

UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Zentrum für Geburtshilfe, Kinder- und Jugendmedizin
Institut für Humangenetik

Institutsdirektor
Prof. Dr. med. Christian Kubisch

The role of RNA binding proteins in neurodevelopmental disorders

Dissertation

zur Erlangung des Doktorgrades PhD
an der Medizinischen Fakultät der Universität Hamburg

vorgelegt von:

Ilaria Mannucci
aus Italien

Hamburg 2021

(wird von der Medizinischen Fakultät ausgefüllt)

**Angenommen von der
Medizinischen Fakultät der Universität Hamburg am: 10.11.2021**

**Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.**

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Hans-Jürgen Kreienkamp

Prüfungsausschuss, zweite/r Gutachter/in: Prof. Dr. Stefan Linder

Prüfungsausschuss, dritte/r Gutachter/in: Prof. Dr. Daniel N. Wilson

Table of content

Summary.....	II
Zusammenfassung.....	III
Chapter 1 Introduction	1
1.1 The RNA world and RNA binding proteins.....	1
1.2 RNA metabolism	1
<i>From transcription to nuclear export</i>	1
<i>mRNA translation</i>	2
<i>mRNA decay</i>	2
<i>P-bodies and Stress Granules</i>	3
<i>Regulation of gene expression by small non-coding RNAs</i>	7
1.3 RNA metabolism and neurodevelopment	9
1.4 The RNA helicase DHX30	11
Aim of the project.....	12
Chapter 2 Germline AGO2 mutations impair RNA interference and human neurological development	14
Chapter 3 Genotype–phenotype correlations and novel molecular insights into the DHX30-associated neurodevelopmental disorders	29
Chapter 4 Discussion.....	78
4.1 Germline AGO2 mutations impair RNA interference and human neurological development.....	78
4.2 The RNA helicase DHX30: role in neurodevelopmental disorders	79
4.3 Do AGO2 and DHX30 operate in similar pathways during neuronal development?.....	85
Conclusion	86
References	87
Declaration of Contribution.....	97
Acknowledgements	98
Eidesstattliche Erklärung	100

Summary

Neurodevelopmental disorders (NDDs) represent a large group of disorders arising from alterations of tightly coordinated processes that regulate development and function of the brain. Different types of syndromes, including intellectual disability, autism spectrum disorders, attention deficit hyperactivity disorders and epilepsy, have been classified as NDDs (Niemi et al., 2018). Due to their relatively high prevalence, NDDs constitute a serious socio-economic problem of health care. The identification of NDD-associated genes is essential for a better understanding of both the genetic contribution to the disease and the potential underlying pathomechanism. The recent introduction of high-throughput technologies such as next generation sequencing (NGS) has substantially improved our understanding of the genetic causes of different forms of NDDs (Vissers et al., 2016). Specifically, NGS-based studies in children affected by NDDs have led to the identification of mutations in genes encoding proteins involved in RNA metabolism as a major cause of NDDs; this includes *AGO2*, a component of the RNA interference pathway, as well as several RNA helicases such as *DHX30*. The *DHX30* gene was initially established as a candidate gene for NDDs (Eldomery et al., 2017; Zheng et al., 2015). The identification of mutations in *AGO2* in 21 patients, and in *DHX30* in more than 35 patients affected by mild to severe NDDs has ultimately provided strong evidence for assessing the indispensability of these genes in the development of the central nervous system. For *DHX30*, the majority of the affected individuals carry heterozygous missense variants within highly conserved helicase core motifs (HCMs) and display global developmental delay, intellectual disability, severe speech impairment and gait abnormalities. Notably, four individuals carrying heterozygous variants which result in haploinsufficiency or truncated proteins present with a milder clinical course, thereby delineating a different, less severe clinical subtype compared to variants affecting the HCMs. However, the physiological role of *DHX30* in cellular RNA metabolism so far remains largely unexplored.

The current PhD project was focused on analyzing the molecular and cellular dysfunction associated with *AGO2* and, as a central topic, *DHX30* missense mutations and on elucidating the physiological role of *DHX30* in cellular RNA metabolism. In ATPase assays, it is shown here that *DHX30* is an RNA-dependent ATPase, and that all *DHX30* missense mutations affecting highly conserved HCMs lead to disrupted ATPase activity. Using purified recombinant protein in helicase assays with a radiolabeled RNA duplex, *DHX30* is formally established here as an ATP-dependent RNA helicase. Again, all tested missense mutations in HCMs disrupt the helicase activity whereas a mutation outside the core region does not affect helicase activity.

Subcellular localization studies confirmed the cytoplasmic diffuse distribution of overexpressed GFP-tagged *DHX30*-wild type (WT). All HCM missense variants, overexpressed as GFP-fusion proteins, stimulated spontaneous stress granule (SG) hyper-assembly, leading to inhibition of global translation. Again, mutations outside HCMs had lower to no ability to induce SGs.

The clear correlation between SG formation and severity of the phenotype observed in the patients motivated further investigation of the role of *DHX30* in SG assembly. Permanent knockdown using CRISPR/Cas9-based gene editing of *DHX30* in a human cell line led to impaired SG formation upon heat stress, uncovering a previously unknown role for *DHX30* in this aspect of RNA metabolism.

Analysis of the nature of *DHX30* missense mutations showed that RFP-tagged *DHX30* WT is recruited to SGs upon co-expression of several GFP-tagged missense variants, possibly suggesting a dominant negative effect of the mutations. However, the observation that both endogenous and overexpressed *DHX30*-WT are recruited to SGs after stress induction (Lessel et al., 2017) together with the fact that *DHX30*-deficient cells show reduced SG formation, actually suggest that mutations affecting HCMs result in a detrimental gain of function with respect to SG formation.

Zusammenfassung

Neurologische Entwicklungsstörungen (*neurodevelopmental disorders*, NDDs) bilden eine große Gruppe von Erkrankungen, bei denen präzise kontrollierte Mechanismen gestört sind, die zur Entwicklung und Funktion des Gehirns beitragen. Verschiedene Syndrome, einschließlich der Intelligenzminderung, Autismus-Spektrum-Störungen, Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung und Epilepsie wurden als neurologische Entwicklungsstörungen klassifiziert (Niemi et al., 2018). Auf Grund ihrer hohen Prävalenz stellen NDDs ein ernstzunehmendes, sozioökonomisches Problem dar. Die Identifizierung von NDD-assoziierten Genen ist essentiell für das Verständnis der genetischen Komponente der Erkrankungen, sowie der zugrundeliegenden Pathomechanismen. Durch neue Hochdurchsatz-Technologien wie *Next Generation Sequencing* (NGS) wurde unser Verständnis der genetischen Ursachen vieler NDDs deutlich verbessert (Vissers et al., 2016). Insbesondere haben NGS-basierte Studien bei betroffenen Kindern viele Mutationen in Genen identifiziert, die für Proteine des RNA-Stoffwechsels kodieren; dazu gehören *AGO2*, eine Komponente des RNA-Interferenzweges, und mehrere RNA-Helikasen wie z.B. *DHX30*. Das *DHX30*-Gen wurde zunächst als ein Kandidatengen für NDDs klassifiziert (Eldomery et al., 2017; Zheng et al., 2015). Die Identifizierung von Mutationen in *AGO2* bei 21 Patienten, und in *DHX30* bei mehr als 35 Patienten, die von leichten bis schweren NDDs betroffen sind, zeigte, dass diese Gene für die Entwicklung des Nervensystems unverzichtbar sind. Bei *DHX30* trägt die Mehrheit der Patienten heterozygote *missense*-Mutationen innerhalb der hochkonservierten Helikase-Kernmotive (HCM) und weist klinische Phänotypen auf, die durch globale Entwicklungsverzögerung, Intelligenzminderung sowie schwere Sprach- und Gangstörungen geprägt sind. Interessanterweise sind heterozygote Varianten von *DHX30*, die zu Haploinsuffizienz oder der Expression eines verkürzten Proteins führen, bei vier Individuen mit einem mildereren klinischen Verlauf assoziiert; dies ist ein Hinweis auf klinisch unterscheidbare Subtypen der Erkrankung.

Diese Arbeit fokussiert sich auf die Analyse der molekularen und zellulären Dysfunktionen, die mit *missense*-Mutationen in *AGO2*, und hauptsächlich in *DHX30* assoziiert sind, sowie auf die Aufklärung der physiologischen Rolle von *DHX30* im zellulären RNA-Stoffwechsel. In ATPase-*assays* wird hier gezeigt, dass *DHX30* eine RNA-abhängige ATPase ist und dass alle Mutationen in den hochkonservierten HCMs die ATPase-Aktivität hemmen. Mit Hilfe von gereinigtem rekombinantes Protein und einer radioaktiv markierten RNA-Duplex wird erstmalig die Aktivität von *DHX30* als ATP-abhängige RNA-Helikase gemessen. Wieder stören alle untersuchten HCM-*missense*-Mutationen diese Aktivität, nicht jedoch eine Mutation außerhalb der HCMs. Die Untersuchung der subzellulären Lokalisation bestätigt die zytoplasmatisch diffuse Verteilung von überexprimiertem GFP-markiertem *DHX30*-Wildtypprotein (*DHX30*-WT). Alle HCM-*missense*-Varianten dagegen induzierten die spontane Bildung von *stress granules* (SGs) und wurden in diesen lokalisiert; dies ging mit einer Hemmung der Translation einher. Varianten außerhalb der HCMs induzierten wenige oder keine SGs.

Die klare Korrelation von Häufigkeit der SG-Bildung mit der Schwere der bei den Patienten beobachteten Phänotypen gab den Anstoß, die Rolle von *DHX30* bei der SG-Assemblierung genauer zu untersuchen. Permanenter *knock-down* von *DHX30* mit Hilfe der Crispr-Technologie in humanen Zellen führte zu einer reduzierten Bildung von SGs nach Hitze-Stress, was auf eine bisher unbekannte Funktion von *DHX30* in diesem Aspekt des RNA-Metabolismus hinweist.

Die weitere Analyse der *DHX30*-*missense*-Mutationen zeigte, dass RFP-markiertes *DHX30*-WT Protein durch die Koexpression von GFP-markierten *missense*-Varianten ebenfalls in SGs rekrutiert wird. Dies deutet auf einen dominant-negativen Effekt der *missense*-Varianten bezüglich der Funktion des WT-Proteins hin. Jedoch zeigt die Beobachtung, dass sowohl endogenes als auch überexprimiertes *DHX30*-WT Protein durch zellulären Stress in SGs rekrutiert wird, im Zusammenhang mit dem Befund, dass *DHX30*-defiziente Zellen

eine reduzierte SG-Bildung aufweisen, dass Mutationen in den HCMs zu einer pathogenen *gain of function* Veränderung im Hinblick auf die SG-Bildung führen.

Chapter 1 | Introduction

1.1 | The RNA world and RNA binding proteins

RNA is the blueprint of life. Indeed, during the evolution of life on Earth, many isolated mixtures of complex organic molecules produced by random chemistry failed to achieve self-replication and therefore went extinct (Cech, 2012). The discovery that RNA can serve as an informational molecule and as a biocatalyst capable of replicating itself, provided firm basis for the plausibility of a primordial RNA world, highlighting the indispensability of RNA since the dawn of time. In all living organisms, RNA still plays active roles in catalyzing biochemical reactions, translating mRNA into proteins and regulating gene expression (Cech, 1986, 2012; Kruger et al., 1982). However, most of the RNA molecules in eukaryotic cells do not act in isolation but rather exist in complex with RNA-binding proteins (RBPs) to form ribonucleoprotein (RNP) complexes. By forming transient dynamic interactions with coding, untranslated and non-coding RNAs, RBPs regulate every aspect of RNA metabolism ranging from transcription to post-transcriptional gene regulation processes such as RNA maturation, nuclear export, subcellular localization, surveillance, translation, storage and degradation (Glisovic et al., 2008; Lukong et al., 2008; Muller-McNicoll & Neugebauer, 2013). Given the central role of RBPs in determining the fate of RNA molecules, it is not surprising that 7.5 % of ~20,500 protein-coding genes in humans code for essential components of RNPs. Although target-RNA recognition studies show that nearly 50% of RBPs act within messenger ribonucleoprotein (mRNP) complexes and 11% are represented by ribosomal proteins, post-transcriptional gene regulation concerns processes involving both coding and non-coding RNAs (ncRNAs) (Gerstberger et al., 2014).

1.2 | RNA metabolism

From transcription to nuclear export

The nascent protein-coding transcripts are generated by RNA polymerase II (Pol II) and are termed precursor mRNAs (pre-mRNAs). Newly transcribed pre-mRNAs undergo a maturation process which consists of multiple nuclear modification steps including 5' end capping, splicing and 3' end processing (e.g. cleavage and polyadenylation). Several observations point to a mutual regulation mechanism between these processing events. For instance, the importance of the 5' end cap (also called m⁷G cap) in determining the life cycle of the RNA does not only rely on its involvement in cap-dependent translation initiation (Filipowicz, 1978) but also on its protective function against 5' to 3' exonuclease cleavage (Shimotohno et al., 1977) and on its role as a platform for recruiting proteins involved in pre-mRNA splicing (Konarska et al., 1984), polyadenylation (Cooke & Alwine, 1996) and nuclear export (Hamm & Mattaj, 1990). Moreover, splicing factors such as the small nuclear ribonucleoproteins U1 snRNP (Gunderson et al., 1997) and U2AF65 snRNP (Millevoi *et al.*, 2002) have been shown to regulate cleavage and polyadenylation either negatively or positively, respectively. Similarly, 3' end processing factors such as the cleavage and polyadenylation specificity factor CPSF (Kyburz et al., 2006) and the poly(A) polymerase PAP (Vagner et al., 2000) were reported to stimulate splicing efficiency. Taken together, these observations support the idea of a functional coupling between splicing and cleavage/polyadenylation (Kaida, 2016).

The fidelity of the maturation process is essential for generating an export-competent mRNA. The formation of translationally competent transcripts strongly relies on the proper coordination between upstream processing events and nuclear export of mRNA (Stewart, 2019; Wickramasinghe & Laskey, 2015). For instance, the recruitment of the cap-binding complex (CBC) by the m⁷G cap is essential for mediating both pre-mRNA splicing and nuclear export of mature mRNA (Ramanathan et al., 2016). In the first case, CBC

promotes mRNA splicing through interaction with the uridine-rich small nuclear ribonucleoproteins (snRNPs) U4/U6·U5 complex (Pabis et al., 2013). In the second case, CBC mediates the directional exit of the mRNA via interaction with nuclear export factors belonging to the transcription export complex (TREX) (Nojima et al., 2007).

mRNA translation

Upon entry in the cytoplasm, CBC-bound mRNPs undergo a remodeling process whereby CBC is replaced by the translation initiation factor eIF4E (Chiu et al., 2004). Eukaryotic translation occurs in three mechanistic steps - initiation, elongation and termination - but the most rate-limiting and highly regulated step is the initiation. Indeed, the initiation process consists of multiple linked stages that require at least nine eukaryotic initiation factors (eIFs). In eukaryotes most mRNAs are initiated by a cap-dependent scanning mechanism whereby a 43S preinitiation complex binds the capped 5' terminus of the mRNA and proceeds in the 5' to 3' direction searching for an appropriate start codon to initiate protein synthesis. The 43S preinitiation complex consists of a ternary complex including eIF2, GTP and initiator transfer methionyl-tRNA (tRNA_i^{Met}) which together with eIF1, eIF1A, eIF3 and eIF5 binds to the free 40S subunit. In order to stop the scanning process and promote stable codon-anticodon base-pairing at a suitable AUG start codon, the 43S complex must undergo conformational rearrangements that result in the formation of a "closed" complex, the 48S initiation complex where the mRNA entry channel is locked onto the mRNA. Joining of the 60S ribosomal subunit to the 48S initiation complex and concomitant displacement of eIFs lead to assembly of the 80S initiation complex. GTP hydrolysis by eIF5B and release of GDP-bound eIF5B make the 80S complex competent to catalyze the first peptide bond (Aylett & Ban, 2017; Gebauer & Hentze, 2004; Jackson et al., 2010; Pestova & Hellen, 2001). At this point, the anticodon of the methionine-carrying tRNA is bound to the start codon of the mRNA within the peptidyl (P) site of the ribosome. The aminoacyl-tRNA whose anticodon is complementary to the next codon is delivered to the aminoacyl (A) site of the ribosome where it binds to the matching codon on the mRNA. Formation of the peptide bond occurs when the methionine is transferred from the initiator-tRNA to the amino group of the aminoacyl-tRNA in the A site leading to an extended peptidyl-tRNA. As the deacylated tRNA translocates to the exit (E) site of the ribosome and the peptidyl-tRNA to the P site, the A site becomes available for binding of the next aminoacyl-tRNA (Dever et al., 2018; Dever & Green, 2012). Translation terminates when the A site of the ribosome encounters a stop codon leading to the release of the nascent polypeptide from the peptidyl-tRNA in the P site. Post-termination complexes (post-TCs) consisting of an mRNA-bound 80S ribosome, a P-site deacylated tRNA and at least one termination factor, undergo a recycling process which allows ribosomes and mRNAs to participate in multiple rounds of translation. During recycling, the ribosome splits into free 60S subunits and mRNA/tRNA-bound 40S subunits first and then the mRNA is released from the 40S subunits (Hellen, 2018; Jackson et al., 2010).

mRNA decay

The control of eukaryotic gene expression strongly relies on the interconnection between translation and mRNA degradation. However, the early notion that translation of an mRNA prevents its degradation (Schwartz & Parker, 1999) has been overcome by more recent observations showing that the interrelationship between these two processes is actually much more complex. Whether mRNA decay is triggered by inhibition of translation, whether translation repression is a prerequisite for mRNA decay or whether changes in translation efficiency are promoted by alterations in mRNA decay are only a few of the possible mechanisms by which translation and mRNA decay influence each other (Roy & Jacobson, 2013). Nevertheless, cap-dependent translation initiation only occurs if the 5' m⁷G cap and the 3' poly(A) tail have been correctly added to the mRNA in the nucleus. These two elements are essential for regulating both

translation and stability as they serve as a binding platform for the cytoplasmic proteins eIF4E and poly(A) binding protein (PABP), respectively (Bicknell & Ricci, 2017; Decker & Parker, 2012; Garneau et al., 2007). Moreover, a variety of regulatory elements located in the 5' untranslated region (UTR) of the mRNA such as GC-rich secondary RNA structures and unfavorable upstream AUG context usually act as negative regulators of translation initiation (Gebauer & Hentze, 2004; Leppek et al., 2018) thereby increasing the likelihood of deadenylation and decapping events. The first step of eukaryotic mRNA decay is the shortening of the poly(A) tail, a process known as deadenylation. Removal of the poly(A) tail is catalyzed by the deadenylase activity of PAN2-PAN3 and the CCR4-NOT complex. Subsequently, the mRNA is decapped by a decapping complex consisting of DCP1 and DCP2 proteins. Decapping enzymes are assisted by several accessory factors including the RNA helicase DDX6, different Sm-like (Lsm) proteins such as those of the Lsm1-7 complex and EDC3 (enhancer of decapping-3, also known as Lsm16). These events leave the mRNA susceptible to be degraded either from the 5' terminus (5'→3' mRNA decay) by the exonuclease XRN1 or from the 3' terminus (3'→5' mRNA decay) by the exosome (Garneau et al., 2007; Grudzien-Nogalska & Kiledjian, 2017; Łabno et al., 2016). To minimize synthesis of aberrant protein products, it is important to promptly identify errors which are not as obvious as the lack of 5' cap or 3' poly(A) tail. For instance, mRNAs containing a premature termination codon or lacking a stop codon or carrying a ribosome stalling sequence are subjected to different mRNA surveillance mechanisms including non-sense-mediated decay (NMD), non-stop decay (NSD) and no-go decay (NGD), respectively (Bicknell & Ricci, 2017; Decker & Parker, 2012; Garneau et al., 2007). Therefore, mRNA surveillance is another example of the synergistic relationship between translation and mRNA stability.

P-bodies and Stress Granules

Proteins involved in mRNA decay and translation repression are enriched in cytoplasmic RNP granules referred to as processing bodies (P-bodies). A conserved core of proteins found in P-bodies include factors involved in deadenylation, decapping and 5'→3' exonucleolytic decay. Additionally, P-bodies also contain proteins participating in NMD, translation initiation/repression and microRNA-mediated mRNA repression (Parker & Sheth, 2007). Some evidence suggests that P-bodies may exist constitutively as preformed structures which increase in number and size upon stress (Kedersha & Anderson, 2007; Kedersha et al., 2005; Teixeira et al., 2005). However, P-bodies are most likely dynamic entities whose assembly is dependent on and proportional to the cytoplasmic accumulation of translationally repressed mRNAs (Decker & Parker, 2012; Garneau et al., 2007; Luo et al., 2018). Similar to other membrane-less RNP granules such as Cajal bodies, nucleoli and stress granules, P-bodies behave like liquid droplets undergoing a physical process called liquid-liquid phase separation (LLPS). LLPS occurs as a consequence of an energetically favorable process where highly concentrated solutions of macromolecules such as RNAs and proteins condense into a dense phase. Similar to the energetically favorable separation of oil and water, this dense phase is physically separated from a dilute phase corresponding to the surrounding cellular context (Alberti et al., 2019; Luo et al., 2018; Riggs et al., 2020). LLPS depends on intramolecular interactions driven by low complexity sequences (LCSs) residing within intrinsically disordered regions (IDRs) and prion-like domains (PLDs) which frequently occur in P-body components (Jonas & Izaurralde, 2013). In yeast, LCSs are typically present within the Glutamine/Asparagine (Q/N)-rich domain of the decapping Sm-like protein Lsm4 and the highly conserved C-terminal domain of Edc3 (known as Yjef-N). Edc3 (as part of the 5' end-bound complex formed by Dcp1/Dcp2 and Dhh1, the yeast homolog of human DDX6) and Lsm4 (as part of the 3' end-bound complex together with Pat1, Xnr1 and Ccr4-Not) have been shown to promote P-body assembly through interactions between their respective LCSs (Decker et al., 2007; Reijns et al., 2008). The Yjef-N domain of Edc3 is highly conserved, suggesting this protein to likely have a similar function in higher eukaryotes. By contrast, the Q/N-rich domain

of Lsm4 is replaced by an RGG domain in higher eukaryotes suggestive of a different mechanism by which Lsm4 promotes P-body formation compared to yeast (Decker et al., 2007). In higher eukaryotes, additional proteins have been shown to contribute to P-body assembly/maintenance. Among these, the enhancer of mRNA decapping protein 4 (EDC4) (Yu et al., 2005), the cytoplasmic polyadenylation element-binding protein 1 (CPEB1) (Wilczynska et al., 2005) and the eukaryotic translation initiation factor 4E transporter (4E-T) (Andrei et al., 2005) whose depletion was shown to lead to P-body loss. Moreover, components of the microRNA (miRNA) repression pathway including the GW182 protein and the Argonaute proteins were also found in mammalian P-bodies (Liu et al., 2005; Rehwinkel et al., 2005). Nonetheless, GW182 was shown to accumulate in discrete cytoplasmic speckles identified as GW-bodies (Eystathioy et al., 2002). The subsequent observation that GW182 colocalized with mRNA decay associated proteins, a hallmark of P-bodies, led to the definition of GW-bodies and P-bodies as two indistinguishable entities (Eystathioy et al., 2003).

Despite the presence of mRNA decay components in P-bodies, these RNA granules are not only required for mRNA degradation but may also have a broader regulatory role as sites of mRNA storage. Under conditions of stress that inhibit translation initiation, mRNAs may be temporarily sequestered and kept translationally silent in P-bodies without necessarily being degraded afterwards. Translationally arrested mRNAs can then recycle from P-bodies to actively translating polysomes (Bregues et al., 2005).

Repression of translation initiation in response to stress also leads to formation of Stress Granules (SGs). Similar to P-bodies, SGs are membrane-less RNP organelles which form through LLPS in a way that is proportional to the pool of untranslating mRNAs. However, whereas P-bodies are constitutively present in certain cell lines and tend to increase in size and number upon stress, SG assembly is exclusively triggered by stress (Kedersha & Anderson, 2007; Kedersha et al., 2005). This observation implies that SGs are transiently formed compartments which disassemble upon removal of the insult (Panas et al., 2016). In response to exogenous and endogenous stressors, the cell diminishes translation rates by prioritizing synthesis of stress-protective proteins in order to promote cell survival (de Nadal et al., 2011). Translationally arrested mRNAs accumulate in cytoplasmic foci where SGs eventually form (Anderson & Kedersha, 2008). However, SGs are not stable depots of untranslated mRNAs, as indicated by the antagonistic effect of different pharmacological inhibitors of protein translation on SG assembly. In fact, they coexist in a dynamic equilibrium with a cluster of actively translating ribosomes called polysomes. SG assembly is enhanced by the disassembly of translating polysomes. Treatment with puromycin destabilizes polysomes by breaking up the ribosome when incorporated into the nascent polypeptide chain. The release of ribosomal subunits from mRNA transcripts disrupts the polysomal mRNP thereby increasing the pool of untranslated mRNAs and the formation of SGs. By contrast, SG assembly is not favored by polysome stabilization. Treatment with cycloheximide blocks translation elongation and freezes ribosomes on mRNAs thereby increasing the pool of polysome-associated mRNAs, a condition that prevents SG assembly (Anderson & Kedersha, 2002; Kedersha et al., 2000; Wolozin & Ivanov, 2019).

The dynamic shuttling of mRNAs between polysomes and SGs relies on proper assembly of both the cap-binding complex eIF4F (consisting of eIF4E, eIF4A and eIF4G) and the eIF2-GTP-tRNA^{Met} ternary complex (Jackson et al., 2010). Two major signaling pathways regulate SG formation by controlling these two early checkpoints in translation initiation (Wolozin and Ivanov, 2019) (**Fig. 1**). Specifically, the recruitment of eIF4F to the mRNA 5' end is regulated by the mechanistic target of rapamycin (mTOR). mTOR is a serine/threonine kinase which regulates cellular growth and metabolism by modulating protein synthesis in response to environmental inputs (Saxton & Sabatini, 2017). Under optimal growth conditions, mTOR phosphorylates the inhibitor of eIF4E, namely eIF4E-binding protein (4E-BP). When phosphorylated, 4E-BP is not able to bind its

target protein thereby allowing formation of the eIF4F complex and proper translation initiation. By contrast, conditions of metabolic stress such as amino acid and/or serum starvation inactivate mTOR resulting in the accumulation of hypophosphorylated 4E-BP. As a consequence, 4E-BP binds eIF4E thereby preventing proper assembly of the eIF4F complex and ultimately leading to inhibition of translation initiation (Thoreen et al., 2012). The resulting translationally arrested pre-initiation complex (PIC) that lacks essential translation initiation components, forms the seed for SG assembly (Wolozin & Ivanov, 2019).

Another early checkpoint of translation initiation is the assembly of the eIF2-GTP-tRNA^{Met} ternary complex. Delivery of a properly assembled eIF2-GTP-tRNA^{Met} ternary complex to the 40S ribosomal subunit is essential for the formation of a translationally competent 43S PIC. Phosphorylation of the α subunit of eIF2 in response to stress prevents eIF2B from reloading the ternary complex with GTP thereby reducing its availability and the overall translation initiation process (Sudhakar et al., 2000). Phosphorylation of eIF2 α on serine 51 is promoted by a family of four serine/threonine kinases that are activated by different types of stress (Wek, 2018). These four kinases play a key role in the context of the integrated stress response, a eukaryotic signaling pathway which promotes cell survival and homeostasis restoration by modulating protein synthesis in response to different stress stimuli. Specifically, general control non-derepressible 2 (GCN2) binds to deacylated tRNAs in response to amino acid deprivation; protein kinase R (PKR) is activated by double-stranded RNA (dsRNA) during viral infection; PKR-like ER kinase (PERK) activation is triggered by the accumulation of unfolded proteins in the endoplasmic reticulum; and heme-regulated eIF2 α kinase (HRI) is activated by heme deprivation and oxidative/osmotic stress (Pakos-Zebrucka et al., 2016). The reduced availability of ternary complexes caused by increased levels of eIF2 α phosphorylation is responsible for the loading of scanning-incompetent PICs at the 5' end of the mRNA. As this ternary complex-deficient PIC is unable to recruit the 60S subunit to form a translationally competent 80S ribosome, it is targeted to SGs. Elongating ribosomes that are already engaged in translation will complete peptide synthesis and disengage from the mRNA (a process known as ribosome run-off). However, the lack of translationally competent PICs assembling at the 5' end of the mRNA will eventually lead to polysome disassembly (Kedersha et al., 2002). The resulting accumulation of untranslated mRNAs carrying non-canonical PICs is essential for SG assembly. In parallel, RBPs involved in different aspects of mRNA metabolism initiate SG formation by interacting with either exposed regions of the untranslated mRNAs and/or components of the non-canonical PICs such as eIF3, eIF4F, PABP and the 40S subunit but not the 60S subunit (Ivanov et al., 2019; Kedersha et al., 2002; Kimball et al., 2003; Panas et al., 2016). Typical SG nucleators are Ras GTPase-activating protein-binding protein (G3BP1) (Tourriere et al., 2003), T-cell intracellular antigen-1 (TIA-1) and TIA-1 related protein (TIAR) (Gilks et al., 2004), tristetrapolin (TTP) (Stoecklin et al., 2004), fragile X mental retardation protein (FMRP) (Mazroui et al., 2002) and cytoplasmic activation/proliferation-associated protein 1 (CAPRIN1) (Solomon et al., 2007). Moreover, ataxin-2 (ATXN2) (Nonhoff et al., 2007) and AGO2 (Leung et al., 2006) were also identified as SG components, and ataxin-2 can be used as a reliable marker for SGs. Similar to P-body constituents, SG nucleators are characterized by prion-like domains (PLDs) or intrinsically disordered regions (IDDs) which mediate the interaction with PICs components (Kedersha et al., 2016). Several candidate proteins containing IDD have been identified within the 40S ribosomal subunit. Under normal conditions, the 40S subunit would be covered by the 60S subunit to form a translationally competent 80S ribosome. However, within translationally arrested PICs, ribosomal proteins located at the 40S subunit interface are exposed to recruit other RBPs via interaction between IDD (Panas et al., 2016). As an example, G3BP1 has been shown to interact with the 40S subunit via its RGG domain (Kedersha et al., 2016). Moreover, PLDs and IDDs are able to undertake different conformations as a consequence of post-translational modifications such as O-GlcNAcylation, methylation, dephosphorylation and deacetylation (Mahboubi & Stochaj, 2017). Therefore, due to their highly dynamic nature, PLDs and IDDs provide a scaffold for the recruitment of

additional proteins to SGs (Gilks et al., 2004; Kedersha et al., 2016) and facilitate the dynamic shuttling of these proteins in and out of SGs (Ivanov et al., 2019). A critical role in several aspects of SG formation is fulfilled by microtubule-associated motor proteins such as kinesins and dyneins. Indeed, their motor activity was found to be essential for mediating the movement of further mRNPs to SGs (Loschi et al., 2009). Overall, SGs consist of a stable core which serve as a platform for the formation of a transient surrounding shell (Wheeler et al., 2016). Canonical SGs consist of a series of defining components including 40S subunit, translation initiation factors eIF4F, eIF3 and PABP, and polyadenylated mRNAs (Ivanov et al., 2019). However, composition and dynamics of SGs may vary depending on the type of cell and the nature of the stress insult. For example, in U2OS cells, hydrogen peroxide induces formation of SGs that lack eIF3b and therefore differ from canonical SGs (Emara et al., 2012). Similarly, in HAP1 cells, UV-induced SGs contain less eIF3b and eIF4G (Aulas et al., 2017). Interestingly, both hydrogen peroxide (Emara et al., 2012) and UV-induced formation of SGs (Ying & Khapersky, 2020) occurs via a eIF2 α -independent mechanism. This indicates that although for many types of stress SG formation relies on the eIF2 α phosphorylation pathway, some stressors promote SG assembly in a eIF2 α -independent fashion (Farny et al., 2009).

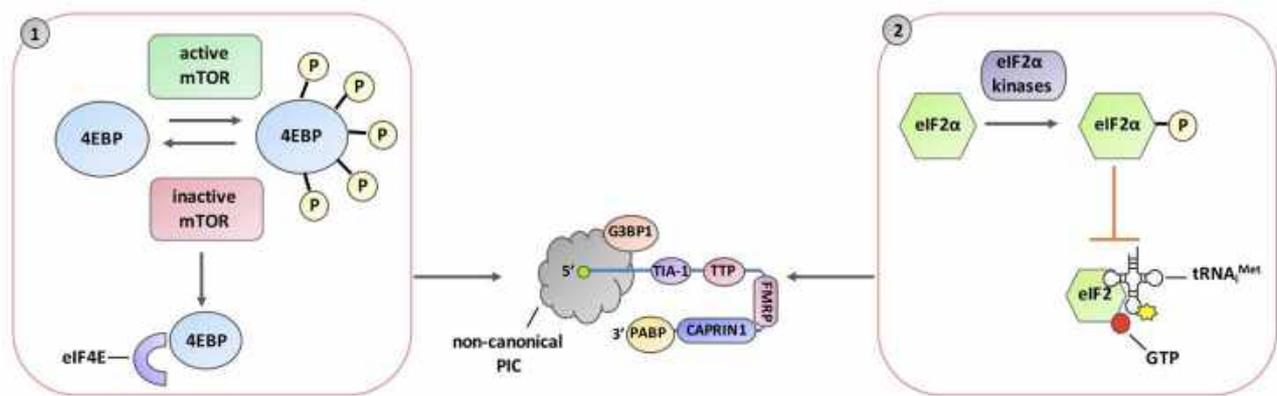


Figure 1 | Regulation of stress granule formation. Two major signaling pathways regulate formation of SGs by interfering with the assembly of key elements of the pre-initiation complex (PIC). (1) Inhibition of mTOR enables hypo-phosphorylated 4EBP1 to bind eIF4E, thereby preventing the assembly of the eIF4F complex on the cap structure of the mRNA. (2) Phosphorylation of eIF2 α prevents reloading of the ternary complex (eIF2-GTP-tRNA^{Met}) with GTP thereby blocking proper delivery of the ternary complex to the 40S ribosomal subunit. Both events result in the accumulation of untranslated mRNAs carrying non-canonical PICs which serve as a binding platform for typical SG nucleators such as G3BP1, TIA-1, TTP, FMRP and CAPRIN1.

Due to their dynamic nature, SGs disassemble upon removal of the stressor. Different factors contribute to SG disassembly. Among these, the ATPase-dependent activity of RNA helicases has recently emerged as a key factor in determining SG dynamics. Using the energy of ATP hydrolysis, RNA helicases unwind RNA structures thereby disrupting mRNA-protein interactions within SGs ultimately leading to their disassembly (Jain et al., 2016). SG clearance may also be mediated by the activity of chaperones such as the heat shock protein HSP70. Its overexpression inhibits SG formation by preventing TIA-1 aggregation (Gilks et al., 2004) and its pharmacological inhibition leads to delayed SG disassembly (Ganassi et al., 2016). Depending on how long they remain assembled, SGs can also be cleared by autophagy. SGs arising from chronic stress or disease mutations are removed by autophagy (Buchan et al., 2013). The valosin-containing protein (VCP) plays a key role in autophagy-dependent clearance of SGs. As an ATPase that catalyzes the extraction of ubiquitinated proteins from complexes, VCP may extract some ubiquitinated proteins from SGs thereby promoting targeting of SGs to autophagy (Buchan et al., 2013). However, the recent observation that ubiquitinated

G3BP1 is targeted by VCP for proteosomal degradation during SG disassembly after removal of heat stress suggests that short-lived SGs are disassembled in an autophagy-independent manner (Gwon et al., 2021).

Several observations indicate that mRNAs might accumulate in both SGs and P-bodies after polysome disassembly suggesting a working model whereby mRNAs cycle between polysomes, SGs and P-bodies. First, pharmacological inhibitors of translation (e.g. puromycin and cycloheximide) that modulate SG assembly have similar effects on P-body formation. Second, the hypothesis of an mRNA cycle is supported by the observation that P-bodies and SGs transiently interact with each other upon oxidative stress. Third, fluorescent recovery after photobleaching experiments (FRAP) highlights the dynamic nature of SG and PB components (Kedersha et al., 2005). Finally, although SGs and P-bodies are distinct and independent entities, as indicated by their composition, several proteins localize to both SGs and P-bodies (Stoecklin & Kedersha, 2013). Different models have been proposed trying to elucidate the mechanism by which mRNAs shuttle between polysomes, SGs and P-bodies but the exact mechanism remains unresolved. One possibility is that in response to defects in translation initiation, mRNAs interact with components of the decapping machinery and accumulate into P-bodies. Within P-bodies, translationally arrested mRNAs can either be targeted for mRNA decay or recycled to actively translating polysomes. However, under conditions of stress that inhibit translation initiation, translationally arrested mRNAs may shift from P-bodies to SGs and be stored until translation is allowed to resume (Decker & Parker, 2012). Another possibility is that in response to stress, the mRNA may be released from polysomes and directly targeted to SGs. In this model, SGs are described as sites of mRNA triage where untranslated mRNAs are sorted and remodeled for storage, reinitiation or transit to P-bodies for mRNA decay (Kedersha et al., 2005). The observation that SGs are defined by components of the translation initiation machinery and P-bodies by factors belonging to the mRNA decay machinery is suggestive of a dynamic integration between molecular, functional and morphological aspects of translation and decay. Moreover, the dynamic relationship between SGs and P-bodies as cytoplasmic entities capable of docking while remaining physically independent, highlights the fundamental role of compartmentalization during mRNP trafficking.

Given their central role in mRNA metabolism, it is not surprising that mRNP granules are closely related to a variety of human diseases (Buchan, 2014). Specifically, SGs have been found associated with cancers, neurodegenerative diseases, viral infections and inflammatory diseases. The relatively simple manipulation of SG dynamics using different chemical compounds targeting SGs offers a potential therapeutic intervention for the treatment of several diseases caused by aberrant SG formation (Wang et al., 2020).

Regulation of gene expression by small non-coding RNAs

SGs and P-bodies contain components of the RNA silencing pathways including the Argonaute proteins AGO1 and AGO2, and miRNAs (Leung et al., 2006; Liu et al., 2005). miRNAs are a class of small non-coding RNAs (ncRNAs) which regulate gene expression through post-transcriptional gene silencing (Leung & Sharp, 2006). Another major class of small ncRNAs is represented by short interfering RNAs (siRNAs). Initially discovered in plants (Hamilton & Baulcombe, 1999) and subsequently shown in *Drosophila* to guide the cleavage of their target RNA (Zamore et al., 2000), siRNAs are nowadays widely used as a synthetic tool to knockdown genes of interest both *in vitro* and *in vivo* (Kole et al., 2012). The mechanism of action of both miRNAs and siRNAs relies on the ability of a short RNA duplex to base-pair match the mRNA of the target gene due to the complementarity of one of the two strands. Depending on whether the source of the RNA duplex is endogenous or exogenous, two different pathways are responsible for carrying out the silencing activity: the miRNA pathway and the RNA interference (RNAi) pathway (**Fig. 2**). Initially discovered by Fire and Mello as a mechanism to manipulate gene expression in *C. elegans* (Fire et al., 1998), RNAi was subsequently found to

be essential in many eukaryotes for protecting the genome from viral infection and transposable elements (Chung et al., 2008; Ding & Voinnet, 2007). Whereas the miRNA pathway leads to gene silencing through translation repression of the target (Hendrickson et al., 2009), RNAi is mediated by the endonucleolytic cleavage of the target (Zamore et al., 2000). Specifically, miRNAs are transcribed in the nucleus by RNA polymerase II as primary transcripts (pri-miRNAs) equipped with a 5' m⁷G cap and a 3' poly(A) tail (Lee et al., 2004). The microprocessor complex formed by the RNase III endonuclease Drosha and the double-stranded RNA-binding protein DGCR8 excises the hairpin at a distance of 11 bp from the binding site thereby generating a precursor miRNA (pre-miRNA) of 60-70 nt in length (Yao et al., 2013). After being transported into the cytoplasm by the Exportin-5/Ran GTP nuclear export machinery (Lund & Dahlberg, 2006), pre-miRNAs are further processed by the RNase III Dicer and its associated dsRNA-binding protein TRBP. The endonucleolytic activity of Dicer cuts out the terminal loop and generates a ~22 bp double-stranded RNA (dsRNA) fragment known as miRNA/miRNA* complex (Hutvagner et al., 2001) which is ready to assemble with AGO proteins to form the RNA-induced silencing complex (RISC). Duplex loading into AGO proteins is a selective process based on the thermodynamic stability of the ends of the duplex (also known as the asymmetry rule). The strand whose 5' end harbors the less thermodynamically stable base-pairing serves as the guide strand and will be retained in AGO, whereas the other strand, the passenger strand will be ejected and eventually degraded (Meijer et al., 2014; Schwarz et al., 2003). Whereas miRNAs are transcribed from the genome (Lee et al., 2004), siRNAs originate either from exogenous dsRNA, viral or synthetic (Zamore et al., 2000), or from transposons, bidirectional transcription of specific chromosomal regions and mRNAs paired to antisense pseudogene transcripts (Ghildiyal & Zamore, 2009). Independent of whether the siRNA precursors are exogenous or endogenous, both will be processed by Dicer into 20-25 bp long dsRNAs characterized by a dinucleotide overhang at the 3' end and a monophosphate group at the 5' end (Zamore et al., 2000). Once loaded onto AGO2, the core component of RISC, the passenger strand of the siRNA duplex is cleaved by AGO2 and ejected. Once the duplex is disrupted, the siRNA or miRNA guide strand remains associated with AGO to form a mature RISC. Although in some organisms, like flies, siRNAs and miRNAs are loaded into different Argonautes, mammalian AGO proteins (Ago1-4) do not discriminate between siRNA and miRNA duplexes during loading (Yoda et al., 2010). However, among the Argonaute proteins, AGO2 is the only one capable of catalyzing the cleavage (Liu et al., 2004). The Argonaute protein subsequently uses the small guide RNA to associate with the complementary target sequence on the mRNA. Depending on the degree of complementarity between the guide RNA and the target mRNA, RNA silencing might occur through either translation repression or target cleavage. miRNAs are usually not fully complementary to their mRNA targets (Ghildiyal & Zamore, 2009). The presence of mismatches in the central part of the miRNA-mRNA duplex prevents the target from being cleaved. Therefore, miRNA-mediated mRNA silencing is thought to occur via translation repression and deadenylation followed by mRNA degradation (Jinek & Doudna, 2009; Olina et al., 2018). The molecular mechanism underlying miRNA-mediated translation repression is not completely understood. However, one possibility is that by interacting with the 3' UTR of the mRNA, miRNAs might interfere with the function of the eIF4F complex thereby inhibiting translation initiation (Fukao et al., 2014). Moreover, AGO proteins have been shown to recruit the GW182 protein to the mRNA target which in turn interacts with PABP and promotes recruitment of deadenylation and decapping enzymes, thereby potentially promoting mRNA degradation (Guo et al., 2010). Unlike miRNAs, siRNAs perfectly pair with their target mRNAs, a condition that enables AGO2 to catalyze the slicing of the target (Liu et al., 2004). Thus, although differing in respect to the source of the small ncRNA, the common point of convergence between the miRNA pathway and the RNAi pathway is the RISC assembly. It is therefore not surprising that AGO proteins, the core components of RISC, are conserved throughout all domains of life (Swarts et al., 2014). The observation that miRNAs are predicted to regulate more than 60% of mammalian genes (Friedman et

al., 2009), further suggests the central role of the small ncRNA silencing machinery in reorganizing the gene expression pattern.

