of human genes that encode proteins undergo pre-mRNA splicing, a complex process by which introns are removed and the exons are aligned and concatenated to form mRNA. While many exons are constitutively spliced together, alternative splicing (AS) is a process during which genes are spliced and otherwise processed in different ways, thus generating numerous transcripts from a single protein-coding gene. AS is a key regulator of gene expression, since precise pre-mRNA splicing is essential for appropriate protein translation. It depends on the presence of consensus cis sequences that define exon-intron boundaries and regulatory sequences recognized by the splicing machinery (Di et al., 2019). Its role in human disease has been widely explored, with an exponential growth on the number of aberrant splicing processes causing human disease (Scotti and Swanson, 2016). Several splicing factors regulating molecular pathways implicated in neurodevelopmental disorders and neural function have been linked to ASD (Smith and Sadee, 2011) (Fig. 4). Further on, over one-third of idiopathic ASD individuals display a disarranging in 30–40% of brain-specific microexons (3–27 nucleotides long exons) (Irimia et al., 2014). Thanks to cell-type specific enrichment and pathway analysis of AS in ASD, Parikshak et al., 2016 (Parikshak et al., 2016), demonstrated that most of the differential splicing events involve exclusion of these microexons. Their results suggest that alterations in transcript structure may affect ASD pathophysiology by changes in the balance of excitation

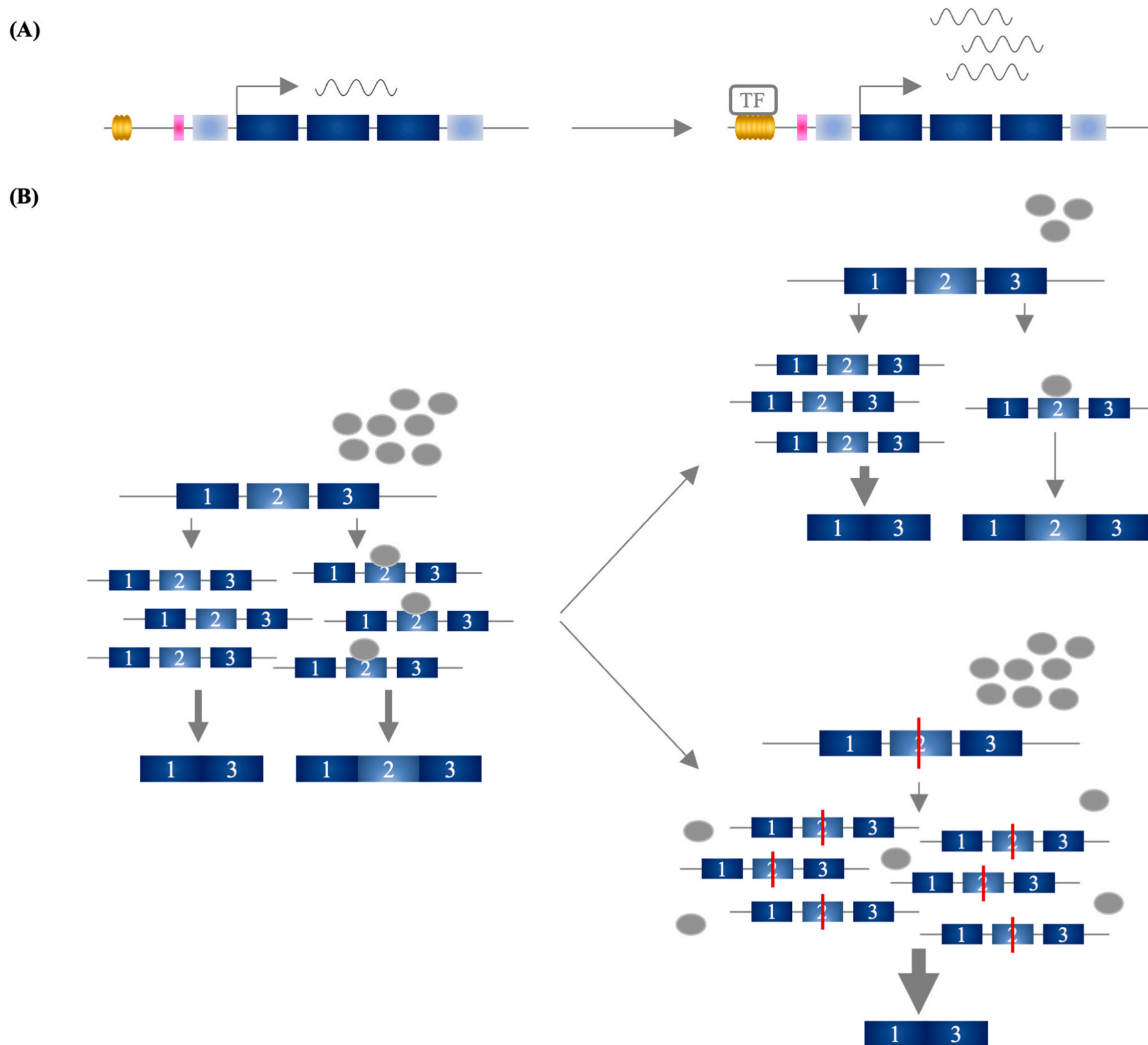


Fig. 4. Examples of noncoding ASD-associated variants interfering tandem repeats and variants affecting proper functioning of the splicing process. Representation of a gene divided into exons (dark blue rectangles). (A) Tandem Repeats. Elongation of tandem repeats (yellow ovals) prompts increased TF binding, and thus, increased mRNA levels/gene expression (represented as undulating lines) (Mitra et al., 2021; Trost et al., 2020). (B) Risk variants are represented as red lines; different exons are enumerated (numbers 1 and 3 represent constitutive exons; number 2 represents an alternative exon with regulatory cis-binding sequence). (Left) Normal levels of trans-acting factor (gray circles) result in typical patterns of spliceoform expression. This means that both spliceoforms (*i.e.*, including the alternative exon and/or both constitutive exons) are produced in regular quantities (represented as gray arrows with the same width). (Top right) Reduced levels of trans-acting factor result in alternative spliced/differentially expressed genes downstream of trans-acting factor (*i.e.*, changes in spliceoform quantities for multiple target RNAs) (Parikshak et al., 2016; Smith and Sadee, 2011; Irimia et al., 2014; Gonatopoulos-Pournatzis et al., 2018). Here, the spliceoform that includes both constitutive exons, but not the alternative one, has a higher expression due to the unavailability of the transcription factor to bind the cis-binding sequence in the alternative exon. (Bottom right) Variants in regulatory cis-binding sequence (harbored in a specific exon/intron) are followed by quantitative/qualitative changes in spliceoform expression of a target RNA (González-Castañeda et al., 2013). Here, the spliceoform including the alternative exon cannot be expressed because trans-acting factors are not able to bind to the mutated cis-binding sequence. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and inhibition and in neuronal activity. Moreover, Pourtnatzis and colleagues (Gonatopoulos-Pournatzis et al., 2018) unraveled 230 genes implicated in regulating these neural microexons, disrupted in autism. Quesnel-Vallières et al., 2016 (Quesnel-Vallières et al., 2016), went a step further and used nSR100/Srrm4 (one of the most important neural microexon regulators) haploinsufficient mice to demonstrate that a single variant in a splicing regulator, and thus the disruption of its target splicing program, is sufficient to reproduce hallmark features of ASD. Last but not least, polymorphisms affecting AS have recently been accounted for an estimated 40–60% of general disease risk in numerous other diseases (Smith and Sadee, 2011). Altogether, these data make it obvious that deregulated RNA splicing is a prominent mechanism underlying ASD etiopathogenesis, and recent findings are likely to represent only the tip of the iceberg.

