



## Article

# Gene Variants Involved in Nonsense-Mediated mRNA Decay Suggest a Role in Autism Spectrum Disorder

Ana Rita Marques <sup>1,2</sup>, João Xavier Santos <sup>1,2</sup>, Hugo Martiniano <sup>1,2</sup>, Joana Vilela <sup>1,2</sup>, Célia Rasga <sup>1,2</sup>,  
Luísa Romão <sup>2,3</sup> and Astrid Moura Vicente <sup>1,2,\*</sup>

- <sup>1</sup> Departamento de Promoção da Saúde e Doenças Não Transmissíveis, Instituto Nacional de Saúde Doutor Ricardo Jorge, Avenida Padre Cruz, 1649-016 Lisboa, Portugal; a.rita.marques@insa.min-saude.pt (A.R.M.); joao.xavier@insa.min-saude.pt (J.X.S.); hugo.martiniano@insa.min-saude.pt (H.M.); joana.vilela@insa.min-saude.pt (J.V.); celia.rasga@insa.min-saude.pt (C.R.)
- <sup>2</sup> BioISI-Biosystems & Integrative Sciences Institute, Faculty of Sciences, University of Lisboa, Campo Grande, C8, 1749-016 Lisboa, Portugal; luisa.romao@insa.min-saude.pt
- <sup>3</sup> Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge, Avenida Padre Cruz, 1649-016 Lisboa, Portugal
- \* Correspondence: astrid.vicente@insa.min-saude.pt

**Abstract:** Autism Spectrum Disorder (ASD) is a heterogeneous neurodevelopmental condition with unclear etiology. Many genes have been associated with ASD risk, but the underlying mechanisms are still poorly understood. An important post-transcriptional regulatory mechanism that plays an essential role during neurodevelopment, the Nonsense-Mediated mRNA Decay (NMD) pathway, may contribute to ASD risk. In this study, we gathered a list of 46 NMD factors and regulators and investigated the role of genetic variants in these genes in ASD. By conducting a comprehensive search for Single Nucleotide Variants (SNVs) in NMD genes using Whole Exome Sequencing data from 1828 ASD patients, we identified 270 SNVs predicted to be damaging in 28.7% of the population. We also analyzed Copy Number Variants (CNVs) from two cohorts of ASD patients (N = 3570) and discovered 38 CNVs in 1% of cases. Importantly, we discovered 136 genetic variants (125 SNVs and 11 CNVs) in 258 ASD patients that were located within protein domains required for NMD. These gene variants are classified as damaging using *in silico* prediction tools, and therefore may interfere with proper NMD function in ASD. The discovery of NMD genes as candidates for ASD in large patient genomic datasets provides evidence supporting the involvement of the NMD pathway in ASD pathophysiology.

**Keywords:** autism spectrum disorder; nonsense-mediated mRNA decay; single nucleotide variants; copy number variants



**Citation:** Marques, A.R.; Santos, J.X.; Martiniano, H.; Vilela, J.; Rasga, C.; Romão, L.; Vicente, A.M. Gene Variants Involved in Nonsense-Mediated mRNA Decay Suggest a Role in Autism Spectrum Disorder. *Biomedicines* **2022**, *10*, 665. <https://doi.org/10.3390/biomedicines10030665>

Academic Editor: Faouzi Baklouti

Received: 30 November 2021

Accepted: 7 March 2022

Published: 13 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

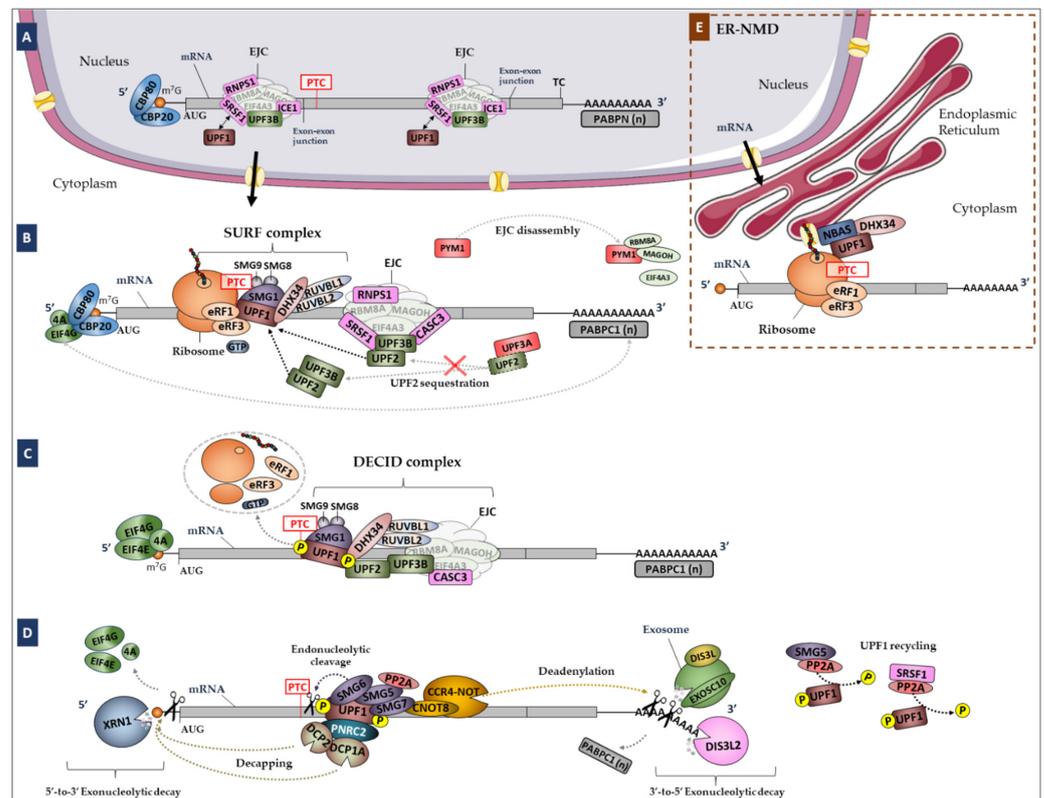
Autism Spectrum Disorder (ASD) is a clinically heterogeneous neurodevelopmental disorder characterized by impaired social communication skills along with repetitive and restricted behaviors and interests [1]. ASD is relatively common, with a median estimated prevalence worldwide of 1–2% [2], and has a major social and economic impacts in families and society. The pathophysiology of this disease is still unclear, precluding the development of effective therapies.

According to heritability studies, genetic factors account for 50 to 80% of the familial ASD risk [3,4], but genetic determinants are still not fully known. Many studies have shown that Copy Number Variants (CNVs) and Single Nucleotide Variants (SNVs) are associated with an increased risk of developing ASD. *De novo*, inherited, rare and common variants in hundreds of genes have been implicated in the etiology of this disease [5–11]. Alterations in gene expression [12] and interactions between genetic and environmental risk factors [13] have also been documented in ASD. These factors are quite variable and

seem to converge in the phenotypic spectrum of autism, but still do not explain all ASD cases. There is also evidence supporting a clinical and genetic overlap between ASD and other Neurodevelopmental Disorders (NDD) [10,14–16], such as Attention Deficit Hyperactivity Disorder (ADHD), Developmental Delay (DD) and Intellectual Disability (ID), and also with Neuropsychiatry Disorders (NPD), including Bipolar Disorder and Schizophrenia [10,17,18]. Considering the high genotypic and phenotypic heterogeneity of ASD, we postulate that important regulatory mechanisms may be modulating genetic expression and clinical presentation. Most ASD candidate genes identified so far are involved in neuronal communication and gene expression regulation, including synaptic proteins, chromatin regulators and transcription factors, and are expressed during brain development [8,10]. Therefore, the discovery of disrupted genes involved in regulatory pathways may be particularly relevant in idiopathic cases.

Nonsense-Mediated mRNA Decay (NMD) is a conserved genetic RNA quality control mechanism in eukaryotes, responsible for targeting mRNAs harboring premature termination codons (PTCs) for degradation (Figure 1). Thus, the NMD pathway is essential to minimizing the expression of truncated proteins that could be toxic to the cell. NMD also regulates the cellular abundance of normal and fully functional mRNAs, targeting ~10% of the mammalian transcriptome [19]. Consequently, NMD plays a critical role during brain development and neuronal differentiation, and contributes to the regulation of synaptic plasticity and cognitive function [20–24]. Disturbance of this control mechanism can lead to several pathologies, including neurodevelopmental disorders [25].

There are several studies suggesting that NMD may be involved in ASD. For instance, mutations in the *UPF3B* gene, a core NMD factor, were identified in patients that display autistic features [14,15,26,27] and missense mutations in this gene were shown to cause NMD impairment leading to altered neuronal differentiation [20,21]. Recently, Kurosaki and colleagues found that loss of the fragile X mental retardation protein (FMRP), the encoded protein of the *FMR1* gene that represents a leading cause of intellectual disability and autism, results in hyperactivated NMD and gene expression misregulation in early stages of neurogenesis [28].



**Figure 1.** Simplified representation of the NMD pathway in mammalian cells. **(A)** The Exon Junction Complex (EJC; composed by eIF4A3, RBM8A and MAGOH) is formed during splicing and deposited 20–24 nucleotides (nts) upstream of exon–exon junctions, remaining associated with the mRNA during its transport to the cytoplasm. **(B)** In the cytoplasm, the mRNA is translated by the ribosome. If the mRNA carries a premature translation termination codon (PTC), during the pioneer round of translation, i.e., while the mRNA is still bound to the cap binding complex CBP80/20, the ribosome stops at the PTC. If the PTC is located more than 50–54 nts upstream of the last exon–exon junction, the ribosome is not able to displace downstream EJCs. Instead, when the ribosome stops at the PTC, UPF1 can interact with eRF3a, inducing premature translation termination that triggers NMD. For that, the SURF complex is formed by UPF1, eRF1, eRF3a, the SMG1c kinase complex and DHX34 associated with RUVBL1 and RUVBL2. At this stage, both PYM1 and UPF3A can act as NMD repressors. **(C)** UPF2 and UPF3B, either diffused in the cytoplasm (in case of EJC-independent NMD) or bound to the EJC, interact with UPF1, favoring its phosphorylation by SMG1 and the formation of the DECID complex. **(D)** Phosphorylated UPF1 recruits SMG6, which cleaves the mRNA near the PTC, and the SMG5-SMG7 dimer, which recruits the decapping complex DPC2-DCP1a through PNRC2 and the deadenylation complex CCR4-NOT through CNOT8. SRSF1 and SMG5-SMG7 independently recruit PP2A, which dephosphorylates UPF1. NMD-targeted mRNAs are further degraded by 5'-to-3' and 3'-to-5' exonucleolytic activities of XRN1 and the exosome, respectively. **(E)** Model of the ER-NMD pathway. The NMD model was adapted from Nogueira et al., 2021 [29]. Legend: NMD, Nonsense-Mediated mRNA Decay; PTC, Premature Termination Codon; EJC, Exon Junction Complex; SURF, SMG1-UPF1-eRF1-eRF3 complex; DECID, Decay Inducing Complex; PP2A, Protein Phosphatase 2A; ER-NMD, NMD response at the Endoplasmic Reticulum; TC, Termination Codon.

In this study, we hypothesized that mutations in genes encoding NMD factors and important NMD regulators (Figure 1) may impair the normal function of this mechanism and be involved in the pathophysiology of ASD. To test this hypothesis, we searched for SNVs and CNVs in 46 key NMD genes in a large cohort of patients diagnosed with ASD. Our results support a role of NMD in ASD, due to the identification of genetic variants located on functional regions of NMD genes in 258 ASD cases. This work reveals novel ASD candidate genes, such as *UPF3A*, *SMG1*, *SMG5*, *SMG7*, *NBAS*, *DHX34* and *ICE1*, which have important functions in the NMD pathway.