AGO proteins have been shown to associate with typical P-body components such as GW182 and DCP1/2 (Liu et al., 2005) and their presence has been detected within SGs (Leung et al., 2006). However, several lines of evidence suggest that miRNA and siRNA-mediated gene silencing does not only occur within these structures. First, miRNA-mediated mRNA repression occurs independent of the appearance of P-bodies in the cell, suggesting that P-body formation is not a prerequisite for the silencing machinery function (Chu & Rana, 2006). Second, AGO2 is recruited to SGs only in the presence of miRNAs. This observation indicates that AGO2 assembles first onto the miRNA in the cytoplasm and then, under conditions of stress, the miRNA-associated AGO2 protein is targeted to SGs (Detzer et al., 2011).

The ability of siRNAs to regulate gene expression through RNAi provides firm basis to exploit this biological mechanism for the regulation of desired target genes and therefore for novel therapeutics. RNAi-based therapeutics rely on the use of synthetic siRNAs or their precursor short-hairpin RNAs (shRNAs) as a tool to knockdown genes of interest for the treatment of cancer, infectious diseases and genetic disorders (Burnett et al., 2011; Karim et al., 2018).

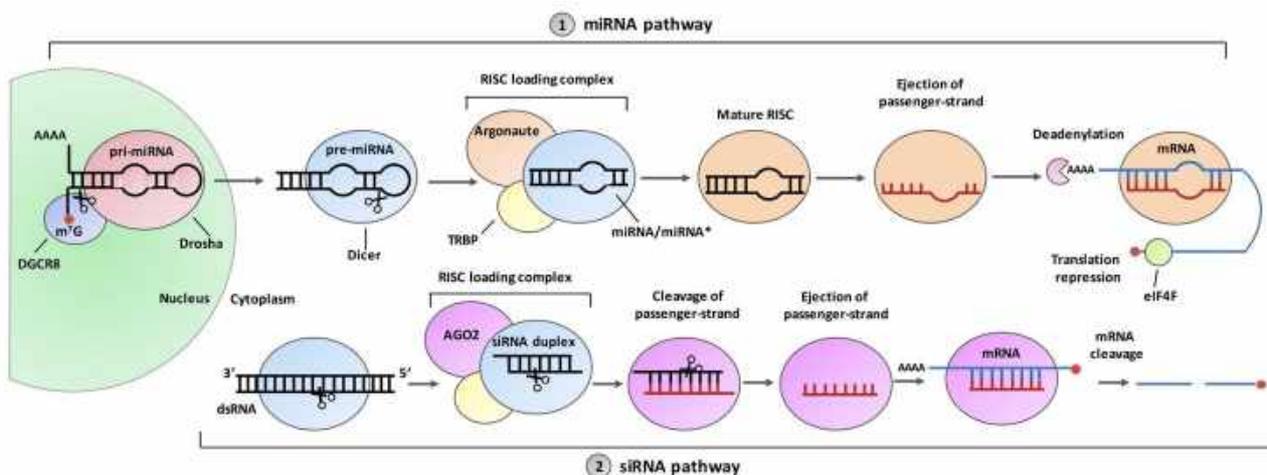


Figure 2 | Small non-coding RNA silencing pathways. Small non-coding RNAs are short duplex RNA molecules capable of silencing the target gene due to the complementarity of one of the two strands. (1) microRNAs (miRNAs) are encoded in the genome and transcribed as primary transcripts (pri-miRNA) which are then cleaved by Drosha-DGCR8 to yield a precursor miRNA (pre-miRNA). In the cytoplasm, Dicer catalyzes further cleavage of the pre-miRNA generating a duplex consisting of a guide strand and a passenger strand (miRNA/miRNA*). After loading onto an Argonaute protein the miRNA strand mediates mRNA silencing through translational repression or deadenylation. (2) Short-interfering RNAs (siRNAs) originate from exogenous long dsRNAs which are processed by Dicer in the cytoplasm to yield siRNA duplexes. Once loaded onto AGO2 within the RNA-induced silencing complex (RISC), the guide siRNA associates with the target mRNA allowing AGO2 to catalyze the cleavage.

1.3 | RNA metabolism and neurodevelopment

Regulation of gene expression at the post-transcriptional level has recently emerged as a key factor for proper development of the central nervous system. The development of the human brain is an extremely delicate process relying on the formation of highly specialized cells like neurons and properly connected neuronal circuits (Sartor et al., 2015). Proper formation of neuronal networks is essential for shaping the physical and behavioral development of the human organism; therefore, tight control of gene expression is

especially critical during early stages of neurodevelopment (Swanger & Bassell, 2011). The morphological complexity of neurons, their ability to change number, size and strength of synapses with gain of experience and the processing of local information strongly rely on the promptness of proteome modifications in axonal and dendritic synaptic compartments (Holt et al., 2019). mRNA trafficking along dendrites and axons and subsequent local mRNA translation at synapses have evolved to fulfill local demand for new proteins within a short time frame, thus playing a key role in axon guidance and growth, dendrite morphogenesis, synaptogenesis and synaptic plasticity (Holt et al., 2019; Swanger & Bassell, 2011). The observation that many forms of long-lasting synaptic plasticity, require dendritic and synaptic protein synthesis suggests that local mRNA translation is important for formation and storage of memories (Sutton & Schuman, 2006). Therefore, it is not surprising that during early steps of neurodevelopment, even small alterations in post-transcriptional gene regulation could negatively impact the ability of the brain to respond to local stimuli resulting in a wide range of neurodevelopmental disorders (NDDs) (Sartor et al., 2015; Tebbenkamp et al., 2014).

As key players in different steps of mRNA metabolism, RNA-binding proteins (RBPs) have a critical role in synaptic development and function and therefore in the regulation of learning and memory processes (Bardoni et al., 2012). The importance of accurate post-transcriptional gene regulation in formation and maintenance of neuronal networks has emerged in a broad range of neurodevelopmental and neurodegenerative diseases. One typical example is the Fragile X syndrome, the most common hereditary form of mental retardation in humans (Hagerman et al., 2017). The presence of a CGG trinucleotide repeat expansion in the 5'UTR of *FMR1*, encoding the ribosome-associated RBP FMRP, was relatively early identified as the cause of the Fragile X syndrome (Kremer et al., 1991). However, it was the recent development of high-throughput technologies and especially next generation sequencing (NGS)-based studies including whole-exome sequencing or whole-genome sequencing approaches, that have significantly contributed to the identification of the genetic basis of several mRNA metabolism related disorders. Given the high susceptibility of the brain to even small changes in mRNA stability and translation during early stages of neurodevelopment, it is not surprising that NGS-based approaches have established mutations affecting RBPs as a major cause of NDDs such as developmental delay (DD) and intellectual disability (ID) (Balak et al., 2019). The most recent examples include mutations in genes encoding proteins involved in different steps of mRNA metabolism such as mRNA decay and translation, SG and P-body assembly, and small non coding RNA-mediated gene silencing.

For instance, germline mutations were identified in genes encoding components of the mRNA decay machinery and regulators of translation repression such as decapping scavenger enzyme (DCPS) (Ng et al., 2015), enhancer of mRNA decapping 3 (EDC3) (Scheller et al., 2018), regulator of nonsense transcript 3B (UPF3B) (Alrahbeni et al., 2015) and the enhancer of decapping/repressor of translation DDX6 (Balak et al., 2019). These mutations have been shown to impair their decapping activity, stabilize specific subsets of mRNAs, reduce the interaction with key binding partners and interfere with their ability to control mRNA levels and assemble P-bodies (Sartor et al., 2015; Weil et al., 2020). Additionally, NDD-associated mutations identified in genes encoding RNA helicases involved in translation such as DDX3X and DHX30 were shown to drive SG assembly and impair global translation (Lessel et al., 2017; Valentin-Vega et al., 2016; Weil et al., 2020). As an additional confirmation of the correlation between RBPs-associated mutations and NDDs, several further pathogenic variants identified in genes involved in different steps of mRNA metabolism have been established as the major cause of these disorders. Examples include the RNA helicase encoding genes *DDX59* (Salpietro et al., 2018), *DHX16*, *DHX34*, *DHX37* and *DDX54* (Paine et al., 2019). Interestingly, the fact that all these genes, in addition to *DDX3X* and *DHX30*, encode DExD/H-box RNA helicases further suggests

the importance of this gene family in neurodevelopment. However, the impact of these mutations on the molecular and cellular function of the respective protein products has not yet been elucidated (Weil et al., 2020). Further examples include genes encoding components of the RNA polymerase II general transcription IID (TFIID) protein complex, such as TATA-box binding protein associated factor *TAF13*. *TAF13* variants interfere with the formation of the TFIID complex thereby causing a syndrome characterized by ID and microcephaly (Tawamie et al., 2017). Additional mutations were identified in genes involved in splicing such as those encoding the heterogeneous nuclear ribonucleoproteins (hnRNP) *HNRNPH2* (Bain et al., 2016) and *SON* (Kim et al., 2016). Whereas no functional analysis was performed in respect to the *HNRNPH2* variants, *SON* mutations were found to downregulate both mRNA and protein expression levels, thereby causing a complex NDD associated with ID and/or developmental delay (DD) and brain malformations. Furthermore, mutations identified in the post-transcriptional repressor Pumilio1 (*PUM1*) were shown to alter neuronal morphology and impair its ability to suppress its mRNA targets (Gennarino et al., 2018).

Although the introduction of NGS-based approaches has significantly contributed to rapidly identify the genetic cause of several NDDs, understanding the molecular mechanism underlying the functional and cellular impact of the identified mutations would be essential for assessing the pathogenicity and potentially identify effective therapeutic intervention.

1.4 | The RNA helicase *DHX30*

DHX30 encodes the DExH-box RNA helicase *DHX30*. As a member of the superfamily 2 (SF2) of RNA Helicases (RHs) containing more than 50 human proteins, *DHX30* owes its naming to the DExH amino acid sequence in its ATP-binding motif II (Umate et al., 2011). Based on similarities to published structures of other SF2 members, it has been proposed that *DHX30* possesses a modular domain organization consisting of a highly conserved helicase core region surrounded by multiple auxiliary domains. Whereas the helicase core region is believed to play an essential role for the catalytic activity of the protein, the function and therefore the biological relevance of the auxiliary domains remain so far unclear. Although *DHX30* was classified as an RH based on sequence similarities to other RHs, it is so far unknown whether this protein actually possesses helicase activity.

The importance of *DHX30* as an essential neurodevelopmental gene was highlighted by the initial observation that *dhx30* knock-out mice exhibit lethal developmental defects in the central nervous system (CNS) (Zheng et al., 2015) and by the subsequent classification of *DHX30* as a candidate gene for human neurodevelopment (Eldomery et al., 2017). Ultimately, 6 *de novo* missense mutations in *DHX30* were identified in 12 unrelated individuals affected by a severe NDD (Lessel et al., 2017). All the identified mutations were shown to impair either ATPase activity or RNA-binding and exhibited an increased propensity to trigger SG formation with concomitant inhibition of global translation. Although these findings provide strong evidence for the indispensability of this gene in the development of the CNS, the exact physiological function of *DHX30* in cellular RNA metabolism remains largely unknown. Our current knowledge of its cellular function relies on the observation that a specific isoform of *DHX30*, equipped with a mitochondrial signal peptide, has a prominent role in mitochondrial ribosome assembly (Antonicka & Shoubbridge, 2015). However, none of the patients described by Lessel et al. (2017) present with typical clinical signs of mitochondriopathies suggesting that the pathological phenotype is determined by others, non-mitochondrial isoforms of *DHX30*. Thus, for a better understanding of the disease, a comprehensive characterization of the molecular and cellular function of non-mitochondrial *DHX30* isoforms in particular will be required.

Aim of the project

Our understanding of the genetic causes of different forms of NDDs has significantly benefited from the recent introduction of high-throughput sequencing technologies such as next generation sequencing (NGS) approaches (Vissers et al., 2016). NGS-based studies in children affected by NDDs have led to the identification of potentially pathogenic mutations in genes encoding proteins involved in RNA metabolism. Analyzing the impact of these mutations on the molecular and cellular function of the affected protein product is essential to clarify whether these mutations are indeed pathogenic.

A growing number of human genetic studies have been reported focusing on the pathological relevance of altered function of RBPs. One recent example is *DHX30*, a member of the DExH-box family of RNA helicases. Several studies have classified *DHX30* as a candidate gene for human NDDs (Eldomery et al., 2017; Zheng et al., 2015). However, it was the identification of six *de novo* missense mutations in *DHX30* in 12 unrelated individuals with NDDs (Lessel et al., 2017) that ultimately confirmed the indispensability of proper function of this RNA helicase for the development of the central nervous system. More recently, the identification of further missense mutations in *DHX30* in similarly affected individuals has provided a strong motivation to further investigate the *DHX30*-related NDD and to shed light on the physiological role of *DHX30* in cellular RNA metabolism, an aspect which has been almost unexplored so far.

Therefore, the main purpose of the current work was to provide further understanding of the *DHX30*-related NDD and novel insights into the role of *DHX30* in the metabolism of RNA. Functional analysis of missense mutations identified in *DHX30* was performed to analyse the molecular and cellular dysfunction associated with these mutations and to infer molecular and functional properties of *DHX30*-WT. Moreover, elucidating the cellular function of *DHX30* was beneficial for a better understanding of the pathomechanism underlying the mutations in *DHX30*.

The observation that miRNAs can regulate formation and function of neuronal networks by modulating local mRNA translation highlights the importance of gene regulation mediated by small ncRNA in the development of the central nervous system (Rajman & Schratt, 2017). The main actor in the process of RNA interference is AGO2 which, among the Argonaute proteins, is the only one capable of catalyzing the cleavage of the target mRNA (Liu et al., 2004). So far, no genetic alterations in *AGO2* have been described to be associated with any human pathology. The recent identification of missense mutations in *AGO2* in patients affected by disturbances in neurological development has strongly motivated further characterization of the dysfunction associated with *AGO2* missense mutations.

Therefore, part of the current PhD work was dedicated to establish experimental conditions aiming to provide molecular insights into how mutations in *AGO2* affect human neurological development and cause a Mendelian disorder.

The Materials and Methods, and Results sections in this thesis are reported as part of the two publications presented in Chapter 2 and Chapter 3.

Chapter 2 | Germline AGO2 mutations impair RNA interference and human neurological development

Lessel D, Zeitler DM, Reijnders MRF, Kazantsev A, Hassani Nia F, Bartholomäus A, Martens V, Bruckmann A, Graus V, McConkie-Rosell A, McDonald M, Lozic B, Tan ES, Gerkes E, Johannsen J, Denecke J, Telegrafi A, Zonneveld-Huijssoon E, Lemmink HH, Cham BWM, Kovacevic T, Ramsdell L, Foss K, Le Duc D, Mitter D, Syrbe S, Merkenschlager A, Sinnema M, Panis B, Lazier J, Osmond M, Hartley T, Mortreux J, Busa T, Missirian C, Prasun P, Lüttgen S, **Mannucci I**, Lessel I, Schob C, Kindler S, Pappas J, Rabin R, Willemsen M, Gardeitchik T, Löhner K, Rump P, Dias KR, Evans CA, Andrews PI, Roscioli T, Brunner HG, Chijiwa C, Lewis MES, Jamra RA, Dymont DA, Boycott KM, Stegmann APA, Kubisch C, Tan EC, Mirzaa GM, McWalter K, Kleefstra T, Pfundt R, Ignatova Z, Meister G, Kreienkamp HJ.

Nat Commun. 2020 Nov 16;11(1):5797. doi: 10.1038/s41467-020-19572-5. PMID: 33199684; PMCID: PMC7670403.

Supplementary information is available for this paper at <https://www.nature.com/articles/s41467-020-19572-5>

ARTICLE

 Check for updates

<https://doi.org/10.1038/s41467-020-19572-5>

OPEN

Germline AGO2 mutations impair RNA interference and human neurological development

Davor Lessel  et al.[#]

ARGONAUTE-2 and associated miRNAs form the RNA-induced silencing complex (RISC), which targets mRNAs for translational silencing and degradation as part of the RNA interference pathway. Despite the essential nature of this process for cellular function, there is little information on the role of RISC components in human development and organ function. We identify 13 heterozygous mutations in AGO2 in 21 patients affected by disturbances in neurological development. Each of the identified single amino acid mutations result in impaired shRNA-mediated silencing. We observe either impaired RISC formation or increased binding of AGO2 to mRNA targets as mutation specific functional consequences. The latter is supported by decreased phosphorylation of a C-terminal serine cluster involved in mRNA target release, increased formation of dendritic P-bodies in neurons and global transcriptome alterations in patient-derived primary fibroblasts. Our data emphasize the importance of gene expression regulation through the dynamic AGO2-RNA association for human neuronal development.

[#]A list of authors and their affiliations appears at the end of the paper.

RNA interference is a major mechanism for post-transcriptional regulation of gene expression. Precursors of microRNAs (miRNAs) are transcribed, processed into mature miRNAs and loaded onto Argonaute (AGO1-4) proteins to form the RNA-induced silencing complex (RISC)¹. Each miRNA may recognize a set of target mRNAs by base pairing², which leads to translational silencing and mRNA degradation in cytoplasmic structures referred to as processing (P-) bodies³⁻⁵. Biallelic loss of *Ago2* leads to early embryonic lethality in mice exhibiting various development defects including anomalies of the central nervous system⁶. Therefore, accurate regulation of gene expression by the RNA interference pathway seems to be of utmost importance for proper development and maintenance of complex neural circuits^{7,8}. However, so far no genetic alterations in the gene encoding for AGO2 have been described which are associated with any human pathology. Thus, it still remains largely elusive how RNA interference, and especially domains of AGO2 or local sequences down to single amino acid residues, regulate human organismal development and function.

Here, we demonstrate that germline *AGO2* mutations affect human neurological development and provide molecular insight into how AGO2 dysfunction causes a human Mendelian disorder.

Results

Identification of patients bearing germline *AGO2* mutations.

During trio whole-exome sequencing of a cohort of 50 children affected by developmental disturbances and neurological manifestations of unknown etiology⁹, we identified a patient bearing a de novo missense mutation p.L192P in *AGO2* (NM_012154.5). Based on the ExAC and gnomAD sequencing data, *AGO2* is one of the most missense-intolerant genes in the human genome, ranked 15 and 30 respectively. Its Z scores of 7.696 (ExAC dataset) and 6.058 (gnomAD dataset) are far higher than the average Z score for genes involved in developmental

disorders¹⁰⁻¹². Only two *AGO2* non-synonymous alterations (p.G88V and p.E186K) with an allele frequency of more than 0.0001 are deposited in ExAC and gnomAD datasets. Moreover, the p.L192P variant was not present in publicly available datasets (dbSNP, ExAC and gnomAD), had a high in silico pathogenic prediction score (CADD_Phred of 27.1) and changed a highly conserved residue (Supplementary Fig. 1). Furthermore, a previous study aiming to identify de novo variants in 20 individuals with sporadic non-syndromic intellectual disability identified a de novo p.L190P in *AGO1*¹³. The leucine at position 192 in *AGO2* corresponds to the leucine at position 190 in *AGO1* (Supplementary Fig. 1). In addition, five individuals affected by neurodevelopmental disturbances bearing a deletion that encompasses both *AGO1* and *AGO3* have been documented previously¹⁴. These findings motivated us to search for further cases carrying heterozygous *AGO2* variants utilizing both the internet-based GeneMatcher tool¹⁵, and direct contact to our network of collaborators.

This approach led to identification of altogether 21 patients, including our index patient, affected by mild to severe global neurodevelopmental delay, who were discovered either by whole-exome sequencing or microarray-based comparative genomic hybridization. Eleven missense mutations, one in-frame deletion (Fig. 1a) and a 235.3-kb deletion involving the first three exons occurred de novo (Fig. 1b). In addition, one of the missense mutations (p.T357M) was transmitted from a similarly affected mother. Five mutations (p.L192P, p.G201V, p.T357M, p.M364T, p.C751Y) were recurrent. None of the mutations were present in publicly available datasets, and similarly to p.L192P, all missense mutations display high in silico pathogenic prediction scores (mean CADD_Phred of 29.1) and change highly conserved residues (Supplementary Data 1; Supplementary Fig. 1). Thus, the genetic data, especially the recurrence of mutations, already provide strong evidence for their pathogenicity.

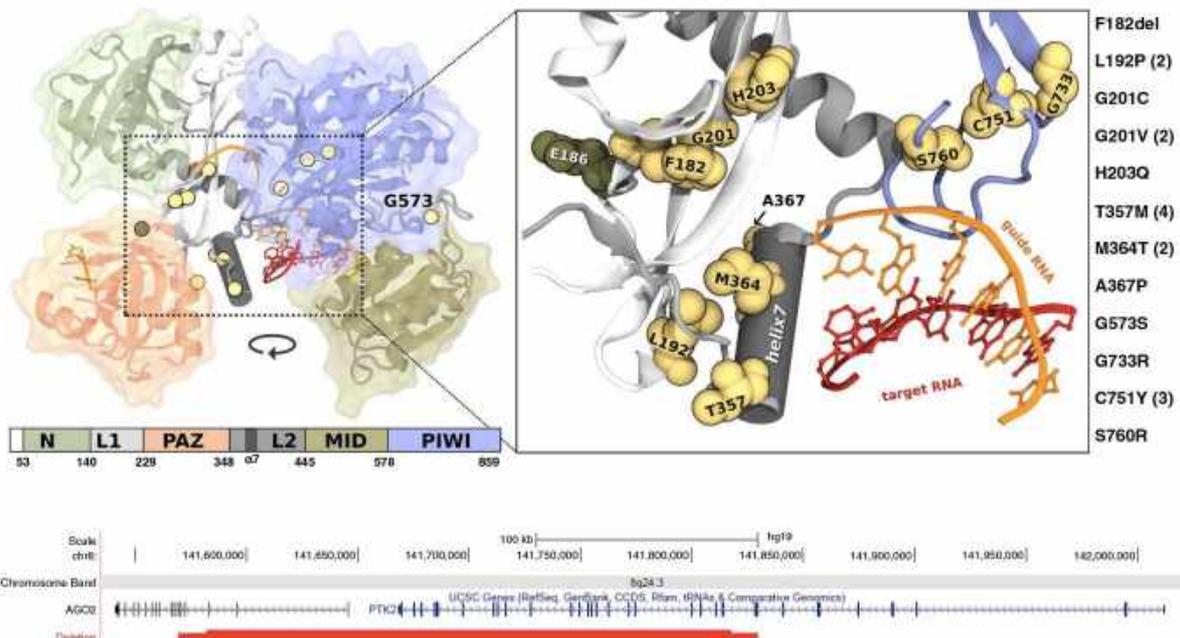


Fig. 1 Location of identified *AGO2* germline mutations. **a** Domain structure of AGO2 and position of the single amino acid mutations using the structure of human AGO2 in complex with a miRNA and a target RNA²². Guide and target RNA are depicted in orange and red, respectively. The recurring mutations are designated in brackets. **b** Genomic region, chr8:hg19:g.(141,582,269-141,817,600)del, of the 235.3-kb deletion identified in case 21, involving the first three *AGO2* exons and the last 23 *PTK2* exons.

Table 1 Summary of clinical findings in individuals bearing AGO2 mutations.

	Amount	Percentage
Neurological signs		
Intellectual disability	21/21	100 %
Motor developmental delay	21/21	100%
Impaired speech development	21/21	100%
Impaired receptive language	13/13	100%
Muscular hypotonia	12/21	57%
Autistic features	9/16	56%
Cerebral MRI abnormalities	9/16	56%
Gait abnormalities	10/18	55%
Attention deficit hyperactivity disorder	8/15	53%
Seizures	8/18	44%
Strabism	7/21	33%
Visual impairment	6/21	29%
Abnormal respiration	5/19	26%
Agressive behavior	4/17	24%
Myopia/Hyperopia	4/21	19%
Craniofacial abnormalities		
Epicanthic folds	11/21	52%
Thin upper lip	11/21	52%
Dental anomalies	9/19	47%
Frontal bossing	9/21	43%
Open mouth appearance	9/21	43%
Deep set eyes	9/21	43%
Upslanting palpebral fissures	6/21	29%
Congenital anomalies of the skull	6/21	29%
Helix anomalies	5/21	24%
Broad nasal bridge	3/21	14%
Other findings		
Neonatal feeding difficulties	12/19	63%
Skeletal anomalies	9/19	47%
Gastroesophageal reflux	7/19	37%
Heart anomalies	6/18	33%

The 21 individuals show overlapping phenotypes, summarized in Table 1, and in more detail in Supplementary Note 1 and Supplementary Fig. 2. All individuals displayed intellectual disability, albeit of variable degree, as well as delayed motor development, impaired speech and receptive language development (13/13). Twelve patients had hypotonia (57%) and ten had gait abnormalities (10/18, 55%). Nine patients (9/16, 56%) showed features of autism spectrum disorder, including stereotypic and hand-flapping behavior, eight patients (8/15, 53%) showed features consistent with attention deficit hyperactivity disorder, and four (4/17, 24%) developed aggressive behavior, predominantly upon entering puberty. Eight patients developed seizures (8/18, 44%). Brain anomalies on MRI, mainly affecting the corpus callosum, were observed in nine patients (9/16, 56%). Vision problems included visual impairment in six (29%), strabismus in seven (33%) and myopia or hyperopia in four (19%) patients. Various breathing abnormalities were observed in five (5/19, 26%) and included central apnea in the postnatal period (observed in both cases carrying the p.L192P mutation), sleep apnea and hypopnea. Craniofacial abnormalities included epicanthic folds (57%), thin upper lip (52%), frontal bossing (43%), open mouth appearance (43%), deep-set eyes (43%), upslanting palpebral fissures (29%), congenital anomalies of the skull (29%) including plagiocephaly (5/21) and scaphocephaly (1/21), various ear helix anomalies (24%) and broad nasal bridge (14%). Dental anomalies were observed in nine (9/19, 47%). Twelve had neonatal feeding difficulties (12/19, 63%). Skeletal anomalies, not including congenital anomalies of the skull, were observed in nine patients (9/19, 47%). Notably, all three

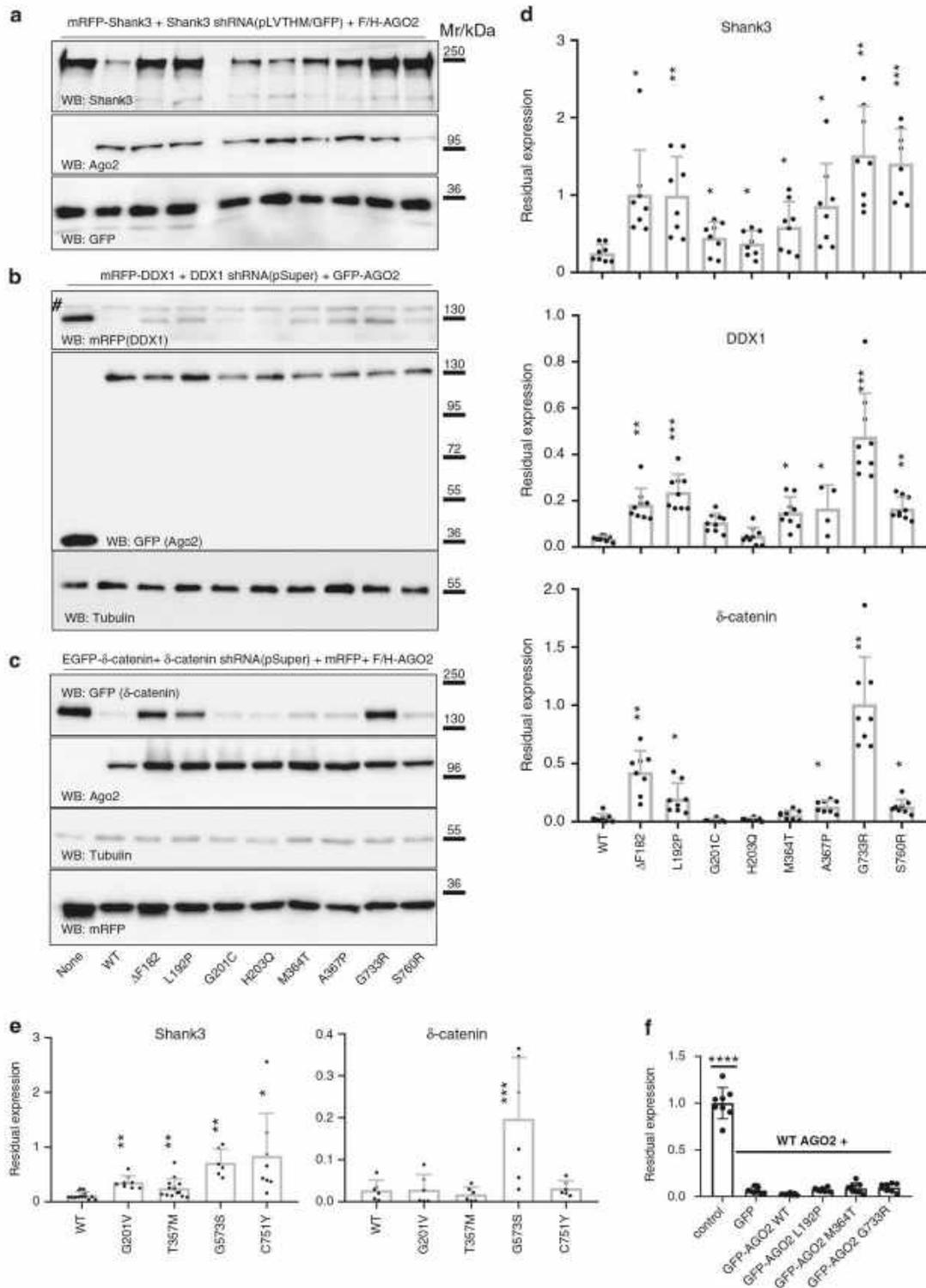
individuals carrying the p.C751Y mutation had bilateral clinodactyly of the 5th finger. Seven individuals experienced gastroesophageal reflux (7/19, 37%). Heart anomalies were observed in six patients (6/18, 33%), whereas three of them had patent foramen ovale.

Spatial clustering of residues affected by germline AGO2 mutations.

AGO2 consists of N-terminal (N), Piwi/Argonaute/Zwille (PAZ), middle (MID) and PIWI (P-element induced wimpy testis) domains which are connected by linker regions L1 and L2 (Fig. 1a). The N domain is indispensable for RNA unwinding during RISC formation¹⁶. The PAZ domain binds to the 3' end of the guide RNA and is involved in RISC activation^{17,18}. The MID domain provides a binding pocket for the 5' end of the guide RNAs¹⁹. The PIWI domain harbors a conserved catalytic core that cleaves the passenger strand, mediates protein-protein interactions needed for the enrollment of GW182, and regulates the interaction between RNA and the MID domain^{20,21}. Binding determinants for miRNA (in the AGO2/miRNA RISC) and the structural changes occurring upon engagement of mRNA targets have been identified through structural analyses^{22,23}. An α -helical segment in the L2 linker (helix-7, residues 359-369) is responsible for the proper positioning of the guide RNA during scanning of target mRNAs; here, residue I365 intercalates between guide RNA bases and induces a conformational change which is required for rapid target recognition²⁴. We identified three mutations in helix-7: p.T357M (four patients), p.M364T (two patients) and p.A367P. All three residues face towards the L1 linker and are in close spatial proximity to L192 in L1, mutated in two patients (p.L192P). Upon target mRNA recognition, helix-7 and the PAZ domain move relative to the rest of the protein by 4 Å, to accommodate the target mRNA strand²². The hinge for this movement is partially localized at the base of L1, where we identified p.F182del, p.G201C, p.G201V (two patients) and p.H203Q; this is a region of the protein which has previously been implicated in the unwinding of RNA duplexes¹⁶. Finally, mutations p.G733R, p.C751Y (three patients, including monozygotic twins) and p.S760R affect residues which are in direct contact to each other, and located at the base of a loop in the PIWI domain which binds to the minor groove in the guide/target duplex²². Thus, we identify three spatial clusters of mutations: at the helix-7/L1 interface, at the hinge region of L1, and in a loop in PIWI which recognizes the guide-target duplex (Supplementary Fig. 3). Only p.G573S affects a residue outside of these clusters at the C-terminal end of the MID domain (Fig. 1a). In addition, the de novo (as determined by qPCR analysis) long-range deletion affecting the first three AGO2 exons (Fig. 1b) is likely to result in haploinsufficiency (Supplementary Fig. 4), similar to deletions affecting both AGO1 and AGO3¹⁴.

AGO2 germline mutations impair shRNA-mediated silencing.

To analyze how single amino acid mutations affect AGO2 function, we deleted AGO2 in human HEK293T cells using CRISPR/Cas9-technology. Complete loss of AGO2 was confirmed by Western blotting (Supplementary Fig. 5). We used three shRNAs which had been established before in the lab and shown to be effective in reducing expression of Shank3, δ -catenin and DDX1, respectively. In line with the previously established essential role of AGO2 in shRNA-mediated gene silencing²⁵, strongly reduced silencing activity was observed for these shRNAs in AGO2 deficient cells, as they did not silence coexpressed Shank3 (Fig. 2a), DDX1 (Fig. 2b) and δ -catenin (Fig. 2c and Supplementary Fig. 6) mRNAs. Re-expression of WT-AGO2 efficiently rescued this phenotype, with residual expression of target proteins Shank3 of about 20%, and DDX1 and δ -catenin of <2% compared to



AGO2-deficient cells. All mutants were significantly less efficient than WT AGO2 in Shank3 silencing, whereas only some of the mutations showed significant effects for DDX1 and δ -catenin (Fig. 2a–e), suggesting target mRNA specific effects of some mutations. We further used this experimental system to investigate whether AGO2 variants may exert a dominant-negative effect

on the function of the WT AGO2 protein. Notably, we observed that knockdown of DDX1 expression by the DDX1 shRNA in the presence of F/H-tagged WT AGO2 was not affected by coexpression of selected mutant GFP-tagged AGO2 (p.L192P; p.M364T or p.G733R; see Fig. 2f). Thus, these data suggest that, at least the analyzed mutants, result in a loss-of-function.

Fig. 2 Germline AGO2 mutations impair shRNA-mediated silencing. **a–c** AGO2-deficient HEK293T cells were transfected with mRFP-Shank3 (**a**), mRFP-DDX1 (**b**), or GFP-tagged δ -catenin (**c**) along with shRNA constructs targeting Shank3 (**a**; in pLVTHM vector coexpressing GFP), DDX1 (**b**; in pSuper) or δ -catenin (**c**; also in pSuper), and Flag/HA-tagged (Shank3, δ -catenin) or GFP-tagged (for DDX1) AGO2 variants as indicated. Efficient transfection (>70% of cells) was verified by fluorescence microscopy using appropriate fluorophores (GFP in **a**) and **b**); RFP in **c**). Cells were lysed and lysates were analysed by Western Blotting using the antibodies indicated. In each case, each experiment was repeated eight times with similar results. **d** Quantification of the representative immunoblots shown in **a–c**. The expression levels are relative to those in control cells without an AGO2 construct (first lane in each immunoblot). **e** Quantification of a second set of AGO2 mutants. For **d, e**, data are means \pm SEM. Data for mutant and control conditions were compared to wt. *, **, *** $p < 0.05$, 0.01, 0.001, respectively; $n = 8$ biologically independent experiments in **d**); one-way-ANOVA, followed by Holm-Sidak's multiple comparisons test; $n = 6–14$ biologically independent experiments in **e** for Shank3; mixed-effects analysis, followed by Holm-Sidak's multiple comparisons test; $n = 6$ biologically independent experiments in **e**) for δ -catenin; one-way ANOVA, followed by Dunnett's multiple comparisons test. **f** Knockdown of DDX1 with DDX1 shRNA was performed as in **b**), but using F/H-tagged AGO2 in combination with GFP, or in combination with GFP-tagged AGO2 variants. Note that the GFP-tagged AGO2 mutants do not interfere with the knockdown capacity of F/H-AGO2 WT. $n = 8$ biologically independent experiments; data are means \pm SD; **** $p < 0.0001$; one-way ANOVA, followed by Dunnett's multiple comparisons test. Source data are provided as a Source Data file.

As the abovementioned experiments were performed by overexpression of AGO2 variants, we next carefully calibrated our experimental system over a wide range of WT AGO2 concentrations. In brief, we transfected different amounts of AGO2 expression vectors into the AGO2 deficient cells, while keeping amounts of cotransfected shRNA (for δ -catenin) and target vector (GFP- δ -catenin) constant. AGO2 expression relative to native HEK293T cells ranged from 54% for WT AGO2, when the lowest concentrations of DNA was used (0.1 μ g), to an about 20-fold overexpression when 1 μ g was transfected into each well of a 12 well plate (Supplementary Fig. 6). Interestingly, the silencing efficiency of WT AGO2 was maintained almost unaltered over this wide range of concentrations. Even at expression levels which are lower than those in unmodified HEK293T cells, efficient downregulation of δ -catenin was observed. Importantly, both at low and at high concentrations of AGO2, differences between WT and mutant forms of AGO2 were maintained, such that even at more than 10-fold overexpression, the p.F182del, p.L192P, and p.G733R variants did not silence as efficiently as AGO2 WT at its lowest concentration (Supplementary Fig. 6).

We conclude from these titration experiments that strong differences between WT and mutant forms of AGO2 are observed at physiological levels of the protein (i.e., at levels found in the non-modified HEK293T cell line).

RISC formation is differentially affected by germline AGO2 mutations. Having established a functional deficit for all single amino acid mutations, we used a spectrum of assays to determine which step of the miRNA pathway might be affected. We tested the ability of AGO2 variants to interact with DICER, the protein responsible for pre-miRNA processing and loading of the mature miRNA onto AGO2 (Supplementary Fig. 7), and with the proteins of the TNRC6 family that associate with AGO proteins during miRNA-guided gene silencing (Supplementary Fig. 8). We further assessed the capacity to bind to an endogenous miRNA (miR19-b) in HEK293T cells (Supplementary Fig. 8), the nuclease or slicing function of the AGO2 variants in *in vitro* cleavage assays using a radiolabeled substrate²⁶ (Supplementary Fig. 9), and the ability to silence a luciferase-based reporter mRNA in HeLa cells (Supplementary Fig. 10). Finally, correct targeting of AGO2 to P-bodies^{3,27} was investigated in U2OS cells (Supplementary Fig. 11). In these assays, most mutants performed similar to WT-AGO2, with the exception of p.G733R, which appeared to be non-functional in almost every aspect tested. Thus, mutations in AGO2 have two functional consequences: first, in the majority of the mutants, basic aspects of RISC formation and AGO2 function are not affected, and second the p.G733R mutant

exhibits a loss-of-function in almost every assay, likely similar to the anticipated effect of the here identified 235.3-kb deletion.

Reduced phosphorylation and altered mRNA target release of germline AGO2 mutations. AGO2 function is regulated by protein kinases whereby phosphorylation at Ser387 by AKT3 favors translational repression of targets over degradation²⁸. Importantly, phosphorylation of a C-terminal cluster (S824–S834) by casein kinase α 1 (CSNK1A1) is associated with the final step of AGO2-mediated mRNA repression, which enhances the release of the target mRNA from the RISC complex^{29,30}. We measured phosphorylation of Ser387 and the C-terminal cluster upon immunoprecipitation of AGO2 from HEK293T cells, followed by quantitative mass spectrometry. Phosphorylation of S387 was not altered, whereas phosphorylation of the C-terminal cluster was strongly reduced in all mutants tested, except for p.H203Q (Fig. 3a). By performing qRT-PCR on RNA samples isolated from AGO2 immunoprecipitates, we further observed that all mutants exhibiting reduced phosphorylation of this cluster also showed an increased association with a set of known target mRNAs, with the exception of p.G733R, which did not bind any mRNAs further supporting a complete loss of function of this mutant (Fig. 3b). These data further corroborate the link between phosphorylation of the S824–S834 cluster and the mRNA release from AGO2^{29,30}, and indicate a slower release of the majority of the identified AGO2 mutants from target mRNAs. Thus, we suggest a model in which a reduced phosphorylation of the C-terminal serine cluster of most of the disease-causing mutants coincides with reduced target release and thus an extended dwelling time of the AGO2 mutants on their targets.