8. Predicting the functional role of non-coding variation

Despite many recent efforts, in silico definition of the functional effect of non-coding variants functional is still a challenging task (Supplementary Fig. 1).

The most evident role of variants in UTRs relies on their impact on the disruption of existing microRNA binding sites or on the creation of new ones (Fig. 2). Prediction algorithms are mostly based on the imperfect base-pairing of miRNAs with their target sites, although some of them take into account the thermodynamic stability of miRNA-mRNA duplex or evolutionary conservation (Dweep et al., 2013). Some

important algorithms and databases are reviewed in Dweep et al., 2013) (Cordaux and Batzer, 2009); Mayr, 2017 (Mayr, 2017) and Riffo-Campos et al.,2016 (Riffo-Campos et al., 2016), with important applications in deciphering aberrant expression and leading to the determination of critical pathways related to ASD (Devanna and Vernes, 2015; Vaishnavi et al., 2014; Noroozi et al., 2021; Xiong et al., 2015). On the other hand, several algorithms focus on 5′UTRs and how they mediate gene expression, although to the best of our knowledge, they have not yet been applied to the study of ASD. Models to predict the impact of the 5′UTR sequence on protein translation examine the specific levels of ribosome loading (Sample et al., 2019) or analyze the utilization of alternative initiation sites for protein synthesis and how this may alter gene expression levels (Wegrzyn et al., 2008).

The vast majority of promoter variation studies in ASD have shown the implication of genetic variation studying genotype specific SNPs (Tansey et al., 2011; Rudie et al., 2012; Yang et al., 2010) and their role in gene expression using functional studies.

The computational mining of genetic data together with expression data in brain tissue, and the use of different databases providing information on TFBS, chromatin and epigenetic marks, is the basis for the development of multiple bioinformatic tools for predicting the effect of variants in non-coding regions (Daoud et al., 2009). For example, Zhang et al., 2019 (Zhang et al., 2019b), combined GWAS data with brain region related enhancer-gene networks, providing novel insights into the pathogenesis of ASD by identifying ASD-associated pathways. Other successful approaches include the use of machine learning algorithms such as ExPecto (Zhou et al., 2018), which can predict the tissue-specific transcriptional effects of variants on a DNA sequence without the need for functional experiments.

On another level, thanks to the use of enhancer browsers and databases, such as VISTA, FANTOM5 or dbSUPER (reviewed in Perenthaler et al., 2019) (Perenthaler et al., 2019)), along with specific computational approaches to predict functional NCREs (non-coding cis-regulatory elements) in silico, Chen et al., 2016 (Chen et al., 2021), developed a machine learning algorithm called FENRIR that enables functional characterization of tissue specific enhancers and allows for the priorization of enhancer-disease associations. Among the 2723 target genes regulated by the FENRIR-predicted ASD-associated enhancers, significant functional enrichment of pathways and processes previously reported to be associated with ASD was detected. These included neuronal development and synaptic functions, chromatin-related functions, and Wnt signaling pathways.

The use of traditional experimental methods to unravel the specific biological pathways and mechanisms by which lncRNAs affect disease onset are time-consuming and expensive, so that several computational approaches have been developed to overcome this limitation (Zhu et al., 2021). Altogether, these algorithms provide a useful tool to identify disease-associated lncRNAs so that they may be used as potential biomarkers and thus improve disease detection, prognosis or even clinical outcome (Khadirnaikar et al., 2020; Sun et al., 2021). In the context of ASD, research has focused on microarray profiling of annotated lncRNAs, as seen in the studies of Ziats and Rennert, 2013) (Ziats and Rennert, 2013), and Wang et al., 2015 (Wang et al., 2015), genome-wide transcriptome analysis (Ghahramani Seno et al., 2011), or specific variants/lncRNAs previously associated with the disease (Vaishnavi et al., 2013). These studies have concluded their investigations by analyzing overlapping genes or performing co-expression network analysis. Moreover, researchers have started to describe specific pipelines facilitating genome-wide expression profiling, including the identification of novel and rare transcripts like noncoding RNAs and novel alternative splicing isoforms from RNA-Seq data (Arrigoni et al., 2016; Ilott and Ponting, 2013). RNA-Seq is an approach to transcriptome profiling that uses deep-sequencing technologies to detect and accurately quantify RNA molecules originating from a genome at a given moment in time. Therefore, novel lncRNAs will continue to be discovered through RNA-seq in different cell and tissue types and under different environmental conditions. However, the relationship between the sequence, expression, and function of these lncRNAs is not yet well understood, making it difficult to predict their specific functions. Nevertheless, as more lncRNAs are studied through targeted experiments, scientists may be able to use computational methods to predict the functions of previously uncharacterized lncRNAs, as is currently performed for protein-coding genes. Additionally, the identification of new and previously unknown features of lncRNAs may lead to the development of new frameworks to better understand these molecules and predict their functions. As far as we are aware, none of the existing studies on lncRNAs in ASD has tried to computationally predict disease-associated lncRNAs, so that future investigation will certainly include these kind of approaches in order to "explore biomarker identifications for possible clinical screening and diagnosis of ASD", as Wang and colleagues already predicted (Wang et al., 2015).

The read length of STRs makes their accurate detection complicated with current technologies, so the development of novel tools to make this possible is needed. Two main algorithms have been used to estimate repeat lengths in ASD: GangSTR, for repeat unit lengths of 1–20 base pairs (Mousavi et al., 2019), and ExpansionHunter Denovo, optimized for detecting repeats longer than 150 base pairs (Dolzhenko et al., 2020). Because of recent estimations that DNM may contribute up to 30% of simplex ASD cases, Mitra et al., 2021 (Mitra et al., 2021), developed an algorithm called MonSTR that allows for the identification of de novo STR variants, and another bioinformatics tool called SISTR that determines the probability of each STR variant to have a deleterious consequence.