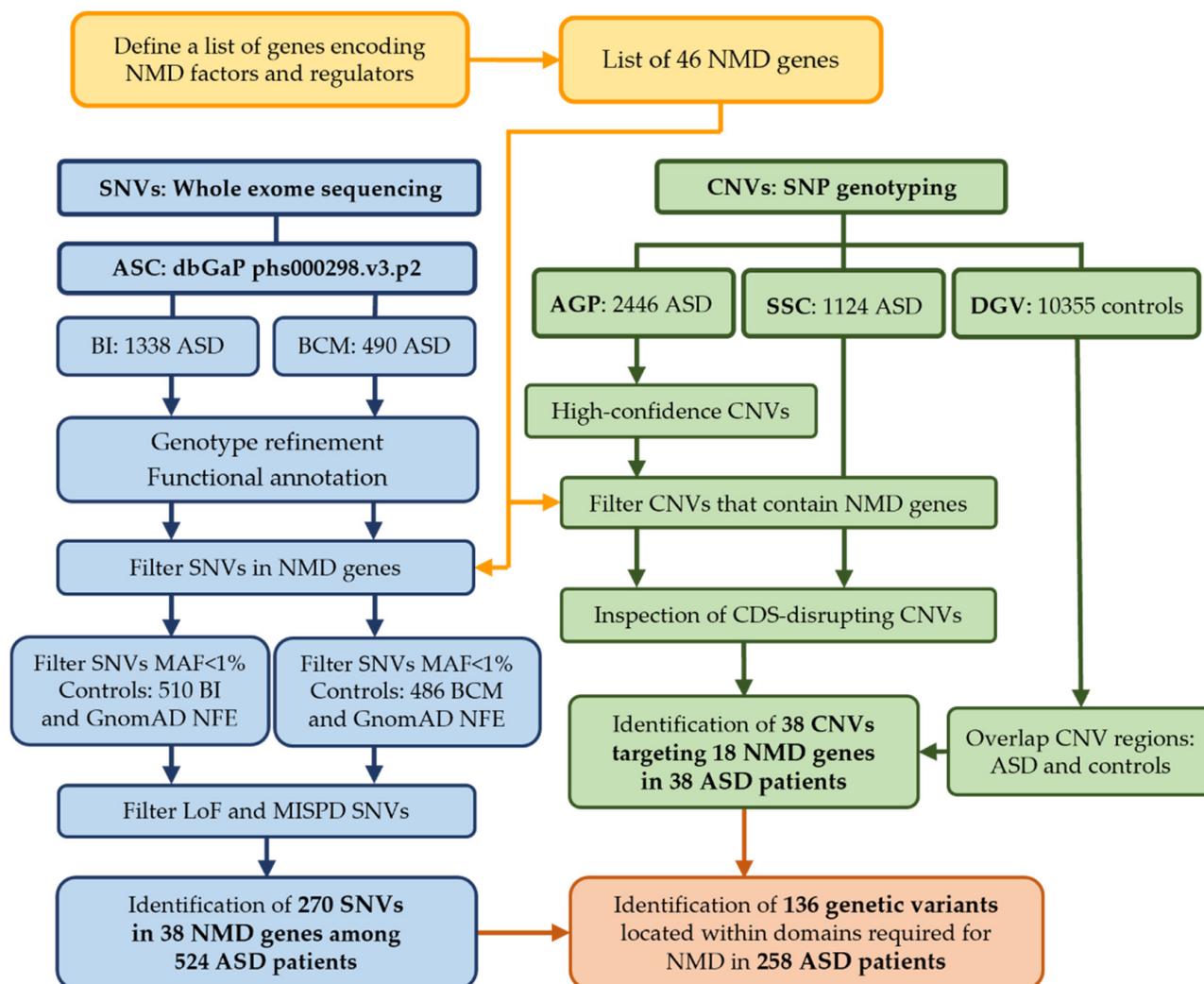
## 2. Materials and Methods

The overall methodology is represented in Figure 2 and is described in detail below.

### 2.1. Identification of Genes Encoding NMD Factors and Regulators

To compile a list of human genes encoding proteins involved in the NMD mechanism, we searched recent reviews in PubMed, using the term “nonsense mediated decay” and article type “Review”. Then we enriched this gene list using the web-based open-access resources AmiGO (<http://amigo.geneontology.org/amigo> [30]) and Reactome (<https://reactome.org/>). Gene Ontology (GO) enrichment analysis was performed by querying the AmiGO database with the term “nonsense mediated decay”, and the output files containing “gene and gene products” results, identified in *Homo sapiens*, were downloaded on 9 September 2020. Reactome was also interrogated using the term “nonsense mediated decay” to retrieve the genes involved in NMD pathway (ID: R-HSA-927802) on 28 September 2020. We identified 122 genes through AmiGO, 116 genes through Reactome (Supplementary Tables S1 and S2) and 33 genes through a literature review, some of which were redundant, as expected. After manually curating this gene list, we ended up with a list of 46 experimentally validated genes that encode proteins involved in NMD (hereafter called NMD genes). Genes were divided into groups according to their functions in NMD: EJC—Exon Junction Complex (EJC) factors and regulators; SURF-DECID—components of SMG1-UPF1-eRF1-eRF3 (SURF) and decay inducing (DECID) complexes; mRNA decay—NMD mRNA decay phase; ER-NMD—NMD response at the Endoplasmic Reticulum (ER);

Regulator—involved in NMD regulation. Gene symbols and names are in accordance with HUGO Gene Nomenclature Committee (HGNC) guidelines [31].



**Figure 2.** Overview of the analysis. The study began with the identification of 46 genes (Table 1) encoding proteins involved in the NMD pathway or in its regulation (yellow) that were used for further analyses. This workflow describes the analysis of SNVs obtained from ASC WES datasets for 1828 ASD patients (1338 BI and 490 BCM, blue) and the analysis of CNVs predicted from SNP genotyping data in 3570 ASD patients from AGP (2446) and SSC (1124) datasets (green). Analysis proceeded separately for CNVs and SNVs to identify variants in NMD genes and ended up with the identification of 270 SNVs (Figure 3) and 38 CNVs (Table 2) in 524 and 38 ASD probands, respectively. Protein domains affected by SNVs or CNVs were then identified (orange), and a total of 136 genetic variants were located within regions required for NMD in 258 ASD patients (Table 3). Legend: SNVs, Single Nucleotide Variants; CNVs, Copy Number Variants; ASC, Autism Sequencing Consortium; AGP, Autism Genome Project; SSC, Simons Simplex Collection; BI, Broad Institute; BCM, Baylor College of Medicine; DGV, Database of Genomic Variant; CDS, coding sequence for protein; MAF, Minor Allele Frequency; NFE, Non-Finnish European; MISPD, missense variants predicted to be damaging and deleterious; LoF, loss-of-function.

## 2.2. ASD Genomic Datasets

To identify genetic variants within NMD genes in ASD patients, we analyzed SNV and CNV data from ASD and control populations published in previous studies.

The SNV dataset was obtained from the Autism Sequencing Consortium (ASC) [32] deposited in dbGaP (ARRA, Autism Sequencing Collaboration; dbGaP Study Accession: phs000298.v3.p2). This dataset contains Whole Exome Sequencing (WES) data from two sources: 490 ASD cases and 486 unrelated controls sequenced using the Solid platform and called with AtlasSNP 2 at Baylor College of Medicine (BCM) [33], and 1338 ASD cases and 510 unrelated controls sequenced using the Illumina platform and called with GATK at Broad Institute (BI) [8]. In total, we analyzed WES for 1828 ASD patients (1451 males and 356 females).

The CNV datasets of ASD patients were obtained from two different studies: 2446 patients (2114 males and 322 females) from The Autism Genome Project (AGP) [6,34] and 1224 patients (967 males and 157 females) from the Simons Simplex Collection (SSC) [35]. We analyzed the CNV calls from the AGP Consortium deposited in dbGaP (Stage I and II, dbGaP Study Accession: phs000267.v5.p2) and a dataset of rare and high-confidence CNVs that passed quality control from the study by Sanders et al., 2011 [5]. Two CNV datasets from unrelated controls (Cooper 2011 and Shaikh 2009) were inspected for variant frequencies (10,355 individuals with no history of neuropsychiatric disease) in the Database of Genomic Variant (DGV: <http://dgv.tcag.ca/dgv/app/home>, last updated on 15 May 2016). ASD patients from the ASC, AGP and SSC met the criteria for autism or ASD using the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) [6,32,34,35].

### 2.3. Sequencing Data Processing, Annotation and SNV Discovery

To exclude artifacts often observed due to sequencing errors, we applied quality filtering and genotype refinement to the raw VCFs using Genome Analysis Toolkit (GATK, version 3.7) for BCM and BI separately [36]. Multi-allelic sites were split into bi-allelic sites, and we only considered variants with a minimum depth filter ( $DP > 8$ ) and genotype quality ( $GQ > 20$ ) for all events. We only considered SNVs in ASD cases that were rare in controls from ASC ( $MAF < 1\%$ ). After evaluation of data quality, variants were functionally annotated with Variant Effect Predictor (VEP, version 86) [37] using the human genome build 37 (GRCh37/hg19) as reference. VEP assigned properties to the variants, including gene name; consequence type; pathogenicity predictions (Polyphen, SIFT, Combined Annotation-Dependent Depletion (CADD) [38]; probability of being loss-of-function intolerant (pLI) [39] and missense Z (mis\_Z) [40] scores); and the allele frequencies observed in controls from the Genome Aggregation Database (gnomAD, v2.2.1). For this work we focused on SNVs in NMD genes identified in ASD patients. Variants with predicted damaging effects were classified as loss-of-function (LoF—variants which include frameshifts, stop gains, stop losses and splicing  $\leq 2$  bp) or probably damaging and deleterious missense (MISPD) variants (defined by Polyphen [41] and SIFT [42]). We further investigated whether these variants were present in controls from the sub-population of Non-Finnish European (NFE) in gnomAD, and further analyzed only the variants that were rare ( $MAF < 1\%$ ) or not detected in gnomAD.

### 2.4. CNV Discovery

CNV discovery was previously performed using Illumina SNP genotyping data as described in the original papers [5,6,34]. Only rare CNVs ( $<1\%$  frequency) called by at least two of three algorithms were considered in this analysis. First, we identified CNVs encompassing NMD genes. Then we manually inspected these CNVs to characterize the extension of the deletions or amplifications encompassing the coding sequence (CDS) of each NMD gene. Finally, we assessed whether these CNVs were also observed in DGV controls.

### 2.5. Protein Domains

There are known conserved protein domains that display important functions in NMD. To assess whether the variants predicted to be damaging in this study impact important

protein domains of NMD genes, SNVs and CNVs were manually curated, and protein domains were identified. We used annotation with the Universal Protein Resource (UniProt) database and reviewed literature to identify protein domains.

### 2.6. ASD Candidate Genes

To examine if NMD genes were previously reported as candidate genes for ASD, we compared our list of NMD genes with a list of ASD susceptibility genes. A list of 1003 ASD candidate genes was downloaded from the manually curated SFARI Gene database (<https://gene.sfari.org/> accessed on 8 March 2021), released on 13 January 2021. This comprehensive database contains up-to-date information on genes linked to ASD risk, and based on the evidence supporting association with ASD, genes are ranked into four categories: syndromic (126 genes), high confidence (207 genes, category 1), strong candidates (211 genes, category 2) and suggestive evidence (506 genes, category 3). Additionally, to identify ASD candidates that are known NMD targets, we compared SFARI gene list with a list of high-confidence neuronal NMD targets obtained from Kurosaki et al., 2021 [28].

### 2.7. Brain Expression of NMD Genes

We used gene expression data to assess whether NMD genes were expressed in adult human brain and during the early stages of human brain development. For that we obtained gene expression data from two open-source resources: The Human Protein Atlas [43] and the Expression Atlas from European Bioinformatics Institute (EMBL-EBI). The Brain Atlas subset within Human Protein Atlas was downloaded and proteins expressed in Human Adult brain were identified (<https://www.proteinatlas.org/>, last updated on 24 February 2021). We also used expression data from The Human Developmental Biology Resource (HDBR) [44] which contains baseline gene expression from different brain regions across a substantial period of early development (4 to 20 post-conception weeks [PCW]) and was deposited in the EMBL-EBI Expression Atlas (E-MTAB-4840, 15 May 2019). EMBL-EBI makes available the normalized counts per gene (TPMs) and defines the following cut-offs: no expression (below cutoff;  $TPM < 0.5$ ), low expression (between 0.5 to 10 TPM), medium expression (between 11 to 1000 TPM) and high expression (above 1000 TPM). We used these criteria to classify gene expression of our candidate genes.

## 3. Results

### 3.1. Genes Encoding Proteins Involved in the NMD Pathway

We identified 46 experimentally validated genes encoding core NMD factors and regulators (Table 1 and Figure 1) through manual curation of a gene list obtained by enrichment analysis and literature review.

All 46 NMD genes are expressed in the human adult brain, as confirmed by the Human Protein Atlas. We also examined the EMBL-EBI Expression Atlas to assess whether these genes are expressed during early brain development, and we confirmed that all genes are expressed in the forebrain, midbrain and hindbrain in the first 4 to 8 post-conception weeks (PCW) and in the cerebral cortex in the age range of 8–17 PCW. Additionally, we observed that the expression levels of these genes vary during different stages of neurodevelopment. For example, the *UPF1*, *DHX34* and *RUVBL2* genes presented lower baseline expression at 8 PCW, whereas *SRSF1* presented higher baseline expression, when compared to the overall transcriptome.