Molecular dynamics simulations suggest an effect of germline AGO2 mutations on AGO2 unwinding function. To further clarify the effect of AGO2 mutations, apart from the clear loss-of-function mutation p.G733R, we performed non-biased molecular dynamics (MD) simulations. Our goal was to gain further insight into the effect of the above-mentioned mutations in apo-AGO2 and AGO2-RNA complexes. As a positive control, we included p.F181A which was shown before in an alanine scanning mutagenesis to reduce the unwinding of both siRNA and miRNA duplexes¹⁶. As a negative control, we included p.E186K, one of the two common non-synonymous AGO2 variants found in public repositories. AGO2 mutations were simulated in five states corresponding to different complexes along RISC formation: (i) apo-AGO2; two AGO2 complexes with guide RNA (core-RISC), namely intercalating (ii; *int* core-RISC) and non-intercalating (iii; *non-int* core-RISC); and two AGO2 complexes with guide-target duplexes (holo-RISC) with fully matched seed region (iv; g2-7

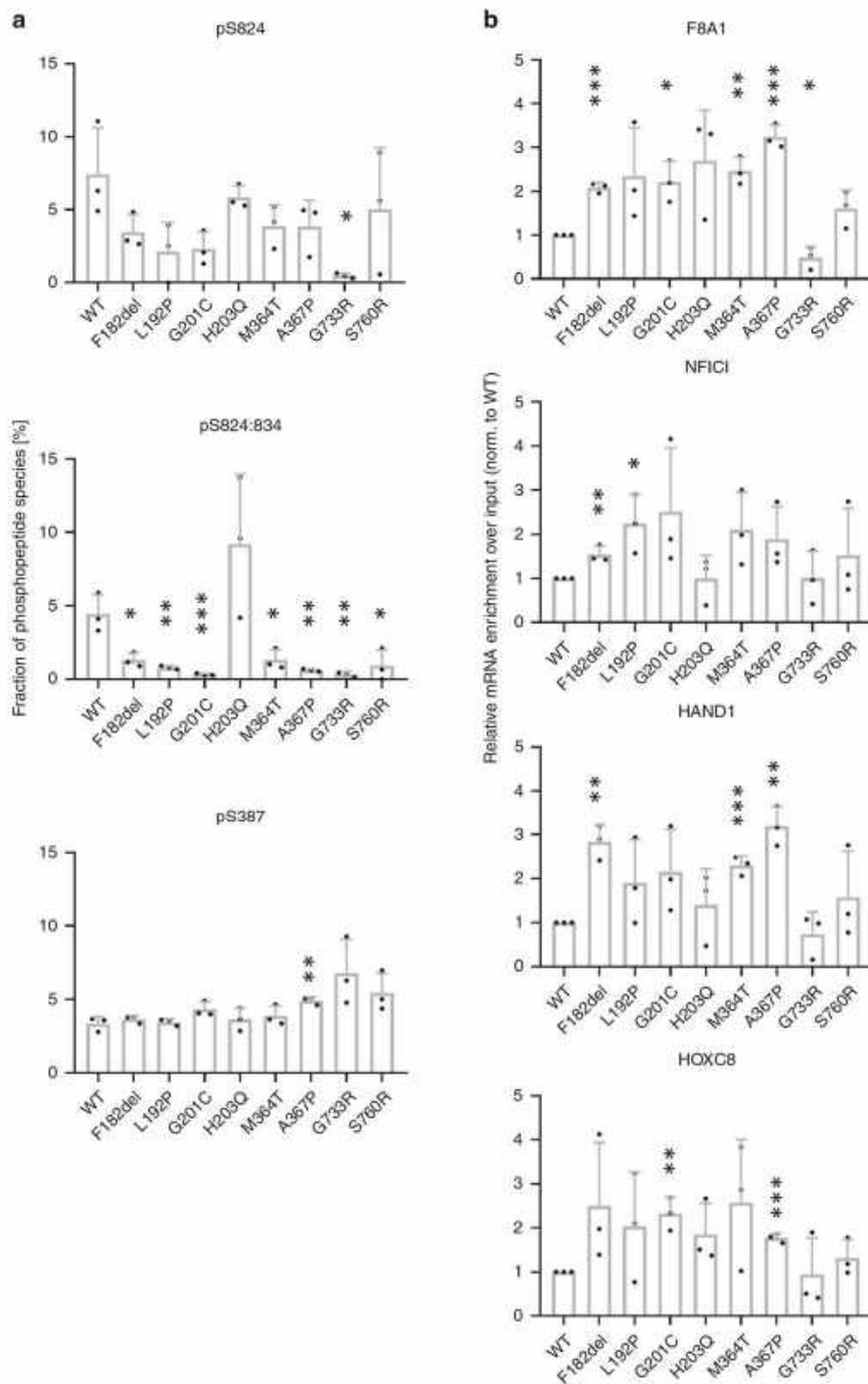


Fig. 3 Reduced target dissociation of AGO2 germline mutants. **a** 293 cells were transfected with F/H-tagged AGO2. After immunoprecipitation, phosphorylation of 5824 alone, of the 5824–5834 cluster and of 5387 was measured by a targeted quantitative mass spectrometry approach (Selected reaction monitoring with isotopically labeled spike-in peptides). The y-axis represents the percentage of individual phosphorylated peptide species assuming the sum of singly, multiply and non-phosphorylated peptides to be 100%. Significance was assessed by two-sided Student's *t* test in relation to WT. *n* = 3 biologically independent experiments. Data are presented as mean + SD. *, **, ****p* < 0.05, 0.01, 0.001, respectively. **b** RNA was isolated from F/H-AGO2 immunoprecipitates and analyzed by qRT-PCR using primers for the genes indicated. The significance was assessed by two-sided Student's *t* test in relation to WT. *n* = 3 biologically independent experiments. Data are presented as mean + SD. *, **, ****p* < 0.05, 0.01, 0.001, respectively. Source data are provided as a Source Data file.

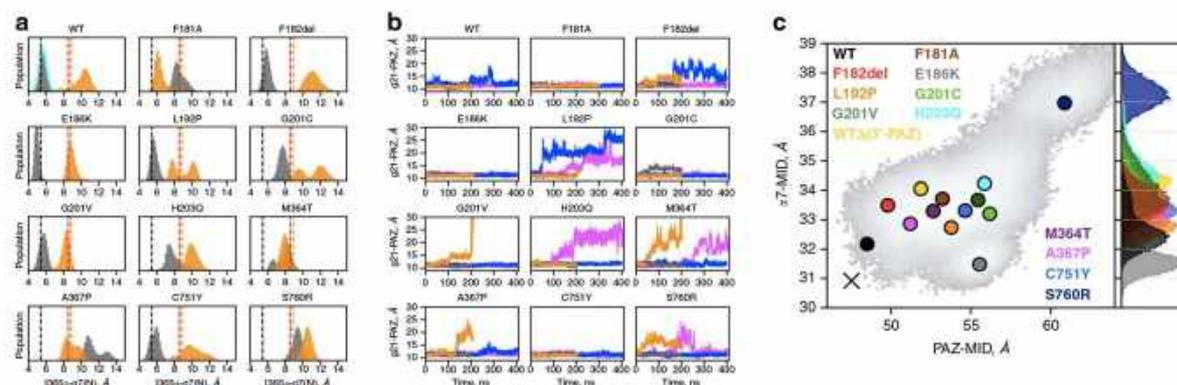


Fig. 4 Molecular dynamic simulation of effects of AGO2 mutations. **a** Distance between the C α atom of I365 and the glycosidic N atom of the g7 residue of the guide (I365 α -g7(N)). Populations of all core-RISC trajectories were normalized to the same bin number (60). Black and orange dashed vertical lines denote the reference X-ray structures of the *int* and *nonint* core-RISC states (4OLA and 4W5N, respectively). The red dashed line denotes the X-ray structure of the helix-7 mutant (5WEA). **b** Distance between the c.o.m. of the 3'-end of the guide and the c.o.m. of the PAZ domain (g21-PAZ) along the trajectories of all RNA-bound states. Note that the relatively short length of the individual MD trajectories could affect the results. Color code of the states in both a and b panels: apo-Ago2—green, *int* core-RISC—gray, *nonint* core-RISC—orange, g2-8 holo-RISC—magenta and g2-7 holo-RISC—blue. **c** Motion of the helix7 and the PAZ domain along the open-closed mode (left panel, gray scale histogram). The histograms are calculated by concatenating the last 100 ns of non-biased trajectory of each variant with mismatched RNA duplex. Colored circles depict the maximum population density of each trajectory. Black cross denotes the maximum population density of the WT AGO2 with guide RNA. Right panel: population histograms on α 7-MID corresponding to the maxima on the left panel. Equivalent analysis of the variants in in complex with a fully matched seed duplex is shown on Fig. S16.

holo RISC) or with a mismatch at position 8 of the guide (v; g2-8, holo RISC). The underlying structures and pdb codes are listed in Supplementary Fig. 12. This large number of simulations (12 AGO2 variants in five different states) was chosen as an initial screening approach to identify possible alterations in mutant forms of AGO2; however, this large number also precluded detailed quantitative measurements. Among the simulated complexes, the mutations appeared not to uniformly affect the population distributions on this conformational mode. Principal component analysis of apo-AGO2 trajectories suggested that only two mutations, p.L192P and p.F182del, affected global protein dynamics. The predominant conformational mode of these two variants deviated from the open-closed mode, previously described as the most pronounced conformational mode of AGO proteins³¹ (Supplementary Fig. 13). Further, analyses of the non-biased MD trajectories of RNA-bound complexes suggested two effects: (i) a compromised interaction in the *int* core-RISC state between helix-7 and g7 which was observed for the mutations p.G201C, p.H203Q, p.M364T, p.A367P and p.S760R, similar to the p.F181A positive control (Fig. 4a); and (ii) a loss of anchoring of the 3'-end of the guide (g21) at the PAZ domain, observed for all patient-derived mutations apart from the p.C751Y in at least one simulated state (Fig. 4b). The first effect (on the helix-7-g7 interaction) is somewhat reminiscent of a previously reported mutation in helix-7, which was shown to affect rapid target recognition by RISC²⁴. The I365 α -g7 distance derived from the 3D structure of this mutant was therefore included here in Fig. 4a for comparison.

Based on these findings, we next addressed the guide-target duplex unwinding function of AGO2. First, we performed 360ns-long 1D metadynamics (MetD) simulations of WT and p.L192P in the duplex-bound complex, in which we enhanced sampling of I365 intercalation between g6 and g7 using I365 δ -g(6,7) distance as a collective variable (see Supplementary Fig. 14a for the definition of collective variables). Here, we used the energy potentials deposited during the MetD simulations to induce intercalation-mediated partial duplex unwinding, as we could not calculate the free energy profiles of intercalation. Therefore, the MetD here provides only a qualitative picture of the possible unwinding mechanism, allowing us to speculate which aspects of

this process are altered by the patient-derived mutations. As a simplified metric of the unwinding progress, we used mean guide-target duplex width ($\langle C1'-C1' \rangle$) (Supplementary Fig. 14b). During unwinding, helix-7 appears to move towards the MID domain and 'squeeze' the RNA duplex at base pairs g4-g7 (Supplementary Fig. 14a, 14c), thereby pushing the two RNA strands apart (Supplementary video 1). In WT-AGO2 the highest duplex width reached is slightly larger than in p.L192P (~ 13 Å vs. ~ 11.5 Å, respectively), as helix-7 in WT-AGO2 shifted closer to the MID domain (Supplementary Figs. 14b and 15). However, the nature of this analysis allowed only for the qualitative comparison between WT and p.L192P. The movements of the PAZ domain and helix-7 towards the opposing AGO2 lobe (MID and PIWI domains) are concerted (Fig. 4c, gray histogram in the background), corroborating previous suggestions^{22,32}.

Population distributions of the non-biased MD trajectories of variants bound to a mismatched duplex (g2-7 holo-RISC) suggest that compared to WT-AGO2, helix-7 shifted further away from the MID domain—an effect that seems to be common among all AGO2 mutants (Fig. 4c; see also Supplemental Fig. 16). Furthermore, the helix-7-MID distance appeared larger also for the p.F181A positive control (~ 1.6 Å larger than WT), but not for the common non-synonymous AGO2 variant p.E186K. Moreover, a similar effect is also observed when we manually removed the guide 3'-end from the PAZ domain of the duplex-bound state (WT Δ (3'-PAZ) in Fig. 4c). This provides a further link between loss of anchoring of the guide 3'-end at the PAZ domain, observed for almost all here identified mutations, and AGO2-mediated guide-target duplex unwinding³³. Taken together, based on our simulations, we hypothesize that the mutations identified in patients, apart from p.G733R, lead to reduced target release due to impaired AGO2 unwinding function. However, one should keep in mind that our data are mostly qualitative, and further, more quantitative simulations will be needed to completely dissect how the mutations in AGO2 affect unwinding.

AGO2 germline mutations lead to aberrant density of dendritic P-bodies. GFP-tagged WT-AGO2 and a representative set of

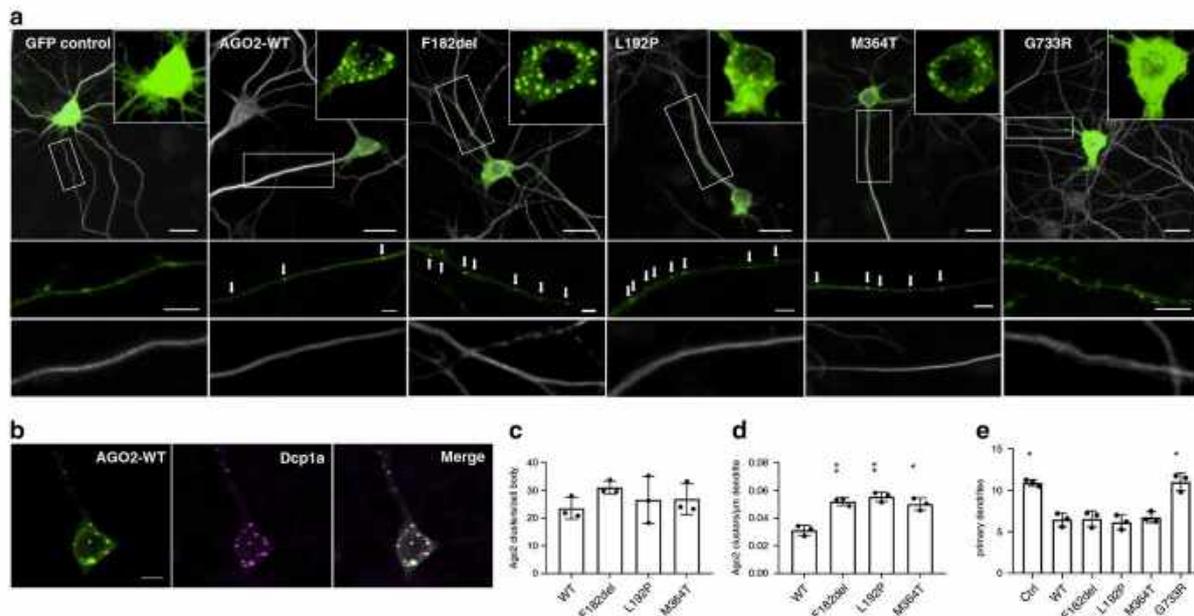


Fig. 5 AGO2 germline mutations lead to an increased number of dendritic P-bodies. **a** Primary cultured murine hippocampal neurons were transfected with GFP control or GFP-AGO2 WT and variants and Tomato-red. Staining for the dendritic marker MAP2 is shown in gray; inserts display the GFP signal in cell bodies. Boxed areas are magnified below for GFP- and MAP2 signals. Arrows indicate dendritic GFP-AGO2 clusters. Similar results were obtained in three independent biological experiments; results are quantified in (c, d, e). Scale bars: 20 μ m in overview pictures; 5 μ m in inserts. **b** Neurons expressing GFP-AGO2 were co-stained for Dcp1a. Similar costaining results were obtained in two biologically independent experiments for wt and F182del, L192P and M364T mutants, with 30 cells analyzed per experimental condition. Scale bar: 5 μ m. **c-e** Quantification of GFP-AGO2 clusters in cell bodies (c), GFP-AGO2 clusters in dendrites (d), and primary dendrites per cell (e). Data for mutant and control conditions were compared to wt. *, ** $p < 0.05$, 0.01, respectively; Brown-Forsythe and Welch one-way-ANOVA, followed by Dunnett's T3 multiple comparisons test; $n = 3$; mean \pm SD). Source data are provided as a Source Data file.

identified mutations were expressed in primary cultured rat hippocampal neurons. Cells were stained for the dendrite marker MAP2 (Fig. 5a), synaptic marker Shank3 (Supplementary Fig. 17), and the P-body marker Dcp1a (Fig. 5b). WT-AGO2 formed multiple clearly isolated punctae in cell bodies and throughout dendrites, consistent with previous observations^{34–36}. This pattern was maintained for p.F182del, p.L192P and p.M364T mutants but not for p.G733R, which appeared entirely diffuse similar to the GFP-control protein (Fig. 5a). In agreement with the data from non-neuronal cells (Supplementary Fig. 11), all GFP-AGO2 puncta were identified as P-bodies by co-staining for Dcp1a (shown for WT in Fig. 5b). Importantly, we found that whereas the number of AGO2-containing granules in neuronal cell bodies was not affected by the mutations (Fig. 5c), the density of dendritic P-bodies was almost doubled upon expression of mutants p.F182del, p.L192P and p.M364T (Fig. 5d). We conclude that the reduced phosphorylation at the C-terminal serine cluster, and the delayed dissociation from mRNA targets leads to an increased presence of AGO2 at dendritic P-bodies. This indicates that mutations in AGO2 lead to impairment of local translation that may result in altered plasticity in response to synaptic activity^{34–36}. Morphological analyses showed that neurons expressing WT, p.F182del, p.L192P and p.M364T exhibited a reduced number of dendrites compared to GFP-control and p.G733R (Fig. 5a, e). It has been shown previously that expression of WT-AGO2 is capable of reducing dendritic complexity³⁷, and again only the p.G733R failed to do so. The number of Shank3 positive clusters in dendrites was similar for all variants, indicating that synaptogenesis is not affected by over-expression of WT-AGO2 and the AGO2 mutants (Supplementary Fig. 17).

Global transcriptome alteration in primary fibroblasts of AGO2 patients. Given the major role of AGO2 in post-transcriptional regulation of gene expression, we next assessed global changes in the transcriptome in patient-derived primary dermal fibroblasts. RNA sequencing was performed of primary dermal fibroblasts obtained from two patients bearing the p.L192P (cases 2 and 3) and one patient with the p.A367P mutation (case 14). We compared the expression patterns of the protein-coding transcripts to fibroblasts from age-matched individuals, who were either unaffected or bear a causative mutation unrelated to AGO2. More than 770 genes were differentially expressed (DE) in case 2 compared to his three age-matched controls, whereas more than 1500 genes were differentially expressed in both cases 3 and 14, each of which was compared to a single age-matched control. 485 DE genes overlapped between cases 2 and 3, who bear the identical de novo p.L192P mutation (Supplementary Data 2 and 3). All three AGO2-patients shared 164 commonly DE genes (Fig. 6a, b, Supplementary Data 4), suggesting common but also mutation-specific effects. Gene ontology enrichment analysis of the commonly DE genes revealed enrichment for terms related to mitosis and cell cycle regulation (Supplementary Data 5), which is in line with the previously described function of AGO2 in regulating accurate chromosome segregation and cell cycle progression^{38,39}.

Discussion

In this study, we present clinical and molecular findings in 21 individuals bearing heterozygous AGO2 mutations, accompanied by *in-depth* functional characterization of identified missense mutations. All individuals exhibited intellectual disability and

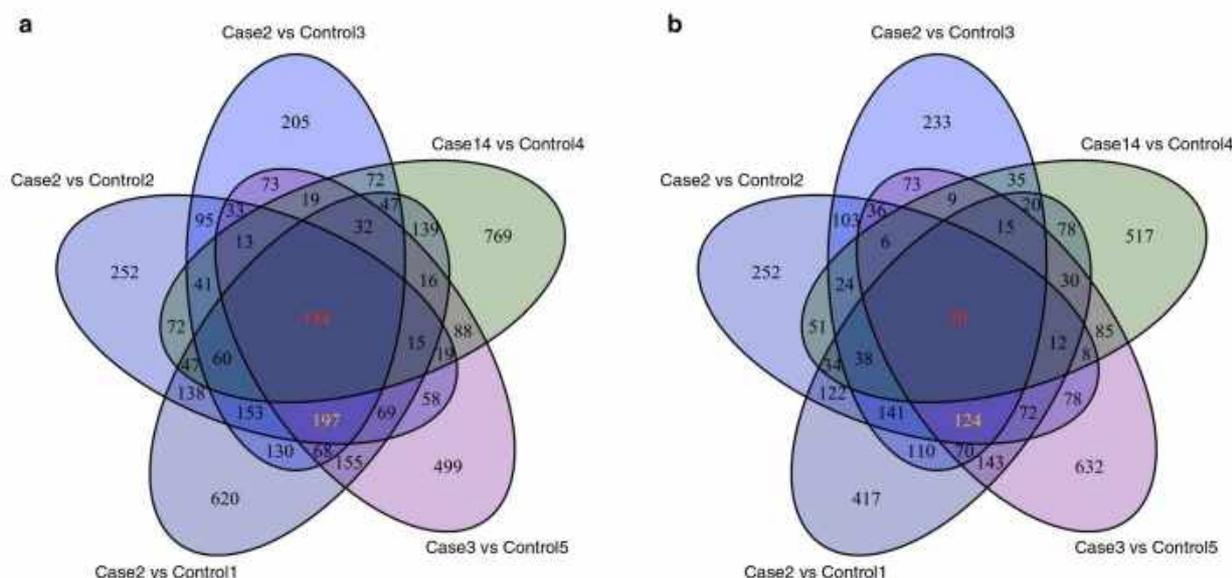


Fig. 6 Global transcriptome alteration in primary fibroblasts of AGO2 patients. **a, b** Venn diagram of upregulated (**a**) and downregulated (**b**) transcripts in fibroblasts isolated from AGO2 patients (cases 2, 3, and 14) compared to five age-matched controls. The number of the upregulated (**a**) or downregulated (**b**) genes in all three cases are marked red, whereas deregulated genes in both cases bearing the p.L192P mutation are marked yellow.

developmental delay, including delayed motor development and impaired speech development. Moreover, more than half of the patients had neonatal feeding difficulties and hypotonia. The majority of the patients presented with behavioral abnormalities including features of autistic spectrum disorder and attention deficit hyperactivity disorder. Some individuals developed aggressive behavior upon entering puberty. Seizures were a further common finding, observed in almost half of the individuals described here. Although most displayed some dysmorphic features, we did not observe a recognizable facial pattern. An increased incidence of skeletal and heart anomalies was noted. Interestingly, five individuals presented with plagiocephaly, a feature which may be related to hypotonia in four individuals. Furthermore, some specific clinical signs and symptoms were observed exclusively in individuals bearing the identical mutation. These include bilateral clinodactyly of the 5th finger as well as misaligned or crowded teeth with large incisors in all three individuals carrying the p.C751Y, and central apnea in the postnatal period in both individuals bearing the p.L192P. We additionally observed some unique clinical signs and symptoms. Individual 3 presented with a differential diagnosis of Pallister–Killian syndrome, individual 14 had blue sclera, individual 15 developed hydronephrosis, and individual 20 presented with precocious puberty. Future clinical reports beyond the large number of the affected individuals identified in this study will both broaden genotype-phenotype correlations and likely identify further mutation-specific clinical features.

Our functional analyses show a clear reduction of activity of the here identified AGO2 mutations in an shRNA-based silencing assay. This corroborated the *in silico* pathogenic predictions. In several cases we observed only subtle changes, which is consistent with the overall extreme intolerance of human AGO2 to mutations. Thus, our data show that even minor deficits in AGO2 function are sufficient to elicit aberrant neurological development. Currently, our characterization of selected mutations supports the view that they lead to a loss-of-function rather than a dominant-negative effect (see Fig. 2f). This is consistent with the

fact that so far only a single homozygous, non-synonymous AGO2 variant carrier (bearing the p.G88V) is documented in the gnomAD dataset.

Notably, one of the mutations identified here, p.G733R, stands out from the others as it fails in almost every functional assay. We assume that the substitution of the rather bulky and positively charged Arg for the small Gly will cause structural perturbations and local unfolding in the PIWI domain, thus leading to a complete loss of function. In this respect, we predict it to be functionally similar to the large chromosomal deletion observed in patient 21. The deletion encompasses part of the AGO2 gene and is likely to lead to haploinsufficiency. It does not only lead to loss of the first 3 AGO2 exons, but also the last 23 exons of PTK2. To our knowledge, no mutation in PTK2 has so far been connected to any human disorder. Several individuals bearing heterozygous loss-of-function mutations in PTK2 have been documented in the gnomAD dataset. Nonetheless, since no additional individuals with similar AGO2 loss-of-function mutations, (either a gross deletion, nonsense or frameshift mutation), are known currently, we cannot deduce to which extent the additional PTK2 deletion might be contributory. It is worth noting that the clinical presentation of this individual somewhat resembles children bearing a deletion that encompasses both AGO1 and AGO3¹⁴.

Currently, the pathomechanism of the p.G573S mutation remains somewhat unclear. It is conceivable that the change from glycine to serine, due to the difference in the physico-chemical properties of these residues, may cause some local structural changes and alter the interface between PIWI and MID domains.

Importantly, our analyses add physiological and pathological significance to recent advancements in structural and mechanistic understanding of ARGONAUTE proteins. The fact that most of the pathogenic variants in AGO2 alter residues either in L1 or helix-7 of L2, corroborate the structural analyses showing that subtle movements at the interface between these two segments determine the kinetics of target recognition^{22,24}. Furthermore, phosphorylation of the serine cluster at residues 824 to 834 has

recently been shown to coincide with the fast release of targets from RISC^{29,30}. In the AGO2 mutants, we observe reduced phosphorylation of this cluster, which together with our MD simulations indicate that the unwinding of guide-target duplexes may be slowed down by mutations, due to reduced movement of helix-7 towards the MID domain of the protein.

The reduced phosphorylation and enhanced binding to target mRNAs of mutant AGO2 coincide with an enhanced appearance of AGO2-positive dendritic P-bodies in neurons, as observed here for p.F182del, p.L192P and p.M364T variants. This may contribute to the neuronal phenotype seen in patients, as dendritic P-bodies and the activity of AGO proteins in dendrites are believed to contribute to synaptic activity^{34–36}. In particular, it is conceivable that changes in dendritic P-bodies will alter the complement of dendritic mRNAs which may be locally translated upon synaptic stimuli. Further work will be needed to determine how this affects neuronal function and morphology, as well as synaptic plasticity, learning and memory.

Our observation that patient-derived fibroblast cells exhibit global alterations of gene expression support the view that the global transcriptome changes due to altered function of AGO2 protein. Interestingly, we observed enrichment for commonly differentially expressed genes with broad roles in mitosis and cell cycle regulation, including cell division, mitotic nuclear division, sister chromatid cohesion, chromosome segregation, microtubule binding and regulation of cell cycle, to highlight only a few. Notably, the causal link of neurodevelopmental disorders and genes encoding for proteins regulating abovementioned processes have been previously established^{40–45}. It would be interesting to delineate which of these biological processes are affected in patient primary fibroblasts. Moreover, given the different organ- and tissue-specific expression patterns, a further emerging question is whether similar transcriptome changes will be observed in iPSC-derived neurons from these patients or even in patient-derived cerebral organoids. In addition, further studies are needed to delineate changes in global miRNA expression and investigate if these pathogenic mutations result in different RNA and miRNA binding sites. These analyses will require extensive further work and will be the main aim of our future studies.

Taken together, our study demonstrates that mutations affecting a core component of the RNAi machinery are associated with altered human neurological development, supporting previous observations that development and function of the nervous system is particularly vulnerable to alterations in gene expression patterns and their regulation⁷.

Methods

Research subjects. Written informed consent for all subjects was obtained in accordance with protocols approved by the respective ethics committees of the institutions involved in this study (approval number by the Ethics Committee of the Hamburg Chamber of Physicians: PV 3802). The authors affirm that the research participants and their legal representatives, and in the case of minors the parents or legal representatives of the human research participants provided informed consent for publication of the images in Supplementary Fig. 2.

Genetic analyses. Some of the investigators presenting affected individuals in this study were connected through GeneMatcher, a web-based tool for researchers and clinicians working on identical genes¹⁵. Whole-exome sequencing (WES) or trio whole-exome sequencing (trio-WES) experiments, data annotation and interpretation were performed in nine different centers with slightly different procedures. Trio-WES in families of cases 4, 6, 7, 8, 12, 13, and 16 was performed at the Radboud University Medical Center in Nijmegen, the Netherlands⁴⁰. Exome capture was performed with the Agilent SureSelect Human All Exon v5 enrichment kit (Agilent Technologies). Whole-exome sequencing was performed on the Illumina HiSeq platform (BGI, Copenhagen, Denmark) and the Illumina Nova-Seq6000 (NSW Health Pathology Randwick Genomics, Sydney, Australia). Data were analysed with BWA (read alignment) and GATK (variant calling) software packages. Variants were annotated using an in-house developed pipeline.

Prioritization of variants was done by an in-house designed 'variant interface' and manual curation. Trio WES in families of cases 11, 15, 19, and 20 and quad-WES in a family of cases 17 and 18 were performed on exon targets isolated by capture using the SureSelect Human All Exon V4 (50 Mb) or the IDT xGen Exome Research Panel v1.0. Massively parallel (NextGen) sequencing was done on an Illumina system with 100 bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19 (for case 11 to GRCh38/UCSC hg38), and analyzed for sequence variants using custom-developed analysis tools⁴⁷. The general assertion criteria for variant classification are publicly available on the GeneDx ClinVar submission page (<http://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/>). Trio-WES in the family of case 2 was performed using a SureSelect Human All Exon 50 Mb V5 Kit (Agilent, Santa Clara, CA, USA), and sequencing was performed on a HiSeq2500 system (Illumina, San Diego, CA, USA). Reads were aligned to the human genome assembly hg19 (UCSC Genome Browser) with Burrows-Wheeler Aligner (BWA, v.0.5.87.5), and detection of genetic variation was performed using SAMtools (v0.1.18), PINDEL (v 0.2.4t), and ExomeDepth (v1.0.0)⁴⁸. Trio-WES in family of case 14 was performed using SureSelect Human All Exon 50 Mb kit (Agilent Technologies, Santa Clara, CA) on a HiSeq2500 system (Illumina, San Diego, CA, USA). In-house developed scripts were applied to detect protein changes, affected splice sites and overlaps to known variations, with filtering against dbSNP build 138, the 1000 Genomes Project data build November 2014 and ExAC Browser (status from August 2019)⁴⁹. DNA of case 9 and his unaffected father was sequenced SureSelect Clinical Research Exome V2 (Agilent Technologies, Santa Clara, CA) on a HiSeq2500 system (Illumina, San Diego, CA, USA). The raw data were analyzed using the Care4Rare analysis pipeline⁵⁰. DNA of case 1 and his parents were extracted from peripheral blood using the Genra Puregene Blood Kit (Genra Systems Inc., Minneapolis, USA). Quality and quantity were checked with NanoDrop Spectrophotometer (NanoDrop Technologies, Wilmington, USA). Genomic DNA was fragmented by sonication to generate fragments of 200–500 base pairs. Library preparation was done using Kapa DNA HTP Library Preparation Kit (KAPA Biosystems, 07138008001). Hybridization of the adapter-ligated DNA was performed to a biotin-labeled probe included in the Nimblegen SeqCap EZ Human Exome Kit (Roche, 06465692001). Libraries were sequenced using the Illumina HiSeq 2500 sequencing system and paired-end 101 bp reads were generated for analysis. Trio-WES in family of case 3 was performed using the SureSelect Human All Exon V4 (50 Mb) kit (Agilent, Santa Clara, CA, USA), and sequencing was performed on a HiSeq2500 system (Illumina, San Diego, CA, USA). Variants were identified using haplotype caller within GATK and FreeBayes. The intersection of the two variant callers were annotated with SnpEff and loaded into a database using the GEMINI framework. Annotations included predicted functional effect (e.g., splice-site, nonsense, missense), protein position, known clinical associations (OMIM, CLINVAR), mouse phenotypes (MGI), conservation score (PhastCons, GERP), and effects protein function (PolyPhen), CADD scores, and population allele frequencies (Exome Variant Server and Exome Aggregation Consortium data). Trio-Wes in family of case 5 was performed on MGISEQ-2000 platform (BGI-Wuhan, Wuhan, China). DNA extraction as well as quality controls were performed using standard methods. Half a microgram of genomic DNA was randomly fragmented by Covaris. Using Agencourt AMPure XP-Medium kit fragments of 150–250 bp were selected. The fragments were then subjected to end-repair, 3' adenylation and adaptors ligation. After PCR amplification and purification, hybridization using BGI Hybridization and Wash kits were used. After a second PCR and recovering step, the double stranded PCR products were heat-denatured and circularized by the splint oligo sequence. The single-strand circle DNA (ssCir DNA) were formatted as the final library. Library was qualified by Qubit ssDNA kit. The library was amplified to make DNA nanoball which have more than 300 copies of one molecular. The DNBS were loaded into the patterned nanoarray and pair-end 100 bases reads were generated in the way of sequenced by combinatorial Probe-Anchor Synthesis (cPAS). The raw data were then transferred to Limbus (Rostock, Germany) and bioinformatic secondary and tertiary analyses were performed using the standard algorithms of GATK. SNVs and CNVs were then presented in a web-based interface and evaluation was performed by experienced scientists. Identified candidate genes were then prioritized based on in-house scoring system.

Chromosomal Microarray Analysis in family 21 was performed using a 4×180 K whole-genome oligonucleotide microarray following the manufacturer's protocol (Agilent Technologies, Santa Clara, CA, USA). Results were interpreted with Cytogenomics software v3.0.1.1 (ADM2 method). A CNV was defined as at least three contiguous oligonucleotides with an abnormal mean log ratio (>0.25 or <−0.25). GRCh37/hg19 was used as the reference sequence.

Plasmids. An expression vector for GFP-tagged human AGO2 (in pEGFP-C1) was obtained from Phil Sharp (MIT) via Addgene (#21981⁵¹). The vector for Flag/HA (F/H) tagged human AGO2 has been described before²⁶. Mutations were introduced by site-directed mutagenesis using the QuikChange II kit (Agilent; CA), using complementary oligonucleotides. Constructs were verified by Sanger sequencing. For expression in neurons, the entire cDNA fragments coding for GFP-AGO2 fusion proteins were subcloned into FUV vector which uses a ubiquitin promoter for driving expression. FUV was obtained via Addgene #14882 from D. Baltimore, Caltech.

Antibodies. The following primary antibodies were used: Chicken anti-MAP2 (antibodies Online; ICC: 1:1000); guinea pig anti-Shank3 (Synaptic Systems # 162 304; ICC: 1:500); rabbit monoclonal anti Dcp1a, Abcam #183709; ICC 1:1000); rabbit anti-Shank and anti-mRFP antisera have been generated by custom immunization by Biogenes GmbH, Berlin, Germany. Anti-Dicer antibody (Bethyl, 1:1000), anti-HA antibody (Covance Research Products, clone 16B12, 1:1000); anti-AGO2 (EMD Millipore; clone 11A9; #MABE253; 1:1000); and anti-TNRC6ABC, clone 7A9 (Merck Millipore) were used for protein detection in Western Blotting. Secondary antibodies: Alexa fluor 633 goat anti-rb, Alexa 405 goat anti-chk IgG (abcam), Abberior star red goat anti guinea pig were used at 1:1000 dilution.

Cell lines and transfections. HEK293T, U2OS and HeLa cells were obtained from ATCC and cultured in Dulbecco's modified Eagle medium (DMEM; ThermoFisher) supplemented with 10% fetal bovine serum (FBS; GE Healthcare) and penicillin-streptomycin (100 U/ml and 100 mg/ml, respectively; ThermoFisher) under 5% CO₂ and at 37 °C. For deletion of the endogenous AGO2 gene, cells were transfected with a Crispr targeting construct in pLentiCrisprV2 (Genscript), carrying a guide sequence encompassing the AGO2 start codon (seq.: 5'-GCCACCATGTACTCGGGAG-3'). Cells were selected with puromycin (2 µg/ml) for several days. Absence of AGO2 expression was verified by Western blotting. For silencing assays, cells were plated on 12 well dishes and transfected 4–6 h later using Turbofect transfection reagent (3.5 µl/well). Cells were transfected with an expression vector for a gene of interest (mRFP-Shank3, mRFP-DDX1, or GFP-tagged δ -catenin) in combination with an shRNA vector targeting the corresponding mRNA. For this, shRNA constructs were generated in pSuper for DDX1 and δ -catenin and in pLVTHM for Shank3 using target sequences AGGAGGAGGACCTTGATAAA (rat DDX1; bp 1633–1651 in NM_053414.1), GCAACTATGTGCGACTTCTA (mouse δ -catenin; 4240–4258 in NM_008729.3) and GGAAGTCCACAGAGGACAAGA (rat Shank3; 3794–3814 in NM_021676.2). For each well either empty vector or AGO2-expression vectors were added to the transfection mix (1 µg/well).

Primary human dermal fibroblast cultures were established from skin biopsies taken from the three patients (cases 2, 3 and 13) and four age-matched controls (controls 1, 2, 3, and 5; aged 2, 2, 2, and 15 years at sampling, respectively). In addition, control 4 (GM01887 aged 7 years at sampling) was obtained from Coriell Institute. Primary fibroblasts were cultured in the same manner as the cell lines mentioned above.

Primary dissociated hippocampal neurons isolated from embryonic (E18) rats were co-transfected after 7 days in vitro (DIV7) with GFP-AGO2 constructs (in the FUW vector) and pmRFP-C1 using the calcium phosphate method⁵². The neurons were fixed at DIV14 and stained for endogenous Shank3 (postsynaptic marker); MAP2 (dendritic marker) and DCP1a (P-body marker). RFP fluorescence was used to verify successful transfection. All animal experiments were approved by, and conducted in accordance with, the guidelines of the Animal Welfare Committee of the University Medical Center (Hamburg, Germany) under permission number Org766.

Microscopy. Confocal images of hippocampal neurons were acquired with a Leica Sp5 confocal microscope using a $\times 63$ objective. Quantitative analysis for images was performed using ImageJ. Three independent experiments were performed for all neuron data. Primary dendrites were counted at a ring within 10 µm distance from the cell body. 12–15 neurons per each condition were counted. For counting postsynaptic Shank3 clusters, 12–15 neurons per each condition with a total of 45 branches were evaluated. For counting dendritic AGO2-clusters, these were counted along entire dendritic branches. Cluster density was obtained by dividing the number of AGO2 clusters/P-bodies by µm of dendrite length. 12–15 neurons with a total of 36–45 dendrites per each condition were evaluated.

Immunoprecipitation (IP). For immunoprecipitation of overexpressed FLAG/HA-tagged AGO2 proteins, anti-FLAG M2 affinity agarose gel (Sigma-Aldrich) was used and washed twice with cold PBS before incubation with lysate. After incubation of 2.5 h at 4 °C on a rotating wheel, the beads with bound proteins were centrifuged for 1 min at 1000 \times g and the supernatant was removed. For qRT-PCR analysis, the affinity matrix was washed with NET buffer (50 mM Tris/HCl pH 7.5, 5 mM EDTA, 0.5% NP-40, 10% Glycerol, 1 mM NaF, 0.5 mM DTT, 1 mM AEBSF) + 300 mM NaCl twice, once with lysis buffer + 450 mM NaCl, once with 600 mM NaCl, once with 450 mM NaCl, followed by washing with PBS once. For Western Blot, Northern Blot and MS analysis, beads were washed three times with NET-lysis buffer with 300 mM NaCl and once with PBS. Subsequent mass spectrometric analysis was performed as described below³⁰.

Targeted quantification of AGO2 phosphorylation by SRM (Selected Reaction Monitoring). Phosphorylation levels of AGO2 variants were quantified by obtaining selected AGO2 phosphopeptides as well as their non-phosphorylated counterparts as stable isotope-labeled and quantified spike-in standards from JPT (Innovative Peptide solutions, Berlin)³⁰. The following ¹³C¹⁵N-labeled peptides: SASFNTDPYVR and SApSFNTDPYVR for detection of phospho-S387; YHLVDKEHDSAEGSHTSGQSNR and YHLVDKEHDPsAEGSHTSGQSNR for detection of pS824; and YHLVDKEHDPsAEGpSHTpSGQpSNR for detection of the pS824/pS828/pS831/pS834 cluster, were used to set up a SRM method on a hybrid triple quadrupole/linear ion trap instrument (QTRAP4500, SCIEX). A spectral library built from DDA (data-dependent analysis) runs of the heavy

peptides was built and imported into the open source software Skyline (MacCoss lab software, Seattle, USA). In Skyline, then a targeted method was built according to the occurrence of precursor charge states +2, +3, +4 during several DDA runs. After manual inspection of MS2 spectra at least 3 transitions were selected for each peptide. The resulting transition list was imported into the instrument software (Analyst 1.6.1) and the following parameters were set: Q1 and Q3 at unit resolution (0.7 m/z half-maximum peak width), dwell time 20 ms, cycle time < 3 s. After annotating peptide retention times from the initial SRM run and setting the following parameters: cycle time: 2 s, retention time window: 5 min, a scheduled SRM method was created in Skyline. Sample preparation of overexpressed and immunoprecipitated Flag/HA-tagged Ago2 variants was performed as follows: after separation on SDS-PAGE AGO2 bands were excised, washed with 50 mM NH₄HCO₃, 50 mM NH₄HCO₃/acetonitrile (3/1), 50 mM NH₄HCO₃/acetonitrile (1/1) and lyophilized. AGO2 variant proteins were then subjected to overnight in gel tryptic digest at 37 °C with ~2 µg trypsin per 100 µl gel volume (Trypsin Gold, mass spectrometry grade, Promega). Importantly, 100 fmol of each heavy peptide were spiked into the digests. After digestion, peptides were extracted twice with 100 mM NH₄HCO₃, followed by 50 mM NH₄HCO₃ in 50% acetonitrile. The combined eluates were lyophilized and reconstituted in 20 µl 1% TFA for LC-MS analysis. The LC-MS/MS system consisted of an UltiMate 3000 RSLCnano System (Thermo Scientific, Dreieich) coupled via a NanoSpray II source (SCIEX) to a QTRAP4500 mass spectrometer. Peptides were separated by reversed-phase chromatography on an Acclaim Pepmap100 C18 nano column (75 µm i.d. \times 150 mm, Thermo Fisher) with a C18 Acclaim Pepmap100 pre-concentration column (100 µm i.d. \times 20 mm, Thermo Fisher) in front. At a flow rate of 300 nL/min, a 60 min linear gradient of 4–40% acetonitrile in 0.1% formic acid was used. SRM measurements resulted in .wiff files which were imported back into Skyline. By calculating the heavy-to-light ratios of the peak areas of the respective transitions absolute quantification of endogenous phosphorylated or non-phosphorylated peptides was facilitated. Relative quantification of phosphorylated AGO2 peptides was performed in Excel by first calculating the absolute amount of either peptide species, followed by adding up the amounts of the non-modified peptide species and the related phosphorylated peptide species. Assuming this sum to represent 100%, it was possible to calculate the percentage of the individual phosphopeptide species.

Quantitative real-time PCR (qRT-PCR). For qRT-PCR analysis of overexpressed FLAG/HA-tagged AGO2 proteins, the affinity matrix after immunoprecipitation was washed with lysis buffer + 300 mM NaCl in total twice, once with lysis buffer + 450 mM NaCl, once with 600 mM NaCl, once with 450 mM NaCl, followed by washing with PBS once. The RNA of Input and IP samples were isolated using TRIzol (Thermo Fisher Scientific) and a second step with chloroform. For cDNA synthesis, 1 µg of the Input and complete RNA yield of the IP samples were first digested with DNaseI (Thermo Fisher Scientific). After the digest, cDNA was synthesized using the First-Strand cDNA synthesis kit (Thermo Fisher Scientific), following the manufacturer's protocol. qRT-PCR was performed with Sso Fast Eva Green Mix (Bio-Rad). *NFIC*, fwd 5'-GACCTGTACCTGGCCTACTTTG, rev 5'-CACACCTGACGTGACAAAGCTC; *F8A1*, fwd 5'-GTTTGGCTCTGGGGA GGAAT, rev 5'-TGGTAACGTTTCAGCCAACGA; *HAND1*, fwd 5'-GGAGTCC GCAGAAAGGGTTAAA, rev 5'-CGGGCAAGGCTGAAAATGAG; *HOXC8*, fwd 5'-CGGAGACGCTCCAAATTCT, rev 5'-GCCTTGCTCTCGCTACTGT. qRT-PCRs were run on a CFX96 cycler (Bio-Rad) and data were analyzed using $\Delta\Delta$ Ct method³⁰.