Lastly, several algorithms have been employed for the study of alternative splicing in ASD, allowing for the prioritization of a large number of variants depending on their potential effect on splicing. These include GeneSplicer, MaxEntScan, BDGP/NNSplice or Human Splicing Finder (Piton et al., 2013; Jaganathan et al., 2019). Unlike the aforementioned algorithms, that largely only focus on local motifs, computational tools that account for long-range specificity determinants have much greater performance. Xiong et al., 2015 (Xiong et al., 2015), developed a machine learning algorithm that analyses DNA sequence features (cis elements) and uses them to predict the number of transcripts with the given exon spliced in. The algorithm's implementation in an ASD cohort resulted in ~200 genes with at least one variant predicted to affect splicing. Additionally, they prioritized 19 genes as more compelling ASD disease candidates because of their known neurological, neurobehavioral or neurodevelopmental phenotypes in humans, with just two of them being previously suggested to have a role in ASD pathology. Further on, obtaining tissue-specific models of alternative splicing from particular relevant tissues implicated in the disease may help to gain greater insight into the etiopathology of ASD. Following this intention, Chen et al., 2021 (Cheng et al., 2021), developed Multi-tissue Splicing (MTSplice), that predicts the effect of genetic variants on splicing in 56 human tissues. They assessed the potential of MTSplice on scoring a large number of DNMs in thousands of ASD simplex families. It was found that DNMs from probands disrupted splicing sites in brain tissues more severely than variants of the control group for all tissues. Jaganathan et al., 2019 (Jaganathan et al., 2019), modeled a deep neural network called SpliceAI that predicts the splice function of each position in the pre-mRNA sequence, so that each variant can be scored for its splice-altering impact evaluating 10,000 nucleotides of the flanking context sequence but also taking into account nucleosome positioning (because of its influence in chromatin state) and exon-intron lengths. Due to the wide window analyzed, its accuracy exceeds other algorithms for splice site detection and allowed them to estimate that that 9%–11% of the pathogenic variants in patients with rare genetic disorders are caused by cryptic splicing, a process previously underappreciated. This shows once and for all this field's great potential to unravel ASD genetic architecture.

9. Conclusion and future directions

Non-coding DNA has traditionally been considered as "junk" DNA. However, thanks to recent WGS studies over the past five years, this theory is being completely refuted. The aim of this review was to discuss and summarize the current knowledge on the various regulatory elements involved in the etiology of ASD.

Currently, we have only just begun to explore the intricate functional mechanisms of non-coding elements, mainly due to limitations in traditional genetic technologies and the intrinsic difficulties in studying and interpreting variants lying within the non-coding genome unannotated regions, their interactions, and their phenotypic outcome. Projects like ENCODE greatly helped researchers to further mine the non-coding genome, providing access to a vast registry of candidate regulatory elements across a large number of cell types, and thus, identifying elements in charge of tissue-specific gene expression. Similar strategies, such as a fully comprehensive study on RBPs and how they bind to specific recognition elements on UTR sequences, maybe even in a cell-type or development stage specific manner, would tremendously increase our understanding of UTRs' biology in ASD.

In other neurodevelopmental diseases, such as Prader-Willi syndrome (PWS) and Angelman Syndrome (AS), changes on epigenetic status of promoters have also been observed. Kosaki et al., 1997 (Kosaki et al., 1997), validated a simple Bisulfite-Treated Methylation-Specific PCR Method, a rapid and cost-effective technique, for diagnosis of PWS and AS. The promoter region of the *SNRPN* harbors a CpG island, methylated in the maternally derived allele, present in patients with PWS, and unmethylated in the paternally derived allele, present in AS patients. Using 2 sets of allele-specific primers, both alleles were successfully distinguished. Once progress is made in unraveling promoter genetic and epigenetic modification in ASD, the use of this kind of molecular tests as an initial evaluation of the disease merits considerable attention. Moreover, promoter-enhancer interactions, tuning promoter regulatory action, are just now starting to gain consideration in the research area and new bioinformatic tools allow to identify enhancers and correlates them with expression of the closest neighboring gene, a huge limitation considering that the most proximal gene is not necessarily the real biological target. Even so, thanks to the implementation of this kind of analyses, Yao et al., 2015 (Yao et al., 2015), identified novel networks with prognostic associations in kidney cancer.

Moreover, as we can conclude from experiments on ASD, the study of gene regulation has largely focused on the control of gene activation/increase on gene level expression, but quite few examples exist on the regulation of gene repression/silencing, a process just as critical. Several studies have analyzed these mechanisms in other neurodegenerative diseases, such as Huntington's Disease or Alzheimer Disease, but, to date, none of them have explored the role of silencers in ASD. Because of the lack of genome-wide computational (or experimental) identification of silencers, Doni Jayavelu et al., 2020 (Doni Jayavelu et al., 2020), developed a support vector machine classifier to predict candidate silencers in human and mouse cell/tissue types, confirming silencer activity for nearly half of the uncharacterized CREs. Large-scale functional validation screens are for sure needed, but new sequencing technologies are likely to increase the number as well as the detection of enhancer-disease associations in polygenic disorders. Also, the swingeing part of the literature focuses on the recognition of variants disrupting TFBS, but early steps in considering a gain in TFBS at promoters and enhancers have been introduced by the FunSeq2 software (Fu et al., 2014).

On another note, miRNAs and lncRNAs have emerged as ubiquitous molecules managing all cellular processes. As such, they have become potential biomarkers or drug targets of numerous diseases (Zhang et al., 2019a; Lekka and Hall, 2018; Cao and Zhen, 2018; Ciccacci et al., 2020). Mundalil-Vasu et al., 2014, Hicks et al., 2016 (Mundalil Vasu et al., 2014), found serum miRNAs differentially expressed in ASD patients compared to controls, with some of them having high values for sensitivity and specificity, and Hicks et al., 2016 (Hicks et al., 2016), detected salivary miRNAs with an accuracy of more than 95% in the detection of ASD. Also, they showed that expression patterns of these miRNAs were significantly correlated with several features of adaptive behavior. Altogether, this data shows how miRNAs would result in successful noninvasive biomarkers for ASD. Furthermore, because of their remarkable stability in different body fluids compared to other nucleic acid molecules, reproducibility and resistance to RNase action and other extreme conditions (including boiling temperatures, high or low pH, long storage, and freeze-thaw cycles) (Kichukova et al., 2017; Mundalil Vasu et al., 2014), future work should further explore the possibility of using a miRNA-based ASD diagnosis and prediction of severity along with existing standard developmental screening tests. Additionally, miRNAs' potential could be seen from a therapeutic perspective. As discussed in this review, both up-regulation and down-regulation of specific mi-RNAs have been linked to ASD. The delivery of miRNAs has been proposed as a potential treatment in other diseases, such as cancer and kidney fibrosis (Lu et al., 2018), so that similar approaches could be performed in autism animal or cell models. To do so, more efficient and better tolerated ncRNAs delivery systems need to be developed before clinical trials can be used to determine specific safe doses and therapeutic potentials.