**Table 1.** List of human genes encoding NMD factors and regulators.

NMD Group	Gene SYMBOL <sup>(1)</sup>	Alternative Name <sup>(2)</sup>	Role in NMD	References
EJC	<i>EIF4A3</i>	eIF4AIII DDX48	RNA helicase eukaryotic initiation factor 4A3 is a core EJC factor that interacts with the Y14-MAGOH heterodimer to provide a stable and direct binding site for the UPF3B protein and activate NMD	[45,46]
	<i>RBM8A</i>	Y14	RNA-binding motif protein 8A is a core EJC factor that interacts with eIF4A3 and MAGOH to provide a stable and direct binding site for the UPF3B protein and activate NMD	[45,46]
	<i>MAGOH</i>	MAGOH1	Mago nashi homolog protein is a core EJC factor that interacts with eIF4A3 and Y14 to provide a stable and direct binding site for the UPF3B protein and activate NMD	[45–47]
	<i>MAGOHB</i>	MGN2	Mago nashi protein homolog B is a paralog of MAGOH that interacts with eIF4A3 and Y14 forming the trimeric EJC core to activate NMD	[47]
	<i>RNPS1</i>		RNA-binding protein S1 is a component of the SR-rich EJCs that enhances NMD in early phase of the pathway	[45,48]
	<i>CASC3</i>	MLN51 BTZ	The peripheral EJC component CASC3 activates NMD and promotes SMG6-dependent endonucleolytic cleavage	[48–50]
	<i>ICE1</i>	KIAA0947	Component of the little elongation complex promotes the association of EJC with UPF3B and activates NMD	[51]
	<i>PYM1</i>	WIGB	Ribosome-associated protein PYM interacts with Y14-MAGOH triggering EJC disassembly which leads to NMD inhibition	[52]
SURF-DECID	<i>UPF1</i>	RENT1 smg-2	Up-frameshift protein 1 is the central component of the NMD pathway; its helicase and ATPase activities are essential to trigger NMD	[46,53,54]
	<i>UPF2</i>	RENT2 smg-3	Up-frameshift protein 2 and UPF3B interact with UPF1 favoring its phosphorylation by SMG1 and formation of DECID complex	[46,53]
	<i>UPF3B</i>	UPF3X	Up-frameshift protein 3B and UPF2 interact with UPF1 favoring its phosphorylation by SMG1 and formation of DECID complex; UPF3B also forms a stable trimeric complex with eRF1-eRF3a to promote dissociation of the termination complexes and triggers NMD	[46,53–55]
	<i>ETF1</i>	eRF1	Eukaryotic release factor 1 is part of the eRF1–eRF3 translation termination complex that associates with UPF1 and SMG1-SMG8-SMG9 to form SURF and activate NMD	[53]
	<i>GSPT1</i>	eRF3a	Eukaryotic release factor 3 is part of the eRF1–eRF3 translation termination complex that associates with UPF1 and SMG1-SMG8-SMG9 to form SURF and activate NMD	[53]
	<i>NCBP1</i>	CBP80	Component of the cap-binding complex (CBC) directly binds to UPF1, promoting the interaction with UPF2 to form SURF and activate NMD	[56–58]
	<i>NCBP2</i>	CBP20	Component of the CBC is essential for the stability of complex	[56–58]

Table 1. Cont.

NMD Group	Gene SYMBOL <sup>(1)</sup>	Alternative Name <sup>(2)</sup>	Role in NMD	References
SURF-DECID	<i>EIF4E</i>	eIF4E	Eukaryotic translation initiation factor 4E binds to UPF1 and activates NMD	[59]
	<i>SMG1</i>	ATX	Suppressor of morphogenesis in genitalia-1 associate with SMG8-SMG9 to form the SMG1c kinase complex that catalyzes UPF1 phosphorylation	[53,60,61]
	<i>SMG8</i>		Suppressor of morphogenesis in genitalia-8 and suppressor of morphogenesis in genitalia-9 are co-factors that regulate SMG1 kinase activity	[53,60,61]
	<i>SMG9</i>			
	<i>DHX34</i>	KIAA0134	RNA helicase DHX34 binds SMG1 and promotes UPF1 phosphorylation, triggering the conversion from the SURF to the DECID complex	[62–64]
	<i>RUVBL1</i>	RVB1	AAA-ATPases RUVBL1 and RUVBL2 form a hetero-hexameric ring promoting the transition from SURF to the DECID complex	[65,66]
	<i>RUVBL2</i>	RVB2		
mRNA decay	<i>SMG5</i>	EST1B	Suppressor of morphogenesis in genitalia-5 and -7 form a heterodimer that binds p-UPF1 and recruit decapping enzymes, the CCR4-NOT complex (through CNOT8) and PP2A	[54,67,68]
	<i>SMG7</i>	EST1C		
	<i>SMG6</i>	EST1A	Endonuclease that interacts both with UPF1 and p-UPF1 and cleaves NMD targets close to the PTC	[54,69–71]
	<i>CNOT8</i>	POP2 CAF1	CCR4-NOT transcription complex subunit 8 is recruited by SMG7 to degrade NMD targets	[72]
	<i>DCP1A</i>		mRNA-decapping enzyme 1A is a decapping activator and together with PNRC2 stimulate the decapping activity of DCP2	[73]
	<i>PNRC2</i>		Proline rich nuclear receptor coactivator 2 binds p-UPF1 and stimulate the decapping activity of DCP2	[73]
	<i>DCP2</i>		Decapping protein engaged in the 5'→3' mRNA degradation	[74]
	<i>MOV10</i>	gb110	RNA helicase contributes to degradation of UPF1-regulated transcripts	[75]
	<i>PPP2CA</i>	PP2AC	Protein phosphatase 2 (PP2A) promotes dephosphorylation of UPF1; both structural ( <i>PPP2R1A</i> ) and catalytic ( <i>PPP2CA</i> ) subunits of PP2A interact with SMG5	[54,67,68]
	<i>PPP2R1A</i>	PP2AA		
	<i>XRN1</i>		Exonuclease involved in 5'→3' mRNA degradation	[74]
	<i>DIS3L</i>	DIS3L1	Core exosome-associated factor involved in the 3'→5' mRNA degradation	[76]
	<i>DIS3L2</i>		Exoribonuclease that degrades mRNA from 3'→5'	[77]
	<i>EXOSC10</i>	PM/Sc1100 Rrp6p	Exosome catalytic subunit involved in the 3'→5' mRNA degradation	[74]
	<i>PARN</i>		Ribonuclease engaged in the 3'→5' mRNA degradation	[74]

Table 1. Cont.

NMD Group	Gene SYMBOL <sup>(1)</sup>	Alternative Name <sup>(2)</sup>	Role in NMD	References
ER-NMD	<i>NBAS</i>	NAG	Protein involved in Golgi-to-endoplasmic reticulum (ER) retrograde transport recruits UPF1 to the membrane of the ER and regulates a subset of NMD targets translated at the ER	[62,78]
	<i>PABPC1</i>	PABP1	Polyadenylate-binding protein 1 inhibits the interaction of UPF1 with eRF3, repressing NMD	[79,80]
Regulator	<i>EIF4G1</i>	EIF4G	Eukaryotic initiation factor 4G inhibits NMD	[81,82]
	<i>UPF3A</i>	UPF3	Up-frameshift protein 3A compete with UPF3B for UPF2-binding and inhibits NMD	[83]
	<i>FMR1</i>	FMRP	Fragile X mental retardation protein binds directly to UPF1 and acts as an NMD repressor	[28]
	<i>EIF3E</i>	INT6 EIF3S6	Eukaryotic translation initiation factor 3 subunit E is a non-core eIF3 subunit that interacts with UPF2 and triggers NMD	[84]
	<i>SRSF1</i>	SFRS1	Serine/arginine-rich splicing factor 1 promotes NMD by enhancing UPF1-binding to the mRNA in the nucleus and it is also involved in UPF1 dephosphorylation	[85]
	<i>SEC13</i>		GNL2 and SEC13 are conserved NMD factors that regulate endogenous NMD targets but their exact role is unknown	[86]
	<i>GNL2</i>	Ngp-1		

<sup>(1)</sup> Gene annotation according to HGNC. <sup>(2)</sup> Previous SYMBOL or alias reported by HGNC database or identifier found in literature review. Legend: NMD, Nonsense-Mediated mRNA Decay; PTC, Premature Termination Codon; EJC, Exon Junction Complex; SURF, SMG-1–Upf1–eRF1–eRF3 complex; DECID, Decay Inducing Complex; ER-NMD, NMD response at the Endoplasmic Reticulum; p-UPF1, phosphorylated UPF1.

### 3.1.1. Evidence for NMD Involvement in ASD Pathophysiology

From the list of 46 NMD genes, we found that seven are previously known ASD candidate genes described in the SFARI gene database, classified as syndromic (*PPP2CA*), high confidence (*UPF3B* and *FMR1*) or suggestive evidence (*EIF4E*, *EIF4G1*, *SMG6* and *UPF2*). Additionally, we overlapped the list of 1003 ASD candidate genes from SFARI with a list of 1277 high-confidence neuronal NMD targets from a previously published study [28] to understand whether the expression of known ASD risk genes might be regulated by NMD. We found that 71 ASD risk genes (Supplementary Table S5) are upregulated upon UPF1 knockdown and bound to hyperphosphorylated UPF1 in neuronal cells (see Kurosaki et al., 2021 [28]). Overall, this analysis indicated that NMD might be important to regulating the expression of ASD risk genes in the brain.

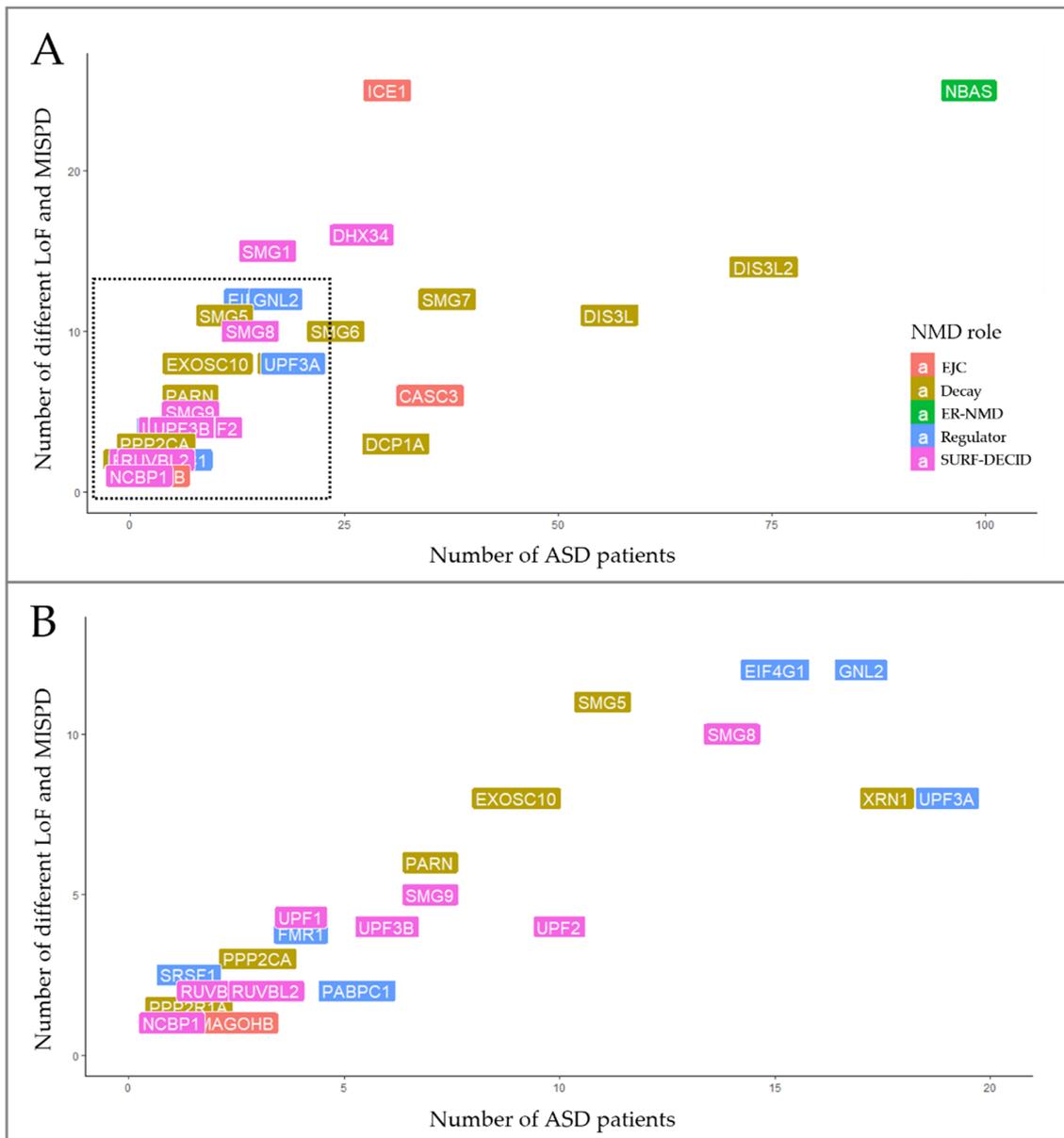
### 3.2. Discovery of SNVs in NMD Genes

We explored WES data from 1828 ASD patients and retrieved a total of 4922 SNVs in the 46 NMD genes. Following variant prioritization, we identified a total of 270 SNVs predicted to be damaging located on 38 NMD genes, in 28.7% (524/1828) of the ASD patients (Supplementary Table S3). These variants were almost equally distributed among males (416/1451, 28.7%) and females (99/356, 27.8%). The number of unique SNVs identified in ASD patients among the 38 NMD genes are represented in Figure 3. These variants were either rare (MAF < 1%) or not observed in gnomAD controls, and 11.5% (31/270) were classified as LoF and 88.5% (239/270) as MISPD. Among LoF variants, we discovered 6 frameshift, 10 nonsense and 15 splicing variants in 20 NMD genes (Supplementary Table S3).