For family survey of case 20, qRT-PCR analysis of AGO2 and *PTK2* mRNA was performed using the kit LightCycler[®] 480 SYBR Green I Master following the manufacturer's protocol (Roche), with detection on a Roche LightCycler 480 Real-Time PCR instrument (Roche Diagnostics Corporation, USA). *SULF1* was used as reference gene for normalization. The primer pairs were: *AGO2*, fwd 5'-GATATG CCTTCAAGCCTCGA, rev 5'-AACTCTCTCGGGCACTTCT; *PTK2*, fwd 5'-TG GGTGAGCTCATCAACAAG, rev 5'-GCCCAAGCATTTTTCAGTCTT; *SULF1*, fwd 5'-CCCCAAGAAGTGGTCACTA, rev 5'-CAGGCAAGACTGCCCTA GAC. Data were analyzed using $\Delta\Delta$ Ct method.

In vitro cleavage assay. The cap-³²P-labeling of target RNA perfect complementary to the endogenous miR-19b, and the in vitro cleavage assay were performed according to described protocols^{53,54}. For this, 25% of the total immunoprecipitate was separated for subsequent analysis by Western Blotting. Thereafter, translation mix was added to a final concentration of 1 \times translation mix and the reaction was started by addition of cap-labeled target RNA. The reaction was incubated for 1 h at 30 °C and stopped by addition of TRIzol (Thermo Fisher Scientific) and chloroform, shaking and centrifuging, followed by precipitation overnight at -20 °C in ethanol with 20 µg glycogen RNA grade (Thermo Fisher Scientific). After pelleting, the RNA was resuspended in RNA sample buffer.

RNA sequencing and gene expression analysis of primary fibroblasts. Total RNA was extracted with the RNeasy mini kit (Qiagen) from primary fibroblasts that were in all in the same passage 8. These included primary fibroblasts of three patients bearing AGO2 mutation (cases 2, 3, and 14), 4 age-matched individuals affected by NDD not related to AGO2 mutation, and an apparently healthy

individual control 4 (GM01887 aged 7 years at sampling) obtained from Coriell Institute. RNA integrity and quality was assessed with Epoch Microplate Spectrophotometer (BioTek) and on 1% agarose gels. RNA purity was checked using the NanoPhotometer[®] spectrophotometer (IMPLEN, CA, USA). RNA concentration was measured using Qubit[®] RNA Assay Kit in Qubit[®] 2.0 Fluorometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Library preparation and transcriptome sequencing were performed at Novogene. In brief, a total amount of 3 µg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext[®] Ultra[™] RNA Library Prep Kit for Illumina[®] (NEB, USA) following the manufacturer's recommendations and index codes were added to attribute sequences to each sample. mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First-Strand Synthesis Reaction Buffer (5× first-strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H⁻). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure was ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150–200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 µl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using HiSeq PE Cluster Kit cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform and 125 bp/150 bp paired-end reads were generated. Quality trimming and adapter cutting were performed using Cutadapt v2.5. Genome mapping in paired-end mode was done using Bowtie2 v2.3.4.1 to the human genome GRCh38.p12. Read counts were obtained using bedtools v2.26.0 summarized per protein-coding gene using annotation version GRCh38.97.

Differential expression (DE) analysis was performed by calculating expression fold-changes for each gene in patients and compared to the age-matched controls. To account for high variations among (human) individuals we set a threshold of twofold difference to extract DE genes. Gene ontology enrichment analysis for molecular function and biological process were obtained using the DAVID tool v6.8.

Molecular dynamics simulations. WT-AGO2 and AGO2 variants were simulated in the apo-AGO2 state and in four AGO2-RNA complexes using classical all-atom MD (Supplementary Fig. 13). In total, 60 non-biased MD trajectories were obtained; the minimal trajectory length in this set is 200 ns, with a cumulative simulation time of ~17 µs. All MD simulations were conducted in NAMD 2.12⁵⁵ using CHARMM36 force field⁵⁶ in TIP3P⁵⁷ water box. Production MD simulations were performed in NVT ensemble using Langevin thermostat at standard conditions with 2 fs integration step. Principal component analysis of Cartesian coordinates of protein Cα atoms was applied to obtain dominant conformational modes using ProDy⁵⁸. To induce helix7-mediated duplex unwinding and to probe helix7 movements upon p.L192P mutation and perturbed guide-PAZ interactions, we performed metadynamics (MetD) simulations (reviewed in ref. 59). In the case of non-tempered MetD of WT-AGO2 and p.L192P in the complex with a mismatched duplex, we aimed at inducing g6-g7 kink by enhancing sampling of I365 intercalation between g6 and g7. The sampling of the collective variable I365g6 (6,7) was not restrained, leading to sampling outside the grid for Gaussian potentials deposition. Thus, the potential of mean force (PMF) from this set of MetD simulations cannot be used. Trajectories from not fully converged MetD simulations can be used to obtain mechanistic insights but with some restrictions as discussed in the original reference⁶⁰. The setup and performance of all simulations are described in detail in Supplementary Methods. Colvar module of NAMD was used for enhanced sampling simulations⁶¹.

Statistical analyses. Statistical analyses were performed with Prism software (GraphPad, San Diego, CA).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The RNA-seq data were submitted to GEO repository under the accession number GSE141099. The mass spectrometry data for quantification of AGO2 phosphorylation have been submitted to the peptidatlas repository under accession PASS01561. The raw whole-exome sequencing and microarray-based comparative genomic hybridization data

that support the findings in affected individual cannot be made publicly available for reasons of patient confidentiality. Qualified researchers may apply for access to these data, pending institutional review board approval. Cells are available upon signing a material transfer agreement. The data supporting the findings of this study are available from the corresponding authors upon reasonable request. Source data are provided with this paper.

Received: 28 December 2019; Accepted: 21 September 2020;

Published online: 16 November 2020

References

- Treiber, T., Treiber, N. & Meister, G. Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nat. Rev. Mol. Cell Biol.* **20**, 5–20 (2019).
- Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Sen, G. L. & Blau, H. M. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nat. Cell Biol.* **7**, 633–636 (2005).
- Gregory, R. L., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631–640 (2005).
- Wu, L. & Belasco, J. G. Let me count the ways: mechanisms of gene regulation by miRNAs and siRNAs. *Mol. Cell* **29**, 1–7 (2008).
- Liu, J. et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–1441 (2004).
- Rajman, M. & Schratt, G. MicroRNAs in neural development: from master regulators to fine-tuners. *Development* **144**, 2310–2322 (2017).
- Schratt, G. microRNAs at the synapse. *Nat. Rev. Neurosci.* **10**, 842–849 (2009).
- Mahler, E. A. et al. Exome sequencing in children. *Dtsch Arztebl Int.* **116**, 197–204 (2019).
- Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
- Samocha, K. E. et al. A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* **46**, 944–950 (2014).
- Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
- Rauch, A. et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* **380**, 1674–1682 (2012).
- Tokita, M. J. et al. Five children with deletions of 1p34.3 encompassing AGO1 and AGO3. *Eur. J. Hum. Genet.* **23**, 761–765 (2015).
- Sobreira, N., Schiettecatte, F., Boehm, C., Valle, D. & Hamosh, A. New tools for Mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. *Hum. Mutat.* **36**, 425–431 (2015).
- Kwak, P. B. & Tomari, Y. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat. Struct. Mol. Biol.* **19**, 145–151 (2012).
- Ma, J. B., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318–322 (2004).
- Gu, S., Jin, L., Huang, Y., Zhang, F. & Kay, M. A. Slicing-independent RISC activation requires the argonaute PAZ domain. *Curr. Biol.* **22**, 1536–1542 (2012).
- Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E. & Weichenrieder, O. Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proc. Natl Acad. Sci. USA* **108**, 10466–10471 (2011).
- Parker, J. S., Roe, S. M. & Barford, D. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* **434**, 663–666 (2005).
- Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434–1437 (2004).
- Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for microRNA targeting. *Science* **346**, 608–613 (2014).
- Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–1040 (2012).
- Klum, S. M., Chandradoss, S. D., Schirle, N. T., Joo, C. & MacRae, I. J. Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition. *EMBO J.* **37**, 75–88 (2018).
- Shang, R. et al. Ribozyme-enhanced single-stranded Ago2-processed interfering RNA triggers efficient gene silencing with fewer off-target effects. *Nat. Commun.* **6**, 8430 (2015).
- Meister, G. et al. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15**, 185–197 (2004).

27. Meister, G. et al. Identification of novel argonaute-associated proteins. *Curr. Biol.* **15**, 2149–2155 (2005).
28. Horman, S. R. et al. Akt-mediated phosphorylation of argonaute 2 downregulates cleavage and upregulates translational repression of MicroRNA targets. *Mol. Cell* **50**, 356–367 (2013).
29. Golden, R. J. et al. An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. *Nature* **542**, 197–202 (2017).
30. Quevillon Huberdeau, M. et al. Phosphorylation of Argonaute proteins affects mRNA binding and is essential for microRNA-guided gene silencing in vivo. *EMBO J.* **36**, 2088–2106 (2017).
31. Gan, H. H. & Gunsalus, K. C. Assembly and analysis of eukaryotic Argonaute-RNA complexes in microRNA-target recognition. *Nucleic Acids Res.* **43**, 9613–9625 (2015).
32. Kong, R. et al. Exploring the RNA-bound and RNA-free human Argonaute-2 by molecular dynamics simulation method. *Chem. Biol. Drug Des.* **90**, 753–763 (2017).
33. Jung, S. R. et al. Dynamic anchoring of the 3'-end of the guide strand controls the target dissociation of Argonaute-guide complex. *J. Am. Chem. Soc.* **135**, 16865–16871 (2013).
34. Oh, J. Y. et al. Activity-dependent synaptic localization of processing bodies and their role in dendritic structural plasticity. *J. Cell Sci.* **126**, 2114–2123 (2013).
35. Cougot, N. et al. Dendrites of mammalian neurons contain specialized P-body-like structures that respond to neuronal activation. *J. Neurosci.* **28**, 13793–13804 (2008).
36. Zeitelhofer, M. et al. Dynamic interaction between P-bodies and transport ribonucleoprotein particles in dendrites of mature hippocampal neurons. *J. Neurosci.* **28**, 7555–7562 (2008).
37. Storchel, P. H. et al. A large-scale functional screen identifies Noval1 and Ncoas3 as regulators of neuronal miRNA function. *EMBO J.* **34**, 2237–2254 (2015).
38. Huang, C., Wang, X., Liu, X., Gao, S. & Shan, G. RNAi pathway participates in chromosome segregation in mammalian cells. *Cell Discov.* **1**, 15029 (2015).
39. Vasudevan, S., Tong, Y. & Steitz, J. A. Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318**, 1931–1934 (2007).
40. de Wolf, B. & Kops, G. Kinetochore malfunction in human pathologies. *Adv. Exp. Med. Biol.* **1002**, 69–91 (2017).
41. Gholkar, A. A. et al. The X-linked-intellectual-disability-associated Ubiquitin ligase Mid2 interacts with Astrin and regulates Astrin levels to promote cell division. *Cell Rep.* **14**, 180–188 (2016).
42. Hempel, M. et al. De novo mutations in CHAMP1 cause intellectual disability with severe speech impairment. *Am. J. Hum. Genet.* **97**, 493–500 (2015).
43. Kline, A. D. et al. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. *Nat. Rev. Genet.* **19**, 649–666 (2018).
44. Walters, G. B. et al. MAP1B mutations cause intellectual disability and extensive white matter deficit. *Nat. Commun.* **9**, 3456 (2018).
45. Lasser, M., Tiber, J. & Lowery, L. A. The role of the microtubule cytoskeleton in neurodevelopmental disorders. *Front. Cell Neurosci.* **12**, 165 (2018).
46. Neveling, K. et al. A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum. Mutat.* **34**, 1721–1726 (2013).
47. Retterer, K. et al. Clinical application of whole-exome sequencing across clinical indications. *Genet. Med.* **18**, 696–704 (2016).
48. Lessele, D. et al. De novo missense mutations in DHX30 impair global translation and cause a neurodevelopmental disorder. *Am. J. Hum. Genet.* **101**, 716–724 (2017).
49. Lessele, D. et al. Analyses of LMNA-negative juvenile progeroid cases confirms biallelic POLR3A mutations in Wiedemann-Rautenstrauch-like syndrome and expands the phenotypic spectrum of PYCR1 mutations. *Hum. Genet.* **137**, 921–939 (2018).
50. Kernohan, K. D. et al. Diagnostic clarity of exome sequencing following negative comprehensive panel testing in the neonatal intensive care unit. *Am. J. Med. Genet. A* **176**, 1688–1691 (2018).
51. Leung, A. K., Calabrese, J. M. & Sharp, P. A. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc. Natl Acad. Sci. USA* **103**, 18125–18130 (2006).
52. Hassani Nia, F., Woike, D., Kloth, K., Kortum, F. & Kreienkamp, H. J. Truncating mutations in SHANK3 associated with global developmental delay interfere with nuclear beta-catenin signaling. *J. Neurochem.* <https://doi.org/10.1111/jnc.15014> (2020).
53. Hauptmann, J. et al. Turning catalytically inactive human Argonaute proteins into active slicer enzymes. *Nat. Struct. Mol. Biol.* **20**, 814–817 (2013).
54. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574 (2002).
55. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
56. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
57. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
58. Bakan, A. et al. Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* **30**, 2681–2683 (2014).
59. Valsson, O., Tiwary, P. & Parrinello, M. Enhancing important fluctuations: rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **67**, 159–184 (2016).
60. Haldar, S. et al. Insights into stability and folding of GNRA and UNGC tetraloops revealed by microsecond molecular dynamics and well-tempered metadynamics. *J. Chem. Theory Comput.* **11**, 3866–3877 (2015).
61. Fiorin, G., Klein, M. L. & Henin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **111**, 3345–3362 (2013).

Acknowledgements

We thank all family members for their participation and collaboration, Hans-Hinrich Hönck (Institute for Human Genetics, UKE Hamburg) for technical assistance, and UKE microscopic imaging facility (umif) for providing assistance with confocal microscopes. This work was funded in part by Werner Otto Stiftung (to D.L. and H.-J.K.), Deutsche Forschungsgemeinschaft (LE4223/1-1 to D.L.; Kr1321/8-2 to H.-J.K.; KJ 488/7-2 to S.K.; SFB960/B3; DFG/ANR 20647-1 to G.M.), postdoctoral fellowships from Forschungsförderungs-fonds Medizin at UKE Hamburg (to F.H.N. and V.M.), the DAAD fellowship (to A.K.), the Clinician Scientist Program Medizinische Fakultät der Universität Leipzig (to D.L.D.), the Dietmar-Hopp-Stiftung (to S.S., 23011236), the National Institute of Neurological Disorders and Stroke (NINDS) under award number K08NS092898, Jordan's Guardian Angels, the Brotman Baty Institute for Precision Medicine (to G.M.M.), the SOLVE-RD program (EU to H.G.B.), the National Research Foundation Singapore under its National Medical Research Council Centre Grant Programme (Project No. NMRC/CG/M003/2017 to E.C.T.), and, in part, under the Care4Rare Canada Consortium funded by Genome Canada and the Ontario Genomics Institute (OGI-147), the Canadian Institutes of Health Research, Ontario Research Fund, Genome Alberta, Genome British Columbia, Genome Quebec, and Children's Hospital of Eastern Ontario Foundation (to K.M.B.), T.R., K.R.D. and C.A.E. are supported through the Australian NHMRC Centre for Research Excellence in Neurocognitive Disorders.

Author contributions

D.L., D.M.Z., F.H.N., A. Bruckmann, V.G., I.M., V.M., I.L., C.S., S.K., Z.I., G.M., and H.-J.K. performed cell experiments and analyzed the data. D.L., M.R.F.R., M.M., B.L., E.-S.T., E.G., J.L., J.D., E.Z.-H., T. Kovacevic, I.R., K.F., D.M., S.S., A.M., M.S., B.P., J.L., T.B., C.M., P.P., S. Lüttgen, J.P., R.R., M.W., T.G., K.L., P.R., H.G.B., C.C., S.L., D.A.D., K.M.B., G.M.M., A.M.-R., T.R., P.L.A., and K.R.D. were involved in patient care and gathered detailed clinical information for the study. D.L. analyzed clinical data. D.L., M.R.F.R., E.-S.T., A.T., H.H.L., B.W.C., D.L.D., M.O., T.H., J.M., R.A.J., A.P.A.S., C.K., E.-C.T., G.M.M., K.M., T.K.L., R.P., C.A.E., K.R.D., and T.R. analyzed and supported the trio-exome sequencing results. A.K. and Z.I. performed non-biased molecular dynamics (MD) simulations. D.L., A. Bartholomäus and Z.I. analyzed the RNA-sequencing data. D.L., Z.I., G.M., and H.-J.K. directed functional analyses. D.L. and H.-J.K. jointly directed the study and drafted the initial manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

A.T. and K.M. are employees of GeneDx, Inc. The other authors declare that they have no conflict of interest.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-19572-5>.

Correspondence and requests for materials should be addressed to D.L. or H.-J.K.

Peer review information *Nature Communications* thanks Ian MacRae, Christopher Woods, Kyowon Jeong and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Davor Lessel¹, Daniela M. Zeitler², Margot R. F. Reijnders^{3,4}, Andriy Kazantsev⁵, Fatemeh Hassani Nia¹, Alexander Bartholomäus^{5,6}, Victoria Martens¹, Astrid Bruckmann², Veronika Graus², Allyn McConkie-Rosell⁷, Marie McDonald⁷, Bernarda Lozic^{8,9}, Ee-Shien Tan¹⁰, Erica Gerkes¹¹, Jessika Johannsen¹², Jonas Denecke¹², Aida Telegrafi¹³, Evelien Zonneveld-Huijssoon¹¹, Henny H. Lemmink¹¹, Breana W. M. Cham¹⁰, Tanja Kovacevic¹¹, Linda Ramsdell¹⁴, Kimberly Foss¹⁴, Diana Le Duc¹⁵, Diana Mitter¹⁵, Steffen Syrbe¹⁶, Andreas Merckenschlager¹⁷, Margje Sinnema⁴, Bianca Panis¹⁸, Joanna Lazier¹⁹, Matthew Osmond²⁰, Taila Hartley²⁰, Jeremie Mortreux^{21,22}, Tiffany Busa²¹, Chantal Missirian^{21,22}, Pankaj Prasun²³, Sabine Lüttgen¹, Ilaria Mannucci¹, Ivana Lessel¹, Claudia Schob¹, Stefan Kindler¹, John Pappas²⁴, Rachel Rabin²⁴, Marjolein Willemsen³, Thatjana Gardeitchik³, Katharina Löhner¹¹, Patrick Rump¹¹, Kerith-Rae Dias^{25,26}, Carey-Anne Evans^{25,26}, Peter Ian Andrews^{27,28}, Tony Roscioli^{25,29,30}, Han G. Brunner^{3,4}, Chieko Chijiwa³¹, M. E. Suzanne Lewis³¹, Rami Abou Jamra¹⁵, David A. Dymant^{19,20}, Kym M. Boycott^{19,20}, Alexander P. A. Stegmann^{3,4}, Christian Kubisch¹, Ene-Choo Tan³², Ghayda M. Mirzaa^{33,34,35}, Kirsty McWalter¹³, Tjitske Kleefstra³, Rolph Pfundt^{3,11}, Zoya Ignatova⁵, Gunter Meister² & Hans-Jürgen Kreienkamp¹

¹Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany. ²Regensburg Center for Biochemistry (RCB), Laboratory for RNA Biology, University of Regensburg, Regensburg, Germany. ³Department of Human Genetics, Radboud University Medical Center, 6500 HB Nijmegen, The Netherlands. ⁴Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands. ⁵Institute of Biochemistry & Molecular Biology, University of Hamburg, Hamburg, Germany. ⁶GFZ German Research Centre for Geosciences, Section Geomicrobiology, Potsdam, Germany. ⁷Division of Medical Genetics, Department of Pediatrics, Duke University, Durham, NC 27707, USA. ⁸University Hospital of Split, Split, Croatia. ⁹University of Split School of Medicine, Split, Croatia. ¹⁰Genetics Service, Department of Paediatrics, KK Women's & Children's Hospital, Singapore, Singapore. ¹¹Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. ¹²Department of Pediatrics, University Medical Center Eppendorf, 20246 Hamburg, Germany. ¹³GeneDx, Gaithersburg, MD 20877, USA. ¹⁴Division of Genetic Medicine, Seattle Children's Hospital, Seattle, WA 98105, USA. ¹⁵Institute of Human Genetics, University of Leipzig Hospitals and Clinics, Leipzig, Germany. ¹⁶Department of General Paediatrics, Division of Pediatric Epileptology, Centre for Paediatrics and Adolescent Medicine, University Hospital Heidelberg, Heidelberg, Germany. ¹⁷Department of Neuropediatrics, University of Leipzig, Leipzig, Germany. ¹⁸Department of Pediatrics, Zuyderland Medical Center, Heerlen and Sittard 6419, the Netherlands. ¹⁹Department of Genetics, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada. ²⁰Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON, Canada. ²¹Département de Génétique Médicale, CHU Timone Enfants, Assistance Publique - Hôpitaux de Marseille AP-HM, Marseille, France. ²²Aix Marseille Univ, INSERM, MMG, U1251 Marseille, France. ²³Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA. ²⁴Department of Pediatrics, New York University Grossman School of Medicine, New York, NY 10016, USA. ²⁵Neuroscience Research Australia (NeuRA), Prince of Wales Clinical School, University of New South Wales, Sydney, Australia. ²⁶NSW Health Pathology Randwick Genetics, Sydney, Australia. ²⁷Department of Neurology, Sydney Children's Hospital, Sydney, Australia. ²⁸School of Women's and Children's Health, University of New South Wales, Sydney, Australia. ²⁹Centre for Clinical Genetics, Sydney Children's Hospital, Sydney, Australia. ³⁰New South Wales Health Pathology Genomics Laboratory Randwick, Sydney, Australia. ³¹Department of Medical Genetics, University of British Columbia, Vancouver, BC V6H 3N1, Canada. ³²Research Laboratory, KK Women's & Children's Hospital, Singapore, Singapore. ³³Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA, USA. ³⁴Department of Pediatrics, University of Washington, Seattle, WA, USA. ³⁵Brotman Baty Institute for Precision Medicine, Seattle, WA 98195, US. [✉]email: d.lessel@uke.de; kreienkamp@uke.de

Chapter 3 | Genotype–phenotype correlations and novel molecular insights into the DHX30-associated neurodevelopmental disorders

Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza-Meireles A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, Busk ØL, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp HJ, Lessel D.

Genome Med. 2021 May 21;13(1):90. doi: 10.1186/s13073-021-00900-3. PMID: 34020708; PMCID: PMC8140440.

RESEARCH

Open Access



Genotype–phenotype correlations and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders

Ilaria Mannucci¹, Nghi D. P. Dang², Hannes Huber³, Jaclyn B. Murry^{4,5}, Jeff Abramson⁶, Thorsten Althoff⁶, Siddharth Banka^{7,8}, Gareth Baynam^{9,10,11}, David Bearden¹², Ana Belezza-Meireles¹³, Paul J. Benke¹⁴, Siren Berland¹⁵, Tatjana Bierhals¹, Frederic Bilan^{16,17}, Laurence A. Bindoff^{18,19}, Geir Julius Braathen²⁰, Øyvind L. Busk²⁰, Jirat Chenbhanich²¹, Jonas Denecke²², Luis F. Escobar²³, Caroline Estes²³, Julie Fleischer²⁴, Daniel Groepper²⁴, Charlotte A. Haaxma²⁵, Maja Hempel¹, Yolanda Holler-Managan²⁶, Gunnar Houge¹⁵, Adam Jackson^{7,8}, Laura Kellogg²⁷, Boris Keren²⁸, Catherine Kiraly-Borri²⁹, Comelia Kraus³⁰, Christian Kubisch¹, Gwenaél Le Guyader^{16,17}, Ulf W. Ljungblad³¹, Leslie Manace Brenman³², Julian A. Martinez-Agosto^{5,33,34,35}, Matthew Might³⁶, David T. Miller³⁷, Kelly Q. Minks¹², Billur Moghaddam²⁷, Caroline Nava²⁸, Stanley F. Nelson^{5,35,38}, John M. Parant², Trine Prescott²⁰, Farrah Rajabi³⁷, Hanitra Randrianaivo³⁹, Simone F. Reiter¹⁵, Janneke Schuurs-Hoeijmakers⁴⁰, Perry B. Shieh⁴¹, Anne Slavotinek²¹, Sarah Smithson¹³, Alexander P. A. Stegmann^{40,42}, Kinga Tomczak⁴³, Kristian Tveten²⁰, Jun Wang², Jordan H. Whitlock³⁶, Christiane Zweier^{30,44}, Kirsty McWalter⁴⁵, Jane Juusola⁴⁵, Fabiola Quintero-Rivera^{4,5,46}, Utz Fischer³, Nan Cher Yeo^{2*}, Hans-Jürgen Kreienkamp^{1*} and Davor Lessel^{1*} 

Abstract

Background: We aimed to define the clinical and variant spectrum and to provide novel molecular insights into the *DHX30*-associated neurodevelopmental disorder.

Methods: Clinical and genetic data from affected individuals were collected through Facebook-based family support group, GeneMatcher, and our network of collaborators. We investigated the impact of novel missense variants with respect to ATPase and helicase activity, stress granule (SG) formation, global translation, and their effect on embryonic development in zebrafish. SG formation was additionally analyzed in CRISPR/Cas9-mediated *DHX30*-deficient HEK293T and zebrafish models, along with in vivo behavioral assays.

(Continued on next page)

* Correspondence: nyeo@uab.edu; kreienkamp@uke.de; d.lesse@uke.de

²Department of Pharmacology and Toxicology, University of Alabama, Birmingham, USA

¹Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Results: We identified 25 previously unreported individuals, ten of whom carry novel variants, two of which are recurrent, and provide evidence of gonadal mosaicism in one family. All 19 individuals harboring heterozygous missense variants within helicase core motifs (HCMs) have global developmental delay, intellectual disability, severe speech impairment, and gait abnormalities. These variants impair the ATPase and helicase activity of DHX30, trigger SG formation, interfere with global translation, and cause developmental defects in a zebrafish model. Notably, 4 individuals harboring heterozygous variants resulting either in haploinsufficiency or truncated proteins presented with a milder clinical course, similar to an individual harboring a de novo mosaic HCM missense variant. Functionally, we established DHX30 as an ATP-dependent RNA helicase and as an evolutionary conserved factor in SG assembly. Based on the clinical course, the variant location, and type we establish two distinct clinical subtypes. *DHX30* loss-of-function variants cause a milder phenotype whereas a severe phenotype is caused by HCM missense variants that, in addition to the loss of ATPase and helicase activity, lead to a detrimental gain-of-function with respect to SG formation. Behavioral characterization of *dhx30*-deficient zebrafish revealed altered sleep-wake activity and social interaction, partially resembling the human phenotype.

Conclusions: Our study highlights the usefulness of social media to define novel Mendelian disorders and exemplifies how functional analyses accompanied by clinical and genetic findings can define clinically distinct subtypes for ultra-rare disorders. Such approaches require close interdisciplinary collaboration between families/legal representatives of the affected individuals, clinicians, molecular genetics diagnostic laboratories, and research laboratories.

Background

RNA helicases (RH) are highly specialized proteins which use ATP hydrolysis for the unwinding of RNA secondary structures and the remodeling of ribonucleoprotein particles (RNPs) [1, 2]. RHs are classified into six known superfamilies based on their sequence and structure [1]. Among these, the large helicase superfamily 2 (SF2) contains more than 50 members in humans [3]. These are designated DDX and DHX proteins based on the consensus amino acid sequence DExD or DExH signature in their ATP-binding motif II (Walker B motif) [3]. All SF2 RNA helicases are built around a highly conserved helicase core region consisting of two domains that resemble the bacterial recombination protein recombinase A (referred to as RecA-1 and RecA-2). Within these two core helicase domains, eight highly conserved sequence elements, helicase core motifs (HCMs) play a role in either RNA binding, or ATP binding and hydrolysis. The roles of SF2 RNA helicases include regulation of splicing, nuclear mRNA export, translation, transcription, facilitation of mRNA decay, microRNA processing, and cytoplasmic transport and storage of RNAs [1]. So far, many of the RHs have been studied in various cancers revealing the role of translation in carcinogenesis [4], and serve as potential biomarkers for diagnosis and prognosis, and novel drug targets [5]. The importance and functional relevance of certain SF2 RHs in human neurodevelopment is demonstrated by the identification of pathogenic germline variants in *DDX3X* [6], *DDX6* [7], *DHX30* [8], and *DDX5* [9] in individuals with neurodevelopmental disorders. Additionally, a paralog-based study implicated a role for *DHX16*, *DHX34*, *DHX37*, and *DDX54*, in human neurodevelopmental disorders and suggested that *DHX8*, *DDX47*, and *DHX58* may also be neurodevelopmental genes [10].

Previously, we reported 12 unrelated individuals with global developmental delay (GDD), intellectual disability (ID) accompanied by severe speech impairment and gait abnormalities, harboring one of six different de novo missense variants located within highly conserved HCMs of *DHX30* [8]. Moreover, a recent study reported gonadal mosaicism in two brothers carrying a de novo missense variant, p.(Ser737Phe), which resides within a HCM [11]. Here, we performed clinical, genetic, and functional analyses to provide further understanding of *DHX30*-related neurodevelopmental disorders through the identification of 25 previously unreported individuals. This systematic clinical and research approach, partially facilitated through social media, establishes novel genotype-phenotype correlations based on *in-depth* functional analyses accompanied by clinical and genetic findings.

Methods

Human subjects and genetic analyses

Written informed consent for all 25 subjects was obtained from the parents or legal guardians in accordance with protocols approved by the respective ethics committees of the institutions involved in this study. Next-generation sequencing-based analyses were performed in various independent research or diagnostic laboratories worldwide, using previously described procedures [8, 12–16]. Trio-whole exome sequencing (WES) was performed in families of subjects 1, 4, 5, 6, 10, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22, and 23. Single WES was performed in subjects 2, 3, 8, 16, and 25. Targeted Sanger sequencing was performed in subject 9, half-sister of subject 8. For subject 7, WES was performed as duo with DNA sample of his mother. Trio-whole genome

sequencing (WGS) in family of the subject 12 was done on an Illumina system using Nextera DNA Flex Library Prep. Reads were aligned to human genome build GRCh38 and analyzed for sequence variants using Cpipe analysis tool [17]. Classification the identified variants was based on the American College of Medical Genetics and Genomics (ACMG) guidelines [18]. Clinical Chromosomal Microarray analysis in family 24 was performed using standardized platforms [19]. Interpretation of identified copy number variants followed ACMG guidelines [20]. Most individuals were enrolled in the present study through the “*DHX30 family support group*” on Facebook: <https://www.facebook.com/groups/1808373282809332>. In such a case the families/legal representatives were asked to provide the contact details of attending physicians in order to obtain objective and accurate clinical and genetic data. Others presented in the University Medical Center Hamburg-Eppendorf, Hamburg, Germany, or were recruited through GeneMatcher [21] and our network of collaborators. For all 25 individuals, clinical data and information on genetic testing were uniformly obtained from attending physicians using a structured clinical summary (Additional file 1) and clinical table (Additional file 2: Table S1).

Cell culture and *in-vitro* assays

Human embryonic kidney 293 T (HEK 293 T) cells and human bone osteosarcoma epithelial (U2OS) cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) as described previously [8]. *DHX30* expression vectors based on pEGFP-C3 (leading to an N-terminal GFP-tag) and pEGFP-N2 (for expression of the mitochondrial form of *DHX30* with a C-terminal GFP-tag) have been described previously [8]. Newly identified missense variants were introduced into both vectors using Quick-Change II site-directed mutagenesis kit (Agilent, Waldbronn, Germany). HEK293T and U2OS cells were transfected with TurboFect or Lipofectamine 2000, respectively, transfection reagent (ThermoFisher Scientific) according to the manufacturer's recommendations. Immunocytochemistry and puromycin incorporation assay in U2OS cells were performed utilizing the following antibodies at manufacturers' recommended dilutions: anti-Puromycin mouse monoclonal (Millipore, #MABE343); goat anti-mouse coupled to Alexa Fluor 555 (ThermoFisher Scientific). A custom made anti-ATXN2 (#8G3, kindly provided by Dr. S. Kindler, Human Genetics, UKE; Hamburg) rat monoclonal antibody was used at a 1:10 dilution as previously described [8]. ATPase assay was performed as previously described [8]. Briefly, after transfection of HEK293T cells with *DHX30* expression vectors, followed by lysis in 1 ml of radioimmunoprecipitation assay buffer (RIPA), the lysates were clarified by centrifugation at 20,000×g for 20

min at 4 °C. GFP-containing proteins were purified from the supernatant by immunoprecipitation using 20 µl of GFP-Trap_A matrix (Chromotek, Munich, Germany). Precipitates were washed twice in RIPA buffer, and twice in phosphate-free ATPase assay buffer (40 mM KCl; 35 mM HEPES pH 7.5; 5 mM MgCl₂; prepared in plastic ware to avoid phosphate contamination). Precipitates were then incubated in 50 µl phosphate-free buffer supplemented with 2 mM ATP and 2 mM DTT at 30 °C for 30 min (for assaying ATPase activity in the absence of exogenous RNA). After brief centrifugation (1 min, 1000×g), the supernatant was removed and precipitated samples were incubated in phosphate-free buffer containing 2 mM ATP; 2 mM DTT, and 100 µg/ml yeast RNA for 30 min at 30 °C (for assaying ATPase activity in the presence of exogenous RNA). The amount of free phosphate released by ATP hydrolysis was determined photometrically using Biomol Green reagent (Enzo Life Sciences, Lörrach, Germany). Subsequently, bead-attached proteins were denatured in SDS-sample buffer, and the amount of *DHX30* protein was determined by western blotting using anti-GFP (Covance). In each case, ATPase activity was normalized to the amount of GFP-tagged *DHX30* protein attached to the GFP-trap matrix.

Helicase assay

6xHis-SUMO-*DHX30* wild-type and mutant proteins were expressed in the *E. coli* BL21 (DE3) pLysSpRARE cells (Novagen, Germany). Proteins were purified from lysates using Ni-NTA beads (Qiagen, Germany) as previously described [22]. To test the RNA unwinding activity of *DHX30*, a [³²P]-labeled RNA duplex was synthesized using the T7 RNA polymerase from a linearized DNA template designed by Tseng-Rogenski and Chang [23]. Helicase activity was measured in 20 µl of reaction mixture containing 0.13 pmol of purified protein (=20 ng of full-length protein), 25 fmol [³²P]-labeled RNA duplex, 17 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1 mM MgCl₂, 2 mM DTT, 1 mM spermidine, 0.3% PEG8000, 5% glycerol, 150 mM KCl, 20 units of RNasin™ Plus (Promega, USA), 1 mM ATP. The mixture was incubated for 1 h at 37 °C, mixed with 2X non-denaturing loading dye and subjected to gel electrophoresis through non-denaturing 8% PAGE (19:1) in 0.5X TBE at 4 °C. Reaction products were visualized by autoradiography. For more information see Additional file 1: Supplementary methods.

Generation of a HEK293T *DHX30* stable knockout line

HEK293T *DHX30*-deficient cells were generated by transfecting a plasmid (pLentiCRISPR v2, GenScript, #52961) encoding a single guide RNA (CGAGTGCTAG CTGATCGCTT) targeting exon 7, the Cas9 endonuclease and a puromycin resistance gene under the control of the EFS promoter. Cells were transfected with

TurboFect transfection reagent (Thermo Scientific) and treated with puromycin for 3 days. Surviving cells were then subjected to single cell sorting using BD FACS Aria™ IIIu Cell Sorter (BD Biosciences). Single-cell clones were grown in 96-well plates for two weeks and then expanded into 6 well dishes. DHX30 knockout efficiency was assessed by Western blotting using an anti-DHX30 rabbit polyclonal antibody (Bethyl, #A302-218A) (1:500).

Stress treatment

HEK293T WT and *DHX30*-deficient cells were plated on glass coverslips coated with poly-L-Lysine. After 24 h, cells were heat stressed at 43.5 °C for 1 h, fixed in 4% paraformaldehyde and permeabilized with 0.1% Triton X-100 (Sigma). Blocking was performed using 10% horse serum (HS). Rat monoclonal anti-ATXN2 was used as a primary antibody (1:10 in 2 % HS in PBS), followed by goat anti-rat IgG coupled to Alexa Fluor 647 (Thermo Fisher Scientific). Coverslips were mounted on glass microscope slides with ProLong Diamond Antifade Mountant with DAPI (Thermo Fisher Scientific). Immunofluorescence images were acquired using a confocal microscope (Leica TCS SP5, 63x/1.25 objective) and processed with ImageJ software.

Construction of Tol2 plasmids

DHX30 cDNA plasmids were assembled using the Tol2 MultiSite Gateway™ kit (Invitrogen, USA). Briefly, the cDNA of the wild-type DHX30 and DHX30 containing respective missense variants were amplified from the pEGFP-C3-DHX30 plasmids, using primers containing the appropriate att site sequences for BP recombination reactions. PCR products were purified and cloned into a pDONR221 donor vector using BP Clonase II enzyme mix following the manufacturer's manual. The resulting middle entry clones pME-DHX30 were purified and verified by direct sequencing. To assemble the final expression plasmids, p5E-tuba1a promoter and pME-DHX30 were cloned into a Tol2-based destination vector, pDest-Tol2CG2 containing *cmlc2:EGFP* transgenesis marker, using LR Clonase II Plus enzyme mix following the manufacturer's instructions. The resultant pTol2pA2-*cmlc2:EGFP;tuba1a:DHX30* vectors were purified and verified by direct sequencing.

Zebrafish maintenance and manipulation

Tol2 transposase mRNA were synthesized using mMES-SAGE mMACHINE™ T7 Transcription Kit (Ambion) per the manufacturer's instructions. Twenty-five nanograms/μl Tol2 mRNA and 25 ng/μl of pTol2pA2-*cmlc2:EGFP;tuba1a:DHX30* DNA were injected into 1-cell stage zebrafish embryos (Danio rerio AB strain). To investigate a potential dominant-negative effect, 25 ng/μl

Tol2 mRNA and 25 ng/μl equal mixture of pTol2pA2-*cmlc2:EGFP;tuba1a:DHX30* with the respective variant DNA were injected into 1-cell stage zebrafish embryos. The embryos were raised and scored for abnormal development 1–7 days post fertilization. Zebrafish were maintained in the Zebrafish Research Facility at the University of Alabama at Birmingham using standard protocols. All fish were maintained at 28 °C and kept at 14-h light and 10-h dark cycle under standard laboratory conditions.

Generation of zebrafish *dhx30* stable knockout line

The zebrafish *dhx30* stable knockout line was generated using CRISPR/Cas9 with sgRNA target sequence 5'-TCAAGTTCAGCTGCACGGAT-3' made by Integrated DNA Technologies (IDT) according to the manufacturer's protocol. The mutant contains an 8-bp deletion that shifts the translational reading frame after amino acid 90 and results in a premature stop codon at amino acid 107, compared to 1173 amino acids for the wild-type (WT) protein. Mutant animals were genotyped and sequenced using primers 5'-ATCTTCACGCCAAAAA CCTG-3' and 5'-GACCACGGTTCAGCTCTCTC-3'. The *dhx30* heterozygous mutants were outcrossed to the parental AB strain for at least two generations before use in experiments to eliminate potential off-target variants. After each assay described below, test animals were individually genotyped using PCR with primers 5'-ATCTTCACGCCAAAAA CCTG-3' and 5'-GACCACGGTTCAGCTCTCTC-3' and high-resolution melting (HRM) analysis as previously described [24].

Stress treatment and zebrafish whole-mount immunostaining

The *dhx30* +/- animals were in-crossed to generate *dhx30* +/+, +/-, and -/- sibling progeny for heat shock and immunostaining analyses. Twenty-four-hour post-fertilization embryos were dechorionated and incubated at 28 °C or 42 °C for 1 h. After treatment, embryos were fixed overnight in cold 4% paraformaldehyde (PFA). Embryos were then dehydrated with acetone at -20 °C for 7 min, washed in PBST [PBS+0.1% Tween 20], and blocked with 10% goat serum for at least 1 h at room temperature. Thereafter, embryos were incubated with rabbit anti-TIAL-1 (Novus Biologicals, NBP1-79932; 1:200) overnight at 4 °C, washed with PBST, and incubated with secondary antibody Alexa Fluor 488-conjugated goat anti-rabbit IgG (Invitrogen, A11034; 1:200) for 2 h at room temperature. Embryos were washed with PBST, incubated with 100 μM DAPI (1:500) to counterstain nuclei for 10 min, and stored in PBS at 4 °C. For imaging, stained embryos were mounted in 1% low melting agarose and imaged using a Nikon A1 inverted confocal microscope at approximately 50-μm Z-stacks at 5.6 μm

intervals. The number of TIAL-1-labeled stress granules per 50 nuclei was quantified using Nikon NIS Element. After imaging, test animals were individually genotyped by PCR and HRM analysis to delineate the *dhx30* genotype.

Behavioral assays

For each behavioral experiment, *dhx30* +/- animals were in-crossed to generate *dhx30* +/+, +/-, and -/- sibling progeny.

Twenty-four-hour sleep-wake activity

For each sleep-wake study, zebrafish larvae at 5-day post-fertilization (dpf) were chosen randomly and placed individually into each well of a flat-bottom 24-well plate. The activity of each larva was tracked for 24 h consisting of 14-h light and 10-h dark using the DanioVision system (Noldus Information Technology). The average swimming distance was measured for 24 h per 1-h time-bins using EthoVision XT software (Noldus).

Social preference assay (SPA)

We adopted and modified a previously described social preference assay (SPA) [25]. Briefly, SPA was performed using a flat-bottom 12-well plate and custom-built removable opaque dividers. The individual "test" animals, whose behaviors were analyzed, were placed in each of the 4 middle wells of the plate, and a WT conspecific of similar age and size was placed in a well either above or below each middle well. The activity of each test larva was tracked using the DanioVision (Noldus Information Technology) system and data analyzed using EthoVision XT software (Noldus). Before data acquisition, animals were given 5-min habituation period. The "baseline" activity of the test fish was then recorded while the opaque dividers were inserted between each well to prevent the animals from seeing each other. The dividers were then removed, allowing each test animal to view one well containing a conspecific animal and one empty well. The fish were given another 5-minute habituation period, followed by a 10-min "post-baseline" recording. For data analyses, wells containing test fish were divided into two 0.5 cm × 2.2 cm zones, one closest to the well containing a conspecific animal and one closest to the empty well. The amount of time spent by a test fish in each zone during the baseline and post-baseline periods was analyzed. The social preference of each test fish was quantified by calculating the social preference index (SPI) = (time spent in zone near the conspecific fish - time spent in zone near the empty well)/time spent in both zones as previously described [25].