Additionally, accumulating attention has been focused on alternative splicing on ASD. Alternative splicing regulates many biological processes, and variants disrupting this finely tuned mechanism lead to disease. Several studies have looked upon exon skipping as a powerful tool to reverse the effect of missplicing in disease (Zuccato et al., 2007). The most widely studied approach is the use of antisense oligonucleotides (AONs) that target the pre-mRNA at splice sites so that the splicing machinery cannot bind effectively, thus inducing exon skipping. For example, Burghes and McGovern, 2010 (Burghes and McGovern, 2010), generated an AON directed against an intron splice silencer thus restoring levels of *SMN2* (survival motor neuron 2) allowing the correction of spinal muscular atrophy. Studies like this one have made great progress towards clinical trials, and similar approaches could be performed taking into account all the information we have summarized here, much more existing and much more yet to come, in order to fully characterize how alternative splicing results in ASD phenotypes.

Moreover, splicing variants are interconnected with common variation, and thus recent investigations have performed splicing quantitative trait locus (sQTL) analysis to detect these associations. For example, Takata et al., 2017 (Takata et al., 2017), detected a significant enrichment of sQTLs among disease-associated loci previously identified by GWAS, especially related to SZP, revealing four regions with a strong linkage disequilibrium between the index SNP and the sQTL, ultimately implying alternative splicing as the underlying mechanism of the association signal. Although multiple ASD GWAS have been performed, more efforts are required to understand the full extent of implication of the SNPs detected, so that sQTL analysis could shed some light into the functional pathways and mechanisms underlying GWAS results, and thus lead to a better comprehension of the genetic architecture of complex disease traits such as ASD.

In addition, the future study of non-coding variants could be helped by the study of the three-dimensional (3D) conformation of the genome that plays a fundamental role in gene regulation. Thus, how chromatin is packed, folded and organized in the nucleus has relevant implications in the control of gene expression and cell differentiation. Proximity ligation-based chromosome conformation capture (3C) techniques as Hi-C, are used to depict the genome 3D organization through the identification of the interactions between distant loci (Lieberman-Aiden et al., 2009). The 3D genome should be taken into consideration together with the identified genetic variants to decipher the pathogenesis of ASD. Thus, the 3D genome can be especially relevant in the case of non-coding variants that affect gene expression in regulatory regions but also in the case of structural variants that can disrupt TADs (topologically associating domains) and change the spatial location of

enhancer elements (Lupiáñez et al., 2016; Spielmann and Mundlos, 2016).

10. Concluding remarks

Future studies on non-coding regions will for sure shed some light, as current research on regulatory regions is gathering lots of useful information of major relevance for future address of the disorder by means of designing molecular therapies or modulating expression of certain genes relevant for the development of the disease. Because of the increasing social burden and the variability in ASD clinical outcome, there is an imperative need in understanding the missing heritability of ASD, part of it hidden in non-coding regions, and associated biomarkers, to further improve and enable early diagnosis, to date, mainly based in behavioral tests, and to allow the use of therapeutic approaches.

Funding

Fundación María José Jove supported this work. SDA was supported by a Xunta de Galicia predoctoral fellowship. CRF was supported by a contract from the REACT-EU.

CRediT authorship contribution statement

S. Dominguez-Alonso: wrote the manuscript and did the review work, participated in the design of this review. **A. Carracedo:** participated in the design of this review, critically revised the work and approved the final content. **C. Rodriguez-Fontenla:** participated in the design of this review, critically revised the work and approved the final content.

Declaration of competing interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability

No data was used for the research described in the article.

Acknowledgements

We thank the project: Instituto de Salud Carlos III (ISCIII)/PI1900809/Cofinanciado FEDER.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2023.104752.

References

Abu-Elneel, K., Liu, T., Gazzaniga, F.S., et al., 2008. Heterogeneous dysregulation of microRNAs across the autism spectrum. Neurogenetics 9 (3), 153–161. https://doi.org/10.1007/s10048-008-0133-5.

An, J.Y., Lin, K., Zhu, L., et al., 2018. Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. Science 362 (6420), eaat6576. https://doi.org/10.1126/science.aat6576.

Ander, B.P., Barger, N., Stamova, B., Sharp, F.R., Schumann, C.M., 2015. Atypical miRNA expression in temporal cortex associated with dysregulation of immune, cell cycle, and other pathways in autism spectrum disorders. Mol. Autism. 6 (1), 37. https://doi.org/10.1186/s13229-015-0029-9.

Arrigoni, A., Ranzani, V., Rossetti, G., et al., 2016. Analysis RNA-seq and noncoding RNA. Methods Mol Biol Clifton NJ 1480, 125–135. https://doi.org/10.1007/978-1-4939-6380-5_11.

Bădescu, G.M., Fîlfan, M., Sandu, R.E., Surugiu, R., Ciobanu, O., Popa-Wagner, A., 2016. Molecular mechanisms underlying neurodevelopmental disorders, ADHD and autism. Romanian J Morphol Embryol Rev Roum Morphol Embryol 57 (2), 361–366.

Brandler, W.M., Antaki, D., Gujral, M., et al., 2018. Paternally inherited cis-regulatory structural variants are associated with autism. Science 360 (6386), 327–331. https://doi.org/10.1126/science.aan2261.

Brum, C.B., Paixão-Côrtes, V.R., Carvalho, A.M., et al., 2021. Genetic variants in miRNAs differentially expressed during brain development and their relevance to psychiatric disorders susceptibility. World J Biol Psychiatry Off J World Fed Soc Biol Psychiatry 22 (6), 456–467. https://doi.org/10.1080/15622975.2020.1834618.

Burghes, A.H.M., McGovern, V.L., 2010. Antisense oligonucleotides and spinal muscular atrophy: skipping along: Figure 1. Genes Dev. 24 (15), 1574–1579. https://doi.org/10.1101/gad.1961710.

C Yuen, R.K., Merico, D., Bookman, M., et al., 2017. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat. Neurosci. 20 (4), 602–611. https://doi.org/10.1038/nn.4524.

Cabezas, O.R., Flanagan, S.E., Stanescu, H., et al., 2017. Polycystic kidney disease with hyperinsulinemic hypoglycemia caused by a promoter mutation in phosphomannomutase 2. J. Am. Soc. Nephrol. 28 (8), 2529–2539. https://doi.org/10.1681/ASN.2016121312.

Campbell, D.B., Datta, D., Jones, S.T., et al., 2011. Association of oxytocin receptor (OXTR) gene variants with multiple phenotype domains of autism spectrum disorder. J. Neurodev. Disord. 3 (2), 101–112. https://doi.org/10.1007/s11689-010-9071-2.

Cao, T., Zhen, X.C., 2018. Dysregulation of miRNA and its potential therapeutic application in schizophrenia. CNS Neurosci. Ther. 24 (7), 586–597. https://doi.org/10.1111/cns.12840.

Cao, X., Yeo, G., Muotri, A.R., Kuwabara, T., Gage, F.H., 2006. Noncoding RNAs in the mammalian central nervous system. Annu. Rev. Neurosci. 29, 77–103. https://doi.org/10.1146/annurev.neuro.29.051605.112839.

Cao, D.D., Li, L., Chan, W.Y., 2016. MicroRNAs: key regulators in the central nervous system and their implication in neurological diseases. Int. J. Mol. Sci. 17 (6), 842. https://doi.org/10.3390/ijms17060842.