Interestingly, the *NBAS* gene stood out for having twenty-five different variants (20 MISPD, 3 nonsense and 2 splicing) identified in 5.4% (98/1828) of the ASD patients (Figure 3). From these, 10 SNVs were located on two predicted functional domains (the secretory pathway protein Sec39 and N-terminal WD40 repeats domains), 10 SNVs were in the C-terminal region, one variant was predicted to affect splicing within N-terminal region and 3 SNVs were located upstream of the Sec39 domain (Figure 4D). The *NBAS* gene is not constrained (pLI = 0, mis\_Z < 0) but is important for the NMD response at the ER [78]; however, the required domains for NMD are still not known. Moreover, we identified 97 unique SNVs in 8 genes that were present in 1 to 4% of the ASD population: *DIS3L2* (4%, 74/1828), *DIS3L* (3%, 56/1828), *SMG7* (2%, 37/1828), *CASC3* (1.9%, 35/1828), *DCP1A* (1.7%, 31/1828), *ICE1* (1.6%, 30/1828), *DHX34* (1.5%, 27/1828) and *SMG6* (1.3%, 25/1828) (Figure 3, Supplementary Table S3).

To evaluate the impact of SNVs in NMD genes identified in ASD cases, we performed a grouped analysis of overall LoF and MISPD variants. Among the five groups of NMD genes, we found that 34.4% (93/270) of the SNVs mapped to thirteen mRNA decay genes, 26.3% (71/270) mapped to twelve SURF-DECID genes, 17% (46/270) mapped to eight regulator genes, 13% (35/270) mapped to four EJC genes and 9.3% (25/270) mapped to one ER-NMD gene (Figure 3 and Supplementary Table S3).

Through the analysis of the protein domains affected by these variants, we identified 125 SNVs within gene regions encoding protein domains known to be important for NMD function, in 13.5% (247/1828) of the ASD patients. Most of these variants were identified in constrained genes with pLI ≤ 0.5, and some of these genes are extremely intolerant to LoF (pLI > 0.9) and missense variants (mis\_Z > 3.09) [39,40]. For some genes, namely, *DCP1A*, *DCP2*, *DIS3L*, *DIS3L2*, *EXOSC10*, *MOV10*, *PARN*, *PPP2R1A*, *PPP2CA* and *XRN1*, we were not able to identify the SNVs that could influence NMD function because the protein domains necessary for NMD are not described.



**Figure 3.** Distribution of rare SNVs predicted to be damaging among 38 NMD genes in ASD patients. (A) Variants in NMD genes were identified in 1 to 5% of ASD patients. (B) Detailed distribution of gene variants in <1% of the ASD population sample.

### 3.2.1. EJC Components and Regulators

We discovered 34 SNVs (2 splicing, 2 frameshift and 30 MISPD) in three genes encoding EJC components, *EIF4A3*, *CASC3* and *ICE1*, in 68 ASD patients (Supplementary Table S3). Three MISPD are located within the N-terminal and the SELOR domain of *CASC3* gene, seven MISPD are located on the functional MIF4G domain of the *ICE1* gene, and two frameshift variants are predicted to cause premature termination of translation upstream the MIF4G domain of the *ICE1* gene (Figure 4A). *CASC3* (also known as BTZ) promotes the degradation of NMD substrates by interacting with EJC via its SELOR domain [49], a mechanism that also requires the N-terminus domain [50]. The C-terminal MIF4G domain of *ICE1* interacts with *eIF4A3* to promote UPF3B-EJC association [51]. Additionally, one MISPD was located in the *EIF4A3* N-terminal domain, which is required to trigger NMD [45]. Importantly, *ICE1*, *CASC3* and *EIF4A3* are constrained genes ( $pLI \geq 0.5$ ).

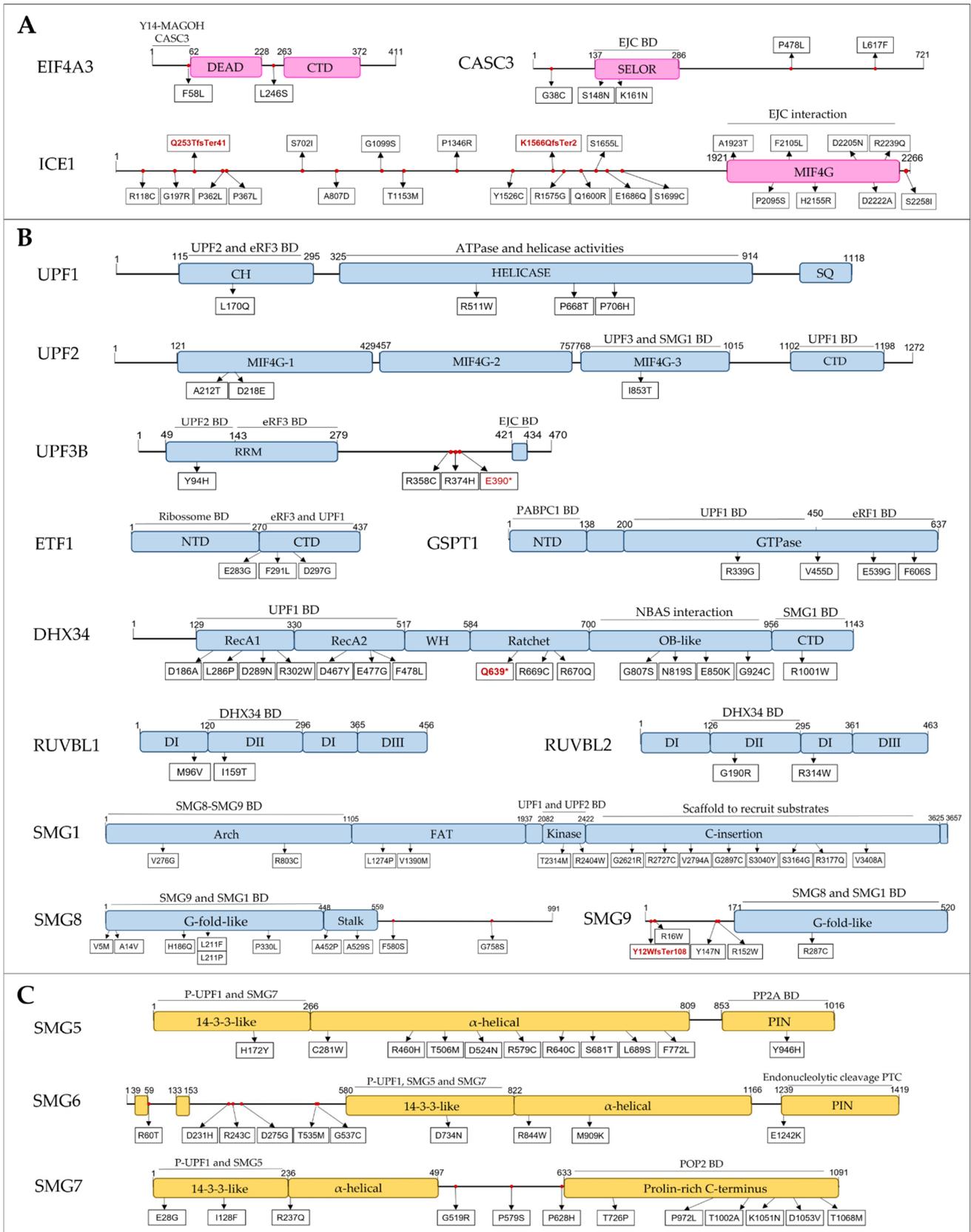
### 3.2.2. SURF and DECID Complexes

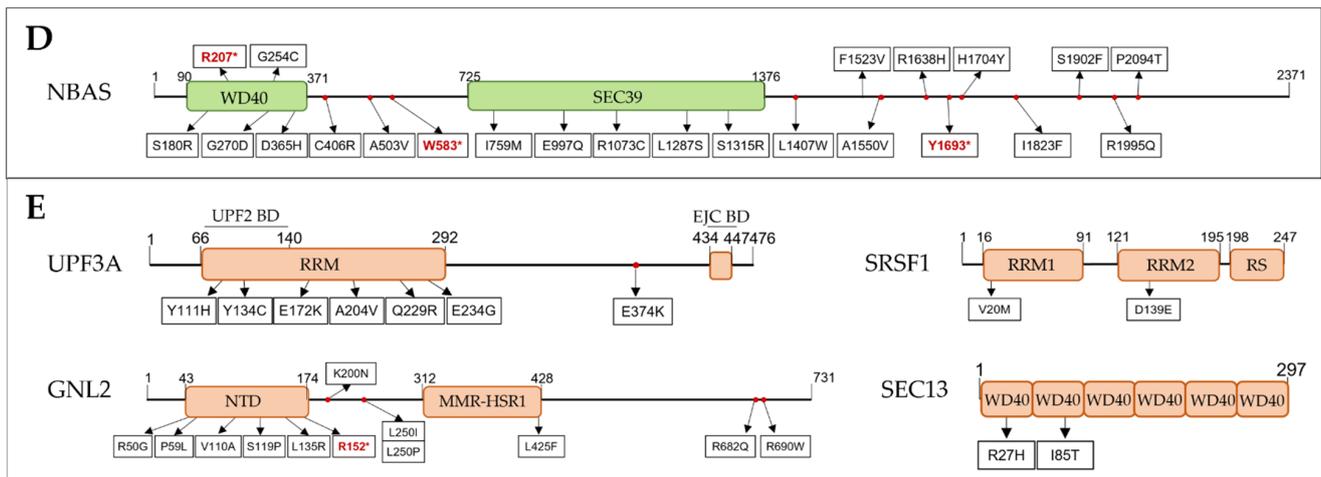
We discovered a total of 12 variants in *UPF1*, *UPF2* and *UPF3B* genes in 20 ASD patients (Supplementary Table S3). From these, we found 1 splicing and 8 MISPD variants located on functional protein domains: 3 within the helicase domain and 1 in the regulatory CH rich domain of *UPF1*, 4 within the MIF4G domains of *UPF2* and 1 located on RNA recognition motif (RRM), necessary for *UPF2*-binding, of *UPF3B* (Figure 4B). Additionally, we discovered three variants in a region between RRM and EJC-binding domains of the *UPF3B* gene. The helicase and CH rich domains of *UPF1* protein are required for ATPase and helicase activities [87] and for *UPF2*-binding [46], respectively. The MIF4G-1 and MIF4G-2 domains of *UPF2* have a key scaffolding role, and the MIF4G-3 domain is required for *UPF3* and *SMG1*-binding, thereby being the essential module for triggering NMD [88].