Statistical analyses

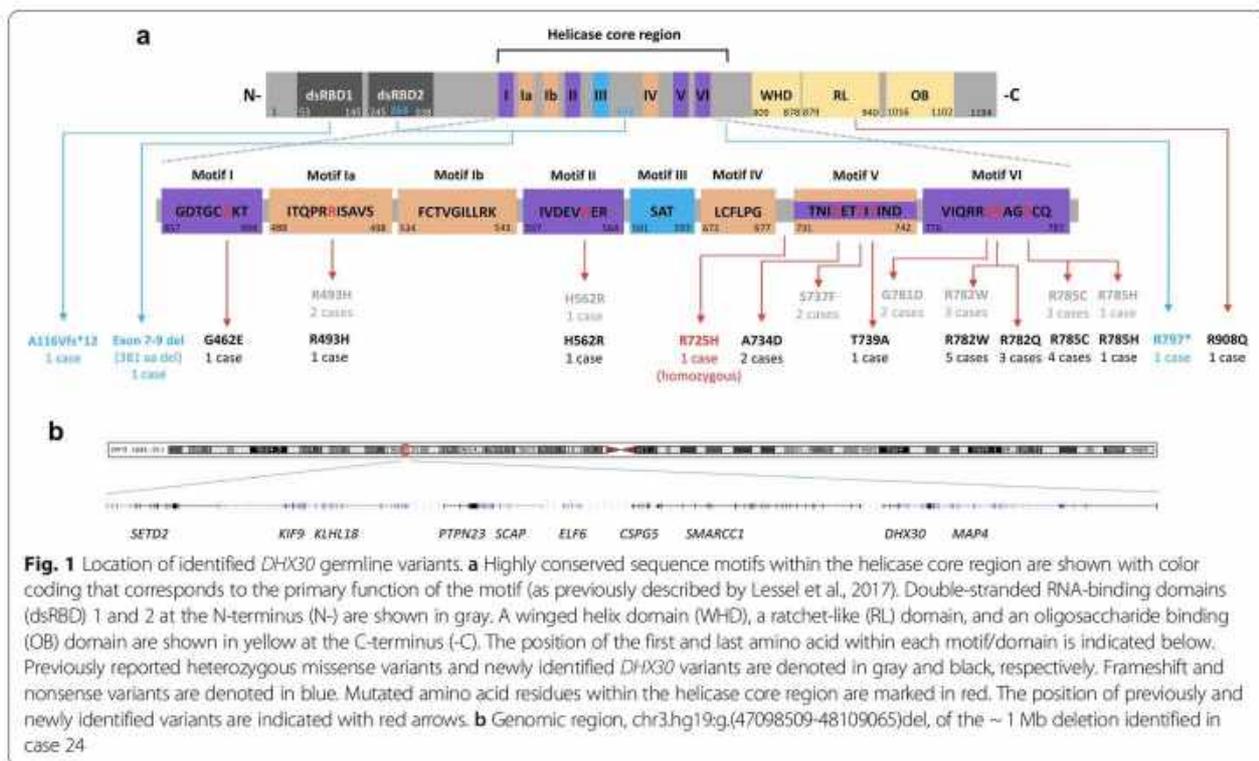
All cell line data (U2OS and HEK293T) are presented as mean ± SD and analyzed by One-Way ANOVA followed by Dunnett's multiple comparisons test or unpaired Student's *t* test as indicated in figure descriptions. All zebrafish-related data are presented as mean ± SEM and analyzed by unpaired Student's *t* test. The percentage of developmental defects observed upon overexpression of *dhx30* was analyzed by the χ^2 test.

Results

Identification of likely causative variants in *DHX30*

We identified 25 individuals carrying likely causative variants in *DHX30* (Fig. 1). Of these, 12 individuals carry a previously reported heterozygous missense variant localizing within highly conserved helicase core motifs (HCMs): p.(Arg493His), p.(His562Arg), p.(Arg782Trp) (5 individuals including two half-sisters indicative of gonadal mosaicism), p.(Arg785Cys) (4 individuals), and p.(Arg785His). Further, 7 individuals have a novel heterozygous missense variant classified as either "likely pathogenic" or "pathogenic" according to The American College of Medical Genetics and Genomics (ACMG) guidelines (Additional file 2: Table S1) [18]. Indeed, each of these variants alters a highly conserved amino acid within a HCM predicted to be responsible for ATP binding and/or hydrolysis (Fig. 1 and Additional file 3: Figure S1). p.(Gly462Glu) identified in a single individual affects motif I, also referred to as Walker A motif, that binds γ phosphate and coordinates, together with motifs II and VI, ATP binding and hydrolysis in other DEXH family members [26, 27]. p.(Ala734Asp) identified in two unrelated individuals, one of which (individual 6) appears to have mosaicism for the variant (Additional file 4: Figure S2), and p.(Thr739Ala) identified in a single individual, both affect motif V which regulates both ATP binding and/or hydrolysis and RNA binding [2, 27]. Three individuals carry p.(Arg782Gln), located within motif VI affecting the identical arginine residue (Arg782), that we previously reported p.(Arg782Trp) [8], which was identified here in five additional individuals.

Moreover, a homozygous variant p.(Arg725His) located within the helicase core region albeit between motifs IV and V, unlike all the missense variants mentioned above, was identified in individual 4 and classified as "variant of uncertain significance". Additionally, a heterozygous *de novo* variant p.(Arg908Gln) was identified in individual 21. This was the only variant not located within the helicase core region and was classified as "likely pathogenic". Predictions based on homology to other SF2 helicases [26, 28] and published structures of the Prp43 [29] and Mle [30] revealed three novel highly conserved C-terminal regulatory domains (CTD). These include a winged helix (WH), a ratchet-like (RL) and an



oligosaccharide binding (OB) fold domain (Fig. 1) with a potential role in coupling ATP hydrolysis to RNA unwinding [31]. Notably, the p.(Arg908Gln) affects a highly conserved residue within the RL domain (Figure S1).

Furthermore, we identified four individuals bearing likely pathogenic loss-of-function variants. A heterozygous de novo frameshift variant, p.(Ala116Valfs*12) in individual 22, a heterozygous nonsense variant, p.(Arg797*) in individual 23 inherited from a mosaic mother, and a *de novo* in-frame deletion encompassing exons 7-9 of *DHX30*, leading to deletion of 381 amino acids, in individual 25. Individual 24 has a large heterozygous de novo deletion (arr[GRCh37] 3p21.31 (47098509_48109065)del) encompassing ten genes including two disease genes previously associated with an autosomal dominant inheritance, *SETD2* [32–34] and *DHX30* (Fig. 1b and Additional file 5: Figure S3), possibly pointing to a dual diagnosis. The whole gene deletion results in haploinsufficiency, whereas the in-frame deletion, frameshift, and nonsense variant, if they were to result in stable proteins, are predicted to lead to loss of functionally important domains (Fig. 1a).

Notably, none of these *DHX30* alterations was present in the gnomAD dataset v2.1.1 (Additional file 2: Table S1) [35], indicating that they are extremely rare in the population and unlikely to be variants unrelated to disease. As previously noted, *DHX30* is one of the most missense-intolerant genes in the human genome [8]. Furthermore, according to the gnomAD v2.1.1 dataset

DHX30 is, with a probability of being loss-of-function intolerant (pLI) score of 1 and a loss-of-function observed/expected upper bound fraction (LOEUF) score of 0.04, extremely loss-of-function intolerant [35]. Additionally, the degree of intolerance to deleterious variants of *DHX30* according to the Residual Variation Intolerance (RVI) score, which quantifies gene intolerance to functional variants, is -1.51 (3.54th percentile) and thus even lower than the average RVI score for genes involved in developmental disorders (0.56; 19.54th percentile) [36, 37].

Clinical spectrum of the *DHX30*-associated neurodevelopmental disorders

All 19 individuals harboring a heterozygous missense variant within a highly conserved motif in the helicase core domain have global developmental delay (GDD), intellectual disability (ID), severe speech impairment, and gait abnormalities, similar to our initial findings [8]. In more detail, all individuals had an intellectual disability, only nine (47%) learned to walk, all with an ataxic gait. The majority had no speech (74%), four individuals spoke only single words, and only individual 6, who is mosaic for the de novo p.(Ala734Asp) variant, spoke simple sentences. It is worth noting that the individuals highly benefit from communication devices (tablets, smartphones, and eye-driven tablet communication systems) which significantly reduced frustration-related

behavior (D.L. personal communication with legal guardians and family members). Additional phenotypic features included muscular hypotonia in eighteen (95%), feeding difficulties in sixteen (84%), microcephaly in thirteen (81%, 13/16), joint hypermobility in fourteen (74%), structural brain anomalies in eleven (65%, 11/17), sleep disturbances in nine (47%), strabismus in eight (42%), autistic features in five (33%, 5/15), and seizures in four (21%) individuals (Table 1, Additional file 2: Table S1 and Additional file 6). Noteworthy, individual 6 had a relatively milder clinical course, with a moderate intellectual disability, independent walking at 2 years and 8 months, and the ability to speak in simple sentences at the age of 15 years. This individual's presentation is similar to that of the four individuals (#22, #23, #24, and #25), who carry either a frameshift or nonsense variant, whole-gene deletion or in-frame deletion, respectively, who all learned to walk in the second year of life, had a mild muscular hypotonia and spoke at least 20 words by the age of 3 years. Although some individuals displayed some dysmorphic features (Additional file 6) we did not

observe a recognizable facial gestalt, similar to our previous findings [8].

Two individuals (#4 and #21) clearly stand out phenotypically. Individual 4 was homozygous for p.(Arg725His), developed early-onset infantile epileptic encephalopathy, and died at 11 months. In contrast, individual 21, who harbors the de novo p.(Arg908Gln) variant, had unremarkable psychomotor development until the age of 8 years when she presented with progressive balance impairment with truncal ataxia. Subsequently, she experienced a decline in motor skills and developed cognitive problems with reduced concentration (Table 1, Additional file 2: Table S1 and Additional file 6).

Effect of novel *DHX30* missense variants on ATPase activity

To corroborate the pathogenicity of the novel missense variants identified in this study, along with the recently reported p.(Ser737Phe) [11], we performed several previously established functional assays [8]. First, we analyzed the ATPase activity of wild-type (WT) and mutant forms of *DHX30*. As

Table 1 Clinical features in 25 individuals bearing pathogenic *DHX30* variants and frequency of these features in previously reported individuals

<i>DHX30</i> variant	Heterozygous missense variants within a HCM (this study)	p.(Ala734Asp) mosaic (this study)	Haploinsufficiency/protein truncating variants (this study)	Homozygous p.(Arg725His) (this study)	Heterozygous p.(Arg908Gln) (this study)	Heterozygous missense variants within a HCM (previous studies: Lessel et al. 2017 and Cross et al. 2020)
Sex	11 females/7 males	Female	1 female/3 males	Male	Female	8 females/6 males
Intellectual disability	18/18	+	4/4	?	–	13/13
Speech ability	14/18 non-verbal 4/18 single words	Simple sentences	20 words to normal speech ability	?	Normal speech ability	11/13 non-verbal 2/13 single words
Motor development delay	18/18	+	4/4 mild	+	–	14/14
Muscular hypotonia	17/18	+	3/4	+	–	14/14
Gait abnormalities	10/18 no independent walking 8/18 ataxic	Ataxic gait	0/4 no independent walking 3/4 ataxic gait	?	Ataxic gait	7/13 no independent walking 6/13 ataxic gait
Feeding difficulties	15/18	+	1/4	+	–	11/14
Microcephaly	13/15	+	0/4	–	–	7/10
Joint hypermobility	13/18	+	1/3	–	–	6/14
Brain MRI anomalies	11/17	–	2/3	+	+	10/14
Sleep disturbance	8/18	+	2/3	+	–	7/12
Strabismus	8/18	–	2/4	–	+	6/14
Autistic features	4/14	+	0/3	?	–	7/12
Seizures	3/18	+	2/3	Severe	–	3/14

+, present; –, absent; ?, too young to evaluate; NA, unknown

previously shown, DHX30-WT acts as an RNA-dependent ATPase, and its ATPase activity is strongly stimulated by the addition of RNA [8]. In contrast, and similar to the previously analyzed mutants [8] all missense variants (p.(Gly462Glu), p.(Arg725His), p.(Ala734Asp), p.(Ser737Phe), p.(Thr739Ala), p.(Arg782Gln), and p.(Arg908Gln)) show a significant reduction in ATPase activity in the presence of exogenous RNA (Fig. 2a). For control experiments, we included two common non-synonymous *DHX30* variants found in public repositories [35]. Namely, p.(Val556Ile) is located within the helicase core region albeit not within a HCM, similar to p.(Arg725His), and p.(Glu948Lys) in the vicinity of p.(Arg908Gln). Notably, in comparison to the missense variants identified in affected individuals, the ATPase activity was not significantly reduced neither for p.(Val556Ile) nor for p.(Glu948Lys) (Fig. 2b).

RNA helicase activity of DHX30 is disrupted by missense variants within the helicase core motifs

DHX30 has been classified as an RH due to the presence of the highly conserved motifs in its helicase core region and sequence similarity to other RHs. To confirm that it indeed possesses RNA helicase activity we established an RNA unwinding assay for recombinant full-length DHX30 purified from bacteria as a His₆-SUMO-tagged protein. As a substrate, we used a synthetic [³²P]-labeled RNA molecule which carries a sequence with a strong propensity to self-anneal and form a double helix. Analysis of this RNA substrate by non-denaturing PAGE resulted in a single band of low electrophoretic mobility, corresponding to the dimer linked by the double helical segment. This dimer could be resolved into a band of higher mobility, the monomer, by pre-incubation at 96 °C (Fig. 2c). To identify the amount of the DHX30-WT necessary to resolve the dimeric form we performed a titration analysis from 1 to 160 ng. In the presence of ATP, 10 ng of DHX30-WT was sufficient to resolve the dimer into the monomeric form, confirming that DHX30 indeed possesses the ATP-dependent RNA helicase activity (Fig. 2c and Additional file 7: Figure S4). We next analyzed the impact of selected missense variants on the helicase activity, each affecting a different helicase core motif (p.(Gly462Glu) in motif I, p.(Arg493His) in motif Ia, p.(His562Arg) in motif II, p.(Ser737Phe) in motif V, and p.(Arg785Cys) in motif VI) along with p.(Arg908Gln) located in the RL domain. All missense variants within a HCM failed to unwind the RNA substrates in this assay, whereas the p.(Arg908Gln) mutant behaved similarly to DHX30-WT (Fig. 2d). It is worth noting that we subsequently failed to purify the p.(Arg725His) mutant protein product. This finding suggests misfolding of this mutant protein, followed by either deposition of the insoluble protein in

inclusion bodies or its direct degradation [38]. Thus, we could not analyze its impact on the helicase activity.

Subcellular localization and effect on global translation of novel DHX30 missense variants

We have previously shown that the expression of mutant forms of DHX30 induces the formation of stress granules, concomitant with a global down-regulation of translation [8]. Therefore, we repeated this analyses for selected novel missense variants. In keeping with the previous results [8], we observed that mutants within a HCM, p.(Gly462Glu), p.(Ala734Asp), p.(Ser737Phe), and p.(Thr739Ala), also strongly accumulated in cytoplasmic foci shown to be stress granules upon co-staining with Ataxin-2 (ATXN2). Expression of the p.(Arg908Gln) mutant, however, resulted in localization to cytoplasmic aggregates that co-stained with Ataxin-2 in only 50% of the transfected cells. In contrast, p.(Arg725His) was mostly diffusely localized throughout the cytoplasm similar to the DHX30-WT (Fig. 3 and Additional file 8: Figure S5). Global translation was measured by incorporation of puromycin into nascent peptide chains, which were visualized with a puromycin-specific antibody. Interestingly, expression of both the HCM mutants and the p.(Arg908Gln) mutant resulted in dramatically decreased puromycin incorporation, suggestive of a global decrease in protein synthesis (Fig. 3). Analogous to the results obtained in the ATPase assay, the two common variants p.(Val556Ile) and p.(Glu948Lys) were diffusely localized throughout the cytoplasm, resembling the DHX30-WT (Additional file 8: Figure S5).

In vivo analyses of selected DHX30 missense variants

Given the somewhat conflicting results of functional analyses of the p.(Arg725His) and p.(Arg908Gln) variants, and in order to gain a better understanding of the impact of *DHX30* missense variants in vivo, we utilized a zebrafish model. Previous studies showed that overexpression of pathogenic alleles in zebrafish results in defective embryonic development [39, 40]. Thus, we overexpressed human wild-type DHX30 cDNA or DHX30 cDNA harboring selected missense variants, p.(Arg493His), p.(Arg725His), p.(Arg785Cys), and p.(Arg908Gln) as well as p.(Val556Ile) and p.(Glu948Lys) in zebrafish using Tol2 transposition. Tol2 mRNA and pTol2pA2-cmlc2:EGFP;tuba1a:DHX30 were co-injected into 1-cell stage zebrafish embryos. For analyses, we selected embryos with strong cmlc2:EGFP expression which indicates a high level of transgene integration in somatic cells. Overexpression of DHX30-WT, p.(Val556Ile) or p.(Glu948Lys) had little or no impact on zebrafish embryonic development: over 88% of embryos displayed normal development and morphology. However, expression of DHX30 harboring one of the missense variants resulted in developmental defects in 75–90% of embryos (Fig. 4 and Additional file 9: Figure S6),

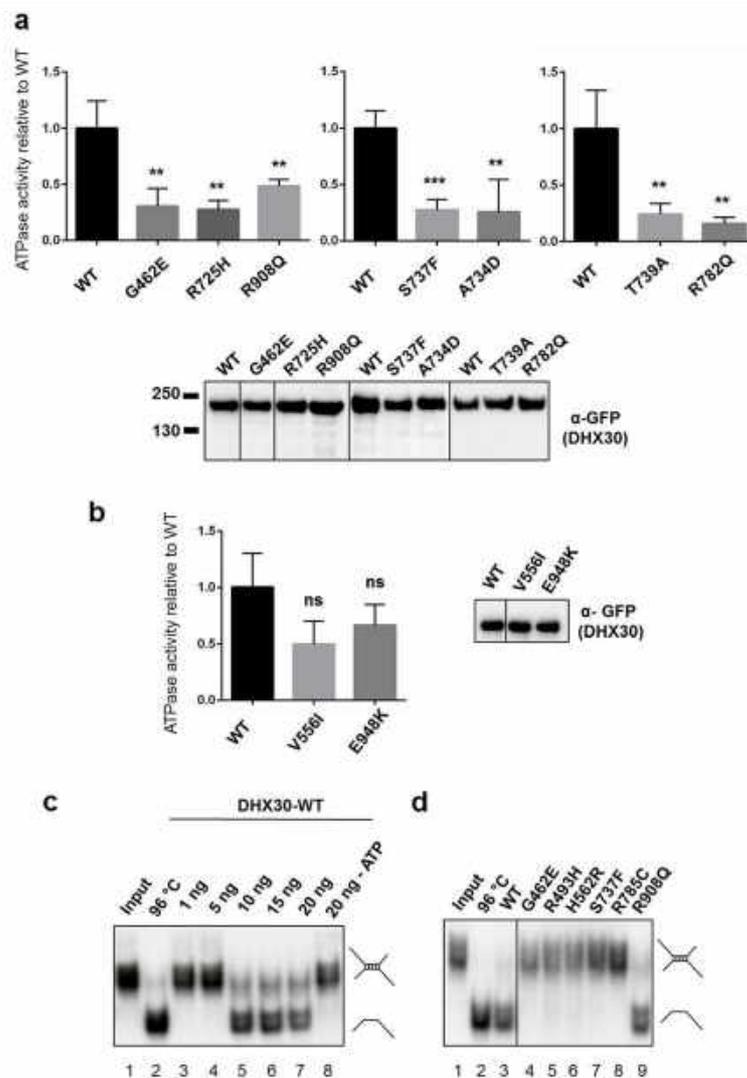


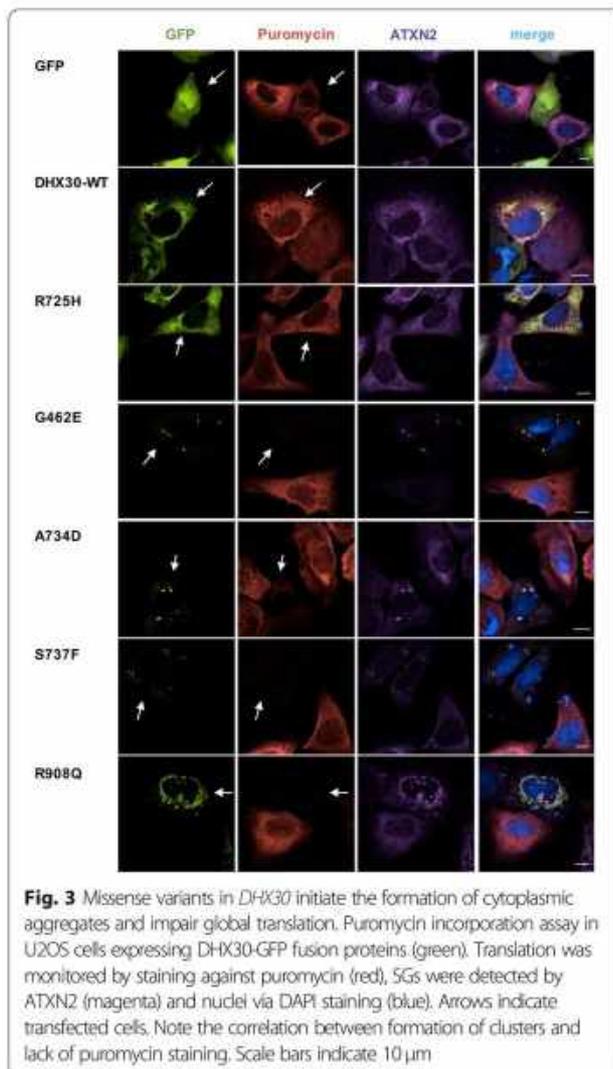
Fig. 2 Protein variants of *DHX30* affect ATPase and helicase activity. **a, b** ATPase assays were performed for *DHX30*-WT, novel *DHX30* missense variants (**a**), and two common polymorphisms, p.(Val556Ile) and p.(Glu948Lys) (**b**) in the presence of exogenous RNA. ATPase activity was calculated by subtracting phosphate values obtained with GFP alone from those obtained with GFP-tagged *DHX30*-WT and mutants. These figures were then normalized on precipitated protein amounts using the intensities of the GFP signal in the western blot. Means \pm standard deviation values are based on 3 replications. **,***: significantly different from *DHX30*-WT (*** p < 0.001, ** p < 0.01; n = 3; One-Way ANOVA, followed by Dunnett's multiple comparisons test). Values were normalized on *DHX30*-WT ATPase activity obtained in the presence of RNA. **c** Increasing amounts of His₆-SUMO-tagged *DHX30* WT protein were incubated with a ³²P-labeled RNA substrate in the presence (lane 3–7) or absence (lane 8) of ATP and analyzed by native PAGE. The position of the RNA duplex and the single-stranded RNA are indicated in the first and second lanes, respectively. Their schematic representation is shown at the right side. **d** Helicase assay was repeated for selected *DHX30* missense variants affecting either conserved motifs within the helicase core region (lane 4–8) or the auxiliary RL domain (lane 9)

suggesting that these mutant alleles interfere with normal embryonic development and supporting the pathogenicity of p.(Arg725His) and p.(Arg908Gln).

Analyses of the nature of the *DHX30* missense variants

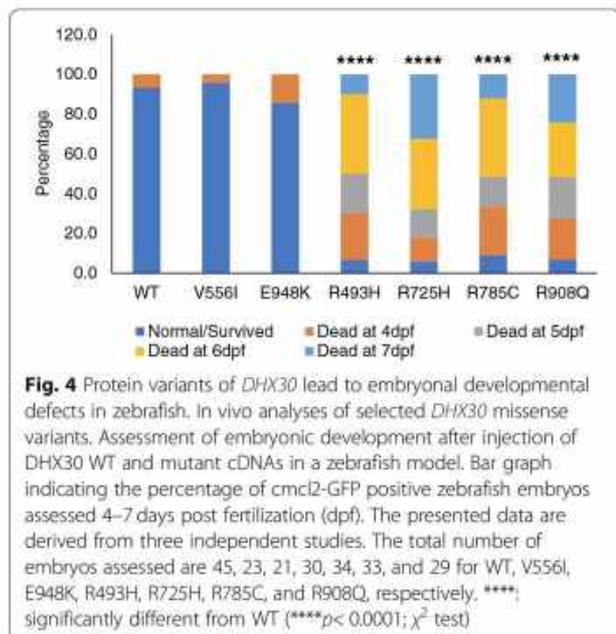
Given the somewhat milder clinical presentation of individuals carrying a whole gene deletion, in-frame deletion, frameshift, or nonsense variant as compared to the individuals harboring a *de novo* missense variant in one of the

HCMs, we further investigated the nature of the latter. First, we analyzed the localization of the RFP-tagged *DHX30*-WT co-expressed with respective GFP-tagged missense variants. Notably, their equimolar expression resulted in each case in *DHX30*-WT being localized in Ataxin-2 positive cytoplasmic clusters (Fig. 5). These data suggest that these missense variants either exert a dominant negative effect on the wild-type or lead to a gain-of-function since both overexpressed *DHX30*-WT and



endogenous *DHX30* are recruited to cytoplasmic clusters only after stress [8].

Next, we analyzed if the *DHX30*-WT can rescue the inability of p.(Arg493His), p.(His562Arg), and p.(Arg785Cys) to unwind RNA. The addition of *DHX30*-WT to p.(His562Arg) and p.(Arg785Cys) efficiently resolved the dimer into the monomeric form even in the presence of increased amounts of the respective mutants. However, we observed only a partial rescue when *DHX30*-WT was added to the p.(Arg493His) variant (Fig. 6a). Our data suggest that the mutants cause a loss of helicase function rather than having a dominant negative effect. Given these somewhat contradictory results, we turned again to the zebrafish model. We co-injected pTol2pA2-cmlc2:EGFP;tuba1a:*DHX30* p.(Arg493His) or p.(Arg785Cys) with wild-type *DHX30* cDNA and assessed embryonic development. Interestingly, co-injection of *DHX30*-WT, at a similar level, partially



rescued the abnormal phenotypes associated with both p.(Arg493His) and p.(Arg785Cys) (Fig. 6b), a finding that could potentially support both loss-of-function and a dominant negative effect as a mechanism underlying disease.

DHX30 deficiency impairs stress granules formation in HEK293T cells and zebrafish

To further characterize the role of *DHX30* we established HEK293T *DHX30* stable knockout lines. CRISPR/Cas9 based knockout (KO) of *DHX30* in HEK293T cells yielded several cell lines with a residual *DHX30* immunoreactivity of less than 10 % (Fig. 7). Given that *DHX30* is recruited to SGs, we wondered whether *DHX30* additionally plays a role in SG formation. Therefore, we assessed the ability of KO cells to induce SGs or cytoplasmic clusters following heat stress treatment. By incubating cells at 43.5 °C, a condition after which endogenous *DHX30* accumulates in SGs [8], we observed that KO cells had a significantly reduced number of SG-positive cells as compared to HEK293T WT cells (Fig. 7). These data suggest a previously unknown role of *DHX30* in SG assembly. Combined with our previous findings (Fig. 3 and 5) these data actually suggest that the HCM missense variants exhibit a gain of function by triggering SG formation which results in global translation inhibition.

Next, we generated a predicted null allele in the single zebrafish *dhx30* ortholog using CRISPR/Cas9. At day five post fertilization, transcript levels of *dhx30* were barely detectable in homozygous mutant animals compared to wild-type, whereas heterozygous siblings

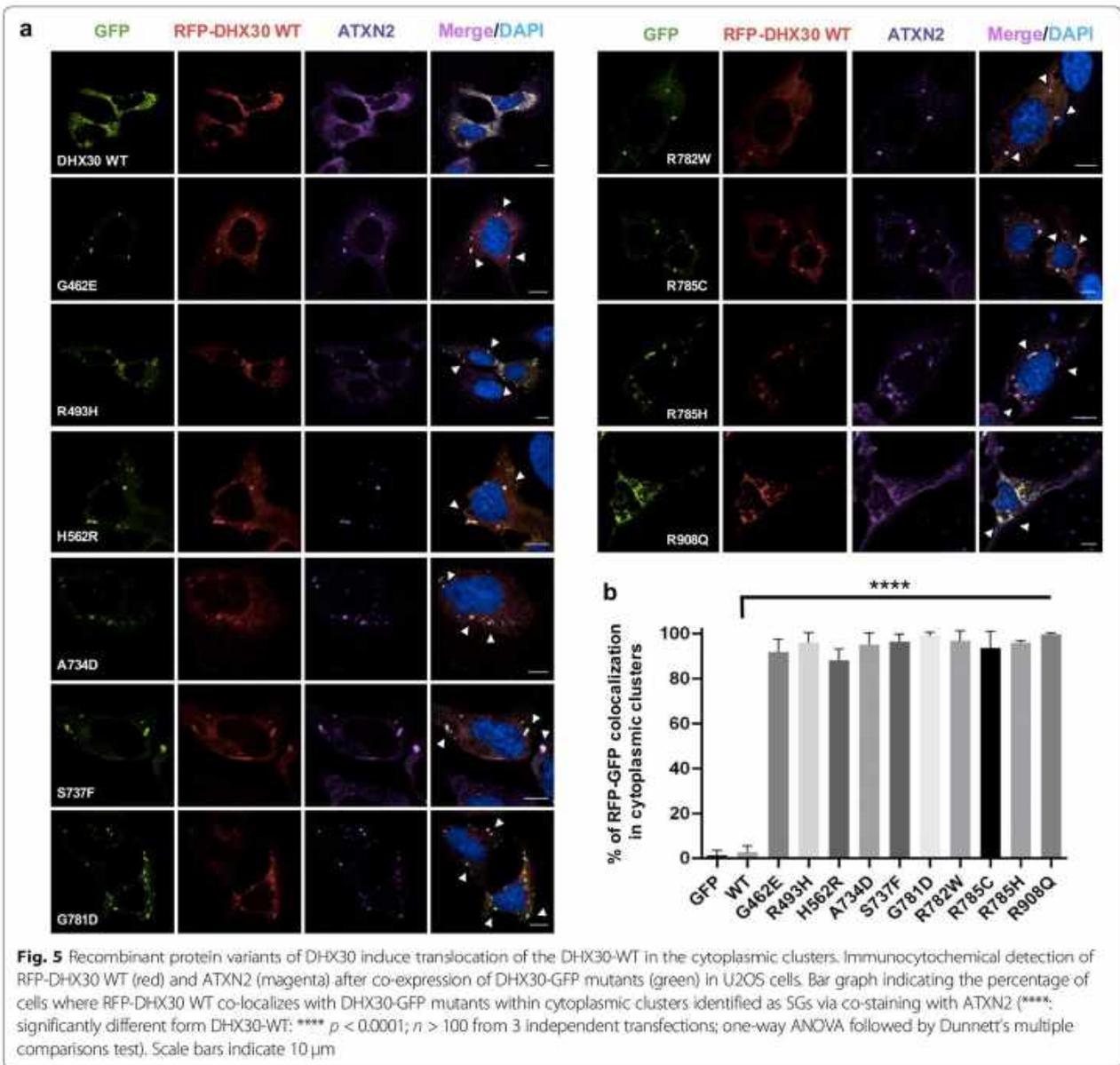


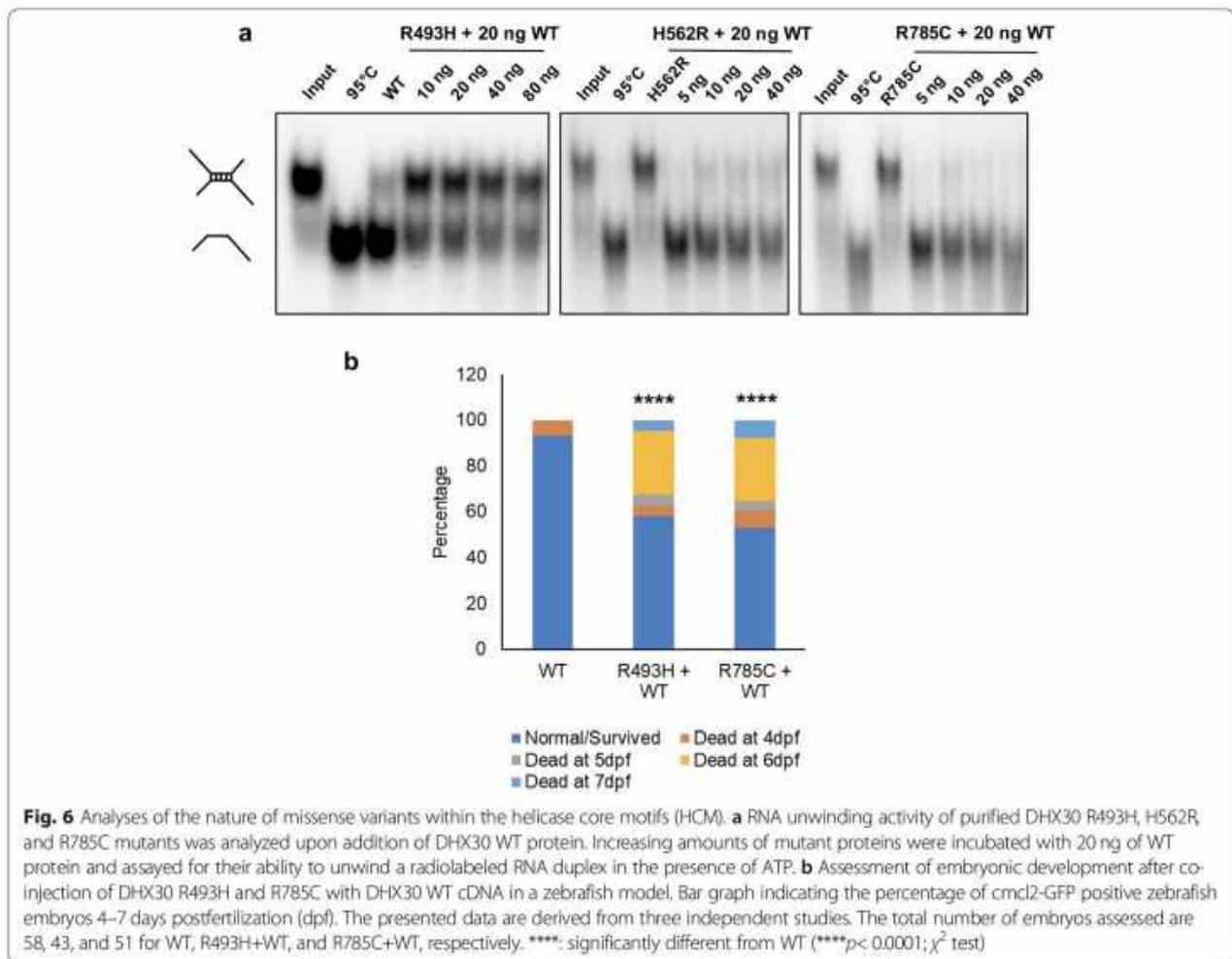
Fig. 5 Recombinant protein variants of DHX30 induce translocation of the DHX30-WT in the cytoplasmic clusters. Immunocytochemical detection of RFP-DHX30 WT (red) and ATXN2 (magenta) after co-expression of DHX30-GFP mutants (green) in U2OS cells. Bar graph indicating the percentage of cells where RFP-DHX30 WT co-localizes with DHX30-GFP mutants within cytoplasmic clusters identified as SGs via co-staining with ATXN2 (****: significantly different from DHX30-WT: **** $p < 0.0001$; $n > 100$ from 3 independent transfections; one-way ANOVA followed by Dunnett's multiple comparisons test). Scale bars indicate 10 μ m

displayed ~ 30% lower *dhx30* expression as compared to wild type, potentially due to nonsense mediated decay of the mutated alleles (Additional file 10: Figure S7). The homozygous mutant animals are viable, fertile, and morphologically indistinguishable from their wild-type and heterozygous siblings (data not shown). Previous studies have demonstrated that during early embryonic development, zebrafish exhibit robust SG formation in response to stress, such as heat shock [41]. Therefore, based on our *in vitro* findings, we first asked whether *dhx30* mutant zebrafish also exhibit impaired SG formation *in vivo*. At 24-h post-fertilization and normal condition, compared to *dhx30*-WT the homozygous mutant exhibited significantly lower number of SGs, determined by

staining for TIAL-1, an established stress granule marker (Fig. 8). Although an increase in SG formation occurred upon heat shock, the number of TIAL-1-labeled SGs remained significantly lower in the homozygous mutants compared to sibling controls (Fig. 8). Thus, these data show that SG formation is compromised in the homozygous mutants and suggests an evolutionarily conserved role for DHX30 in SG assembly.

Dhx30-deficient zebrafish display altered behavioral activity

We next examined whether *dhx30*-deficient zebrafish exhibit abnormal sleep-wake activity and social behaviors, similar to those recently observed in a zebrafish

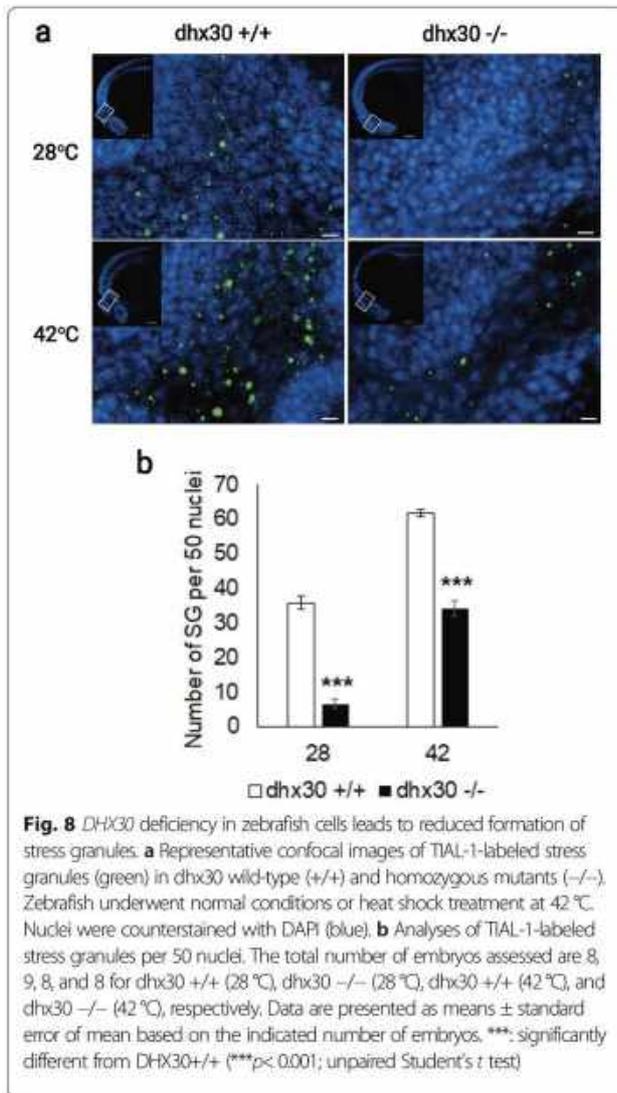


model of the NR3C2-related neurodevelopmental disorder [25]. We first analyzed sleep-wake behaviors in 5-day-old *dhx30* KO mutants. Compared to wild-type and heterozygous siblings, the homozygous mutants displayed significantly less activity during the day and more nocturnal activity (Fig. 9a–c), mimicking somewhat the sleep disturbances in individuals affected by a *DHX30*-related neurodevelopmental disorder. Additionally, using an established social preference assay, we observed that the wild-type and heterozygous animals showed the previously described social behavior of preferring to stay close to conspecific fish of similar age and size, whereas the homozygous animals did not show this preference (Fig. 9d–e). There were no obvious dysmorphic phenotypes in the homozygous mutant animals compared to their wild-type and heterozygous siblings. We propose, therefore, that the mutant phenotype was not simply due to developmental delay but influenced by abnormalities in complex neural circuitry. Taken together, our data indicate that *dhx30* KO zebrafish have a social behavioral deficit with altered sleep-wake activity, which is

consistent with findings in *DHX30*-related neurodevelopmental disorders.

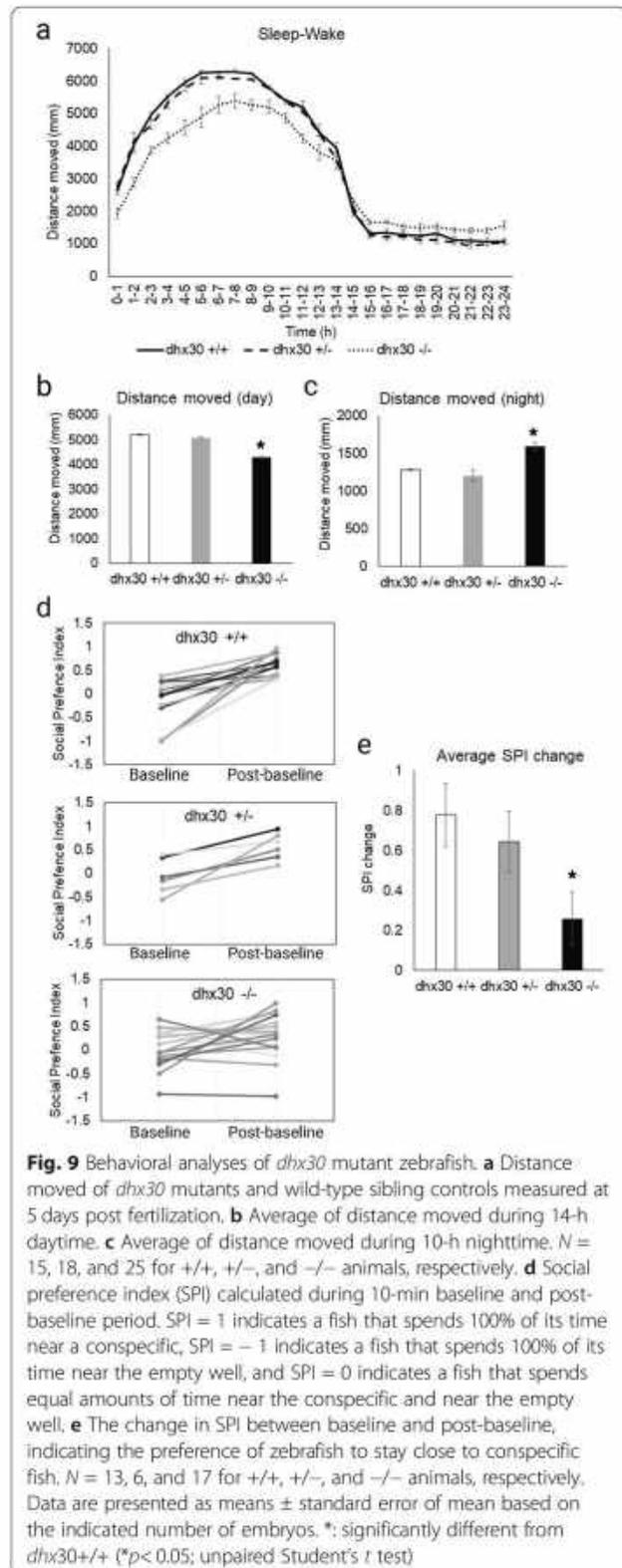
Discussion

Our study has allowed further delineation of the clinical spectrum of *DHX30*-related neurodevelopmental disorders through analysis of 25 novel affected individuals, partially facilitated by the use of a social media-based family support group. Individuals harboring heterozygous missense variants affecting highly conserved residues within a HCM present with global developmental delay, intellectual disability, muscular hypotonia, severe gait abnormalities (if walking is acquired), and remain non-verbal or speak only single words. We also identified microcephaly as an additional common feature. Individuals with either a mosaic missense variant within a HCM, or with variants resulting in haploinsufficiency or with protein-truncating variants all learned to walk in the second year of life, had a mild muscular hypotonia, and spoke at least 20 words by the age of 3 years. Therefore, based on the clinical and molecular findings we



selected HCM missense variants interfere with normal zebrafish embryonic development.