Cesana, M., Cacchiarelli, D., Legnini, I., et al., 2011. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 147 (2), 358–369. https://doi.org/10.1016/j.cell.2011.09.028.

Chen, X., Zhou, J., Zhang, R., et al., 2021. Tissue-specific enhancer functional networks for associating distal regulatory regions to disease. Cell Syst 12 (4), 353–362.e6. https://doi.org/10.1016/j.cels.2021.02.002.

Cheng, J., Çelik, M.H., Kundaje, A., Gagneur, J., 2021. MTSplice predicts effects of genetic variants on tissue-specific splicing. Genome Biol. 22 (1), 94. https://doi.org/10.1186/s13059-021-02273-7.

Ciccacci, C., Latini, A., Colantuono, A., et al., 2020. Expression study of candidate miRNAs and evaluation of their potential use as biomarkers of diabetic neuropathy. Epigenomics 12 (7), 575–585. https://doi.org/10.2217/epi-2019-0242.

Cogill, S.B., Srivastava, A.K., Yang, M.Q., Wang, L., 2018. Co-expression of long non-coding RNAs and autism risk genes in the developing human brain. BMC Syst. Biol. 12 (S7), 91. https://doi.org/10.1186/s12918-018-0639-x.

Constantin, L., 2017. The role of MicroRNAs in cerebellar development and autism spectrum disorder during embryogenesis. Mol. Neurobiol. 54 (9), 6944–6959. https://doi.org/10.1007/s12035-016-0220-9.

Cordaux, R., Batzer, M.A., 2009. The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. 10 (10), 691–703. https://doi.org/10.1038/nrg2640.

Daoud, H., Bonnet-Brilhault, F., Védrine, S., et al., 2009. Autism and nonsyndromic mental retardation associated with a de novo mutation in the NLGN4X gene promoter causing an increased expression level. Biol. Psychiatr. 66 (10), 906–910. https://doi.org/10.1016/j.biopsych.2009.05.008.

Delobel-Ayoub, M., Ehlinger, V., Klapouszczak, D., et al., 2015. In: Triche, E.W. (Ed.), Socioeconomic Disparities and Prevalence of Autism Spectrum Disorders and Intellectual Disability, PLOS ONE, vol. 10, e0141964. https://doi.org/10.1371/journal.pone.0141964, 11.

Devanna, P., Vernes, S.C., 2015. A direct molecular link between the autism candidate gene RORa and the schizophrenia candidate MIR137. Sci. Rep. 4 (1), 3994. https://doi.org/10.1038/srep03994.

Devanna, P., Chen, X.S., Ho, J., et al., 2018. Next-gen sequencing identifies non-coding variation disrupting miRNA-binding sites in neurological disorders. Mol. Psychiatr. 23 (5), 1375–1384. https://doi.org/10.1038/mp.2017.30.

Di, C., Syafrizayanti, Zhang, Q., et al., 2019. Function, clinical application, and strategies of Pre-mRNA splicing in cancer. Cell Death Differ. 26 (7), 1181–1194. https://doi.org/10.1038/s41418-018-0231-3.

Dolzhenko, E., Bennett, M.F., Richmond, P.A., et al., 2020. ExpansionHunter Denovo: a computational method for locating known and novel repeat expansions in short-read sequencing data. Genome Biol. 21 (1), 102. https://doi.org/10.1186/s13059-020-02017-z.

Doni Jayavelu, N., Jajodia, A., Mishra, A., Hawkins, R.D., 2020. Candidate silencer elements for the human and mouse genomes. Nat. Commun. 11 (1), 1061. https://doi.org/10.1038/s41467-020-14853-5.

Drivas, T.G., Lucas, A., Ritchie, M.D., 2021. eQTpLot: a user-friendly R package for the visualization of colocalization between eQTL and GWAS signals. BioData Min. 14 (1), 32. https://doi.org/10.1186/s13040-021-00267-6.

Dweep, H., Sticht, C., Gretz, N., 2013. In-silico algorithms for the screening of possible microRNA binding sites and their interactions. Curr. Genom. 14 (2), 127–136. https://doi.org/10.2174/1389202911314020005.

D'haene, E., Vergult, S., 2021. Interpreting the impact of noncoding structural variation in neurodevelopmental disorders. Genet. Med. 23 (1), 34–46. https://doi.org/10.1038/s41436-020-00974-1.

Fu, Y., Liu, Z., Lou, S., et al., 2014. FunSeq2: a Framework for Prioritizing Noncoding Regulatory Variants in Cancer, p. 15. Published online.

Geschwind, D.H., State, M.W., 2015. Gene hunting in autism spectrum disorder: on the path to precision medicine. Lancet Neurol. 14 (11), 1109–1120. https://doi.org/10.1016/S1474-4422(15)00044-7.

Ghahramani Seno, M.M., Hu, P., Gwadry, F.G., et al., 2011. Gene and miRNA expression profiles in autism spectrum disorders. Brain Res. 1380, 85–97. https://doi.org/10.1016/j.brainres.2010.09.046.

Gonatopoulos-Pournatzis, T., Wu, M., Braunschweig, U., et al., 2018. Genome-wide CRISPR-cas9 interrogation of splicing networks reveals a mechanism for recognition of autism-misregulated neuronal microexons. Mol Cell 72 (3), 510–524.e12. https://doi.org/10.1016/j.molcel.2018.10.008.

González-Castañeda, R.E., Sánchez-González, V.J., Flores-Soto, M., Vázquez-Camacho, G., Macías-Islas, M.A., Ortiz, G.G., 2013. Neural restrictive silencer factor and choline acetyltransferase expression in cerebral tissue of Alzheimer's Disease patients: a pilot study. Genet. Mol. Biol. 36 (1), 25–36. https://doi.org/10.1590/S1415-47572013000100005.

Göpferich, M., George, N.O., Muelas, A.D., et al., 2020. Single cell 3'UTR analysis identifies changes in alternative polyadenylation throughout neuronal differentiation and in autism. Neuroscience. https://doi.org/10.1101/2020.08.12.247627.

Graf, J., Kretz, M., 2020. From structure to function: route to understanding lncRNA mechanism. Bioessays 42 (12), 2000027. https://doi.org/10.1002/bies.202000027.

Grove, J., Ripke, S., Als, T.D., et al., 2019. Identification of common genetic risk variants for autism spectrum disorder. Nat. Genet. 51 (3), 431–444. https://doi.org/10.1038/s41588-019-0344-8.

Haberle, V., Stark, A., 2018. Eukaryotic core promoters and the functional basis of transcription initiation. Nat. Rev. Mol. Cell Biol. 19 (10), 621–637. https://doi.org/10.1038/s41580-018-0028-8.

Hannan, A.J., 2018. Tandem repeats mediating genetic plasticity in health and disease. Nat. Rev. Genet. 19 (5), 286–298. https://doi.org/10.1038/nrg.2017.115.