Regarding the *SMG1c* kinase complex, we found 15 SNVs (14 MISPD and 1 splicing) located on four functional *SMG1* domains: Arch, FAT, Kinase and C-Insertion domains; 8 SNVs located within two functional *SMG8* domains: G-fold-like and Stalk domains; and two SNVs (1 frameshift and 1 MISPD) predicted to affect the G-fold-like domain of *SMG9* (Figure 4B). From these, we identified 23 ASD patients with 19 SNVs located in domains required to form the *SMG1*-*SMG8*-*SMG9* complex and to activate NMD through *UPF1* phosphorylation: Arch, *SMG1* kinase and C-Insertion domains of *SMG1* protein and G-fold domains of *SMG8* and *SMG9* proteins [61]. Among SURF components, we discovered 1 splicing and 3 MISPD variants located on C-terminal domain of the *ETF1* gene, necessary for eRF3a and *UPF1* binding, and 4 MISPD variants within the GTPase domain of *GSPT1* gene, required for eRF1 and *UPF1* binding (Figure 4B). Moreover, we identified 1 splicing, 1 nonsense and 14 MISPD variants located within four functional protein domains of *DHX34* gene: helicase, ratchet, OB-like and C-terminal domains; and two MISPD variants in *RUVBL1* and *RUVBL2* genes located on DII domains (Figure 4B). *DHX34* binds *UPF1* and *SMG1* through its helicase and C-terminal domains, respectively, promoting NMD [64], and its OB-like domain is required for its interaction with NBAS [78]. The *RUVBL1*-2 DII domains are important for *DHX34* binding [66], and thus are involved in NMD. Importantly, *UPF1*, *UPF2*, *UPF3B*, *SMG1*, *ETF1*, *GSPT1*, *RUVBL1* and *RUVBL2* are constrained genes ( $pLI \geq 0.5$ ).

### 3.2.3. mRNA Decay

Thirty-three variants were discovered in *SMG5*, *SMG6* and *SMG7* genes in 60 ASD patients (Supplementary Table S3). Of these, 12 MISPD variants were located in functional protein domains essential for NMD: 2 variants in the 14-3-3-like and PIN domains of *SMG5*, 2 variants in the 14-3-3-like and PIN domains of *SMG6* and 8 variants in the 14-3-3-like and proline-rich C-terminus domains of *SMG7* (Figure 4C). The 14-3-3-like domain of *SMG6* and the *SMG5*-*SMG7* complex associates with phosphorylated *UPF1* [54,89], the PIN domain of *SMG6* promotes endonucleolytic cleavage near the PTC of a nonsense mRNA [71] and the *SMG7* proline-rich C-terminus interacts with CNOT8 to recruit the CCR4-NOT complex. *SMG6* and *SMG7* are constrained genes ( $pLI \geq 0.5$ ). Additionally, we found 47 variants in genes encoding exoribonucleases involved in NMD [74] in 164 ASD patients. Among these, 35 MISPD variants were located in functional protein domains: CDS1, CDS2 and RNB domains (*DIS3L* and *DIS3L2*); PMC2NT, DNA\_pol\_A\_exo1 and HRDC domains (*EXOSC10*); N-terminus, XRN\_D2\_D3 domains and Xnr1 SH3-like domain (*XRN1*); and CAF1 domain (*PARN*). We also found eight LoF variants in these genes: 3 nonsense (*DIS3L2*, *EXOSC10* and *PARN*), 3 frameshift (*EXOSC10*, *DIS3L2* and *XRN1*) and 2 splicing (*PARN*, *EXOSC10*). Moreover, we identified 13 SNVs within *MOV10*, *PPP2R1A*, *PPP2CA*, *DCP1A* and *DCP2* genes in 41 ASD patients, of which two were located on HEAT and PP2A subunit C binding domains (*PPP2R1A*) and three were located on CH and helicase domains (*MOV10*). Although these genes are involved in the degradation of NMD targets [74], the required domains to exert their function in NMD are not defined.





**Figure 4.** Schematic domain architecture of proteins involved in NMD, and representations of the protein alterations encoded by MISPD and nonsense/frameshift variants identified in ASD subjects. The binding regions for interacting proteins are indicated. Proteins were grouped by NMD groups: (A) EJC, (B) SURF-DECID, (C) mRNA decay, (D) ER-NMD and (E) Regulator. Legend: MISPD, missense variants predicted to be damaging and deleterious; BD, Binding Domain; CH, cysteine and histidine-rich domain; SQ, serine- and glutamine-rich domain; RS, arginine and serine-rich domain; RRM, RNA recognition motif; EJC, Exon Junction Complex; PIN, PilT N-terminus; P-UPF1, phosphorylated UPF1; WH, Winged-Helix; CTD, C-terminal domain; NTD, N-terminal domain.

### 3.2.4. NMD Regulators

This group includes the genes encoding proteins known to regulate NMD, but either they do not belong to the core machinery, or their exact functions are still unknown. We discovered eight SNVs within *UPF3A* gene in 19 ASD patients: 1 splicing variant upstream of the EJC-binding domain, 1 MISPD within a region of unknown function and 6 MISPD variants located on the RRM domain, including the region needed for UPF2 binding (Figure 4E). Additionally, one variant was identified in the RRM2 domain of *SRSF1*, which was described to be important for UPF1 dephosphorylation through interactions with SMG7 and PP2A [85], and nine variants (1 nonsense and 8 MISPD) were discovered within functional domains of *GNL2* and *SEC13* (Figure 4E).

### 3.3. CNVs Encompassing NMD Genes in ASD Patients

To further assess the contribution of large deletions and duplications of the 46 NMD genes to ASD, we analyzed CNVs from 3570 ASD patients. We found 38 CNVs encompassing 18 NMD genes in 1% (38/3570) of the ASD patients (34 males and 4 females), of which 8 were CNV losses and 30 CNV gains (Table 2).

We further characterized the extent of the deletion or duplication observed for each gene and analyzed whether the same variant was observed in 10,355 DGV controls (Supplementary Table S4). The majority of CNVs disrupting whole genes or important functional domains were exclusive in cases or observed in only one or two controls. Importantly, we discovered two ASD patients carrying CNVs (partial duplications) that included two regions required for NMD: the PIN domain of SMG6 protein and the MIF4G domains of UPF2 protein [71,88]. Moreover, we identified 10 patients with CNVs encompassing complete genes (*RBM8A*, *UPF3B*, *UPF3A*, *GSPT1*, *NCBP2* and *DHX34*) that may lead to dosage imbalances and thus influence NMD. Both *SMG6* and *UPF2* are highly constrained genes ( $pLI \geq 0.9$ ); and *RBM8A*, *UPF3B* and *GSPT1* genes are also constrained ( $pLI \geq 0.5$ ). The numbers of CNVs and SNVs disrupting genes or gene regions involved in NMD function are shown in Table 3.

**Table 2.** NMD genes disrupted by CNVs identified in a population of 3570 ASD patients.

NMD Group	Gene	Location	CNV Type	Gene Region	Protein Domains Affected	ASD N <sup>(1)</sup>
EJC	<i>RBM8A</i>	1q21.1	Deletion	Whole gene	All domains	1
	<i>RBM8A</i>	1q21.1	Duplication	Whole gene	All domains	3
SURF-DECID	<i>UPF2</i>	10p14	Duplication	Partial	MIF4G domains	1
	<i>UPF3B</i>	Xq24	Duplication	Whole gene	All domains	1
	<i>GSPT1</i>	16p13.13	Duplication	Whole gene	All domains	1
	<i>NCBP2</i>	3q29	Deletion	Whole gene	All domains	1
	<i>DHX34</i>	19q13.32	Deletion	Whole gene	All domains	1
	<i>RUVBL2</i>	19q13.33	Duplication	Partial	DI domain	1
mRNA decay	<i>DIS3L</i>	15q22.31	Duplication	Whole gene	All domains	1
	<i>DIS3L2</i>	2q37.1	Deletion	Partial	Part of CSD2 and RNB domains	1
	<i>DIS3L2</i>	2q37.1	Duplication	Partial	RNB and C-terminal S1 domain	1
	<i>EXOSC10</i>	1p36.22	Deletion	Partial	PMC2NT, EXO1 and HRCB domains	1
	<i>EXOSC10</i>	1p36.22	Duplication	Partial	EXO and HRCB domain	1
	<i>MOV10</i>	1p13.2	Duplication	Whole gene	All domains	2
	<i>PARN</i>	16p13.12	Deletion	Partial	All domains	1
	<i>PPP2R1A</i>	19q13.41	Duplication	Partial	PP2A subunit B binding	1
	<i>SMG6</i>	17p13.3	Duplication	Partial	PIN domain	1
	<i>XRN1</i>	3q23	Duplication	Partial	XRN1 SH3-like domain	13
ER-NMD	<i>NBAS</i>	2p24.3	Duplication	Partial	Sec39-like domain	1
Regulator	<i>FMR1</i>	Xq27.3	Deletion	Partial	KH2, NES and RGG domains	1
	<i>UPF3A</i>	13q34	Duplication	Whole gene	All domains	2
	<i>UPF3A</i>	13q34	Duplication	Partial	EJC-binding domain	1

<sup>(1)</sup> ASD N—number of ASD patients carrying CNVs.

**Table 3.** Numbers of SNVs or CNVs identified within protein domains important for NMD in 258 ASD patients. A total of 136 genetic variants in 23 NMD genes were identified.

NMD Group	Gene	Location	SNVs <sup>(1)</sup>		CNVs <sup>(1)</sup>	N ASD <sup>(2)</sup>	pLI	mis_Z	SFARI <sup>(3)</sup>
			LoF	MISPD					
EJC	<i>EIF4A3</i>	17q25.3		1		1	1.00	4.02	1q21.1 region
	<i>RBM8A</i>	1q21.1			2	4	0.57	2.16	
	<i>CASC3</i>	17q21.1		3		3	0.61	1.32	
	<i>ICE1</i>	5p15.32	2	7		10	1.00	0.73	
SURF-DECID	<i>UPF1</i>	19p13.11		4		4	1.00	5.63	category 3 category 1
	<i>UPF2</i>	10p14	1	3	1	11	1.00	3.19	
	<i>UPF3B</i>	Xq24	1	3	1	7	0.98	1.84	
	<i>ETF1</i>	5q31.2	1	3		6	1.00	4.39	
	<i>GSPT1</i>	16p13.13		4	1	7	1.00	3.32	
	<i>DHX34</i>	19q13.32	2	12	1	25	0.00	−0.08	
	<i>RUVBL1</i>	3q21.3		1		1	1.00	3.39	
	<i>RUVBL2</i>	19q13.33		1	1	2	1.00	3.11	
	<i>SMG1</i>	16p12.3	1	12		13	1.00	3.30	
	<i>SMG8</i>	17q22		6		8	0.01	1.73	
<i>SMG9</i>	19q13.31	1	1		4	0.00	1.60		
mRNA decay	<i>SMG5</i>	1q22		2		2	0.01	1.13	category 3
	<i>SMG6</i>	17p13.3		2	1	6	0.98	0.18	
	<i>SMG7</i>	1q25.3		8		14	1.00	2.19	
ER-NMD	<i>NBAS</i> <sup>(4)</sup>	2p24.3	5	20	1	99	0.00	−0.87	
Regulator	<i>UPF3A</i>	13q34	1	2	2	11	0.00	−0.60	
	<i>SRSF1</i>	17q22		1		1	0.98	3.96	
	<i>SEC13</i> <sup>(4)</sup>	3p25.3		2		2	0.02	0.62	
	<i>GNL2</i> <sup>(4)</sup>	1p34.3	1	11		17	0.00	0.28	

<sup>(1)</sup> Numbers of SNVs and CNVs located within regions necessary for NMD function. <sup>(2)</sup> Numbers of ASD patients carrying SNVs and CNVs. <sup>(3)</sup> Genes and CNVs associated with ASD, described in SFARI Gene database (<https://gene.sfari.org/> accessed on 8 March 2021). <sup>(4)</sup> These genes do not have the described NMD domains. Legend: SNVs, Single Nucleotide Variants; CNVs, Copy Number Variants; pLI, probability of being loss-of-function intolerant; mis\_Z, missense Z scores.

#### 4. Discussion

Despite enormous efforts in ASD research since its first description, the pathophysiology of this disorder is still unclear. In this study, we tested the hypothesis that genes encoding proteins involved in NMD (Figure 1), a regulatory mechanism that ensures the degradation of PTC-containing transcripts and that controls the expression of some naturally occurring transcripts [19], play a role in ASD.