We have previously suggested that the missense variants within HCM might have a more severe effect than a loss of one gene copy [8]. This hypothesis is now supported by identification of four affected individuals carrying variants that result in either haploinsufficiency or a truncated protein, all of whom presented with a milder phenotype as compared to the individuals harboring missense variants within HCM. To gain further insight into the nature of these variants we determined that *DHX30*-WT can rescue the inability of selected HCM missense variants to unwind an RNA duplex, and that co-injection of *DHX30*-WT with selected HCM missense variants can partially ameliorate the observed zebrafish phenotypes. These data point to loss-of-function effects of the HCM mutants on a molecular level. However, co-expression of HCM missense variants together



with the DHX30-WT resulted in recruitment of DHX30-WT into SG's, a finding that might possibly suggest a dominant negative effect. It is worth noting that DHX30-WT, as well as the endogenous protein, are recruited to the SG's after stress induction [8]. Thus, HCM missense variants might actually result in a detrimental gain-of-function by inducing SG formation with concomitant global translation impairment even without endogenous or exogenous stressors.

To gain further clarity we focused on the relation of DHX30 to SG formation. Using CRISPR/Cas9 based technology, we established two *DHX30* knockout models. Analyses of both, *DHX30*-deficient HEK293T cells and zebrafish, revealed an impairment of SG formation upon heat stress, pointing to an essential and evolutionary conserved role of DHX30 in SG assembly. These findings provide a molecular explanation for the above-mentioned phenotypic differences, as they strongly suggest that pathogenic missense HCM variants, in addition to the loss of ATPase or RNA-binding activity and with impaired helicase function, exert a selective gain-of-function by triggering SG formation. This is in line with our hypothesis that due to SG hyper-assembly these pathogenic variants generate a chronic condition of impaired translation [8]. Noteworthy, impaired translation due to aberrant SG formation is associated with a broad variety of neurodegenerative and neurodevelopmental diseases [8, 42]. Furthermore, repeat expansion underlying *C9orf72*-associated neurodegenerative disorders has recently been suggested to result in chronic cellular stress due to aberrant SG formation [43].

Beyond providing the molecular explanation for the genotype-phenotype correlation of these two subtypes we additionally performed in vivo behavioral modeling of zebrafish *dhx30* KO's. Zebrafish exhibit all the hallmarks of mammalian sleep by utilizing neurotransmitters known to coordinate sleep and wake states in humans [44]. Analysis of *dhx30*-deficient animals revealed a compromised sleep/wake behavior, as they were less active during the day but more active and slept less at night than *dhx30*-WT animals. This is partially reminiscent of the sleep disturbances observed in almost half of DHX30-affected individuals. Additionally, homozygous *dhx30* KO animals displayed altered social behavior as manifested by their performance in the social preference assay, e.g., showing reduced preference for conspecifics as compared to *dhx30*-WT zebrafish. The observed social behavioral deficits and altered sleep-wake activity are similar to the findings in zebrafish models of other neurodevelopmental disorders [25, 45], and to some extent recapitulate the clinical findings in individuals affected by the *DHX30*-related neurodevelopmental disorder.

Furthermore, we present here two individuals who clearly stand out both in terms of their clinical presentation and their identified *DHX30* variant. Individual 4 with an early-lethal infantile epileptic encephalopathy carries a homozygous missense variant, p.(Arg725His), and individual 21 with a de novo p.(Arg908Gln) variant shows late-onset progressive ataxia. Trio-WES analysis performed in both individuals identified these *DHX30* variants as the only candidates (Supplementary Data). As

Table 2 Summary of functional analyses of missense variants

<i>DHX30</i> variant	p.(Gly462Glu), p.(His562Arg), p.(Ala734Asp), p.(Ser737Phe), p.(Thr739Ala), p.(Gly781Asp) p.(Arg782Gln) p.(Arg782Trp), p.(Arg785Cys), p.(Arg785His)	p.(Arg493His)	p.(Arg725His)	p.(Arg908Gln)	p.(Val556Ile)	p.(Glu948Lys)
Location in DHX30	Helicase core motifs I, II, V, or VI (nucleotide-interacting motifs)	Helicase core motif Ia (nucleic acid-binding)	Helicase core region, between motifs IV and V	Ratchet-like domain	Helicase core region, between motifs Ib and II	C-terminal region
gnomAD v2.1.1	Not identified	Not identified	Not identified	Not identified	0/39/282352	1/49/282090
ATPase activity	Reduced	Similar to wt*	Reduced	Reduced	Similar to wt	Similar to wt
RNA binding capacity	n.d.	Reduced*	n.d.	n.d.	n.d.	n.d.
Helicase activity	Reduced**	Reduced	n.d.***	Similar to wt	n.d.	n.d.
Cellular localization	Stress granules	Stress granules	Cytoplasmic, similar to wt	Cytoplasmic aggregates	Cytoplasmic, similar to wt	Cytoplasmic, similar to wt
Puromycin incorporation	Impaired	Impaired*	Similar to wt	Impaired	n.d.	n.d.
Zebrafish development	Impaired**	Impaired	Impaired	Impaired	Similar to wt	Similar to wt

n.d., not determined; *, Lessel et al. 2017; **, only selected variants analyzed; ***, unable to purify the protein

these variants occurred outside the HCM motifs, we included two similarly located common non-synonymous *DHX30* variants found in gnomAD, p.(Val556Ile) and p.(Glu948Lys), in our functional analysis for comparison. However, these two latter variants behaved similarly to *DHX30*-WT in all assays performed (Table 2).

For the variant p.(Arg725His), located within the helicase core region but not within a HCM, we observed a reduced ATPase activity. However, unlike HCM missense variants it does not trigger SG hyper-assembly. When attempting to analyze its impact on the helicase activity we consistently failed to purify the p.(Arg725His) mutant protein product. We therefore suggest that this biallelic variant leads to a loss-of-function, likely due to misfolding. The fact that it was inherited from unaffected heterozygous parents indicates that its effect is somewhat milder as compared to the variants identified here resulting in haploinsufficiency or protein-truncation, and that similar heterozygous missense variants may not contribute to disease.

The *de novo* missense variant p.(Arg908Gln) affects a highly conserved residue within the RL domain. This variant impairs the ATPase but not helicase activity of *DHX30*, suggesting that the RL domain is required for the coupling of helicase activity to ATP hydrolysis. Whereas this variant leads to formation of aberrant cytoplasmic aggregates, which cannot be eliminated by co-expression of *DHX30*-WT, not all of these aggregates/foci could be confirmed to be translationally silent SGs.

The functional characterization of both variants identified differences to HCM missense variants, which may potentially explain the genotype-phenotype correlation. Additionally, both individuals presented with clinical signs and symptoms observed in other affected individuals, suggestive of a phenotypic continuum. We cannot exclude, however, the possibility that they carry additional variants with phenotypic consequences, which were undetected by trio-whole exome sequencing. Identification of similarly affected individuals carrying similar variants is required to establish their causality.

Conclusions

The identification of 25 affected individuals has expanded the clinical and genetic spectrum of the *DHX30*-associated neurodevelopmental disorder. Our data suggest the existence of clinically distinct subtypes correlating with location and nature of pathogenic variants. Our study highlights the usefulness of social media-based family support groups as a resource in defining ultra-rare disorders as well as the need for *in-depth* functional characterization of potentially pathogenic variants to understand their biological consequences. We confirmed that *DHX30* is an ATP-dependent RNA helicase, and showed that *DHX30* is essential for stress granule

assembly in cellular and *in vivo* models. Missense variants in helicase core motifs lead to a loss of ATPase and helicase activity, concomitant with a gain-of-function with respect to SG formation, and a severe phenotype. In contrast, *DHX30* loss-of-function variants are associated with a milder phenotype. Additional studies are required to further delineate the variety of clinical outcomes underlying different *DHX30* variants as well as the roles of *DHX30* in various aspects of RNA metabolism.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-021-00900-3>.

Additional file 1. Supplementary methods.

Additional file 2: Table S1. Summary on clinical features of individuals bearing pathogenic *DHX30* variants.

Additional file 3: Figure S1. Identified missense variants affect highly conserved amino acids.

Additional file 4: Figure S2. *De novo* mosaicism in individual 6.

Additional file 5: Figure S3. Whole gene deletion in individual 24.

Additional file 6. Clinical reports of here presented individuals.

Additional file 7: Figure S4. *DHX30* WT acts as an ATP-dependent RNA helicase.

Additional file 8: Figure S5. Recombinant protein variants of *DHX30* induce the formation of cytoplasmic clusters.

Additional file 9: Figure S6. Representative images of zebrafish embryos.

Additional file 10: Figure S7. Generation of zebrafish CRISPR-Cas9-mediated *dhx30* stable knockout line.

Acknowledgements

We thank all affected individuals and their family members/legal guardians for their participation and collaboration, Hans-Hinrich Hönck (Institute for Human Genetics, UKE Hamburg) for technical assistance, and UKE microscopic imaging facility (umif) for providing assistance with confocal microscopes.

Authors' contributions

IM, N.D.P.D., H.H., J.W., J.M.P., U.F., N.C.Y., H.-J.K., and D.L. generated and analyzed the functional data. Zebrafish experiments were designed and performed in the laboratory of N.C.Y., D.L. and H.-J.K. supervised the study. D.L. wrote the manuscript. J.B.M., J.A., T.A., S.B., G.B., D.B., A.B., P.J.B., S.B., T.B., F.B., L.A.B., G.J.B., Ø.L.B., J.C., J.D., L.F.E., C.E., J.F., D.G., C.A.H., M.H., Y.H.-M., G.H., A.J., L.K., B.K., C.K., C.K., C.K., G.L.G., U.W.L., L.M.B., J.A.M.-A., M.M., D.T.M., K.Q.M., B.M., C.N., S.F.N., T.P., F.R., H.R., S.F.R., J.S.-H., P.B.S., A.S., S.S., A.P.A.S., K.To., K.T.v., J.H.W., C.Z., K.M., J.J., and F.Q.-R. identified and collected affected individuals. All authors read and approved the final manuscript.

Funding

This work was funded in part by Werner Otto Stiftung (to D.L. and H.-J.K.) and Deutsche Forschungsgemeinschaft (LE4223/1-1 to D.L.; Kr1321/9-1 to H.-J.K.), by startup funds from University of Alabama, Birmingham (to N.C.Y.), by NIH U54 OD030167 (to J.P.M.), by the UCLA Pathology Translational Research Fund (to J.B.M. and F.Q.-R.) and by the UCLA California Center for Rare Diseases (to S.F.N.). Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The raw next-generation sequencing and microarray-based comparative genomic hybridization data that support the findings in affected individuals cannot be made publicly available for reasons of patient confidentiality. Qualified researchers may apply for access to these data, pending

institutional review board approval (contact D.L., d.lesse@uke.de). Cells (contact H.-J.K., kreienkamp@uke.de) and zebrafish (contact N.C.Y., nyeo@uab.edu) are available upon signing a material transfer agreement. All other data generated or analyzed during this study are included in the main text and/or the additional files.

The newly identified *DHX30* variants have been deposited to the Leiden Open (source) Variation Database (LOVD) [46] (<https://databases.lovd.nl/shared/variants/DHX30/unique>) with the following variant numbers #0000763353 to #0000763362:

<https://databases.lovd.nl/shared/variants/0000763353> [47]

<https://databases.lovd.nl/shared/variants/0000763354> [48]

<https://databases.lovd.nl/shared/variants/0000763355> [49]

<https://databases.lovd.nl/shared/variants/0000763356> [50]

<https://databases.lovd.nl/shared/variants/0000763357> [51]

<https://databases.lovd.nl/shared/variants/0000763358> [52]

<https://databases.lovd.nl/shared/variants/0000763359> [53]

<https://databases.lovd.nl/shared/variants/0000763360> [54]

<https://databases.lovd.nl/shared/variants/0000763361> [55]

<https://databases.lovd.nl/shared/variants/0000763362> [56]

Declarations

Ethics approval and consent to participate

The study was performed in accordance with the Declaration of Helsinki protocols. Written informed consent for all 25 subjects was obtained from the parents or legal guardians in accordance with protocols approved by the Ethics Committee of the Hamburg Chamber of Physicians: PV 3802 and the University of California, Los Angeles (UCLA) IRB: 11-001087. Zebrafish were maintained according to protocols by the University of Alabama Zebrafish Research Facility (ZRF) Animal Resources Program which maintains full Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC) accreditation and is assured with the Office of Laboratory Animal Welfare (OLAW). All zebrafish studies followed protocols approved by the University of Alabama Institutional Animal Care and Use Committee (IACUC): APN22158.

Consent for publication

A written consent was obtained from the parents or legal guardians to publish the details of all 25 affected individuals.

Competing interests

K.M. and J.J. are employees of GeneDx, Inc. The remaining authors declare that they have no competing interests.

Author details

¹Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany. ²Department of Pharmacology and Toxicology, University of Alabama, Birmingham, USA. ³Department of Biochemistry, Theodor Boveri Institute, Biocenter of the University of Würzburg, 97070 Würzburg, Germany. ⁴Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. ⁵UCLA Clinical Genomics Center, University of California Los Angeles, Los Angeles, CA, USA. ⁶Department of Physiology, University of California Los Angeles, Los Angeles, CA, USA. ⁷Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Health Innovation Manchester, Manchester, UK. ⁸Division of Evolution & Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. ⁹Faculty of Medicine and Health Sciences, University of Western Australia, Perth, WA, Australia. ¹⁰Western Australian Register of Developmental Anomalies, King Edward Memorial Hospital, Perth, Australia. ¹¹Telethon Kids Institute, Perth, Australia. ¹²Division of Child Neurology, Department of Neurology, University of Rochester School of Medicine, Rochester, NY, USA. ¹³Clinical Genetics Department, University Hospitals Bristol and Weston, Bristol, UK. ¹⁴Joe DiMaggio Children's Hospital, Hollywood, FL, USA. ¹⁵Department of Medical Genetics, Haukeland University Hospital, 5021 Bergen, Norway. ¹⁶Department of Medical Genetics, Centre Hospitalier Universitaire de Poitiers, Poitiers, France. ¹⁷Laboratoire de Neurosciences Cliniques et Expérimentales-INSERM U1084, Université de Poitiers, Poitiers, France. ¹⁸Department of Clinical Medicine (K1), University of Bergen, Bergen, Norway. ¹⁹Department of Neurology, Haukeland University Hospital, Bergen,

Norway. ²⁰Department of Medical Genetics, Telemark Hospital Trust, Skien, Norway. ²¹Division of Medical Genetics, Department of Pediatrics, University of California San Francisco, San Francisco, CA, USA. ²²Department of Pediatrics, University Medical Center Eppendorf, 20246 Hamburg, Germany. ²³Peyton Manning Children's Hospital, Ascension Health, Indianapolis, IN, USA. ²⁴Department of Pediatrics, Southern Illinois University School of Medicine, Springfield, IL 62702, USA. ²⁵Department of Pediatric Neurology, Amalia Children's Hospital and Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands. ²⁶Division of Neurology, Department of Pediatrics, Ann and Robert H. Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ²⁷Kaiser Permanente Sacramento, Sacramento, USA. ²⁸Département de Génétique, Hôpital La Pitié-Salpêtrière, Assistance Publique-Hôpitaux de Paris, Paris, France. ²⁹Genetic Services of Western Australia, Perth, Western Australia 6008, Australia. ³⁰Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany. ³¹Department of Pediatrics, Vestfold Hospital, 3116 Tønsberg, Norway. ³²Department of Genetics, Kaiser Permanente Northern California, Oakland, USA. ³³Semel Institute of Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA. ³⁴Department of Pediatrics, Division of Medical Genetics at David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. ³⁵Department of Human Genetics at David Geffen School of Medicine University of California Los Angeles, Los Angeles, CA, USA. ³⁶Department of Medicine, Hugh Kaul Precision Medicine Institute, University of Alabama at Birmingham, 510 20th St S, Birmingham, AL 35210, USA. ³⁷Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA, USA. ³⁸Center for Duchenne Muscular Dystrophy, University of California Los Angeles, Los Angeles, CA, USA. ³⁹UF de Génétique Médicale, GHSR, CHU de La Réunion, Saint Pierre, La Réunion, France. ⁴⁰Department of Human Genetics, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands. ⁴¹Department of Neurology at David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. ⁴²Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands. ⁴³Department of Neurology, Boston Children's Hospital, Boston, MA, USA. ⁴⁴Department of Human Genetics, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, Switzerland. ⁴⁵GeneDx, Gaithersburg, MD 20877, USA. ⁴⁶Department of Pathology and Laboratory Medicine, School of Medicine, University of California Irvine, Irvine, CA, USA.

Received: 2 September 2020 Accepted: 28 April 2021

Published online: 21 May 2021

References

- Jankowsky E. RNA helicases at work binding and rearranging. *Trends Biochem Sci.* 2011;36(1):19–29. <https://doi.org/10.1016/j.tibs.2010.07.008>.
- Linder P, Jankowsky E. From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol.* 2011;12(8):505–16. <https://doi.org/10.1038/nrm3154>.
- Umate P, Tuteja N, Tuteja R. Genome-wide comprehensive analysis of human helicases. *Commun Integr Biol.* 2011;4(1):118–37. <https://doi.org/10.4161/cib.13844>.
- Heerma van Voss MR, van Diest PJ, Raman V. Targeting RNA helicases in cancer: The translation trap. *Biochim Biophys Acta Rev Cancer.* 2017;1868(2): 510–20. <https://doi.org/10.1016/j.bbcan.2017.09.006>.
- Cai W, Xiong Chen Z, Rane G, Satendra Singh S, Choo Z, Wang C, et al. Wanted DEAD/H or Alive: Helicases Winding Up in Cancers. *J Natl Cancer Inst.* 2017;109(6):djw278. <https://doi.org/10.1093/jnci/djw278>.
- Snijders Blok L, Madsen E, Juusola J, Gilissen C, Baralle D, Reijnders MR, et al. Mutations in *DDX3X* Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am J Hum Genet.* 2015;97(2):343–52. <https://doi.org/10.1016/j.ajhg.2015.07.004>.
- Balak C, Benard M, Schaefer E, Iqbal S, Ramsey K, Ernoult-Lange M, et al. Rare De Novo Missense Variants in RNA Helicase *DDX6* Cause Intellectual Disability and Dysmorphic Features and Lead to P-Body Defects and RNA Dysregulation. *Am J Hum Genet.* 2019;105(3):509–25. <https://doi.org/10.1016/j.ajhg.2019.07.010>.
- Lesse D, Schob C, Kury S, Reijnders MRF, Harel T, Eldomery MK, et al. De Novo Missense Mutations in *DHX30* Impair Global Translation and Cause a Neurodevelopmental Disorder. *Am J Hum Genet.* 2017;101(5):716–24. <https://doi.org/10.1016/j.ajhg.2017.09.014>.

9. Shamseldin HE, Rajab A, Alhashem A, Shaheen R, Al-Shidi T, Alamro R, et al. Mutations in DDX59 implicate RNA helicase in the pathogenesis of orofacioidigital syndrome. *Am J Hum Genet.* 2013;93(3):555–60. <https://doi.org/10.1016/j.ajhg.2013.07.012>.
10. Paine I, Posey JE, Grochowski CM, Jhangiani SN, Rosenheck S, Kleyner R, et al. Paralog Studies Augment Gene Discovery: DDX and DHX Genes. *Am J Hum Genet.* 2019;105(2):302–16. <https://doi.org/10.1016/j.ajhg.2019.06.001>.
11. Cross LA, McWalter K, Keller-Ramey J, Henderson LB, Amudhavalli SM. A report of gonadal mosaicism in DHX30-related neurodevelopmental disorder. *Clin Dysmorphol.* 2020;29(3):161–4. <https://doi.org/10.1097/MCD.0000000000000316>.
12. Lessel D, Zeitler DM, Reijnders MRF, Kazantsev A, Hassani Nia F, Bartholomaeus A, et al. Germline AGO2 mutations impair RNA interference and human neurological development. *Nat Commun.* 2020;11(1):5797. <https://doi.org/10.1038/s41467-020-19572-5>.
13. Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA.* 2014;312(18):1880–7. <https://doi.org/10.1001/jama.2014.14604>.
14. Lessel D, Gehbauer C, Bramswig NC, Schluth-Bolard C, Venkataramanappa S, van Gassen KLI, et al. BCL11B mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain.* 2018;141(8):2299–311. <https://doi.org/10.1093/brain/awy173>.
15. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med.* 2016;18(7):696–704. <https://doi.org/10.1038/gim.2015.148>.
16. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. *Nature.* 2017;542(7642):433–8. <https://doi.org/10.1038/nature21062>.
17. Sadedin SP, Dashnow H, James PA, Bahlo M, Bauer DC, Lonie A, et al. Cpipe: a shared variant detection pipeline designed for diagnostic settings. *Genome Med.* 2015;7(1):68. <https://doi.org/10.1186/s13073-015-0191-x>.
18. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405–24. <https://doi.org/10.1038/gim.2015.30>.
19. Ji J, Lee H, Argiropoulos B, Dorrani N, Mann J, Martinez-Agosto JA, et al. DYRK1A haploinsufficiency causes a new recognizable syndrome with microcephaly, intellectual disability, speech impairment, and distinct facies. *Eur J Hum Genet.* 2015;23(11):1473–81. <https://doi.org/10.1038/ejhg.2015.71>.
20. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST, Working Group of the American College of Medical Genetics Laboratory Quality Assurance C. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med.* 2011;13(7):680–5. <https://doi.org/10.1097/GIM.0b013e3182217a3a>.
21. Sobreira N, Schiettecatte F, Boehm C, Valle D, Hamosh A. New tools for Mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. *Hum Mutat.* 2015;36(4):425–31. <https://doi.org/10.1002/humu.22769>.
22. Guenther UP, Handoko L, Lagerbauer B, Jablonka S, Chari A, Alzheimer M, et al. IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1). *Hum Mol Genet.* 2009;18(7):1288–300. <https://doi.org/10.1093/hmg/ddp028>.
23. Tseng-Rogenski SS, Chang TH. RNA unwinding assay for DExD/H-box RNA helicases. *Methods Mol Biol.* 2004;257:93–102. <https://doi.org/10.1385/1-592-59-750-5.093>.
24. Thomas HR, Percival SM, Yoder BK, Parant JM. High-throughput genome editing and phenotyping facilitated by high resolution melting curve analysis. *PLoS One.* 2014;9(12):e114632. <https://doi.org/10.1371/journal.pone.0114632>.
25. Ruzzo EK, Perez-Cano L, Jung JY, Wang LK, Kashef-Haghighi D, Hartl C, et al. Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. *Cell.* 2019;178(4):850–66 e826. <https://doi.org/10.1016/j.cell.2019.07.015>.
26. Tanner NK, Linder P. DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell.* 2001;8:251–62.
27. Caruthers JM, McKay DB. Helicase structure and mechanism. *Curr Opin Struct Biol.* 2002;12(1):123–33. [https://doi.org/10.1016/S0959-440X\(02\)00298-1](https://doi.org/10.1016/S0959-440X(02)00298-1).
28. Buttner K, Nehring S, Hopfner KP. Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol.* 2007;14(7):647–52. <https://doi.org/10.1038/nsmb1246>.
29. Tauchert MJ, Fournann JB, Luhrmann R, Ficner R. Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *Elife.* 2017;6. <https://doi.org/10.7554/eLife.21510>.
30. Prabu JR, Muller M, Thoma AW, Schussler S, Bonnae F, Becker PB, et al. Structure of the RNA Helicase MLE Reveals the Molecular Mechanisms for Uridine Specificity and RNA-ATP Coupling. *Mol Cell.* 2015;60(3):487–99. <https://doi.org/10.1016/j.molcel.2015.10.011>.
31. Murakami K, Nakano K, Shimizu T, Ohto U. The crystal structure of human DEAH-box RNA helicase 15 reveals a domain organization of the mammalian DEAH/RHA family. *Acta Crystallogr F Struct Biol Commun.* 2017;73(6):347–55. <https://doi.org/10.1107/S2053230X17007336>.
32. Luscan A, Laurendeau I, Malan V, Francannet C, Odent S, Giuliano F, et al. Mutations in SETD2 cause a novel overgrowth condition. *J Med Genet.* 2014;51(8):512–7. <https://doi.org/10.1136/jmedgenet-2014-102402>.
33. Lumish HS, Wynn J, Devinsky O, Chung WK. Brief Report: SETD2 Mutation in a Child with Autism, Intellectual Disabilities and Epilepsy. *J Autism Dev Disord.* 2015;45(11):3764–70. <https://doi.org/10.1007/s10803-015-2484-8>.
34. O'Roak BJ, Vives L, Fu W, Egerton JD, Stanaway IB, Phelps IG, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science.* 2012;338(6114):1619–22. <https://doi.org/10.1126/science.1227764>.
35. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581(7809):434–43. <https://doi.org/10.1038/s41586-020-2308-7>.
36. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.* 2013;9(8):e1003709. <https://doi.org/10.1371/journal.pgen.1003709>.
37. Hempel M, Cremer K, Ockeloen CW, Lichtenbelt KD, Herkert JC, Denecke J, et al. De Novo Mutations in CHAMP1 Cause Intellectual Disability with Severe Speech Impairment. *Am J Hum Genet.* 2015;97(3):493–500. <https://doi.org/10.1016/j.ajhg.2015.08.003>.
38. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol.* 2004;22(11):1399–408. <https://doi.org/10.1038/nbt1029>.
39. Davis EE, Frangakis S, Katsanis N. Interpreting human genetic variation with in vivo zebrafish assays. *Biochim Biophys Acta.* 1842:2014:1960–70.
40. Finckbeiner S, Ko PJ, Carrington B, Sood R, Gross K, Dolnick B, et al. Transient knockdown and overexpression reveal a developmental role for the zebrafish enosf1b gene. *Cell Biosci.* 2011;1:32.
41. Zampedri C, Tinoco-Cuellar M, Carrillo-Rosas S, Diaz-Tellez A, Ramos-Balderas JL, Pelegri F, et al. Zebrafish P54 RNA helicases are cytoplasmic granule residents that are required for development and stress resilience. *Biol Open.* 2016;5(10):1473–84. <https://doi.org/10.1242/bio.015826>.
42. Wolozin B, Ivanov P. Stress granules and neurodegeneration. *Nat Rev Neurosci.* 2019;20(11):649–66. <https://doi.org/10.1038/s41583-019-0222-5>.
43. Zhang YJ, Gendron TF, Ebbert MTW, O'Raw AD, Yue M, Jansen-West K, et al. Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nat Med.* 2018;24(8):1136–42. <https://doi.org/10.1038/s41591-018-0071-1>.
44. Sorribes A, Thornorsteinsson H, Arnardottir H, Johannesdottir I, Sigurgeirsson B, de Polavieja GG, et al. The ontogeny of sleep-wake cycles in zebrafish: a comparison to humans. *Front Neural Circuits.* 2013;7:178.
45. Patowary A, Won SY, Oh SJ, Nesbitt RR, Archer M, Nickerson D, et al. Family-based exome sequencing and case-control analysis implicate CEP41 as an ASD gene. *Transl Psychiatry.* 2019;9(1):4. <https://doi.org/10.1038/s41398-018-0343-z>.
46. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v2.0: the next-generation in gene variant databases. *Hum Mutat.* 2011;32(5):557–63. <https://doi.org/10.1002/humu.21438>.
47. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borzi C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuur-Hoijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kriekamp H-J, Lessel D. NM_138615.2(DHX30): c.1385G>A (p.

- Gly462Glu). Variant #0000763353. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763353>.
48. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.2174G>A (p. Arg725His). Variant #0000763354. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763354>.
49. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.2217C>A (p. Ala734Asp). Variant #0000763355. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763355>.
50. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.2215A>G (p. Thr739Ala). Variant #0000763356. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763356>.
51. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.2345G>A (p. Arg782Gln). Variant #0000763357. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763357>.
52. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.2723G>A (p. Arg908Gln). Variant #0000763358. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763358>.
53. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.347_360del (p. Ala116Valfs*12). Variant #0000763359. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763359>.
54. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): c.2389C>T (p. Arg797*). Variant #0000763360. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763360>.
55. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): g.47098509_48109065del. Variant #0000763361. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763361>.
56. Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, L. Busk Ø, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp H-J, Lessel D. NM_138615.2(DHX30): g.47882366_47884746del. Variant #0000763362. LOVD 3.0. <https://databases.lovd.nl/shared/variants/0000763362>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Additional information for: Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders

Mannucci et al.

Additional file 1: Supplementary methods.

Expression of 6xHis-SUMO-DHX30 constructs

6xHis-SUMO-DHX30 wild-type and mutant constructs were transformed into *E. coli* BL21 (DE3) pLysS pRARE competent cells. Colonies were inoculated into 100 mL of superbroth (SB) medium supplemented with 2% glucose, 0.1% kanamycin and 0.1% chloramphenicol and incubated overnight at 37°C under shaking at 200 rpm. The 100 ml bacterial pre-culture was transferred into 2 L of SB medium and incubated for 3h at 37°C under shaking at 180 rpm until an OD₆₀₀=1 was reached. Then, protein expression was induced by adding 1 mM of IPTG and cells were cultured overnight at 14°C under shaking at 180 rpm.

Cells were collected by centrifugation and cell pellets were resuspended in 50 mL of lysis buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 5 mM MgCl₂, 0.01% NP-40, 10% glycerol, 5 mM β-mercaptoethanol, 0.1 mM AEBSF, 0.5 mg/L leupeptin/pepstatin A, 2 mg/L aprotinin). Solubilized samples were additionally sonified with the Brandson 250 sonifier. Insoluble cell debris was removed by ultracentrifugation with a 45Ti Beckman rotor at 25 000 rpm for 1h at 4°C.

Protein purification. Cell lysates were incubated for 2h at 4°C on Ni-NTA beads (Qiagen, #30210) previously equilibrated with lysis buffer. Then, beads were washed three times with 50 times bed volume of washing buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 5 mM MgCl₂, 0.01% NP-40, 10% glycerol, 40 mM Imidazole, 5 mM β-mercaptoethanol, 0.1 mM AEBSF, 0.5 mg/L leupeptin/pepstatin A, 2 mg/L aprotinin). Bound proteins were eluted with 8 times bed volume of elution buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 5 mM MgCl₂, 10% glycerol, 300 mM imidazole, 5 mM β-mercaptoethanol.). The elution fractions were subsequently analyzed for the presence of the target protein by SDS-PAGE and Coomassie blue staining. Afterwards the fractions of interest were pooled and sample was concentrated using concentrators (Vivaspin™ 500, MWCO 10 000). 2 - 5 mg of the concentrated sample was further purified by size exclusion chromatography (Superdex 75) with the Äkta™ purifier system (GE Healthcare). Fractions were analyzed for the presence of the target protein by SDS-PAGE and Coomassie staining. Samples were quantified from the Coomassie stained gel using the ImageJ software.

In vitro synthesis of RNA molecules. To test the RNA unwinding activity of DHX30, a ³²P-Labeled RNA duplex was synthesized using the T7 RNA polymerase from a linearized DNA template designed by (Tseng-Rogenski & Chang, 2004). The *in vitro* transcription reaction mix was prepared as follows: 1X transcription buffer (40 mM Tris-HCl pH 7.9, 1 mM Spermidine, 26 mM MgCl₂, 0,01% Triton X, 5mM DTT), NTPs (GTP 8 mM, ATP 5 mM, CTP 5 mM, UTP 2 mM, 50 μCi of ³²P-UTP), 3 μM DNA template, 3 μM top strand primer (5'-TAATACGACTCACTATAG-3'), 7 U of T7 RNA polymerase. Transcription reactions were incubated for 2h at 37° C. RNA was precipitated by adding 0.1 volumes of 3M NaAc (pH 5.5) and 3 volumes of ethanol for 30 min at -20°C. Then, samples were centrifuged at 13 000 g, 30 min at 4°C. RNA pellets were rinsed with 70 % ethanol and resuspended in H₂O.

RNA samples were mixed with 2X denaturing RNA loading dye (1X Tris-Borate-EDTA pH 8.3, 95% formamide, 0.1% bromophenol blue, 0.1% xylene cyanol FF) and analyzed on 8% UREA-PAGE in 1X TBE. Radioactive signals were detected by autoradiography and the RNA band was excised from the gel.

The RNA product was extracted from the gel in RNA extraction buffer (200 mM Tris-Hcl pH 7.0, 0.1% SDS, 1 mM EDTA). Subsequently, RNA was precipitated as described above and the RNA pellets were resuspended in 100 mM KCl.

To promote the formation of the RNA duplex, the RNA sample was boiled at 95°C for 5 min and cooled down over 2h. The sample was mixed with 2X non-denaturing loading dye (1X TBE, 20% glycerol, 0.1% bromophenol blue, 0.1% xylene cyanol FF) and separated on 8% native PAGE. The RNA duplex was excised from the gel and purified as described before. The RNA concentration was determined by measuring absorbance at 260 nm.

RNA unwinding assay

The helicase activity was measured in 20 μ l of reaction mixture containing 0.13 pmol of purified protein (=20 ng of full length protein), 25 fmol radioactively labeled RNA duplex, 17 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1 mM MgCl₂, 2 mM DTT, 1 mM spermidine, 0.3% PEG8000, 5% glycerol, 150 mM KCl, 20 units of RNasin™ Plus (Promega), 1 mM ATP. The mixture was incubated for 1h at 37°C and subsequently mixed with 2X non-denaturing loading dye and subjected to gel electrophoresis through non-denaturing 8% PAGE (19:1) in 0.5X TBE at 4°C. Reaction products were visualized by autoradiography.

Additional file 2: Table S1: Summary on clinical features of individuals bearing pathogenic DHX30 variants.

Clinical findings	Individual 1		Individual 2		Individual 3		Individual 4		Individual 5	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Sex										
Age at last examination (years)	6 6/12	4 years 2 months	23	10 months, deceased at 11 months					11	
Intellectual disability	+	+	+	+	+	+	+	+	+	+
Age of first words (years)	-	-	-	-	-	-	-	-	NA	NA
Speech ability	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	4 words	4 words
Motor development delay	+	+	+	+	+	+	+	+	+	+
Muscular hypotonia	+	+	+	+	+	+	+	+	+	+
Age of walking (years)	-	2y 6 m	4y 6m	too young to evaluate	too young to evaluate	too young to evaluate	too young to evaluate	too young to evaluate	-	-
Gait abnormalities	no independent walking	ataxic	ataxic	ataxic	ataxic	ataxic	ataxic	ataxic	no independent walking	no independent walking
Autistic features	-	-	Uncertain. Poor eye contact	-	Uncertain. Poor eye contact	-	Uncertain. Poor eye contact	-	+	+
Sleep disturbance	+	+	+	+	+	+	+	+	-	-
Seizures	-	-	No seizures ever noted, normal EEG; vacant episodes have been reported.	-	No seizures ever noted, normal EEG; vacant episodes have been reported.	-	No seizures ever noted, normal EEG; vacant episodes have been reported.	+	-	-
Feeding difficulties	+	+	+	+	+	+	+	+	+	+
Strabismus	+	-	-	-	-	-	-	-	+	+
Joint hypermobility	+	-	+	-	+	-	+	-	+	+
Microcephaly	+	+	NA	+	NA	+	+	-	+	+
Cerebral MRI anomalies	Mostly frontal cerebral atrophy	-	Abnormal white matter change around the posterior horns of both lateral ventricles.	-	Abnormal white matter change around the posterior horns of both lateral ventricles.	-	Abnormal white matter change around the posterior horns of both lateral ventricles.	central inflammatory brain lesions	cortical and subcortical atrophy and delayed myelination	
DHX30 alteration (NM_138615.2)	c.1385G>A	c.1478G>A	c.1685A>G	c.1478G>A	c.1685A>G	c.1478G>A	c.1685A>G	c.2174G>A	c.2201C>A	
Mode of inheritance	p.(Gly462Glu)	p.(Arg493His)	p.(His562Arg)	p.(Arg493His)	p.(His562Arg)	p.(Arg493His)	p.(His562Arg)	p.(Arg725His)	p.(Ala734Asp)	
gnomAD v2.1.1	de novo	unknown (heterozygous)	unknown (heterozygous)	unknown (heterozygous)	unknown (heterozygous)	unknown (heterozygous)	unknown (heterozygous)	homozygous	de novo	
ACMG classification	not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified	
ACMG criteria	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	uncertain significance	pathogenic	
ACMG criteria	PS2, PM1, PM2, PP2, PP3	PS1, PS3, PM1, PM2, PP2	PS1, PS3, PM1, PM2, PP2	PS1, PS3, PM1, PM2, PP2	PS1, PS3, PM1, PM2, PP2	PS1, PS3, PM1, PM2, PP2	PS1, PS3, PM1, PM2, PP2	PM2, PP2	PS2, PM1, PM2, PP2, PP3	

(continued on next page)

Individual 6		Individual 7		Individual 8		Individual 9		Individual 10	
Female	Male	Female, half-sister of #9	Female, half-sister of #8	Female	Female	Female	Female	Female	Female
15	33	14	14	6	6	33 months	33 months	33 months	33 months
moderate	+	+	+	+	+	+	+	+	+
2 y	No words	5 y	5 y	-	-	30 months	30 months	30 months	30 months
simple sentence	non-verbal	single words	single words	non-verbal	non-verbal	two words	two words	two words	two words
+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+
2 y 8 m	12	10	10	-	-	-	-	-	-
ataxic	ataxic	unsteady	unsteady	no independent walking	no independent walking	no independent walking	no independent walking	no independent walking	no independent walking
+as toddler	+: interactive	-	-	-	-	+	+	+	+
+	+	-	-	-	-	+	+	+	+
+	-; single febrile seizure	+	+	No seizures ever noted, normal EEG; vacant episodes have been reported.	No seizures ever noted, normal EEG; vacant episodes have been reported.	+	+	+	+
+	+	-	-	+	+	+	+	+	+
-	+: L esotropia	-	-	-	-	Rotatory nystagmus	Rotatory nystagmus	Rotatory nystagmus	Rotatory nystagmus
+	+	-	-	-	-	+	+	+	+
+	+	+	+	+	+	NA	NA	NA	NA
-	mild diffuse prominence of the cerebrospinal fluid space and mild asymmetric prominence of right superior ophthalmic vein	ND	ND	partial agenesis of the corpus callosum	partial agenesis of the corpus callosum	delayed myelination, delayed corpus callosum development, and L hippocampal atrophy	delayed myelination, delayed corpus callosum development, and L hippocampal atrophy	delayed myelination, delayed corpus callosum development, and L hippocampal atrophy	delayed myelination, delayed corpus callosum development, and L hippocampal atrophy
c.2201C>A	c.2215A>G	c.2344C>T	c.2344C>T	c.2344C>T	c.2344C>T	c.2344C>T	c.2344C>T	c.2344C>T	c.2344C>T
p.(Ala734Asp)	p.(Thr739Ala)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)
de novo mosaic	unknown (heterozygous not maternally inherited)	unknown (heterozygous, possibly inherited from mosaic mother)	unknown (heterozygous, possibly inherited from mosaic mother)	unknown (heterozygous, possibly inherited from mosaic mother)	unknown (heterozygous, possibly inherited from mosaic mother)	de novo	de novo	de novo	de novo
not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified
pathogenic	likely pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic
PS2, PM1, PM2, PP2, PP3	PM1, PM2, PP2, PP3	PS1, PS3, PM1, PM2, PP2, PP3	PS1, PS3, PM1, PM2, PP2, PP3	PS1, PS3, PM1, PM2, PP2, PP3	PS1, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3

(continued on next page)

Individual 11		Individual 12		Individual 13		Individual 14		Individual 15	
Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
7	8	8	8	8	4	4	4	2 years 7 months	
+	+	+	+	+	+	+	+	+	+
-	-	-	4	4	-	-	-	-	-
non-verbal	non-verbal	non-verbal	single words	single words	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal
+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+
-	4	4	-	-	-	-	-	-	-
no independent walking	ataxic	ataxic	no independent walking						
N/A	+	+	-	-	-	-	-	N/A; always cheerful	-
-	+	+	-	-	-	-	-	-	-
+	-	-	-	-	-	-	-	-	-
+	+	+	+	+	+	+	+	+	+
When fatigued	+	+	-	-	-	-	-	+	+
+	+	+	+	+	+	+	+	-	-
+	-	-	+	+	+	+	+	+	+
delayed myelination	periventricular leukomalacia	periventricular leukomalacia	-	-	-	-	-	ND	ND
c.2344C>T	c.2344C>T	c.2344C>T	c.2345G>A						
p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Trp)	p.(Arg782Gln)						
<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>
not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified	not identified
pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic
PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PM1, PM2, PP2, PP3						

(continued on next page)

Individual 16		Individual 17		Individual 18		Individual 19		Individual 20	
Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
4	3	7 years 6 months	3	2y5m	16 y 10mo				
+	+	+	+	+	+	+	+	+	+
-	-	-	-	-	-	-	-	-	-
non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal	non-verbal
+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+
-	-	3	-	1y10m	20-21 months				
no independent walking	ataxic	ataxic	no independent walking	unsteady (now using braces)	ataxic				
Inconsistent eye contact	-	-	-	-	+				
-	+	+	-	+	+				
+	-	-	-	-	-				
-	+	+	+	-	+				
-	-	-	-	-	-				
+	-	+	+	+	+				
NA	-	-	+	-	+				
abnormal brain with diffuse abnormal subcortical and periventricular white matter signal abnormality in a symmetric pattern involving the temporal frontal and parietal lobes with associated abnormal cerebral volume loss	decreased in white matter volume		+(under opercularization of the sylvian fissures)						
c.2353 C>T	c.2353 C>T	c.2353 C>T	c.2353 C>T	c.2353 C>T	c.2354G>A				
p.(Arg785Cys)	p.(Arg785Cys)	p.(Arg785Cys)	p.(Arg785Cys)	p.(Arg785Cys)	p.(Arg785His)				
unknown (heterozygous)	de novo	de novo	de novo	de novo	de novo				
not identified	not identified	not identified	not identified	not identified	not identified				
pathogenic	pathogenic	pathogenic	pathogenic	pathogenic	pathogenic				
PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3	PS1, PS2, PS3, PM1, PM2, PP2, PP3				

(continued on next page)

Individual 21	Individual 22	Individual 23	Individual 24	Individual 25
Female	Female	Male	Male	Male
11 years and 4 months	5 years, 10 months	3	3	2
-	+	+	+	+
1 year	~1 year	1.5 years	1	17 months
normal speech ability, slurred speech	normal speech ability, slurred speech	20 words	100 words	25-30 words
-	mild	mild	mild	mild
-	-	+	+	+
1 year	1	16 months	1	18 months
ataxic gait, starting at the age of 8-9 years	unsteady	N/A	unsteady	ataxic
-	-	N/A	-	-
-	+	N/A	-	+
-	Non-epileptic absences	N/A	-	+
-	-	intermittent picky eating	-	-
+	Bilat intermittent exotropia	Left eye exotropia	-	-
-	-	N/A	-	+
-	-	-	-	-
Cerebellar atrophy, progressive	-	ND	T2 hyperintensities consistent with hypomyelination, Chiari type I malformation	ventriculomegaly with mild reduction of white matter
c.2723G>A	c.347_360del	c.2389C>T	arr[GRCCh37] chr3:(47098509_48109065)del	[GRCCh37] chr3:(47882366-47884746)del
p.(Arg908Gln)	p.(Ala116Valfs*12)	p.(Arg797*)	whole gene deletion	Exon 7-9 deletion
de novo	de novo	heterozygous, inherited from mosaic mother	assumed de novo	assumed de novo
not identified	not identified	not identified	not identified	not identified
likely pathogenic	likely pathogenic	likely pathogenic	likely pathogenic	likely pathogenic
PS2, PM1, PM2, PP2	PS2, PM2, PP3	PS2, PM2, PP3	PS2, PM2, PP3	PS2, PM2, PP3

**Additional information for:
Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-
associated neurodevelopmental disorders**

Mannucci *et al.*

Additional file 4

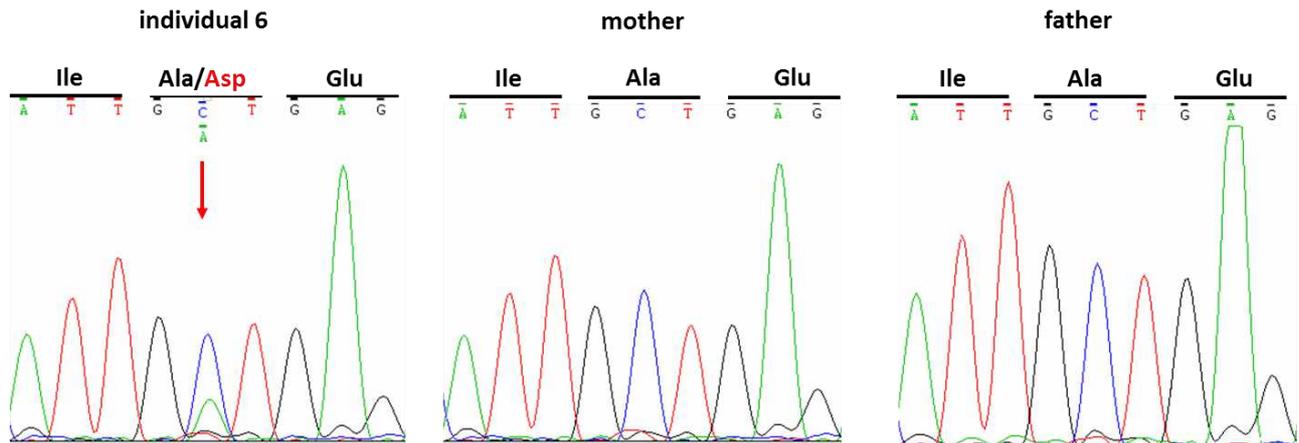


Fig. S2. *De novo* mosaicism in individual 6. Sanger sequence electropherograms of parts of *DHX30* after PCR amplification of genomic DNA of the affected individual 6 and her parents, confirming *de novo* mosaicism. The amino acid translation is shown in the three-letter code above the DNA sequence. The red arrow indicates the variant at c.2201C>A, p.(Ala734Asp) present only in the DNA sample of the affected individual.