Harrison, A.J., Gamsiz, E.D., Berkowitz, I.C., Nagpal, S., Jerskey, B.A., 2015. Genetic variation in the oxytocin receptor gene is associated with a social phenotype in autism spectrum disorders. Am J Med Genet B Neuropsychiatr Genet 168 (8), 720–729. https://doi.org/10.1002/ajmg.b.32377.

Hicks, S.D., Middleton, F.A., 2016. A comparative review of microRNA expression patterns in autism spectrum disorder. Front. Psychiatr. 7 https://doi.org/10.3389/fpsyt.2016.00176.

Hicks, S.D., Ignacio, C., Gentile, K., Middleton, F.A., 2016. Salivary miRNA profiles identify children with autism spectrum disorder, correlate with adaptive behavior, and implicate ASD candidate genes involved in neurodevelopment. BMC Pediatr. 16 (1), 52. https://doi.org/10.1186/s12887-016-0586-x.

Huang, F., Long, Z., Chen, Z., et al., 2015. Investigation of gene regulatory networks associated with autism spectrum disorder based on MiRNA expression in China. In: Li, X.J. (Ed.), PLoS One 10 (6), e0129052. https://doi.org/10.1371/journal.pone.0129052.

Ilott, N.E., Ponting, C.P., 2013. Predicting long non-coding RNAs using RNA sequencing. Methods 63 (1), 50–59. https://doi.org/10.1016/j.ymeth.2013.03.019.

Irimia, M., Weatheritt, R.J., Ellis, J.D., et al., 2014. A highly conserved program of neuronal microexons is misregulated in autistic brains. Cell 159 (7), 1511–1523. https://doi.org/10.1016/j.cell.2014.11.035.

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J.F., et al., 2019. Predicting splicing from primary sequence with deep learning. Cell 176 (3), 535–548.e24. https://doi.org/10.1016/j.cell.2018.12.015.

Jin, P., Alisch, R.S., Warren, S.T., 2004. RNA and microRNAs in fragile X mental retardation. Nat. Cell Biol. 6 (11), 1048–1053. https://doi.org/10.1038/ncb1104-1048.

Khadirnaikar, S., Chatterjee, A., Kumar, P., Shukla, S., 2020. A greedy algorithm-based stem cell LncRNA signature identifies a novel subgroup of lung adenocarcinoma patients with poor prognosis. Front. Oncol. 10, 1203. https://doi.org/10.3389/fonc.2020.01203.

Kichukova, T.M., Popov, N.T., Ivanov, I.S., Vachev, T.I., 2017. Profiling of circulating serum MicroRNAs in children with autism spectrum disorder using stem-loop qRT-PCR assay. Folia Med (Plovdiv). 59 (1), 43–52. https://doi.org/10.1515/folmed-2017-0009.

Kosaki, K., McGinniss, M.J., Veraksa, A.N., McGinnis, W.J., Jones, K.L., 1997. Prader-Willi and Angelman syndromes: diagnosis with a bisulfite-treated methylation-specific PCR method. Am. J. Med. Genet. 73 (3), 308–313. https://doi.org/10.1002/(SICI)1096-8628(19971219)73:3<308::AID-AJMG15>3.0.CO;2-N.

Lekka, E., Hall, J., 2018. Noncoding RNAs in disease. FEBS Lett. 592 (17), 2884–2900. https://doi.org/10.1002/1873-3468.13182.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., et al., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326 (5950), 289–293. https://doi.org/10.1126/science.1181369.

LoParo, D., Waldman, I.D., 2015. The oxytocin receptor gene (OXTR) is associated with autism spectrum disorder: a meta-analysis. Mol. Psychiatr. 20 (5), 640–646. https://doi.org/10.1038/mp.2014.77.

López Castel, A., Cleary, J.D., Pearson, C.E., 2010. Repeat instability as the basis for human diseases and as a potential target for therapy. Nat. Rev. Mol. Cell Biol. 11 (3), 165–170. https://doi.org/10.1038/nrm2854.

Lu, M., Xing, H., Xun, Z., et al., 2018. Exosome-based small RNA delivery: progress and prospects. Asian J. Pharm. Sci. 13 (1), 1–11. https://doi.org/10.1016/j.ajps.2017.07.008.

Luo, T., Liu, P., Wang, X., et al., 2019. Effect of the autism-associated lncRNA Shank2-AS on architecture and growth of neurons. J. Cell. Biochem. 120 (2), 1754–1762. https://doi.org/10.1002/jcb.27471.

Lupiáñez, D.G., Spielmann, M., Mundlos, S., 2016. Breaking TADs: how alterations of chromatin domains result in disease. Trends Genet. 32 (4), 225–237. https://doi.org/10.1016/j.tig.2016.01.003.

Manolio, T.A., Collins, F.S., Cox, N.J., et al., 2009. Finding the missing heritability of complex diseases. Nature 461 (7265), 747–753. https://doi.org/10.1038/nature08494.

Maraganore DM, de Andrade M, Elbaz A, Farrer MJ, Ioannidis JP. Collaborative Analysis of -Synuclein Gene Promoter Variability and Parkinson Disease. :10.

Marti, M., Millan, M.I.P., Young, J.I., Walz, K., 2020. Intellectual disability, the long way from genes to biological mechanisms. J Transl Genet Genomics 4 (2), 104–113. https://doi.org/10.20517/jtgg.2020.10.

Masini, E., Loi, E., Vega-Benedetti, A.F., et al., 2020. An overview of the main genetic, epigenetic and environmental factors involved in autism spectrum disorder focusing on synaptic activity. Int. J. Mol. Sci. 21 (21), 8290. https://doi.org/10.3390/ijms21218290.

Mathelier, A., Shi, W., Wasserman, W.W., 2015. Identification of altered cis-regulatory elements in human disease. Trends Genet. 31 (2), 67–76. https://doi.org/10.1016/j.tig.2014.12.003.

Mayr, C., 2017. Regulation by 3′-untranslated regions. Annu. Rev. Genet. 51 (1), 171–194. https://doi.org/10.1146/annurev-genet-120116-024704.

Mitra, I., Huang, B., Mousavi, N., et al., 2021. Patterns of de novo tandem repeat mutations and their role in autism. Nature 589 (7841), 246–250. https://doi.org/10.1038/s41586-020-03078-7.

Mor, M., Nardone, S., Sams, D.S., Elliott, E., 2015. Hypomethylation of miR-142 promoter and upregulation of microRNAs that target the oxytocin receptor gene in the autism prefrontal cortex. Mol. Autism. 6 (1), 46. https://doi.org/10.1186/s13229-015-0040-1.

Mousavi, N., Shleizer-Burko, S., Yanicky, R., Gymrek, M., 2019. Profiling the genome-wide landscape of tandem repeat expansions. Nucleic Acids Res. 47 (15) https://doi.org/10.1093/nar/gkz501 e90-e90.

Mundalil Vasu, M., Anitha, A., Thanseem, I., et al., 2014. Serum microRNA profiles in children with autism. Mol. Autism. 5 (1), 40. https://doi.org/10.1186/2040-2392-5-40.