NMD depends on multiple genes, and through literature review and enrichment analysis we defined a group of 46 experimentally validated NMD factors and regulators, of which seven were previously known as candidate risk genes for ASD. Some of the proteins encoded by these 46 genes, such as *UPF1*, are essential for the overall NMD response, whereas others participate in alternative NMD branches and are involved in the regulation of different mRNA targets. We confirmed that these 46 genes are expressed in the adult human brain and during early neurodevelopment, when the early stages of cortex development occur, and major brain regions are established [90]. This suggests that NMD may be modulated during neurodevelopment, a period of relevance for ASD. Moreover, we identified 71 neuronal ASD risk genes that are experimentally validated NMD targets [28], suggesting that NMD may regulate the expression of ASD risk genes during neurodevelopment. These observations are consistent with previous findings showing that selective mRNA decay is critical for specifying the developmental fate of human embryonic cell lineages [22].

Our study aimed at identifying genetic variants in these 46 genes that might influence the NMD response, including SNVs predicted *in silico* to be damaging and CNVs that might influence gene expression or disrupt gene structure, leading to novel transcripts [91], comparable to what was observed by Nguyen et al., 2013 [92]. Our search for SNVs was carried out in publicly available large ASD datasets with exome sequences, so that we could explore gene variants directly affecting known functional domains involved in the NMD response, namely the variants located in exonic regions or affecting splicing. We additionally inspected CNVs targeting NMD genes in the well characterized AGP and SSC datasets. The comprehensive search for rare SNVs and CNVs in 46 NMD genes among ASD patients led to the discovery of 270 SNVs predicted to be damaging within 38 genes in 28.7% of the ARRA WES population and 38 CNVs located on 18 genes in 1% of the AGP and SSC population. Both SNVs and CNVs were either rare or absent from control datasets from gnomAD or DGV, respectively. In particular, CNVs targeting NMD genes were rare in ASD patients but also extremely rare in controls (MAF < 0.1%).

Analysis of the protein domains affected by SNVs and CNVs revealed that 136 of these genetic variants, identified in 258 ASD patients, may have an impact on NMD function. Most of these variants mapped to constrained genes, indicative of intolerance to LoF and/or missense variation, therefore supporting a damaging impact in gene function. Some of the variants identified in these genes were located on domains previously known to influence NMD activity and/or implicated in ASD. For instance, we discovered SNVs within a region of *UPF3B* gene, located between RRM and EJC-binding domains (Figure 4), where mutations are known to impair NMD function [21]. Baird et al., (2018) previously reported that *ICE1* depletion leads to an increased abundance of NMD targets, including the ASD candidate gene *ANXA1* [8,51,93]. Our findings thus suggest that the variants predicted to be damaging may influence *ICE1* function in NMD, which in turn would affect the expression of NMD targets in ASD patients. The *UPF3A* gene is known to play an important role during development, and inactivation of the *UPF3A* gene results in hyperactivated NMD in mice [83]. The genetic variants identified in the *UPF3A* gene that were predicted to be damaging may influence *UPF3A* function and NMD activity during neurodevelopment, leading to altered target gene expression. Like *UPF3B*, mutations in *UPF3A* may contribute to ASD pathophysiology, because they exhibit antagonistic effects upon the same mechanism [83]. The *RBM8A* gene (also known as *Y14*) has an equally important role in interneuron development [24]. CNVs targeting the 1q21.1 region where

the *RBM8A* gene is located have previously been associated with neurodevelopmental disorders, including ASD [92,94].

Taken together, our research identified gene variants that may interfere with proper NMD function in 13.3% (SNVs) and 0.4% (CNVs) of the ASD patients analyzed in this study. Although these variants are individually rare in ASD cases, overall, they were identified in over 13% of the patients and are located on regions encoding protein domains required for NMD. Some of the NMD genes with relevant variants identified in this study, were previously implicated in ASD or target ASD candidate genes, reinforcing the notion that NMD dysfunction may contribute to the disease. In this work we focused on exonic SNVs with putative direct impact on the protein domains required for efficient NMD. While regulatory regions in NMD genes are not yet fully characterized, they can also contain genetic variants with an impact on NMD function and in the future need to be evaluated in ASD datasets with whole genome sequences. Future studies to experimentally validate the functional consequences of the identified variants on protein function are now mandatory. Additionally, future work should address the expression levels of NMD targets both in neurons and other cell types, to evaluate whether these predicted pathogenic variants influence NMD and to assess if this effect is specific to neuronal cells.

ASD is characterized by a wide spectrum in clinical presentation. Given the large number of targets for the NMD genes we assessed here, several of which are ASD candidates, the variants identified may influence ASD phenotypic heterogeneity. In this study, we did not explore the association of phenotypic presentation within the autism spectrum with the NMD gene variants. This was not possible because homogeneous clinical information was not available in a sufficient number of subjects from these datasets to draw firm conclusions, highlighting the pressing need for making phenotypic data available and interoperable for large disease datasets. There is also a growing body of evidence implicating genes and molecular pathways associated with ASD in other NDD, and even extending common genetic factors to several NPD [10,16–18]. Both the clinical and genetic boundaries between ASD and other NDD have many overlaps, with some risk genes being more predominant for ASD and others conferring risk for both ASD and other NDD [10,14–16,92,94]. It is therefore not surprising that some of the NMD genes identified in this study have previously been associated with other disorders. For instance, loss-of-function mutations in *UPF2*, *UPF3B*, *SMG8* and *SMG9* genes, and CNVs targeting *UPF2*, *UPF3A*, *SMG6*, *EIF4A3*, *RNPS1* and *RBM8A* genes, have been implicated in a variety of neurodevelopmental disorders, including DD, ID, ADHD and TAR syndrome [14,15,27,92,94–96]. NMD is a regulator of many biological processes from the early stages of development to adulthood and occurs in different tissues and cell types, and variants that impair NMD function may contribute to many pathophysiological mechanisms. Future work will need large datasets with extensive genomic and phenotypic information to ascertain the convergence among brain diseases, and determine the contributions of the NMD pathway, as a critical regulator of synaptic plasticity, neural development and neural stem cell differentiation decisions, to neurodevelopmental disorders.

## 5. Conclusions

An efficient NMD function is extremely important to regulate the expression of NMD targets and is essential during neurodevelopment and neuronal differentiation, and for brain function throughout life. Overall, our study provides novel evidence for a contribution of the NMD pathway to ASD. This work has identified, in a significant number of patients, genetic variants that can interfere with NMD function and may contribute to ASD through the misregulation of NMD target gene expression, particularly during neurodevelopment. Moreover, the branched nature of NMD suggests that diverse genetic alterations may have impacts on different NMD pathways and influence the expression of multiple NMD targets, contributing to the phenotypic heterogeneity of ASD. Further studies are needed to experimentally validate the functional impact of the variants found in this study, so that we can better understand the role of the NMD pathways in ASD. A full appreciation

of these regulatory mechanisms in ASD and other NDD will constitute an opportunity for the development of much needed therapeutic interventions for these conditions.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/biomedicines10030665/s1>, Table S1: List of Genes obtained by querying “nonsense mediated decay” in AmiGO database; Table S2: List of proteins involved in NMD pathway (R-HSA-927802) annotated by Reactome; Table S3: Rare Variants identified in NMD genes in ASD patients; Table S4: Rare, high-confidence CNVs targeting NMD genes in ASD samples; Table S5: List of ASD risk genes from Sfari database that are experimentally validated NMD targets.

**Author Contributions:** Conceptualization, A.R.M., L.R. and A.M.V.; data curation: A.R.M., J.X.S. and H.M.; formal analysis, A.R.M.; investigation: A.R.M., L.R. and A.M.V.; methodology: A.R.M., L.R. and A.M.V.; supervision: L.R., A.M.V.; visualization, A.R.M., J.X.S., L.R. and A.M.V.; writing—original draft, A.R.M., L.R. and A.M.V.; writing—review and editing, A.R.M., J.V., H.M., C.R., L.R. and A.M.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Fundação para a Ciência e a Tecnologia (UIDB/04046/2020 and UIDP/04046/2020 Centre grants to BioISI and PAC-POCI-01-0145-FEDER-016428 MEDPERSYST to A.M.V.) and by National Institute of Health Doutor Ricardo Jorge. A.R.M., J.V. and J.X.S. are recipients of a fellowship from BioSys PhD programme PD65-2012 (A.R.M. Ref: PD/BD/113773/2015; J.X.S. Ref: PD/BD/114386/2016; J.V. Ref: PD/BD/131390/2017) from Fundação para a Ciência e a Tecnologia (Portugal).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing not applicable.

**Acknowledgments:** The authors would like to thank the Autism Sequencing Consortium (ASC), the Autism Genome Project (AGP) and the Simons Foundation Autism Research Initiative (SFARI) Simplex Collection (SSC) for facilitating access to the datasets used for the analysis described in this manuscript. The datasets were obtained from the study Sanders et al., 2011 and dbGaP through the accession numbers phs000267.v5.p2 and phs000298.v3.p2. We acknowledge the clinicians and organizations that contributed to samples used in this study. Finally, we are grateful to the many families whose participation made this study possible.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; Am. Psychiatr. Publ.: Washington, DC, USA, 2013.
2. Chiarotti, F.; Venerosi, A. Epidemiology of Autism Spectrum Disorders: A Review of Worldwide Prevalence Estimates Since 2014. *Brain Sci.* **2020**, *10*, 274. [[CrossRef](#)] [[PubMed](#)]
3. Colvert, E.; Tick, B.; McEwen, F.; Stewart, C.; Curran, S.R.; Woodhouse, E.L.; Gillan, N.; Hallett, V.; Lietz, S.; Garnett, T.; et al. Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry* **2015**, *72*, 415–423. [[CrossRef](#)] [[PubMed](#)]
4. Sandin, S.; Lichtenstein, P.; Kuja-Halkola, R.; Hultman, C.; Larsson, H.; Reichenberg, A. The Heritability of Autism Spectrum Disorder. *JAMA* **2017**, *318*, 1182–1184. [[CrossRef](#)] [[PubMed](#)]
5. Sanders, S.; Ercan-sencicek, G.; Hus, V.; Luo, R.; Murtha, M.T.; Moreno-De-Luca, D.; State, M. Multiple recurrent de novo copy number variations (CNVs), including duplications of the 7q11.23 Williams-Beuren syndrome region, are strongly associated with autism. *Neuron* **2011**, *70*, 863–885. [[CrossRef](#)]
6. Pinto, D.; Delaby, E.; Merico, D.; Barbosa, M.; Merikangas, A.; Klei, L.; Thiruvahindrapuram, B.; Xu, X.; Ziman, R.; Wang, Z.; et al. Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am. J. Hum. Genet.* **2014**, *94*, 677–694. [[CrossRef](#)]
7. Iossifov, I.; O’Roak, B.J.; Sanders, S.J.; Ronemus, M.; Krumm, N.; Levy, D.; Stessman, H.A.; Witherspoon, K.T.; Vives, L.; Patterson, K.E.; et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **2014**, *515*, 216–221. [[CrossRef](#)]