**Additional information for:
Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-
associated neurodevelopmental disorders**

Mannucci *et al.*

Additional file 5

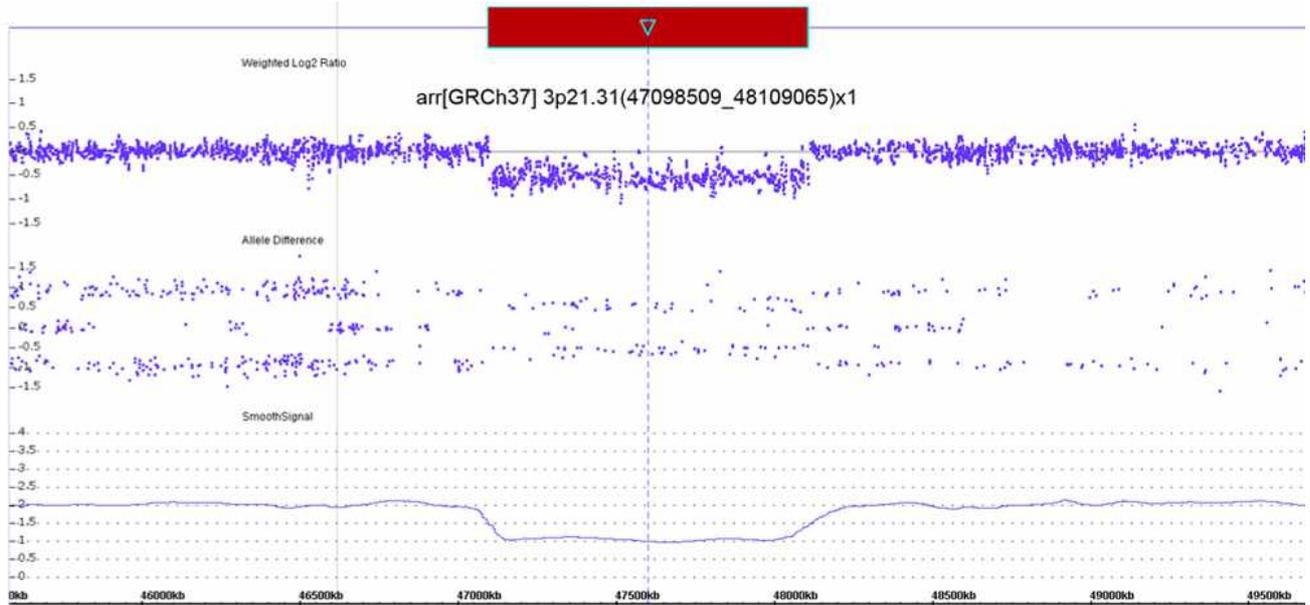


Fig. S3. Whole gene deletion in individual 24. Adapted from Chromosome Analysis Suite 3.3 (ChAS 3.3) showing loss of oligonucleotide probes at 3p21.31. Each dot represents one single nucleotide polymorphism that are distributed on the x-axis which shows the genomic positions.

Additional information for: Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders

Mannucci *et al.*

Additional file 6: Clinical reports:

Individual 1 is a 9-year-old female, the first child of non-consanguineous parents of Caucasian European ancestry. Her older maternal half-brother, mother and maternal grandmother were regarded to have familial hypermobility / Ehlers-Danlos syndrome. Otherwise family history was inconspicuous. Pregnancy and birth were uncomplicated. At 6 months of age her elbow joint was dislocated without any trauma. She could roll over spontaneously at 9 months of age. At 18 months she was able to sit on her knees with malrotated abducted lower legs. At 2 years of age she could stand with support. She was hypermobile and could fall asleep with her head between legs. She was referred for a genetic evaluation at 3 years of age due to muscular hypotonia, joint hypermobility, feeding difficulties, delayed milestones of development and an occipitofrontal head circumference (OFC) of 45 cm (-4.0SD), i.e. -3 cm below 2.5 percentile. Clinically she was suspicious of absences. However, EEG gave normal results and brain MRI revealed somewhat cerebral atrophy, increased subarachnoidal space and marked ventricles. At clinical investigation at 3.5 years of age she had an OFC of 45.5 cm (-3.9 SD), she had a marked muscular hypotonia, joint hypermobility, was non-verbal, had a convergent strabismus of 20 degrees, and a tendency to stand on her toes with support of parent. She started using a walker at the 5 6/12 years of age. At the last clinical examination at 6 years of age she had an OFC of 45.5 cm (-4.6 SD), was non-verbal, could stand alone and use a walker. Antiepileptics, Sodium valproic acid, had some effect on sleep and daily function. Conventional chromosome analysis of lymphocytes, Array Comparative Genomic Hybridization (array-CGH), Multiplex Ligation-dependent Probe Amplification (MLPA) for *ATRX*, Ehlers-Danlos type III&IV, Ehlers-Danlos type VI and Prader-Willi/Angelman, next-generation sequencing (NGS) panel for Ehlers-Danlos syndrome (9 genes), all gave normal results. Trio whole exome sequencing (trio-WES) with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.1385G>A, p.(Gly462Glu) in *DHX30* (NM_138615.2).

Individual 2 is a 4-year-old male, the second of two children of unaffected, non-consanguineous parents of African American ancestry. Family history was non-contributory. He was born at 38 6/7 weeks gestation by repeat cesarean to a 36 year old G2P1-2 mother weighing 2940gm (<1.5SD) with a length of 45.7cm (<2SD) and OFC of 34.3cm (<0.3SD). The pregnancy was complicated by maternal anemia. He had mild feeding issues after birth and was in the NICU for a total of 5 days and was then discharged home. He was born with post-axial polydactyly of both hands and one foot. All of the additional digits had a bony component, and were surgically removed at 3-months. Psychomotor development was severely delayed. Physical exam at 4 years old identified an acquired microcephaly, with an OFC of 47cm (<2.5SD). His length was 93.5cm (<2SD) and his weight was 12.2kg (<2.5SD). He had no significantly dysmorphic features. At 4 years old, he was able to babble, but had no specific words. He followed a few simple commands. He had a history of aspiration and required thickened liquids and mechanically soft foods. He could use a fork and a spoon, but would often over-stuff his mouth. He learned to walk at 2 ½ years, but did not yet run and exhibited ataxia and toe-walking. He could assist in dressing, but was not yet toilet trained. He had a normal brain MRI at 21 months of age. He has had no seizures. A chromosomal microarray at 19 months of age identified a maternally inherited 15q13 duplication which was considered to be a variant of unknown significance. He also had a normal Fragile X test at that time. WES at 3 years of age revealed a heterozygous variant c.1478G>A, p.(Arg493His) in *DHX30* (NM_138615.2).

Individual 3 is the first daughter of unaffected and non-consanguineous couple with unremarkable family history. She was born following a normal pregnancy with a birth weight of 3231 grams. There were no significant neonatal problems reported; and her parents feel that her early development was within normal limits for the first 6 months. Following this, she started displaying signs of developmental delay and muscular hypotonia, which has been understood as regression. She has also had a history of bronchomalacia, frequent upper respiratory tract infections and vomiting, and very poor eye contact. Her growth pattern was age appropriate. She was able to sit up at 1 year, crawled at 20 months, and learned to walk with 54 months. She presented with stereotypic hand movements; and due to these features, the diagnosis of Rett Syndrome was

considered at age 2 years. At the clinical examination at the age of 6 and a half years she was noted to have some distinctive features, such as thick hair, thick vermilion of lower and upper lips, deep set eyes. She also had generalized hypotonia and brisk reflexes on her lower limbs. She was making very slow developmental progresses, and never actually developed expressive speech. She had some swallowing difficulties, problems with chewing and history of severe regurgitation. Neuro-metabolic work-up has been within normal limits. A brain MRI revealed abnormal white matter changes in the posterior horns of both lateral ventricles. She had several EEG's, with no specific features. She did not suffer from epilepsy. Presently at 25 years, she no major medical problems. She has a friendly personality and is always smiling. She does not speak; hence she communicates using an iPad (iGaze), which is working very well for her. She still has some difficulties swallowing and is dependent for feeding. Her gait is unsteady. Her mother suspects a propensity for ear infections (she had several pseudomonas infections). Moreover, she was investigated for Postural tachycardia syndrome (PoTS), as she was suffering from tachycardia and sweets episodes; but a cardiac assessment gave normal results. Over the years, conventional chromosome analysis of lymphocytes, array-CGH, genetic testing for Rett syndrome, SMA, Pitt-Hopkins syndrome and myotonic dystrophy, chromosome 15 deletion for Prader-Willi Syndrome, Fragile-X testing, all gave normal results. Without a diagnosis, in 2013 she was recruited to the DDD study. WES revealed a heterozygous variant c.1685A>G, p.(His562Arg) in *DHX30* (NM_138615.2).

Individual 4 was a boy, the third child of unaffected, consanguineous parents born at 40 weeks of gestation, with a birth weight of 3220 grams (-0.9SD). Regarding family history, the child's father had one paternal and one maternal uncle with epilepsy. He was admitted to pediatric resuscitation unit because of seizures and myoclonia of the left hemibody, with consciousness disorders and a hypertonia of the left hemibody. The electroencephalogram (EEG) showed a slow sleep path slightly pathological. Then, there were repeated hospital admissions. Two weeks later, the clones in the lower left limb resolved. But at the age of 4 months, he again presented myoclonia. The EEG showed centrottemporal focal spots of spikes. The brain MRI revealed a diffusion restriction affecting the sub-cortical white substance of the left on the pre central convolution and the posterior part of the upper frontal convolution; associated with meningeal lesions evoking a localized inflammatory process. At the clinical examination at 5 months, he showed no eye contact, an axial hypotonia, with agitation between the crises. The EEG showed a very slow track with abundant diffuse wave spikes and a right rhythmic delta burst, and also a diffuse wave peaks, which predominate at the right and left centro-temporal regions. He was admitted to hospital for gastroenteritis at 6 months. At 10 months, he was admitted to the pediatric emergency room for increased seizures after a vaccination one week earlier, with onset of more frequent contact ruptures and clonus of the right upper limb for prolonged periods. Feeding difficulties were observed over the following days. Three weeks after, he was again hospitalized with decreased movements of the eyeballs; a much altered contact, a major axial hypotonia, and repeated opisthotonos attacks as well as an oral facial dyskinesia and myoclonias of the left hand which disappeared completely in deep sleep and increased at stimulation. The EEG showed a hypersarrhythmia traversed from spikes to type of repetitive paroxysms predominant on the right. He deceased at the age of 11 months. Standard metabolic screening gave normal results. Conventional chromosome analysis of lymphocytes, array-CGH, genetic testing of mtDNA for MELAS, MERRF and NARP, for eIF2B-related disorders, as well as analyses of *SCN1A* and *GABRG2* gave normal results. The trio-WES, with DNA samples of both unaffected parents and the proband, revealed a homozygous variant c. 2174G>A, p.(Arg725His) in *DHX30* (NM_138615.2). No other candidate variant was identified.

Individual 5 is an 11-year-old female of unaffected and non-consanguineous couple with unremarkable family history. She has 2 healthy brothers. She was born after an uncomplicated pregnancy and a normal delivery at 41 weeks and 3 days of gestation with a birth weight of 2930 g (-1.7 SD) and a birth height of 48.5 cm (-1 SD) and a OFC of 33.5 cm (-1.8 SD). Development was apparently normal up to 4 months with a smile response and the start of voluntary grasping before 3 months. Parents describe deterioration of her condition at around 4 months following an episode of chickenpox infection. A brain MRI at 11 months revealed cortical and subcortical atrophy, and delayed myelination. She later developed hand stereotypies of the middle line and autistic features. At the last clinical examination at the age of 11 years we saw a girl with some dysmorphic features including a short forehead, likely related to her microcephaly, arched eyebrows, a slight synophris, upslanted palpebral fissures, anteverted nares, short columella, and smooth philtrum. There is a delay in the fall of the lacteal teeth. Further, she has micrognathia and high palate. Skin is thin with an apparent venous

network and a slight hypertrichosis was noticed. She did not yet acquire autonomous walking, and is wheelchair-bound using it by advancing her arms and then both legs together. Fine motor skills are mostly absent. She can speak only 4 words: “mom” “dad” “yes” and “no”, but can communicate a few words in sign language. She had repeated pulmonary aspirations and prefers grinded food. She has a general hypotonia, significant hyperlaxity with genu recurvatum of the knees and pes valgus. There is no pyramidal or extrapyramidal syndrome. She developed postnatal microcephaly with -5 SD, a severe delay in weight (-3 SD) and height (-5 SD) gain. She developed strabismus, further ophthalmological and auditive examination gave normal results. Chromatography analysis revealed ammonemia and AICAR-SAICAR accumulation. Array-CGH and Fragile X analysis gave normal results. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2201C>A, p.(Ala734Asp) in *DHX30* (NM_138615.2) which was confirmed by Sanger sequencing.

Individual 6 is a 15-year-old female, the first of three children of healthy non-consanguineous parents of Aramean descent. The family history was unremarkable; in particular no developmental disorders have been reported and both younger siblings are healthy. The girl was born after an uneventful pregnancy at term with normal birth measurements (weight 3290 g (-0.5 SD), length 54 cm (1.0 SD), OFC 35 cm (0 SD)). Soon after birth she was admitted to the ICU because of cyanotic events and hypotonia; she was discharged after a few days. Hypotonia persisted, and she developed feeding difficulties and failure to thrive. A situation resembling Acute Life- Threatening Event (ALTE) occurred at the age of 3 months, and a unique febrile seizure followed at the age of 6 months. Motor and speech development were severely delayed; she started to walk without support at 2 years and 8 month, but gait was unstable and she fell down many times over the next years. She spoke first at 2 years and the language development was not appropriate for her age. Receptive language was much better. During toddlerhood, feeding difficulties switched to hyperphagia, she had no feeling of satiety. Thus, she gained weight and became obese (30.9 kg, +1.4 SD) at the age of 11 years. In addition she developed aggressive behavior with autistic features and sleep disturbances. She attended a school for children with special needs. Analysis of blood and urine showed slight and unspecific elevation of glycosaminoglycan's, but extensive metabolic work up did not detect any additional abnormal results. ECG, EEG, echocardiography, abdominal ultrasound, and brain MRI were unremarkable. Hearing test and eye examination gave normal results. At the last examination at the 15 years of age, we saw a friendly and shy young woman who answered simple questions with simple sentences. She loved to play football although her gait was still ataxic. She was able to read single letters and write single words. She was well- integrated in her family as well as at school; her parents reported that she helps with household chores. She has not shown aggressive and autistic behaviors anymore, but developed an anxiety disorder. Her height was 151.5 cm (-1.1 SD), weight was 49 kg (+0.1 SD) and OFC was microcephalic with 52 cm (-1.6 SD). Conventional chromosome analysis of lymphocytes gave normal results. Trio-WES with DNA samples of both unaffected parents and the proband, revealed mosaicism for a *de novo* heterozygous variant c.2201C>A:p.(Ala734Asp) in *DHX30* (NM_138615.2), identified in 36 out of 174 reads (21%). The *de novo* occurrence of the mosaicism was confirmed by Sanger sequencing.

Individual 7 is a 33-year old male who was born by normal delivery at term following an uneventful pregnancy. The family history was negative for similarly affected individuals and his non-consanguineous parents were of European and Hispanic ethnicity. He developed normally until 8 months of age, when he started to regress. He has had severe developmental delays. He started to walk at age 12 years, but has not developed speech. Dystonia and chorea commenced at 10 years of age and he has also developed tics. At 33 years of age, he used a wheelchair for ambulation and he was non-verbal, although he could communicate with sounds. He did not consistently obey commands and his verbal comprehension was difficult to assess, but he was interactive and affectionate with his family. He has had dysphagia and required a g-tube for liquids, but can eat solids by mouth. Other medical problems have included gait ataxia, hypotonia, cortical blindness, sleep disturbances, constipation and urinary incontinence. He suffered a single febrile seizure at two years of age. On examination at 28 years of age, height was 162.6 cm (3rd centile) and weight was 38.56 kg (<1st centile). His head circumference was 53 cm (8th centile). He demonstrated dystonia and tics with writhing movements of his arms. There were prominent supraorbital ridges, mildly overfolded right ear helix, a narrow nasal bridge and a small and narrow jaw. He had left esotropia and leukocoria. He had wrist and finger hypermobility and a thoracic scoliosis concave to the left. His 4th and 5th toes were small and curled. Neurological examination showed increased tone in all limbs, with brisk reflexes. An MRI of the brain showed mild diffuse prominence of the cerebrospinal fluid space and mild asymmetric prominence of right superior ophthalmic vein. Metabolic

investigations, including ammonia, carnitine levels and acylcarnitine profile, serum amino acids, urine organic acids, testing for congenital disorders of glycosylation, copper, ceruloplasmin and pantothenic acid levels were non-diagnostic. Conventional chromosome analysis of lymphocytes, fragile X syndrome testing and single nucleotide polymorphism (SNP) array gave normal results. Methylation studies for Prader-Willi syndrome/Angelman syndrome were negative and mitochondrial testing was also unrevealing. WES, performed as a duo with his mother, identified a heterozygous variant c.2215A>G, p.(Thr739Ala) in *DHX30* (NM_138615.2). His father was unavailable for testing.

Individuals 8 and 9 are half-sisters born to the same mother but have different fathers. The mother has no physical health problems and does not have any history of developmental delay or intellectual disability. She has a history of anxiety, depression and substance abuse. Her examination did not reveal any evidence of abnormal skin pigmentation or body asymmetry that would indicate a possible mosaic disorder. There is no other relevant family history.

Individual 8 was born by spontaneous vaginal delivery at term. She was able to sit with support from the age of 2 years. She achieved independent walking at the age of 10 years. She had limited vocalization and had 10-15 words by the age of 5 years. Currently at the age of 14 years, she can speak only single words. At the age of 12 years, she started having short-lasting and self-correcting vacant episodes that were associated with head dropping, eye rolling and loss of body tone. She had one particularly severe episode where she went absent and then her left arm lost power. There were no observed focal seizures. She was commenced on sodium valproate however this did not recur and she has since been weaned from this medication successfully. She continues to have absent episodes but these are thought to be non-epileptic in nature as they are short-lived and she can be easily roused from them. She has a round face, brachycephaly, large ears, bilateral epicanthic folds, open mouthed expression, everted lips and micrognathia. She has cold plethoric hands. Her growth parameters at 2 years 4 months were: height 87.5cm (-0.25 SD), weight 9.2kg (-3.33 SD) and OFC 46.5cm (-0.95 SD). Her OFC at 7 years 10 months was 48.3cm (-2.65 SD). An EEG and cranial MRI were both unremarkable. She also had normal CSF cell count and biochemistry, VLCFA, CSF amino acids and TORCH screen.

Individual 9 is the younger half-sister of individual 8, born by spontaneous vaginal delivery at term. She did not require resuscitation but was admitted to SCBU for 5 days due to poor feeding and hypernatraemic dehydration. There were no obvious neonatal seizures however she is noted to be jittery, possibly secondary to neonatal opiate withdrawal. She attained social smile on time. At 9 months of age she was noted to be vocalizing minimally. She was noticed to be poorly fixing but an ophthalmological examination did not identify any abnormalities. At 2 years 11 months, she had no speech, was sitting with support but not independently. At last clinical review, age 6 years, she has a standing frame but her mobility mainly consists of rolling around. She is able to get into a crawling position and sit herself up for a very brief period. Her speech remains unintelligible as she shouts and has no obvious words. She had experienced two chest infections requiring admission to hospital. At this time she was noted to have short-lasting vacant episodes. It is unclear clinically if these episodes could be behavioural although seizures remain a possibility. She has not required any anti-epileptic medication. She has severe reflux requiring treatment with omeprazole and domperidone. She has plagiocephaly, a bifid uvula, high palate and single left palmar crease. Her feet and external genitalia were normal. She continued to have poor visual fixation and was also noted to have cold peripheries, 2-3 mild syndactyly and small nails on the 5th toe bilaterally. Her growth parameters at 9 months were: height 66.4cm (-1.55 SD), weight 6.25kg (-2.30 SD), OFC 41cm (-2.11 SD). At 2 years 11 months, her height was 76.9cm (-4.36 SD), weight was 7.65kg (-6.73 SD) and OFC was 44cm (-2.83 SD). An awake and sleep EEG were both normal but cranial MRI showed partial agenesis of the corpus callosum with possible frontal atrophy. Urine organic and amino acids revealed no abnormality. She was thought to have the same condition as her elder half-sister.

In both girls, Angelman testing (FISH 15q and 15q methylation studies), 7-dehydrocholesterol and white cell enzymes gave normal results. Chromosomal microarray analysis in individual 8 revealed no abnormalities. In individual 9, chromosomal microarray revealed a possible duplication at 22q12.2:q12.3(31592382-32217094). This gain encompasses *RNF185*, *LIMK2*, *PIK3IP1*, *PATZ1*, *DRG1*, *EIF4ENIF1*, *SF11*, *PISD*, *PRR14L* and *DEPDC5* genes. However, as both the affected girls were thought to have the same condition, this gain was thought to be co-incidental. WES was performed in individual 8 as part of the Deciphering Developmental

Disorders (DDD) study and identified heterozygous variants in *DHX30* (chr3:g.47889727C>T, ENST00000445061 c.2344C>T, p.Arg782Trp), *RHOBTB2* (chr8:g.22864764C>T, ENST00000519685 c.1072C>T, p.Arg358Ter) and *KAT6A* (chr8:g.41798484C>G, ENST00000396930 c.2915G>C, p.Arg972Pro).

Variants in *RHOBTB2* have been associated to an autosomal dominant epileptic encephalopathy (OMIM #618004) however all pathogenic variants are missense ones, and a gain-of-function mechanism has been postulated. The p.Arg358Ter variant in individual 8 inserts a premature stop codon in exon 7 out of 12, which is likely to undergo nonsense-mediated decay. This variant does appear in healthy controls at low frequency (0.000008 mean allele frequency). There are also seven copy number losses encompassing *RHOBTB2* on the database of genomic variants which lends further support to loss-of-function not being the pathogenic mechanism. Thus, this is thought to be a low frequency benign variant. Pathogenic variants in *KAT6A* cause autosomal dominant mental retardation (OMIM #616268) with nearly all described variants being frameshift or nonsense. The missense variant in individual 8 is predicted to be benign by in silico tools and is also seen infrequently in healthy controls (mean allele frequency 0.00001) hence it is believed to be benign. Targeted Sanger sequencing for the c.2344C>T *DHX30* variant identified it to be present in the affected half-sister (individual 9). However, targeted Sanger sequencing did not show presence of this variant in the DNA samples extracted from either peripheral blood or saliva of the mother of the two children. Hence we suspect this is a case of gonadal mosaicism given both girls have separate fathers.

Individual 10 is a 3 years old female who was the product of an uncomplicated pregnancy. She was delivered to an unaffected Caucasian couple at 39 weeks of gestation with a birth weight of 3110 grams (-0.6 SD) and length of 48 cm (-1.4 SD) and a OFC of 33.5 cm (-0.9 SD). APGARs were 8 and 9. She was noted to have a heart murmur and an Echocardiogram at 3 days of life showed a small anterior/apical ventricular Septal Defect. The Newborn screening was normal. She had a healthy sister and a healthy paternal half-brother. The maternal half-uncle had cerebral palsy and her paternal aunt had severe epilepsy and intellectual disability. She was referred to the clinical genetics because of developmental delay as well as feeding problems as early as 7 months of age. Hypotonia, mild facial dysmorphism and generalized joint hypermobility were noted. She had rotatory nystagmus which improved in time. The brain MRI at 13 months of age showed delayed myelination (6-8 months) and delayed corpus callosum development with left hippocampal atrophy. At 3 years old age, she has two words and is able to sit up, roll and is starting to crawl. She is also pulling herself up but is not walking. She has two words and is starting to use a communication device. She can finger-feed herself. Her weight is just below the 5th percentile and her length is at the 7th percentile. Chromosomal microarray analysis, DNA testing for Prader Willi Syndrome and Spinal Muscular Atrophy gave normal results. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2344C>T, p.(Arg782Trp) in *DHX30* (NM_138615.2).

Individual 11 is a 7-year-old female, product of an uncomplicated pregnancy following a natural conception. Family history is negative for a similar phenotype; the patient's younger sister is unaffected, and younger brother was diagnosed with high functioning autism spectrum disorder; her parents are healthy of Caucasian descent, consanguinity was denied. She was born by a normal spontaneous vaginal delivery with a birth weight of 2.67 kg. due to poor sucking she required syringe feeding in the first few days of life. She has been followed by Pediatric Neurology, Pediatric Gastroenterology, and Medical Genetics for intractable complex partial epilepsy; profound global developmental delay and intellectual disability (non-verbal); athetoid cerebral palsy (wheelchair dependent); and failure to thrive (G-tube dependent). Brain MRI with spectroscopy performed at 8 months and 27 months of age demonstrated symmetric patchy signal abnormality with T2 hyperintensity throughout the bilateral cerebral white matter indicative of delayed myelination or nonspecific leukoencephalopathy, and mild white matter volume loss with relative thinning of the corpus callosum. Her OFC at the age of 5 years was 46cm (-4 SD). She was hospitalized at 6 years old for septic shock in setting of febrile illness. At the last clinical examination at 7 years of age she presented with progressively worsening mental status alteration, vomiting and constipation. Her weight was 13.6 kg (-5 SD) and height of 103 cm (-4 SD). No dysmorphic facial features were observed, she was highly hypertonic with athetoid extruded tongue posture. Metabolic screening gave normal results. Molecular genetic work-up included comparative genomic hybridization and Prader-Willi methylation testing with normal results. Trio-WES with DNA samples of both

unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2344C>T, p.(Arg782Trp) in *DHX30* (NM_138615.2).

Individual 12 is an 8-year old female, the eldest child of non-consanguineous Australian parents of Caucasian descent. Her siblings and both parents are well and healthy and there is no other family history of note. She was delivered at 40 weeks after an induced rapid and uncomplicated labour. Her birth weight was 3610g (+0.4 SD), her birth length was 49cm (-1.2 SD) and her OFC was around 36-37cm (+0.9 / +1.6SD). She was in good condition at birth and there was no complication in the immediate newborn period. She had feeding problems that have improved but she still chokes on crackers, and she has ongoing constipation issues. Her general health is otherwise good. She has strabismus and cortico-visual deficit, her vision has been improving over time. She has not had any obvious seizures. She has trouble regulating her temperature, she is often cold and wakes up at night when cold. She has inappropriate response to pain (laugh) and she needs strong stimulus (loud noise, strong light) to respond. She smiled at around 12 weeks of age, rolled both ways at 4 months, sat up at 12 months, crawled at 13 months and walked at 4 years of age. Her gait is unsteady and ataxic and she has numerous falls. She has global developmental delay and is non-verbal. She is quite social but doesn't care to interact with people. She is generally placid. She has a high pain tolerance and is not aware of danger. She holds her hands in a particular position with the thumbs abducted. At her last examination, her growth parameters were all around the 50th percentile. She had deep set eyes (family trait), furnished eyelashes, thin straight eyebrows, strabismus subtle facial asymmetry, posteriorly rotated ears with small antitragus, short antihelix, and everted lower lip; she has joint hypermobility of her hips, shoulders and fingers and A brain MRI showed minor periventricular leukomalacia. Plasma amino acids, urine metabolic screen, carbohydrate deficient transferrin, plasma acid carnitine profile, plasma and CFS amino acid and plasma and CFS pyruvate and lactate were all normal. Chromosome microarray testing detected a small maternally inherited, unlikely significant, 2q12.2q12.3 duplication. Methylation studies for Angelman syndrome did not detect any anomaly. Genetic testing for lysosomal disorders indicated that she is a carrier of Tay-Sachs disease (normal levels of beta-hexosaminidase A activity) and Pompe disease. Genetic testing for an in-house absent speech gene panel did not detect any likely causative variant. Trio-WGS with DNA samples of the proband and both unaffected parents, revealed a *de novo* heterozygous variant c.2344C>T, p.(Arg782Trp) in *DHX30* (NM_138615.2).

Individual 13 is an 8 year old male, the only child of non-consanguineous parents of Mexican origin. Regarding family history, mother has 2 brothers, both of whom had seizures during childhood and have normal cognition. The father has 2 brothers and 1 sisters. Father's brother's daughter did not talk at the age of 4, walked by age 2 and has a brother who did not talk at the age of 2. Regarding the perinatal history, the pregnancy was uncomplicated and he was born at full term. Birth weight was 3.72 kg. He was healthy in the newborn period except for transient hypoglycemia treated with bottle feeding. It was noted at 6 months that he was microcephalic and had delayed motor milestones. At 6 years of age regarding expressive language development he was able to say mom, dad, and name of dog. He was not able to point or wave. He could do a high five. In regard to gross motor development, he was able to roll, pull to stand, army crawl, and walk with walker. He could crouch down and balance himself. Regarding fine motor skills he was able to do a raking grasp. He could put things in his mouth. He was not toilet trained. He had dysphagia. He could eat finely chopped foods and had difficulty with chewing. In regard to social skills: he was very social. In regard to academics/cognitive ability: he could recognize things that he likes. He was able to identify four colors. At the last comprehensive examination his OFC was 47.5 cm (-2.05 SD), height was 112 cm (-1.77 SD) and weight was 16.2 kg (-2.96 SD). Facial features were elongated. Head shape exhibited microcephaly. Forehead revealed prominent brow. There were epicanthal folds. He had arched eyebrows and long eyelashes. He had temporal muscle wasting. He had a small mouth, down turned. Mental examination revealed a social smile. He was nonverbal. He had a normal cranial nerve examination. He had central and axial hypotonia. He had increased tone in the Achilles. He had normal reflexes. He had joint hypermobility. He was able to walk with an ataxic gait with two hand assist. Diagnostic studies included an MRI of the brain which was normal, one EEG revealed bifrontal and left central parietal sharp waves during sleep but subsequent ones were normal. Hypotonia panel demonstrated normal results for SMA, myotonic dystrophy, Prader Willi, Maternal UPD 14, and Angelman syndrome. Lactate, pyruvate and urine organic acids were normal. Fragile X screen normal. Chromosomal microarray was normal. Cornelia de Lange testing revealed a VUS in the RAD21 gene that was determined to not cause CDLS or any genetic condition. Trio Autism/ID Xpanded Panel with DNA samples

of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2345 G>A, p.(Arg782Gin) in *DHX30*. (NM_138615.2).

Individual 14 is now a 6-year old male product of a full term normal spontaneous vaginal delivery, the third child of unaffected, non-consanguineous parents of Hispanic and German ancestry. His family history includes two full siblings: an 8-year old brother with speech delay, an 11-year old healthy sister, and two half-siblings through the father: a 10-year old brother with learning disabilities/ADHD and a 13-year old sister with scoliosis. Father is healthy and mother experiences anxiety and depression. At about two months of age, his grandmother noted that he rarely cried and he had significant head lag and was "limp" by her account. His pediatrician formally evaluated him at 6-months of age and requested an ultrasound of his head due to his having a small fontanelle. He received an MRI of the brain as well as an encephalography, which was normal. His Newborn Screen was consistent with sickle cell trait. He also had poor weight gain, acid reflux, and constipation, and was referred to GI, and subsequently gained weight on a high calorie mix. He had high tolerance to pain and reportedly never cried. He was receiving physical therapy and occupational therapy. Developmentally, he began rolling at about 4.5 months, but does not crawl or walk. He presented at the genetics unit at 4-years of age for severe developmental delay and profound generalized hypotonia and dysmorphic features suggestive of a genetic etiology. He could sit up with assistance, but was unable to hold his bottle and had some head control. He could grasp objects and occasionally reached for things. He lacked speech but made some babbling, and cooing, without consonant sounds. His physical features included microcephaly, bitemporal narrowing, a tall forehead, a high arched palate, and a pointed chin with no dimple. Other notable features were a single palmar crease on his left hand, overlapping second and third toes on right foot, hypermobility in his knees, and decreased calf musculature. He had diminished bulk, and was severely hypotonic throughout his body. On vertical suspension, he slipped through the examiner's hands, and on horizontal suspension was not able to lift his head. He showed a severe lag when pulled from a laying to a sitting position. In a prone position, he had some head control and was able to lift his head occasionally. He could not elicit upper extremity reflexes. He would withdraw upon experiencing light touch. He was unable to sit, even when held in a sitting position. Prior negative testing included a Prader-Willi syndrome methylation assay. Metabolic investigations, including carnitine, lactate/pyruvate, urine organic acids, and plasma amino acids were non-diagnostic. Given his significant hypotonia, testing for Spinal Muscular Atrophy, creatine kinase level, and myotonic dystrophy was pursued; none was informative. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2345 G>A, p.(Arg782Gin) in *DHX30*. (NM_138615.2).

Individual 15 is a 2 year old male, the only child of unaffected, non-consanguineous healthy Dutch Parents of Caucasian European descent. Pregnancy was uncomplicated with normal screening ultrasound. The boy was born at 41 weeks and 6 days. There was meconium in the amniotic fluid. He had APGAR score 4/7/7 after 1, 5 and 10 minutes respectively. Oxygen was supplied shortly and he received antibiotics because of suspected perinatal infection. Birth weight was 3320 grams (-1 SD). Hernia inguinalis was noted. Because of hypotonia, tube feeding was started for the first few days. Psychomotor development was delayed, with laughing at 8 weeks. Physiotherapy was started at age 5 months because of delayed motor milestones with persistent axial hypotonia including head lag. Rolling over at age 9 months. He has gastro-oesophageal reflux and frequent choking when drinking fluids. He experienced recurrent otitis. He never cries and had a high threshold for pain. He has a cheerful behavior. Furthermore, he has a trigger finger. He does not experience seizures and electroencephalogram (EEG) at age 1.5 y was normal. At the last physical examination at age 2 years and 7 months, height is 93 cm (-0.31 SD), weight 14 kg (-0.23 SD), head circumference 47 cm (-1.55 SD). There is a broad and high forehead, low position of the ears, overfolded helix, simple ears, flat midface, brachycephaly, clinodactyly of the 4th and 5th toe, and hypoplastic toe nails of the 5th toe on both sides. He is mildly bradyphrenic and vocalizes but does not speak. There are feeding difficulties. He cannot sit without support, there is still some head lag and axial hypotonia, strabism with saccadic ocular movements and mild generalized chorea and poor fine motor skills.). Metabolic investigations were unremarkable. Brain imaging has not been performed. There was a normal CGG repeat length of the *FMR1* repeat. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2345 G>A, p.(Arg782Gin) in *DHX30*. (NM_138615.2).

Individual 16 is a 4 year female- Family history was unremarkable. She has two older, unaffected half siblings (different fathers). She was delivered at 39 weeks by induction because of absent fetal movements for 2 days, confirmed on ultrasound. Birth weight was 3900 grams (+1.33 SD) and birth length was 48.3 cm (-1.3 SD). She did not cry, but did make grunting noises. As she got older her development fell further behind. She did not cry after vaccinations and she had a "vacant" look and poor eye contact. At 9 months of age she began rolling over, at 11 months she could only sit with support, and she made no attempts to crawl. At the clinical examination at 10 months she had a pronounced global developmental delay. She had significant hypotonia but began showing better eye contact. She was found to have a bicuspid, stenotic aortic valve and a moderately dilated ascending aorta (z score 3.9) and mild insufficiency. She developed staring spells, and a Video EEG showed focal slowing and epileptiform discharges. An MRI showed diffuse subcortical and periventricular white matter signal abnormality in a symmetric pattern involving the temporal frontal and parietal lobes with associated abnormal cerebral volume loss suggestive of a leukodystrophy but not typical for periventricular leukomalacia. Organic acids, VLCFA, pipercolic acid and arylsulfatase enzyme testing were normal. Over the next several years, she developed few skills. She said a few words, but none consistently. She had a grand mal seizure age 3, Keppra was started, and she had only one further seizure. At age 4, weight was 15.9 kg (48 %ile) Height: 101.6 cm (52 %ile), She had mild ptosis, hypotonia, and did not talk, stand for very long, or take steps. Because of recurrent ear infections tonsillectomy and adenoidectomy, and tympanostomy tube surgeries were performed at the age of 4. WES revealed a heterozygous variant c.2353 C>T, p.(Arg785Cys) in *DHX30* (NM_138615.2).

Individual 17 is a 7.5-year-old girl, the only child of unaffected, non-consanguineous American parents of Caucasian European descent. Family history was non-contributory. The pregnancy was uncomplicated with normal screening ultrasounds. She was born at 39 weeks of gestation by caesarean section due to breech presentation. Her birth weight was 4054 gram (+1.7 SD), birth length was 53.34 cm (+1SD), and OFC was 38.1 cm (+2.7 SD). Muscular hypotonia was first noted at 8-week old. Since then, she was regularly in physical therapy, and she has had no period of regression. Her milestones of motor development were delayed and has poor fine motor skills. She was able to sit independently at 13 months and to walk without assistance at 37 months of age. She presents symptoms of ataxic gait. She is unable to toilet train, and she was unable to feed herself, but occasionally she can grab food with her fingers to feed herself. She is non-verbal and has been on speech therapy regularly since 1-year-old. Her progress: At one-year-old she could respond to pictures/flashcards. At three-year-old she was able to use a communicator with eye gaze technology. At seven-year-old she was able to transition to a touch screen communication tablet. Overall, she is described as a pleasant, quiet, smiling child with no behavioral concerns. But she is easily irritated by noises. Sleep disturbance is rare. She presents repetitive finger snapping and teeth grinding on a regular basis, which causes trouble in chewing. Physical examination revealed no obvious dysmorphic features or developmental features, except joint hypermobility and low muscle tone, which has improved progressively with physical therapy. Her reflex response findings overall were unremarkable. OFC at last examination at the age of 7years and 8months was 53 cm (+0.9 SD). MRI was performed at 13-month, 17-month, and 37-month-old with similar findings: generalized decreased in white matter volume in a symmetric fashion. The spectroscopy tracing did not reveal any specific abnormality. Ethmoid sinus and left mastoid air cell opacification were noted. She occasionally blanks out or stares into space. Electroencephalogram (EEG) at two-year-old was normal during wake and sleep stages, ruling out absence seizures, and she has not had a history of seizures. She underwent array CGH, Prader-Willi analysis, extensive metabolic analyses (including plasma amino acids, lactate and pyruvate, etc), Canavan disease panel, Krabbe disease test, and lysosomal disease test: all were negative. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2353 C>T, p.(Arg785Cys) in *DHX30* (NM_138615.2).