Nagarajan, R., Hogart, A., Gwye, Y., Martin, M.R., LaSalle, J.M., 2006. Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. Epigenetics 1 (4), 172–182. https://doi.org/10.4161/epi.1.4.3514.

Nakata, M., Kimura, R., Funabiki, Y., Awaya, T., Murai, T., Hagiwara, M., 2019. MicroRNA profiling in adults with high-functioning autism spectrum disorder. Mol. Brain 12 (1), 82. https://doi.org/10.1186/s13041-019-0508-6.

Nguyen, L.S., Lepleux, M., Makhlouf, M., et al., 2016. Profiling olfactory stem cells from living patients identifies miRNAs relevant for autism pathophysiology. Mol. Autism. 7 (1), 1. https://doi.org/10.1186/s13229-015-0064-6.

Noroozi, R., Dinger, M.E., Fatehi, R., Taheri, M., Ghafouri-Fard, S., 2021. Identification of miRNA-mRNA network in autism spectrum disorder using a bioinformatics method. J. Mol. Neurosci. 71 (4), 761–766. https://doi.org/10.1007/s12031-020-01695-5.

Ornoy, A., Weinstein- Fudim, L., Ergaz, Z., 2016. Genetic syndromes, maternal diseases and antenatal factors associated with autism spectrum disorders (ASD). Front. Neurosci. 10 https://doi.org/10.3389/fnins.2016.00316.

Ozkul, Y., Taheri, S., Bayram, K.K., et al., 2020. A heritable profile of six miRNAs in autistic patients and mouse models. Sci. Rep. 10 (1), 9011. https://doi.org/10.1038/s41598-020-65847-8.

Padhi, E.M., Hayeck, T.J., Mannion, B., et al., 2020. De novo mutation in an enhancer of EBF3 in simplex autism. Genomics. https://doi.org/10.1101/2020.08.28.270751.

Parikshak, N.N., Swarup, V., Belgard, T.G., et al., 2016. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. Nature 540 (7633), 423–427. https://doi.org/10.1038/nature20612.

Perenthaler, E., Yousefi, S., Niggl, E., Barakat, T.S., 2019. Beyond the exome: the non-coding genome and enhancers in neurodevelopmental disorders and malformations of cortical development. Front. Cell. Neurosci. 13, 352. https://doi.org/10.3389/fncel.2019.00352.

Piton, A., Jouan, L., Rochefort, D., et al., 2013. Analysis of the effects of rare variants on splicing identifies alterations in GABAA receptor genes in autism spectrum disorder individuals. Eur. J. Hum. Genet. 21 (7), 749–756. https://doi.org/10.1038/ejhg.2012.243.

Popov, N.T., Madjirova, N.P., Minkov, I.N., Vachev, T.I., 2012. Micro RNA HSA-486-3P Gene Expression Profiling in the Whole Blood of Patients with Autism. Biotechnol. Biotechnol. Equip. 26 (6), 3385–3388. https://doi.org/10.5504/BBEQ.2012.0093.

Quesnel-Vallières, M., Dargaei, Z., Irimia, M., et al., 2016. Misregulation of an activity-dependent splicing network as a common mechanism underlying autism spectrum disorders. Mol Cell 64 (6), 1023–1034. https://doi.org/10.1016/j.molcel.2016.11.033.

Riffo-Campos, Á., Riquelme, I., Brebi-Mieville, P., 2016. Tools for sequence-based miRNA target prediction: what to choose? Int. J. Mol. Sci. 17 (12), 1987. https://doi.org/10.3390/ijms17121987.

Rudie, J.D., Hernandez, L.M., Brown, J.A., et al., 2012. Autism-associated promoter variant in MET impacts functional and structural brain networks. Neuron 75 (5), 904–915. https://doi.org/10.1016/j.neuron.2012.07.010.

Sample, P.J., Wang, B., Reid, D.W., et al., 2019. Human 5′ UTR design and variant effect prediction from a massively parallel translation assay. Nat. Biotechnol. 37 (7), 803–809. https://doi.org/10.1038/s41587-019-0164-5.

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Hultman, C., Larsson, H., Reichenberg, A., 2017. The heritability of autism spectrum disorder. JAMA 318 (12), 1182. https://doi.org/10.1001/jama.2017.12141.

Sarachana, T., Zhou, R., Chen, G., Manji, H.K., Hu, V.W., 2010. Investigation of post-transcriptional gene regulatory networks associated with autism spectrum disorders by microRNA expression profiling of lymphoblastoid cell lines. Genome Med. 2 (4), 23. https://doi.org/10.1186/gm144.

Schepici, Cavalli, Bramanti, Mazzon, 2019. Autism spectrum disorder and miRNA: an overview of experimental models. Brain Sci. 9 (10), 265. https://doi.org/10.3390/brainsci9100265.

Scotti, M.M., Swanson, M.S., 2016. RNA mis-splicing in disease. Nat. Rev. Genet. 17 (1), 19–32. https://doi.org/10.1038/nrg.2015.3.

Siwach, P., 2008. Tandem repeats in human disorders: mechanisms and evolution. Front. Biosci. 13, 4467. https://doi.org/10.2741/3017.

Smith, R.M., Sadee, W., 2011. Synaptic signaling and aberrant RNA splicing in autism spectrum disorders. Front. Synaptic Neurosci. 3 https://doi.org/10.3389/fnsyn.2011.00001.

Spielmann, M., Mundlos, S., 2016. Looking beyond the genes: the role of non-coding variants in human disease. Hum. Mol. Genet. 25 (R2), R157–R165. https://doi.org/10.1093/hmg/ddw205.

Steri, M., Idda, M.L., Whalen, M.B., Orrù, V., 2018. Genetic variants in mRNA untranslated regions. WIREs RNA 9 (4). https://doi.org/10.1002/wrna.1474.

Sun, Y., Zhao, H., Zhou, G., Guan, T., Wang, Y., Gao, J., 2021. Random distributed logistic regression framework for predicting potential lncRNA−disease association. In: Wang, Z. (Ed.), J. Mol. Cell Biol. 13 (5), 386–388. https://doi.org/10.1093/jmcb/mjab005.

Takata, A., Matsumoto, N., Kato, T., 2017. Genome-wide identification of splicing QTLs in the human brain and their enrichment among schizophrenia-associated loci. Nat. Commun. 8 (1), 14519 https://doi.org/10.1038/ncomms14519.

Talebizadeh, Z., Butler, M.G., Theodoro, M.F., 2008. Feasibility and relevance of examining lymphoblastoid cell lines to study role of microRNAs in autism. Autism Res. 1 (4), 240–250. https://doi.org/10.1002/aur.33.

Tansey, K.E., Hill, M.J., Cochrane, L.E., Gill, M., Anney, R.J., Gallagher, L., 2011. Functionality of promoter microsatellites of arginine vasopressin receptor 1A (AVPR1A): implications for autism. Mol. Autism. 2 (1), 3. https://doi.org/10.1186/2040-2392-2-3.