8. De Rubeis, S.; He, X.; Goldberg, A.P.; Poultney, C.S.; Samocha, K.; Cicek, A.E.; Kou, Y.; Liu, L.; Fromer, M.; Walker, S.; et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **2014**, *515*, 209–215. [[CrossRef](#)]
9. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444. [[CrossRef](#)]
10. Satterstrom, F.K.; Kosmicki, J.A.; Wang, J.; Breen, M.S.; De Rubeis, S.; An, J.-Y.; Peng, M.; Collins, R.; Grove, J.; Klei, L.; et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **2020**, *180*, 568–584.e23. [[CrossRef](#)]
11. Asif, M.; Martiniano, H.F.D.M.C.; Marques, A.R.; Santos, J.X.; Vilela, J.; Rasga, C.M.; Oliveira, G.; Couto, F.; Vicente, A.M. Identification of biological mechanisms underlying a multidimensional ASD phenotype using machine learning. *Transl. Psychiatry* **2020**, *10*, 43. [[CrossRef](#)]
12. Parikshak, N.; Swarup, V.; Belgard, T.; Irimia, M.; Ramaswami, G.; Gandal, M.; Hartl, C.; Leppä, V.; Ubieta, L.D.L.T.; Huang, J.; et al. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* **2016**, *540*, 423–427. [[CrossRef](#)] [[PubMed](#)]
13. Xavier Santos, J.; Rasga, C.; Moura Vicente, A. Exposure to Xenobiotics and Gene-Environment Interactions in Autism Spectrum Disorder: A Systematic Review. In *Autism Spectrum Disorder—Profile, Heterogeneity, Neurobiology and Intervention*; IntechOpen: London, UK, 2021. [[CrossRef](#)]
14. Tarpey, P.S.; Raymond, F.L.; Nguyen, L.S.; Rodriguez, J.; Hackett, A.; Vandeleur, L.; Smith, R.; Shoubridge, C.; Edkins, S.; Stevens, C.; et al. Mutations in *UPF3B*, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat. Genet.* **2007**, *39*, 1127–1133. [[CrossRef](#)] [[PubMed](#)]
15. Laumonnier, F.; Shoubridge, C.; Antar, C.; Nguyen, L.S.; Van Esch, H.; Kleefstra, T.; Briault, S.; Fryns, J.P.; Hamel, B.; Chelly, J.; et al. Mutations of the *UPF3B* gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. *Mol. Psychiatry* **2010**, *15*, 767–776. [[CrossRef](#)] [[PubMed](#)]
16. Wang, T.; Hoekzema, K.; Vecchio, D.; Wu, H.; Sulovari, A.; Coe, B.P.; Gillentine, M.A.; Wilfert, A.B.; Perez-Jurado, L.A.; Kvarnang, M.; et al. Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. *Nat. Commun.* **2020**, *11*, 4932. [[CrossRef](#)]
17. Forsyth, J.K.; Nachun, D.; Gandal, M.; Geschwind, D.H.; Anderson, A.E.; Coppola, G.; Bearden, C.E. Synaptic and Gene Regulatory Mechanisms in Schizophrenia, Autism, and 22q11.2 Copy Number Variant-Mediated Risk for Neuropsychiatric Disorders. *Biol. Psychiatry* **2020**, *87*, 150–163. [[CrossRef](#)]
18. The Brainstorm Consortium; Anttila, V.; Bulik-Sullivan, B.; Finucane, H.K.; Walters, R.K.; Bras, J.; Duncan, L.; Escott-Price, V.; Falcone, G.J.; Gormley, P.; et al. Analysis of shared heritability in common disorders of the brain. *Science* **2018**, *360*, eaap8757. [[CrossRef](#)]
19. Wittmann, J.; Hol, E.M.; Jäck, H.-M. hUPF2 Silencing Identifies Physiologic Substrates of Mammalian Nonsense-Mediated mRNA Decay. *Mol. Cell. Biol.* **2006**, *26*, 1272–1287. [[CrossRef](#)]
20. Jolly, L.; Homan, C.; Jacob, R.; Barry, S.; Gecz, J. The *UPF3B* gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth. *Hum. Mol. Genet.* **2013**, *22*, 4673–4687. [[CrossRef](#)]
21. Alrahbeni, T.; Sartor, F.; Anderson, J.; Miedzybrodzka, Z.; McCaig, C.; Müller, B. Full UPF3B function is critical for neuronal differentiation of neural stem cells. *Mol. Brain* **2015**, *8*, 33. [[CrossRef](#)]
22. Lou, C.-H.; Dumdie, J.; Goetz, A.; Shum, E.Y.; Brafman, D.; Liao, X.; Mora-Castilla, S.; Ramaiah, M.; Cook-Andersen, H.; Laurent, L.; et al. Nonsense-Mediated RNA Decay Influences Human Embryonic Stem Cell Fate. *Stem Cell Rep.* **2016**, *6*, 844–857. [[CrossRef](#)]
23. Notaras, M.; Allen, M.; Longo, F.; Volk, N.; Toth, M.; Jeon, N.L.; Klann, E.; Colak, D. UPF2 leads to degradation of dendritically targeted mRNAs to regulate synaptic plasticity and cognitive function. *Mol. Psychiatry* **2020**, *25*, 3360–3379. [[CrossRef](#)] [[PubMed](#)]
24. McSweeney, C.; Dong, F.; Chen, M.; Vitale, J.; Xu, L.; Crowley, N.; Luscher, B.; Zou, D.; Mao, Y. Full function of exon junction complex factor, Rbm8a, is critical for interneuron development. *Transl. Psychiatry* **2020**, *10*, 379. [[CrossRef](#)] [[PubMed](#)]
25. Fernandes, R.; Nogueira, G.; da Costa, P.J.; Pinto, F.; Romão, L. Nonsense-Mediated mRNA Decay in Development, Stress and Cancer. In *The mRNA Metabolism in Human Disease*; Advances in Experimental Medicine and Biology; Springer: Cham, Switzerland, 2019; Volume 1157, ISBN 9783030199661. [[CrossRef](#)]
26. Tejada, M.I.; Villate, O.; Ibarluzea, N.; De La Hoz, A.B.; Martínez-Bouzas, C.; Beristain, E.; Martinez, F.; Friez, M.J.; Sobrino, B.; Barros, F. Molecular and Clinical Characterization of a Novel Nonsense Variant in Exon 1 of the *UPF3B* Gene Found in a Large Spanish Basque Family (MRX82). *Front. Genet.* **2019**, *10*, 1074. [[CrossRef](#)] [[PubMed](#)]
27. Domingo, D.; Nawaz, U.; Corbett, M.; Espinoza, J.L.; Tatton-Brown, K.; Coman, D.; Wilkinson, M.F.; Gecz, J.; Jolly, L.A. A synonymous *UPF3B* variant causing a speech disorder implicates NMD as a regulator of neurodevelopmental disorder gene networks. *Hum. Mol. Genet.* **2020**, *29*, 2568–2578. [[CrossRef](#)] [[PubMed](#)]
28. Kurosaki, T.; Imamachi, N.; Pröschel, C.; Mitsutomi, S.; Nagao, R.; Akimitsu, N.; Maquat, L.E. Loss of the fragile X syndrome protein FMRP results in misregulation of nonsense-mediated mRNA decay. *Nat. Cell Biol.* **2021**, *23*, 40–48. [[CrossRef](#)] [[PubMed](#)]

29. Nogueira, G.; Fernandes, R.; García-Moreno, J.F.; Romão, L. Nonsense-mediated RNA decay and its bipolar function in cancer. *Mol. Cancer* **2021**, *20*, 72. [[CrossRef](#)] [[PubMed](#)]
30. Carbon, S.; Ireland, A.; Mungall, C.J.; Shu, S.; Marshall, B.; Lewis, S.; The AmiGO Hub; The Web Presence Working Group. AmiGO: Online access to ontology and annotation data. *Bioinformatics* **2009**, *25*, 288–289. [[CrossRef](#)]
31. Tweedie, S.; Braschi, B.; Gray, K.; Jones, T.E.M.; Seal, R.L.; Yates, B.; Bruford, E.A. Genenames.org: The HGNC and VGNC resources in 2021. *Nucleic Acids Res.* **2021**, *49*, D939–D946. [[CrossRef](#)]
32. Buxbaum, J.D.; Daly, M.J.; Devlin, B.; Lehner, T.; Roeder, K.; State, M.W. The Autism Sequencing Consortium: Large-Scale, High-Throughput Sequencing in Autism Spectrum Disorders. *Neuron* **2012**, *76*, 1052–1056. [[CrossRef](#)]
33. Liu, L.; Sabo, A.; Neale, B.M.; Nagaswamy, U.; Stevens, C.; Lim, E.; Bodea, C.A.; Muzny, D.; Reid, J.G.; Banks, E.; et al. Analysis of Rare, Exonic Variation amongst Subjects with Autism Spectrum Disorders and Population Controls. *PLoS Genet.* **2013**, *9*, e1003443. [[CrossRef](#)]
34. Pinto, D.; Pagnamenta, A.T.; Klei, L.; Anney, R.; Merico, D.; Regan, R.; Conroy, J.; Magalhaes, T.R.; Correia, C.; Abrahams, B.S.; et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **2010**, *466*, 368–372. [[CrossRef](#)] [[PubMed](#)]
35. Fischbach, G.D.; Lord, C. The Simons Simplex Collection: A Resource for Identification of Autism Genetic Risk Factors. *Neuron* **2010**, *68*, 192–195. [[CrossRef](#)] [[PubMed](#)]
36. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
37. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol.* **2016**, *17*, 122. [[CrossRef](#)]
38. Rentzsch, P.; Witten, D.; Cooper, G.M.; Shendure, J.; Kircher, M. CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **2019**, *47*, D886–D894. [[CrossRef](#)] [[PubMed](#)]
39. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O’Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [[CrossRef](#)]
40. Samocha, K.; Robinson, E.; Sanders, S.; Stevens, C.; Sabo, A.; McGrath, L.; Kosmicki, J.A.; Rehnström, K.; Mallick, S.; Kirby, A.; et al. A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* **2014**, *46*, 944–950. [[CrossRef](#)]
41. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248–249. [[CrossRef](#)]
42. Kumar, P.; Henikoff, S.; Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **2009**, *4*, 1073–1081. [[CrossRef](#)]
43. Sjöstedt, E.; Zhong, W.; Fagerberg, L.; Karlsson, M.; Mitsios, N.; Adori, C.; Oksvold, P.; Edfors, F.; Limiszewska, A.; Hikmet, F.; et al. An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science* **2020**, *367*, eaay5947. [[CrossRef](#)]
44. Lindsay, S.J.; Xu, Y.; Lisgo, S.N.; Harkin, L.F.; Copp, A.J.; Gerrelli, D.; Clowry, G.J.; Talbot, A.; Keogh, M.J.; Coxhead, J.; et al. HDBR Expression: A Unique Resource for Global and Individual Gene Expression Studies during Early Human Brain Development. *Front. Neuroanat.* **2016**, *10*, 86. [[CrossRef](#)] [[PubMed](#)]
45. Gehring, N.H.; Kunz, J.B.; Neu-Yilik, G.; Breit, S.; Viegas, M.H.; Hentze, M.W.; Kulozik, A.E. Exon-Junction Complex Components Specify Distinct Routes of Nonsense-Mediated mRNA Decay with Differential Cofactor Requirements. *Mol. Cell* **2005**, *20*, 65–75. [[CrossRef](#)] [[PubMed](#)]
46. Chamieh, H.; Ballut, L.; Bonneau, F.; Le Hir, H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.* **2008**, *15*, 85–93. [[CrossRef](#)] [[PubMed](#)]
47. Singh, K.K.; Wachsmuth, L.; Kulozik, A.E.; Gehring, N.H. Two mammalian MAGOH genes contribute to exon junction complex composition and nonsense-mediated decay. *RNA Biol.* **2013**, *10*, 1291–1298. [[CrossRef](#)] [[PubMed](#)]
48. Mabin, J.W.; Woodward, L.A.; Patton, R.D.; Yi, Z.; Jia, M.; Wysocki, V.H.; Bundschuh, R.; Singh, G. The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. *Cell Rep.* **2018**, *25*, 2431–2446.e7. [[CrossRef](#)]
49. Bono, F.; Ebert, J.; Lorentzen, E.; Conti, E. The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA. *Cell* **2006**, *126*, 713–725. [[CrossRef](#)] [[PubMed](#)]
50. Gerbracht, J.V.; Boehm, V.; Britto-Borges, T.; Kallabis, S.; Wiederstein, J.L.; Ciriello, S.; Aschemeier, D.U.; Krüger, M.; Frese, C.K.; Altmüller, J.; et al. CASC3 promotes transcriptome-wide activation of nonsense-mediated decay by the exon junction complex. *Nucleic Acids Res.* **2020**, *48*, 8626–8644. [[CrossRef](#)]
51. Baird, T.D.; Cheng, K.C.-C.; Chen, Y.-C.; Buehler, E.; Martin, S.E.; Inglese, J.; Hogg, J.R. ICE1 promotes the link between splicing and nonsense-mediated mRNA decay. *eLife* **2018**, *7*, e33178. [[CrossRef](#)]
52. Gehring, N.H.; Lamprinaki, S.; Kulozik, A.E.; Hentze, M.W. Disassembly of Exon Junction Complexes by PYM. *Cell* **2009**, *137*, 536–548. [[CrossRef](#)]
53. Kashima, I.; Yamashita, A.; Izumi, N.; Kataoka, N.; Morishita, R.; Hoshino, S.; Ohno, M.; Dreyfuss, G.; Ohno, S. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* **2006**, *20*, 355–367. [[CrossRef](#)]