Theuns, J., Brouwers, N., Engelborghs, S., et al., 2006. Promoter mutations that increase amyloid precursor-protein expression are associated with alzheimer disease. Am. J. Hum. Genet. 78 (6), 936–946. https://doi.org/10.1086/504044.

Toma, C., Torrico, B., Hervas, A., et al., 2015. Common and rare variants of microRNA genes in autism spectrum disorders. World J. Biol. Psychiatr. 1–11. https://doi.org/10.3109/15622975.2015.1029518. Early Online.

Trost, B., Engchuan, W., Nguyen, C.M., et al., 2020. Genome-wide detection of tandem DNA repeats that are expanded in autism. Nature 586 (7827), 80–86. https://doi.org/10.1038/s41586-020-2579-z.

Trost, B., Thiruvahindrapuram, B., Chan, A.J.S., et al., 2022. Genomic architecture of autism from comprehensive whole-genome sequence annotation. Cell 185 (23), 4409–4427.e18. https://doi.org/10.1016/j.cell.2022.10.009.

Tuncay, I.O., Parmalee, N.L., Khalil, R., et al., 2022. Analysis of recent shared ancestry in a familial cohort identifies coding and noncoding autism spectrum disorder variants. NPJ Genomic Med 7 (1), 13. https://doi.org/10.1038/s41525-022-00284-2.

Turner, T.N., Coe, B.P., Dickel, D.E., et al., 2017. Genomic patterns of de novo mutation in simplex autism. Cell 171 (3), 710–722.e12. https://doi.org/10.1016/j.cell.2017.08.047.

Usdin, K., 2008. The biological effects of simple tandem repeats: lessons from the repeat expansion diseases: Table 1. Genome Res. 18 (7), 1011–1019. https://doi.org/10.1101/gr.070409.107.

Vaishnavi, V., Manikandan, M., Tiwary, B.K., Munirajan, A.K., 2013. Insights on the functional impact of MicroRNAs present in autism-associated copy number variants. In: Bardoni, B. (Ed.), PLoS One 8 (2), e56781. https://doi.org/10.1371/journal.pone.0056781.

Vaishnavi, V., Manikandan, M., Munirajan, A.K., 2014. Mining the 3′UTR of autism-implicated genes for SNPs perturbing MicroRNA regulation. Dev. Reprod. Biol. 12 (2), 92–104. https://doi.org/10.1016/j.gpb.2014.01.003.

Wang, K., Zhang, H., Ma, D., et al., 2009. Common genetic variants on 5p14.1 associate with autism spectrum disorders. Nature 459 (7246), 528–533. https://doi.org/10.1038/nature07999.

Wang, Y., Zhao, X., Ju, W., et al., 2015. Genome-wide differential expression of synaptic long noncoding RNAs in autism spectrum disorder. Transl. Psychiatry 5 (10). https://doi.org/10.1038/tp.2015.144 e660-e660.

Wegrzyn, J.L., Drudge, T.M., Valafar, F., Hook, V., 2008. Bioinformatic analyses of mammalian 5'-UTR sequence properties of mRNAs predicts alternative translation initiation sites. BMC Bioinf. 9 (1), 232. https://doi.org/10.1186/1471-2105-9-232.

Werling, D., Brand, H., An, J.Y., et al., 2019. Limited contribution of rare, noncoding variation to Autism Spectrum Disorder from sequencing of 2,076 genomes in quartet families. Eur. Neuropsychopharmacol 29, S784–S785. https://doi.org/10.1016/j.euroneuro.2017.08.010.

Williams, S.M., An, J.Y., Edson, J., et al., 2019. An integrative analysis of non-coding regulatory DNA variations associated with autism spectrum disorder. Mol. Psychiatr. 24 (11), 1707–1719. https://doi.org/10.1038/s41380-018-0049-x.

Willingham, A.T., Orth, A.P., Batalov, S., et al., 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science 309 (5740), 1570–1573. https://doi.org/10.1126/science.1115901.

Wu, Y.E., Parikshak, N.N., Belgard, T.G., Geschwind, D.H., 2016. Genome-wide, integrative analysis implicates microRNA dysregulation in autism spectrum disorder. Nat. Neurosci. 19 (11), 1463–1476. https://doi.org/10.1038/nn.4373.

Xiong, H.Y., Alipanahi, B., Lee, L.J., et al., 2015. The human splicing code reveals new insights into the genetic determinants of disease. Science 347 (6218), 1254806. https://doi.org/10.1126/science.1254806.

Yang, S.Y., Cho, S.C., Yoo, H.J., et al., 2010. Association study between single nucleotide polymorphisms in promoter region of AVPR1A and Korean autism spectrum disorders. Neurosci. Lett. 479 (3), 197–200. https://doi.org/10.1016/j.neulet.2010.05.050.

Yao, L., Shen, H., Laird, P.W., Farnham, P.J., Berman, B.P., 2015. Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. Genome Biol. 16 (1), 105. https://doi.org/10.1186/s13059-015-0668-3.

Zhang, S.F., Gao, J., Liu, C.M., 2019a. The role of non-coding RNAs in neurodevelopmental disorders. Front. Genet. 10, 1033. https://doi.org/10.3389/fgene.2019.01033.

Zhang, L., Liu, L., Wen, Y., et al., 2019b. Genome-wide association study and identification of chromosomal enhancer maps in multiple brain regions related to autism spectrum disorder: Zhang et al./Autism spectrum disorder. Autism Res. 12 (1), 26–32. https://doi.org/10.1002/aur.2001.

Zhao, Y.T., Kwon, D.Y., Johnson, B.S., et al., 2018. Long genes linked to autism spectrum disorders harbor broad enhancer-like chromatin domains. Genome Res. 28 (7), 933–942. https://doi.org/10.1101/gr.233775.117.

Zhou, J., Theesfeld, C.L., Yao, K., Chen, K.M., Wong, A.K., Troyanskaya, O.G., 2018. Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. Nat. Genet. 50 (8), 1171–1179. https://doi.org/10.1038/s41588-018-0160-6.

Zhu, R., Wang, Y., Liu, J.X., Dai, L.Y., 2021. IPCARF: improving lncRNA-disease association prediction using incremental principal component analysis feature selection and a random forest classifier. BMC Bioinf. 22 (1), 175. https://doi.org/10.1186/s12859-021-04104-9.

Ziats, M.N., Rennert, O.M., 2013. Aberrant expression of long noncoding RNAs in autistic brain. J. Mol. Neurosci. 49 (3), 589–593. https://doi.org/10.1007/s12031-012-9880-8.

Zuccato, C., Belyaev, N., Conforti, P., et al., 2007. Widespread disruption of repressor element-1 silencing transcription factor/neuron-restrictive silencer factor occupancy at its target genes in huntington's disease. J. Neurosci. 27 (26), 6972–6983. https://doi.org/10.1523/JNEUROSCI.4278-06.2007.