54. Okada-Katsuhata, Y.; Yamashita, A.; Kutsuzawa, K.; Izumi, N.; Hirahara, F.; Ohno, S. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res.* **2012**, *40*, 1251–1266. [[CrossRef](#)] [[PubMed](#)]
55. Neu-Yilik, G.; Raimondeau, E.; Eliseev, B.; Yeramala, L.; Amthor, B.; Deniaud, A.; Huard, K.; Kerschgens, K.; Hentze, M.W.; Schaffitzel, C.; et al. Dual function of UPF3B in early and late translation termination. *EMBO J.* **2017**, *36*, 2968–2986. [[CrossRef](#)] [[PubMed](#)]
56. Ishigaki, Y.; Li, X.; Serin, G.; Maquat, L.E. Evidence for a Pioneer Round of mRNA Translation: mRNAs Subject to Nonsense-Mediated Decay in Mammalian Cells Are Bound by CBP80 and CBP20. *Cell* **2001**, *106*, 607–617. [[CrossRef](#)]
57. Hosoda, N.; Kim, Y.K.; Lejeune, F.; Maquat, L.E. CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* **2005**, *12*, 893–901. [[CrossRef](#)] [[PubMed](#)]
58. Hwang, J.; Sato, H.; Tang, Y.; Matsuda, D.; Maquat, L.E. UPF1 Association with the Cap-Binding Protein, CBP80, Promotes Nonsense-Mediated mRNA Decay at Two Distinct Steps. *Mol. Cell* **2010**, *39*, 396–409. [[CrossRef](#)] [[PubMed](#)]
59. Rufener, S.C.; Mühlemann, O. eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* **2013**, *20*, 710–717. [[CrossRef](#)]
60. Yamashita, A.; Izumi, N.; Kashima, I.; Ohnishi, T.; Saari, B.; Katsuhata, Y.; Muramatsu, R.; Morita, T.; Iwamatsu, A.; Hachiya, T.; et al. SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes Dev.* **2009**, *23*, 1091–1105. [[CrossRef](#)]
61. Deniaud, A.; Karuppasamy, M.; Bock, T.; Masiulis, S.; Huard, K.; Garzoni, F.; Kerschgens, K.; Hentze, M.W.; Kulozik, A.E.; Beck, M.; et al. A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. *Nucleic Acids Res.* **2015**, *43*, 7600–7611. [[CrossRef](#)]
62. Longman, D.; Hug, N.; Keith, M.; Anastasaki, C.; Patton, E.E.; Grimes, G.; Cáceres, J.F. DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and *Caenorhabditis elegans*. *Nucleic Acids Res.* **2013**, *41*, 8319–8331. [[CrossRef](#)]
63. Hug, N.; Cáceres, J.F. The RNA Helicase DHX34 Activates NMD by Promoting a Transition from the Surveillance to the Decay-Inducing Complex. *Cell Rep.* **2014**, *8*, 1845–1856. [[CrossRef](#)]
64. Melero, R.; Hug, N.; López-Perrote, A.; Yamashita, A.; Cáceres, J.F.; Llorca, O. The RNA helicase DHX34 functions as a scaffold for SMG1-mediated UPF1 phosphorylation. *Nat. Commun.* **2016**, *7*, 10585. [[CrossRef](#)] [[PubMed](#)]
65. Izumi, N.; Yamashita, A.; Iwamatsu, A.; Kurata, R.; Nakamura, H.; Saari, B.; Hirano, H.; Anderson, P.; Ohno, S. AAA+ Proteins RUVBL1 and RUVBL2 Coordinate PIKK Activity and Function in Nonsense-Mediated mRNA Decay. *Sci. Signal.* **2010**, *3*, ra27. [[CrossRef](#)] [[PubMed](#)]
66. López-Perrote, A.; Hug, N.; González-Corpas, A.; Rodríguez, C.F.; Serna, M.; García-Martín, C.; Boskovic, J.; Fernandez-Leiro, R.; Cáceres, J.F.; Llorca, O. Regulation of RUVBL1-RUVBL2 AAA-ATPases by the nonsense-mediated mRNA decay factor DHX34, as evidenced by Cryo-EM. *bioRxiv* **2020**, *9*, e63042. [[CrossRef](#)]
67. Anders, K.R.; Grimson, A.; Anderson, P. SMG-5, required for *C.elegans* nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. *EMBO J.* **2003**, *22*, 641–650. [[CrossRef](#)]
68. Ohnishi, T.; Yamashita, A.; Kashima, I.; Schell, T.; Anders, K.; Grimson, A.; Hachiya, T.; Hentze, M.; Anderson, P.; Ohno, S. Phosphorylation of hUPF1 Induces Formation of mRNA Surveillance Complexes Containing hSMG-5 and hSMG-7. *Mol. Cell* **2003**, *12*, 1187–1200. [[CrossRef](#)]
69. Huntzinger, E.; Kashima, I.; Fauser, M.; Saulière, J.; Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **2008**, *14*, 2609–2617. [[CrossRef](#)]
70. Nicholson, P.; Josi, C.; Kurosawa, H.; Yamashita, A.; Mühlemann, O. A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Res.* **2014**, *42*, 9217–9235. [[CrossRef](#)]
71. Eberle, A.B.; Lykke-Andersen, S.; Mühlemann, O.; Jensen, T.H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* **2009**, *16*, 49–55. [[CrossRef](#)]
72. Loh, B.; Jonas, S.; Izaurralde, E. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.* **2013**, *27*, 2125–2138. [[CrossRef](#)]
73. Lai, T.; Cho, H.; Liu, Z.; Bowler, M.W.; Piao, S.; Parker, R.; Kim, Y.K.; Song, H. Structural Basis of the PNRC2-Mediated Link between mRNA Surveillance and Decapping. *Structure* **2012**, *20*, 2025–2037. [[CrossRef](#)]
74. Lejeune, F.; Li, X.; Maquat, L.E. Nonsense-Mediated mRNA Decay in Mammalian Cells Involves Decapping, Deadenylation, and Exonucleolytic Activities. *Mol. Cell* **2003**, *12*, 675–687. [[CrossRef](#)]
75. Gregersen, L.H.; Schueler, M.; Munschauer, M.; Mastrobuoni, G.; Chen, W.; Kempa, S.; Dieterich, C.; Landthaler, M. MOV10 Is a 5' to 3' RNA Helicase Contributing to UPF1 mRNA Target Degradation by Translocation along 3' UTRs. *Mol. Cell* **2014**, *54*, 573–585. [[CrossRef](#)] [[PubMed](#)]
76. Staals, R.; Bronkhorst, A.W.; Schilders, G.; Slomovic, S.; Schuster, G.; Heck, A.; Rajmakers, R.; Pruijn, G.J.M. Dis3-like 1: A novel exoribonuclease associated with the human exosome. *EMBO J.* **2010**, *29*, 2358–2367. [[CrossRef](#)] [[PubMed](#)]
77. Da Costa, P.J.; Menezes, J.; Saramago, M.; García-Moreno, J.; Santos, H.A.; Carvalho, M.D.G.; Arraiano, C.M.; Viegas, S.C.; Romão, L. A role for DIS3L2 over natural nonsense-mediated mRNA decay targets in human cells. *Biochem. Biophys. Res. Commun.* **2019**, *518*, 664–671. [[CrossRef](#)] [[PubMed](#)]

78. Longman, D.; Jackson-Jones, K.A.; Maslon, M.; Murphy, L.C.; Young, R.S.; Stoddart, J.J.; Hug, N.; Taylor, M.S.; Papadopoulos, D.K.; Cáceres, J.F. Identification of a localized nonsense-mediated decay pathway at the endoplasmic reticulum. *Genes Dev.* **2020**, *34*, 1075–1088. [[CrossRef](#)]
79. Ivanov, P.; Gehring, N.H.; Kunz, J.B.; Hentze, M.; Kulozik, A.E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **2008**, *27*, 736–747. [[CrossRef](#)]
80. Peixeiro, I.; Inácio, A.; Barbosa, C.; Silva, A.L.; Liebhaber, S.A.; Romão, L. Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. *Nucleic Acids Res.* **2011**, *40*, 1160–1173. [[CrossRef](#)]
81. Fatscher, T.; Boehm, V.; Weiche, B.; Gehring, N.H. The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. *RNA* **2014**, *20*, 1579–1592. [[CrossRef](#)]
82. Joncourt, R.; Eberle, A.B.; Rufener, S.C.; Mühlemann, O. Eukaryotic Initiation Factor 4G Suppresses Nonsense-Mediated mRNA Decay by Two Genetically Separable Mechanisms. *PLoS ONE* **2014**, *9*, e104391. [[CrossRef](#)]
83. Shum, E.Y.; Jones, S.H.; Shao, A.; Dumdie, J.; Krause, M.D.; Chan, W.-K.; Lou, C.-H.; Espinoza, J.L.; Song, H.-W.; Phan, M.H.; et al. The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. *Cell* **2016**, *165*, 382–395. [[CrossRef](#)]
84. Morris, C.; Wittmann, J.; Jäck, H.-M.; Jalinet, P. Human INT6/eIF3e is required for nonsense-mediated mRNA decay. *EMBO Rep.* **2007**, *8*, 596–602. [[CrossRef](#)] [[PubMed](#)]
85. Aznarez, I.; Nomakuchi, T.T.; Tetenbaum-Novatt, J.; Rahman, M.A.; Fregoso, O.; Rees, H.; Krainer, A.R. Mechanism of Nonsense-Mediated mRNA Decay Stimulation by Splicing Factor SRSF1. *Cell Rep.* **2018**, *23*, 2186–2198. [[CrossRef](#)] [[PubMed](#)]
86. Casadio, A.; Longman, D.; Hug, N.; Delavaine, L.; Baier, R.V.; Alonso, C.R.; Cáceres, J.F. Identification and characterization of novel factors that act in the nonsense-mediated mRNA decay pathway in nematodes, flies and mammals. *EMBO Rep.* **2015**, *16*, 71–78. [[CrossRef](#)] [[PubMed](#)]
87. Lee, S.R.; Pratt, G.A.; Martinez, F.J.; Yeo, G.W.; Lykke-Andersen, J. Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. *Mol. Cell* **2015**, *59*, 413–425. [[CrossRef](#)]
88. Clerici, M.; Deniaud, A.; Boehm, V.; Gehring, N.H.; Berger-Schaffitzel, C.H.; Cusack, S. Structural and functional analysis of the three MIF4G domains of nonsense-mediated decay factor UPF2. *Nucleic Acids Res.* **2014**, *42*, 2673–2686. [[CrossRef](#)]
89. Jonas, S.; Weichenrieder, O.; Izaurralde, E. An unusual arrangement of two 14-3-3-like domains in the SMG5–SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes Dev.* **2013**, *27*, 211–225. [[CrossRef](#)]
90. O’Rahilly, R.; Müller, F. Significant features in the early prenatal development of the human brain. *Ann. Anat.* **2008**, *190*, 105–118. [[CrossRef](#)]
91. Gamazon, E.; Stranger, B.E. The impact of human copy number variation on gene expression. *Brief. Funct. Genom.* **2015**, *14*, 352–357. [[CrossRef](#)]
92. Nguyen, L.S.; Kim, H.-G.; Rosenfeld, J.A.; Shen, Y.; Gusella, J.F.; Lacassie, Y.; Layman, L.C.; Shaffer, L.G.; Géczy, J. Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders. *Hum. Mol. Genet.* **2013**, *22*, 1816–1825. [[CrossRef](#)]
93. Correia, C.T.; Conceição, I.C.; Oliveira, B.; Coelho, J.; Sousa, I.; Sequeira, A.F.; Almeida, J.; Café, C.; Duque, F.; Mouga, S.; et al. Recurrent duplications of the annexin A1 gene (*ANXA1*) in autism spectrum disorders. *Mol. Autism* **2014**, *5*, 28. [[CrossRef](#)]
94. Albers, C.A.; Paul, D.S.; Schulze, H.; Freson, K.; Stephens, J.C.; Smethurst, P.A.; Jolley, J.D.; Cvejic, A.; Kostadima, M.; Bertone, P.; et al. Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit *RBM8A* causes TAR syndrome. *Nat. Genet.* **2012**, *44*, 435–439. [[CrossRef](#)] [[PubMed](#)]
95. Alzahrani, F.; Kuwahara, H.; Long, Y.; Al-Owain, M.; Tohary, M.; AlSayed, M.; Mahnashi, M.; Fathi, L.; Alnemer, M.; Al-Hamed, M.H.; et al. Recessive, Deleterious Variants in *SMG8* Expand the Role of Nonsense-Mediated Decay in Developmental Disorders in Humans. *Am. J. Hum. Genet.* **2020**, *107*, 1178–1185. [[CrossRef](#)] [[PubMed](#)]
96. Johnson, J.L.; Stoica, L.; Liu, Y.; Zhu, P.J.; Bhattacharya, A.; Buffington, S.A.; Huq, R.; Eissa, N.T.; Larsson, O.; Porse, B.T.; et al. Inhibition of Upf2-Dependent Nonsense-Mediated Decay Leads to Behavioral and Neurophysiological Abnormalities by Activating the Immune Response. *Neuron* **2019**, *104*, 665–679.e8. [[CrossRef](#)] [[PubMed](#)]