Individual 18 is a now 3-year-old male. He is the first child of unaffected, non-consanguineous Caucasian parents. There was no family history of seizures, developmental problems, neuromuscular disorders, or other neurologic issues in the family. The pregnancy was uncomplicated aside from polyhydramnios noted in the last month of the pregnancy. He was born via Caesarean section at 40 weeks due to a face presentation. There were no other perinatal or neonatal complications, and he went home from the hospital in a normal amount of time. Developmentally, all milestones were delayed. He did not roll until 6 months, and started to sit up independently at about a year of age. At his last visit at age 3, he could pull to stand and was working on walking with a walker but could not yet walk independently. His fine motor skills were somewhat delayed

but he was able to feed himself using a spoon. He had chronic dysphagia and had trouble with chewing and swallowing, though this improved over time. He could copy some sounds but could not yet produce intelligible speech. He could partially communicate needs using a picture board, and does produce some verbal cues that his parents can understand. He was noted to be friendly, social, and interactive with family. He has always made slow, gradual progress and has never regressed. His notable medical problems are strabismus, suspected mild central visual impairment, dysphagia, and failure to thrive. There has never been any evidence of seizures. At his visit at 16 months, he was microcephalic (3%, $Z=-1.95$) with weight at the 10% for age and height at the 12% for age. Over time, he had had poor growth, and at his last visit at age 3 he was noted to be at the 2nd percentile for height. He also had poor weight gain, with his weight also falling below the 3rd percentile over time, though at the last visit his weight gain had improved after intensive effort by his parents and he was at the 30% for weight. At his last visit at age 3 his head circumference was at the 2nd percentile for age ($Z= -2.05$). Facial features were nondysmorphic, and he had no notable birthmarks or other malformations. On neurologic exam, he was alert and appropriately attentive to the examiner. He had mild strabismus. He had mild axial and truncal hypotonia and mild symmetric proximal weakness. He could reach for objects without ataxia. Deep tendon reflexes were normal and symmetric, though he had bilateral upgoing toes on Babinski testing. An MRI scan at the age of 1 year showed mild-under operculization of the Sylvian fissures, but was otherwise unremarkable. EEG was not performed. He had an extensive metabolic workup done which was unrevealing including CK, CMP, serum amino acids, urine organic acids, acylcarnitine profile, tsh, carnitine, and pyruvate all of which were normal. He had a normal chromosome microarray. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2353 C>T, p.(Arg785Cys) in *DHX30* (NM_138615.2).

Individual 19 is a 2y8m old female, the second child of unaffected, non-consanguineous parents of Cuban and African-American ancestry on the maternal side, and Irish, Native American and Cape Verdean ancestry on the paternal side. Family history is non-contributory. There were no medical complications during the pregnancy, and she was delivered at full term via induced vaginal delivery due to poor fetal movement. Her mother and father were ages 21 and 20, respectively, when she was born. Phototherapy was needed for hyperbilirubinemia, but otherwise the neonatal course was uncomplicated. She passed the newborn hearing screen, but later had borderline hearing tests and is awaiting brainstem auditory evoked potential test to rule out hearing loss. The patient was first evaluated by neurology at 20 months for concerns of diffuse low muscle tone and global developmental delays. Her neurologic examination was notable for diffuse axial and appendicular hypotonia with preserved deep tendon reflexes. She did not start to combat crawl until 10 months, and started walking at 22 months. MRI of the brain done at 29 months showed no structural abnormalities and normal myelination for age. She has never had any seizures. Clinical examination at 25 months was notable for a weight of 9.895 kg (-1.6 SD), length of 79.9 cm (-2.2 SD), and OFC 47.0 cm (-1.2SD). She had a 5 mm café au lait macule on the right torso and an irregular hypopigmented lesion on the right buttock, and epicanthal folds. Her delays have persisted. She began using SMO braces at 26 months. Her fine motor skills are behind, and she can only stack two blocks, one on top of the other. Beginning around 30 months, autistic features were first noted. Currently, at 2y8m, she has had some language regression and cannot remember words that she had learned earlier. She only uses one word consistently, which is the name of her brother. She has a happy demeanor and does not cry. She has a very high pain tolerance, has bruxism, and stomps her feet repetitively. Although she appears happy, she has a limited repertoire of emotions, makes only intermittent eye contact, and generally appears somewhat indifferent and disinterested regarding interactions with either peers or toys. In addition to this limited social-emotional reciprocity, she also has feeding difficulties, perseverates, and is strong willed. Initial genetic testing included a G-banded karyotype, and chromosomal microarray; both were negative. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2353 C>T, p.(Arg785Cys) in *DHX30* (NM_138615.2).

Individual 20 is a 16-year-old male, the first child of healthy, non-consanguineous parents of European descent. He has a 10-year old-sister with ADHD and dyslexia, but otherwise normal development. A cousin of the father had dyslexia and developmental delay, remaining family history was unremarkable. He was born after an uneventful pregnancy at 39 weeks of gestation with a weight of 3050 g (-1 SD), a length of 54 cm (0.91 SD) and a head circumference of 36 cm (0.54 SD). In the first months of life feeding difficulties and neck asymmetry were reported, and at age 8 months microcephaly was noted. He had psychomotor developmental delay. Age of sitting was at 9-10 months and age of walking at 20-21 months, gait was unstable and ataxic. He

is non-verbal. MRI at age 16 months and EEG were normal. Sleeping difficulties slightly improved with Melatonin. He has mild constipation and received tympanostomy tubes. He has a short attention span, is very active and occasional has temper tantrums. At the last physical examination at 16 years and 10 months, his height was 174 cm (-0.62 SD), his weight was 69 kg (0.21 SD), and his OFC was 51.5 cm (-3.39 SD). Facial dysmorphism included overfolded helices, a low forehead, narrow palpebral fissures, a short philtrum, a high and narrow palate, prominent incisors, a prominent jaw, broad thumbs, gynecomastia and pes valgus. He was friendly but with short attention span, did not speak and showed stereotypic hand movements. Karyotyping, chromosomal microarray analysis, Angelman syndrome methylation testing and sequencing of UBE3A, MECP2, TCF4, CDKL5 and ARX were normal. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2354G>A, p.(Arg785His) in *DHX30* (NM_138615.2). Additionally, a hemizygous, maternally inherited variant of unknown significance c.10688G>A, p.(Gly3563Asp) was identified in *HUWE1* (NM_031407.5).

Individual 21 is a 15 year old female, the first child of unaffected, non-consanguineous parents of Latvian and Norwegian descent. She has a healthy 11 year old brother, and unremarkable family history. The mother experienced hyperemesis and vomiting during the entire pregnancy. The girl was born with elective Caesarean section at gestational week 39. Birth weight and length were 3120 g (-0.5 SD) and 47 cm (-1.9 SD), respectively. She was operated twice, at 3 months and 3 years of age, for inguinal hernias, and had bilateral auricular surgery at age 8 years because of prominent ears with uneven size and unilateral irregular cartilage of the antihelix. Psychomotor development was reportedly normal. Age of sitting was at 6-7 months and age of walking at 12 months, she spoke first words at 11 months. After an unremarkable psychomotor development, a progressive balance impairment with midline ataxia was noted from the age of 8 years. Subsequently, she developed reduced motor skills and cognitive problems with reduced concentration and fatigue. Neurological examination revealed midline and appendicular ataxia with nystagmus on lateral gaze, dysdiadochokinesia, dysmetria and intentional tremor, broad-based ataxic gait and negative Romberg's test. Metabolic screening was normal. Muscle biopsy histology and testing for mitochondrial respiratory chain defects gave normal results. MRIs showed progressive cerebellar atrophy. Although she had no clinical seizures, EEG showed epileptiform, spike-and-slow-wave activity localized in the left central temporoparietal region. She was given levetiracetam for a period, but this was discontinued following a normal EEG. Audiometry was normal. At the age of 15 years her height was 172cm, her weight was 53kg (-0.3 SD) and her OFC was 55.5 cm (+0.6 SD). Karyotyping with G-banding (46,XX) and array-CGH showed normal results. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a *de novo* heterozygous variant c.2606G>A, p.(Arg908Gln) in *DHX30* (NM_138615.2). Additionally, a *de novo* heterozygous variant of unknown significance c.1535A>C, p.(Glu512Ala) in *KLB* (NM_175737.3) was identified, a gene variants in which have so far not been connected to a human phenotype.

Individual 22 is a 5 year 10 month old female with non-consanguineous parents. She was born after an uncomplicated pregnancy and delivery. Family history is non-contributory. She was born at 41 weeks, birth weight of 4110 g (+1.2 SD), length of 52 cm (-0.1 SD), and OFC of 36 cm (+0.6 SD). Her neonatal course was unremarkable. Her parents became worried when she began to walk around age 1 year because she was unsteady on her feet. She still has an unsteady gait, especially when she is tired. Developmental delay is global and most pronounced for expressive language where her parents estimate that she lags 1-2 years behind her peers. She is starting in ordinary school with extra help. She is a willful child who has some difficulty interacting with other children and often seeks the company of adults. She is very active and needs adult supervision at all times when awake. She wanders at night and ends up in her parents' bed. Periodically need of melatonin. She has generally been somatically healthy. She had an adenotonsillectomy at age 3 years and has had fewer upper airway infections subsequently. A clinical suspicion of brief complex partial seizures has not been confirmed, she has short non-epileptic absences, no AED. Her vision and hearing is normal. She has a bilateral, intermittent, exotropic squint. Ophthalmological exam performed under general anaesthesia at age 3 years was unremarkable. She is neither hypotonic nor dysmorphic. Her height is on the >99 percentile (+2,7 SD), weight 86 perc and OFC 68 perc. Cranial MRI at age 3 years was unremarkable. Array CGH gave normal results. An trio-based, NGS-DDg2p panel revealed a *de novo* heterozygous variant c.347_360del, p.(Ala116Valfs*12) in *DHX30* (NM_138615.2). No other relevant variants were detected.

Individual 23 is a 3-year-old male, the third child of non-consanguineous Caucasian parents. His brother has a history of visual processing disorder and migraines. His sister has a history of strabismus, ptosis, and delayed walking at 20 months of age. Pregnancy was naturally conceived with pre-existing maternal hypertension treated with amlodipine. Worsening maternal hypertension prompted delivery induction at 37 weeks gestational age. Maternal fever at delivery prompted 24 hours of antibiotics, and after discharge at day of life 3 he required readmission for mild hypothermia. He was noted to have global developmental delay. He started laughing and smiling at 2-4 months of age. He is affectionate with good eye contact with his family, but he does engage in some repetitive behaviors, flaps his arms when excited, and lines objects up. He started sitting at 8 months of age, army crawling at 12 months, pulling to stand at 12 months, and walking at 16-17 months. He had an immature pincer grasp at age 3 years. He did not babble much in early infancy. He said his first word around 18 months. He uses more than 20 words at age 3 years, however his speech is largely not intelligible. Evaluation demonstrated normal growth parameters (height 85th %ile, weight 40 %ile, head circumference 44 %ile) with hooded eyelids, left eye exotropia, faint synophrys, widely spaced teeth, underdeveloped left pectoralis muscle, and low axial tone. Genetic evaluation included a normal array CGH and fragile X analysis. Trio-WES with DNA samples of both unaffected parents and the proband, revealed a heterozygous variant c.2389C>T (p.Arg797*) in *DHX30* (NM_138615.2) inherited from a mosaic mother. His mother's history was not notable for any concerns apart from hypertension. His brother was testing and was found not to carry the variant. His sister's testing remains pending.

Individual 24 is now a 5-year old male, the product of an unremarkable pregnancy and delivery, and the second child of unaffected, non-consanguineous parents of Asian ancestry. He has an unaffected 7-year old brother. At 6-months of age, he was reported to display low muscle tone. His developmental history includes: rolling at 6 months, sitting at 9 months, babbling at 11 months, and walking independently after 12 months of age. After starting to walk, his mother noted he did not progress like his healthy older brother. His first words occurred around 1 year of age, and his vocabulary consisted of 100 words. To this day, he cannot climb stairs. His preference is to run over walking, but he falls frequently. While he has significant language delay, given his sociability with other children, his pediatrician had low concern for autism. He displayed significant drooling though he had adenoid hypertrophy. He first presented at 2 years and 11 months of age to the Neuromuscular Clinic for mild hypotonia, global developmental delay, speech delay, with motor impairment, and no evidence of muscular or peripheral nerve disease. Brain MRI detected T2 hyperintensities consistent with Chiari type 1 malformation with hypomyelination in the parietal occipital lobes; no cerebellar or cerebral atrophy was observed. At three years of age, he last presented to Pediatrics. He was receiving speech, occupational and physical therapies and was reported to have made progress. While reportedly bilingual, upon examination, he did not talk, but he could follow commands, but was not responsive or as communicative as a typically developing child. Physical examination revealed large hyperpigmented patches on his back and buttocks, but lacked cafe au lait spots. He lacked any history of a seizure disorder. Also upon examination, his disposition was smiling and responsive, alert and active. He did not display gross motor defects or localized weakness. His muscle tone was normal, but a little uncoordinated with relaxation. He was able to sit and did not display head lag. He was able to walk unassisted and could rise from the floor unassisted. While he had some decreased facial tone, he had no overt dysmorphic features. At last ascertainment, he was enrolled in Special Education classes. Genetic testing included Fragile X, which was negative, and chromosomal microarray analysis, which detected a likely pathogenic, heterozygous deletion of ~1 Mb at 3p21.31, involving the first 15 exons of *SETD2*, as well as *KIF9*, *KLHL18*, *PTPN23*, *SCAP*, *ELF6*, *CSPG5*, *SMARCC1*, *DHX30* (NM_138615.2) and *MAP4*. Out of these, apart *DHX30* only variants in *SETD2* causing Luscan-Lumish syndrome (OMIM#616831) have been associated with a human phenotype following autosomal dominant inheritance. The literature makes note of altogether seven individuals, with somewhat different phenotype, some of which inherited the *SETD2* variant from unaffected parents. Follow-up testing using DNA samples from both unaffected parents was performed and neither parent carried the deletion, confirming a *de novo* status in the proband.

Individual 25 is a 2 year old male referred for evaluation and management of genetic risk associated with global developmental delay. Family history is positive for maternal anxiety, maternal grandmother with depression and a maternal aunt with autism spectrum disorder who also has an affected daughter. Maternal two brothers with autism spectrum disorder. Pregnancy was complicated by prenatal diagnosis of ventriculomegaly and at birth the patient presented perinatal depression requiring mechanical ventilation. At

four months of age the patient showed signs of delayed developmental milestones and he presented seizure activity that responded well to medical treatment. Patient did not sit by himself until 8 months of age, did not crawl, and walked at 18 months of age. He was not verbal until the age of 17 months with a very limited vocabulary by two years of age. He is a happy but hyperkinetic child with a very short attention span for age. On physical examination the patient showed malar hypoplasia, midface hypoplasia, epicanthal folds, almond shaped eyes, stellate pattern of the iris, posteriorly rotated ears, small upturned nose, significant hypotonia and hypermobility. Head MRI confirmed persistent ventriculomegaly with mild reduction of white matter. Next generation sequencing showed an intragenic deletion of exons 7-9 of the *DHX30* (NM_138615.2). *De novo* status was confirmed by Sanger sequencing.

Additional information for:

Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders

Mannucci *et al.*

Additional file 7

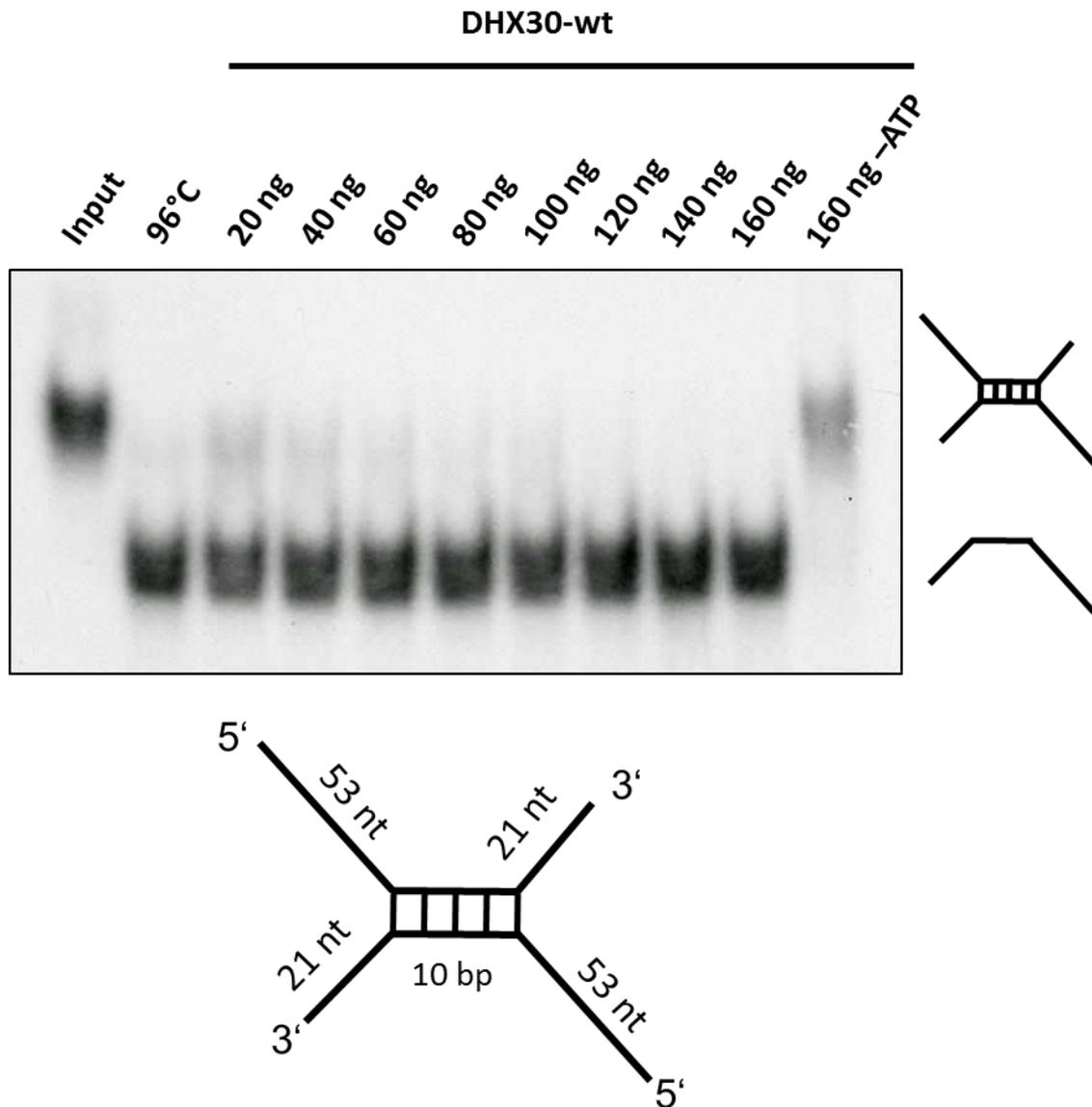


Fig. S4. DHX30 WT acts as an ATP-dependent RNA helicase. Top: Increasing amounts of His6-SUMO-tagged DHX30 WT protein were incubated with a ³²P-labelled RNA substrate in the presence (lane 3-7) or absence (lane 8) of ATP and analyzed by native PAGE. The position of the RNA duplex and the single-stranded RNA are indicated in the first and second lane, respectively. Their schematic representation is shown at the right side. Bottom: RNA duplex containing a central GC sequence flanked by single-stranded regions of 53 nucleotides at the 5' end and 21 nucleotides at the 3' end.

Additional information for:

Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders

Mannucci *et al.*

Additional file 8

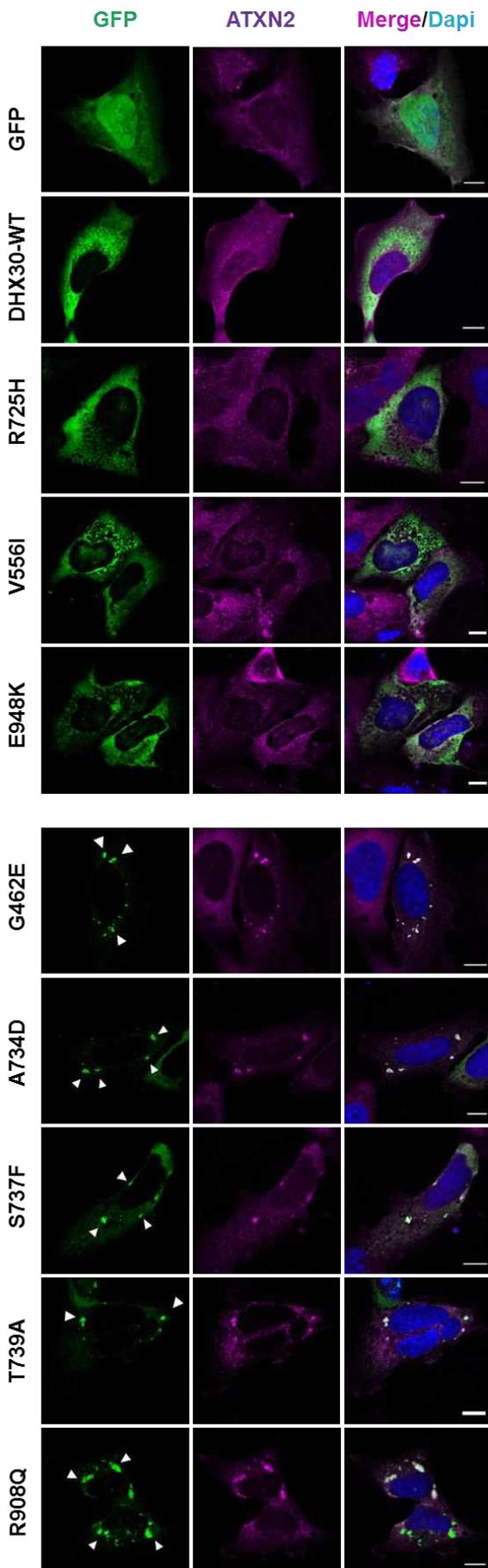
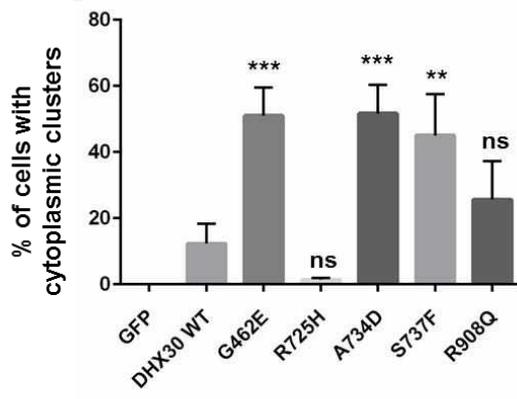
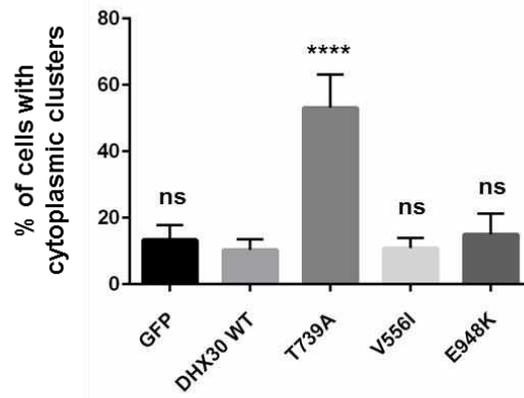
a**b**

Fig. S5. Recombinant protein variants of DHX30 induce the formation of cytoplasmic clusters. (a) Immunocytochemical detection of DHX30-GFP fusion proteins (GFP, green) and endogenous ATXN2 (magenta) in transfected U2OS cells. Upper panel: wild-type DHX30-GFP preferentially resides throughout the cytoplasm and GFP accumulates in nuclei, similar to recombinant protein variants of DHX30 harboring amino acid substitutions V556I, R725H and E948K (upper panel). Lower panel: recombinant protein variants of DHX30 harboring amino acid substitutions in the helicase core region G462E, A734D, S737F and T739A induce the genesis of cytoplasmic foci containing endogenous SG-marker ATXN2 (arrowheads), Notably, the R908Q amino acid substitution lead to the formation of clusters co-localizing with the SG-marker ATXN2 in only 50% of transfected cells. Nuclei are identified via DAPI staining (blue). Scale bars indicate 10 μ m. **(b)** Bar graph indicating the percentage of transfected cells, in which recombinant proteins induce the emergence of clusters. (**, ***, ****: significantly different from DHX30-WT: ** p < 0.01; *** p <0.001; **** p <0.0001; n > 100 from 3 independent transfections; One-Way ANOVA followed by Dunnett's multiple comparisons test).

Additional information for:

Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders

Mannucci et al.

Additional file 9

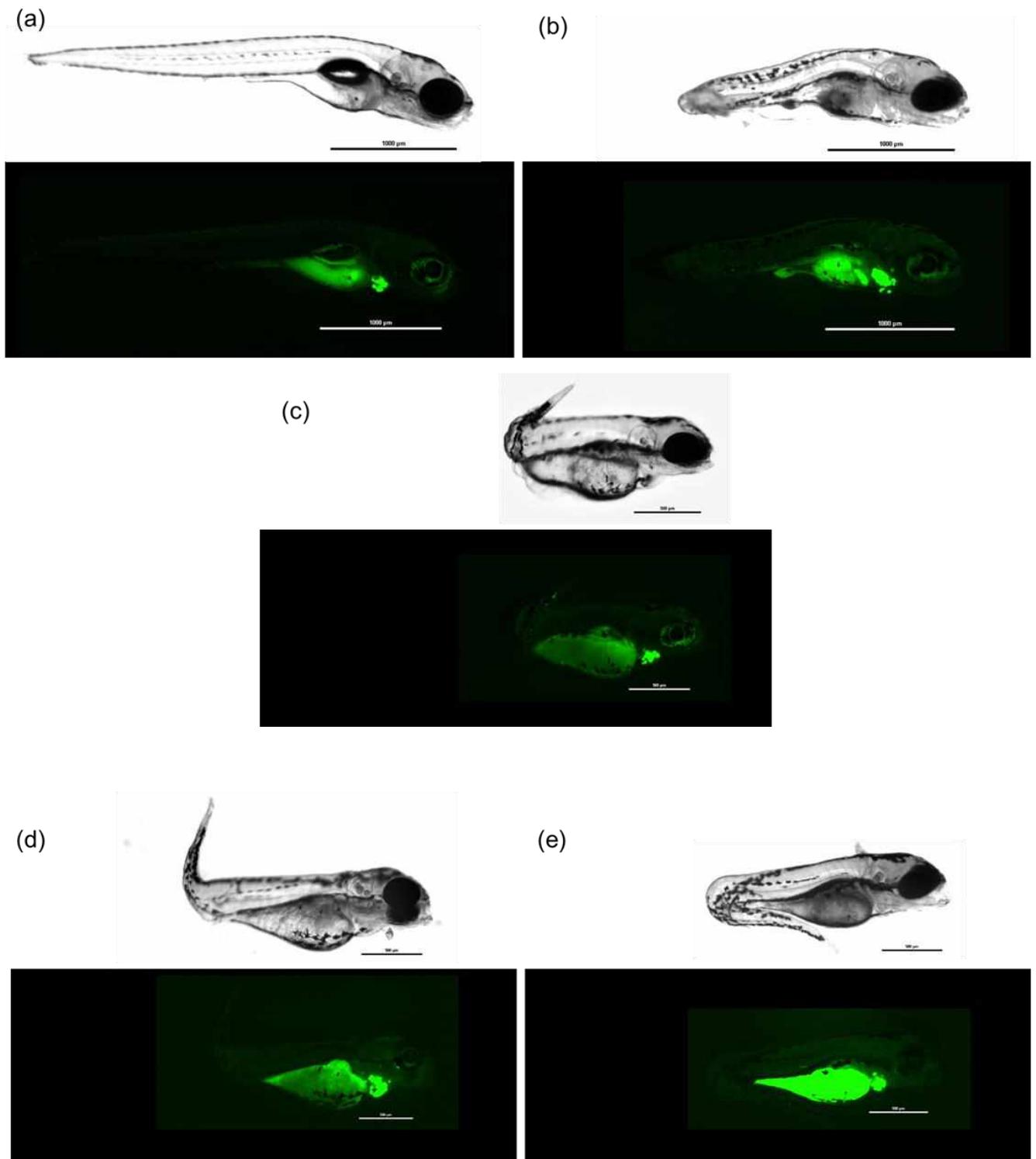


Fig.S6. Representative images of zebrafish embryos. (a) Injected with Tol2 mRNA and pTol2pA2-cmlc2:EGFP;tuba1a:DHX30 wild-type or (b) DHX30 harboring R493H, (c) R725H, (d) R785C, or (e) R908Q. Scale bars show 1000μm or 500 μm per unit. Embryos injected with wild-type DHX30 showed apparently normal development at day 7. Embryos injected with mutated DHX30 showed sign of severe developmental defects before day 7.

**Additional information for:
Genotype–phenotype correlations, and novel molecular insights into the *DHX30*-associated neurodevelopmental disorders**

Mannucci *et al.*

Additional file 10

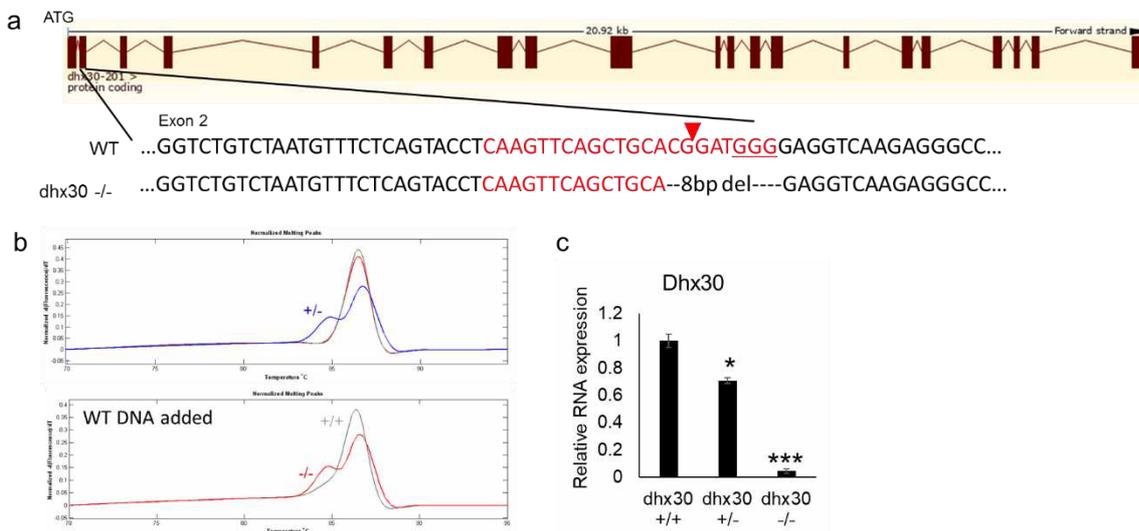


Fig. S7. Generation of zebrafish CRISPR-Cas9-mediated *dhx30* stable knockout line. (a) Shown are genomic regions of zebrafish *dhx30* targeted by CRISPR-Cas9. Red indicates gRNA-binding site with the protospacer-adjacent motif (underlined) in wild-type (WT) sequence. The mutant animals carry an 8 bp-deletions generated by CRISPR-Cas9. (b) *Dhx30*-targeted PCR products were analyzed by high-resolution melting analysis (HRM) to distinguish wild-type (+/+), heterozygous (+/-), and homozygous (-/-) animals. Two different melting peaks were shown in heterozygous PCR product (top). To distinguish wild-type and homozygous animals, wild-type DNA samples are mixed with DNA from the ‘test’ animals. Homozygous DNA hence become heterozygous-like, resulting in two melting peaks (bottom). (c) Analyses of *dhx30* transcript levels in *dhx30* mutant animals at 5 days post fertilization. Data are presented as means \pm standard error of mean and are based on 3 replications. *, ***: significantly different from *DHX30*+/+ (* p <0.05, *** p <0.001; n =3; One-way ANOVA, followed by the Holm-Sidak multiple comparison test).

Chapter 4 | Discussion

4.1 | Germline AGO2 mutations impair RNA interference and human neurological development

Biallelic disruption of *Ago2* during murine embryogenesis leads to early developmental abnormalities in the central nervous system (CNS) and an embryonic lethal phenotype in mice (Liu et al., 2004). Given the central role of AGO2 as a part of RISC, these observations further highlight the importance of proper regulation of gene expression by RNAi in development and maintenance of neuronal circuits. The essential role of miRNAs in development and function of the CNS has been extensively investigated (Rajman & Schratt, 2017). Whereas deletions encompassing both *AGO1* and *AGO3* have been reported to cause neurodevelopmental disturbances in five individuals (Tokita et al., 2015), no genetic alterations in *AGO2* were so far found associated with any human pathology. Recently, our study by Lessel and colleagues (Lessel et al., 2020) identified altogether 21 patients affected by mild to severe global neurodevelopmental delay with heterozygous, *de novo* mutations in *AGO2*, including eleven missense mutations, one *in-frame* deletion and a 235.3-kb deletion in *AGO2*. Similar to the large deletions encompassing *AGO1* and *AGO3*, the 235.3-kb deletion affecting the first 3 exons of *AGO2* is likely to lead to haploinsufficiency. All *AGO2* mutations were found to impair shRNA-mediated gene silencing. The observation that upon co-expression with AGO2 WT, selected AGO2 mutants do not interfere with its knockdown capacity points to a loss of function rather than a dominant negative underlying pathomechanism. Careful dissection of the RNAi pathway led to the observation that the majority of *AGO2* mutations (with the exception of p.G733R which failed in almost every functional assay) do not affect basic aspects of RISC formation such as the ability of AGO2 to interact with DICER or bind endogenous miRNAs. Instead, mutations appear to delay mRNA target release from AGO2. With the exception of p.G573S that alters the C-terminal end of the middle (MID) domain, most of the mutated residues are distributed at the interface between the linker region L1 and the helix-7 of the linker region L2. Interestingly, helix-7 was shown to have a key role in ensuring the timeliness of miRNA silencing by modulating the dynamics of target recognition (Klum et al., 2018). Phosphorylation of a C-terminal serine cluster of AGO2 has been shown to facilitate mRNA target release from RISC (Golden et al., 2017; Quévillon Huberdeau et al., 2017). The reduced phosphorylation of this cluster observed in the AGO2 mutants might be the cause of their extended dwelling time on the target. This hypothesis is further supported by molecular dynamics simulations showing that *AGO2* mutations interfere with target release by slowing down the ability of AGO2 to unwind guide-target duplexes. Ultimately, the slower release of mRNA targets by mutant forms of AGO2 is associated with an increased number of dendritic P-bodies in rat hippocampal neurons, and with global transcriptome alterations in primary fibroblasts of *AGO2* patients. Interestingly, four amino acid residues in AGO1, F180, G199, L190 and T355 were very recently found mutated at the equivalent positions to those found by us in AGO2 (p.F182del and p.G201C at the base of L1, p.L192P in L1 and p.T357M in helix-7 of L2) in similarly affected patients (Schalk et al., 2020). Thus, alterations in the dynamic interaction of AGO proteins with miRNAs and their mRNA targets appear to be a common pathway in the pathogenesis of neurodevelopmental disorders.

4.2 | The RNA helicase DHX30: role in neurodevelopmental disorders

The pathological relevance of altered RNA helicase (RH) function in human neurodevelopment has been recently brought to the attention of the scientific community by a growing number of human genetic studies. Pathogenic germline variants have been identified in *DDX3X* (Snijders Blok et al., 2015), *DDX6* (Balak et al., 2019) and *DDX59* (Salpietro et al., 2018) in individuals affected by neurodevelopmental disorders (NDDs). Moreover, the potential implication of paralog genes encoding members of the DExD/H-box RNA Helicase family such as *DHX16*, *DHX34*, *DHX37*, *DDX54* in human NDDs suggested that variant alleles in other components of this family, including *DHX8*, *DDX47* and *DHX58*, might also be genetic alterations underlying NDDs (Paine et al., 2019). However, the majority of the reported alterations was never scrutinized by an experimental approach and therefore the molecular and cellular consequences of the identified mutations remain so far largely unknown.

One exception is represented by DHX30, which, similar to the other helicases mentioned above, is a member of the superfamily 2 (SF2) of RHs. SF2 contains more than 50 human members characterized by the consensus amino acid sequence DExD or DExH in their NTP-binding motif II and thus termed DDX and DHX proteins, respectively (Umate et al., 2011). Crystal structure analysis of other family members including the *S. cerevisiae* spliceosome associated-helicase Prp43 and the *Drosophila* RNA helicase Mle (male-less) provides the framework for functional and genetic analysis of all DEAH helicases (He et al., 2010; Prabu et al., 2015). Based on these structures, a modular domain organization has been proposed for DHX30 (**Fig. 3**). The catalytic module consists of two domains that resemble the bacterial recombination protein recombinase A, RecA-1 and RecA-2 (aa 444-808). These two domains are packed tightly around the ATP or ADP molecule (Caruthers & McKay, 2002; Fairman-Williams et al., 2010; He et al., 2010; Singleton et al., 2007). The helicase core region is flanked by two dsRBDs (aa 53-338) at the N-terminus, and a winged-helix domain (WHD) (aa 809-878), a ratchet-like domain (RL) (aa 879-940) and an oligosaccharide-binding (OB)-like domain (aa 1016-1102) at the C-terminus. Whereas the C-terminal extension is highly conserved in DEAH-box helicases, the conservation of the N-terminal extension is dramatically lower, exemplified by the presence in Mle of two dsRBDs which are instead missing in Prp43 (Prabu et al., 2015; Tauchert et al., 2017). In Prp43 the WHD interacts extensively with RecA-1. It is followed by a seven-helical bundle corresponding to the RL domain (He et al., 2010). This domain was originally named based on structural homology with the DNA helicase Hel308 (Buttner et al., 2007). However, as recent studies on Mle revealed that this domain has no ratcheting function, I will refer to it as the ratchet-like domain. The carboxyl-terminal domain (CTD) of Mle also includes the OB-like domain which contains a helix and a five-stranded β -barrel motif. The OB-like domain makes intimate contact with the dsRBD2, RecA-2 and the WHD (Prabu et al., 2015). Specifically, together with RecA-2 and the WHD, it forms the 5' portion of the RNA-binding channel (Prabu et al., 2015; Tauchert et al., 2017). Within the helicase core region eight highly conserved motifs have been predicted to mediate either ATP binding/hydrolysis or RNA recognition. Motifs I and II, also referred to as Walker A and Walker B motifs, respectively, are involved in ATP-binding (Walker et al., 1982). Namely, motif I bears a loop structure called the P loop that forms a pocket to bind the phosphates of the ATP molecule. Motif II participates in ATP hydrolysis by interacting with the β and γ phosphates through a coordinated Mg^{2+} ion (Tanner & Linder, 2001). Motif VI binds the γ phosphate and coordinates, together with motif I and II, ATP binding and hydrolysis in other DExH family members (Caruthers & McKay, 2002; Tanner & Linder, 2001). Motifs Ia, Ib and IV are likely involved in RNA-binding while motif III has been suggested to couple ATP binding and hydrolysis with conformational changes needed for RNA unwinding (Tanner & Linder, 2001). Motif V might have a role in both RNA binding and interaction with ATP (Lessel et al., 2017). Whereas structure and

properties of the catalytic core of DExH-box helicases have been well characterized, the mechanisms by which the auxiliary domains regulate the function of individual DExH proteins are poorly understood.

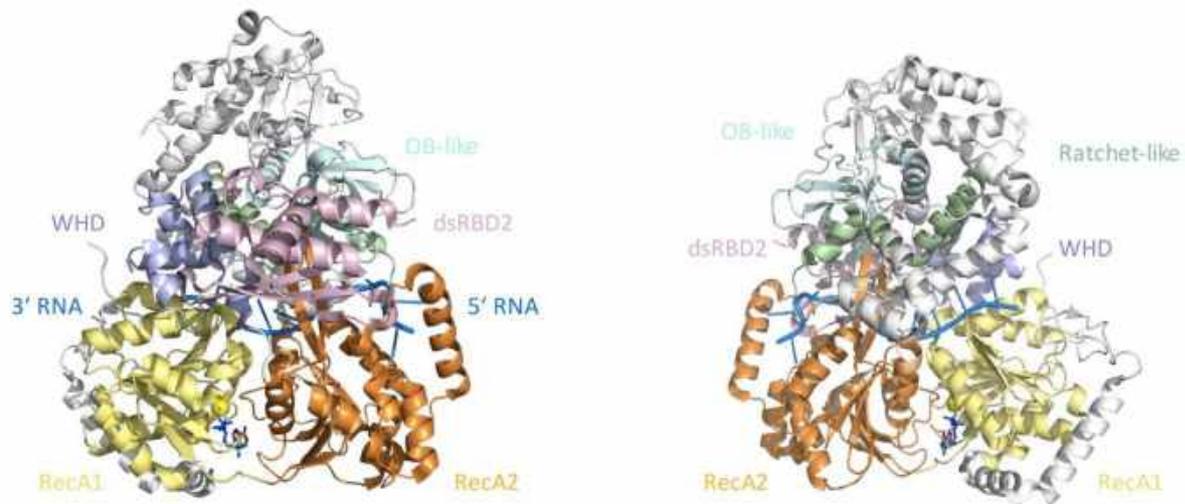


Figure 3 | Crystal structure of the Mle-RNA-ADP-AIF₄ complex. A truncated version of Mle lacking the last 130 amino acids was purified and subjected to proteolysis yielding a stable 120 kDa fragment lacking dsRBD1. This fragment was crystallized in the presence of a U₁₅ RNA and ADP-AIF₄ (aluminium fluoride) to mimic the transition state of the ATPase reaction (Prabu et al., 2015). The structure is represented in two orientations rotated by 180° around the vertical axis. The N-terminal domain containing the dsRBD2 is colored in pink, the two RecA domains that are tightly packed around the ADP molecule, in yellow and orange, the WHD, the Ratchet-like and the OB domains at the C-terminus in violet, green and blue marine, respectively. The RNA molecule is depicted in blue. Coordinates and structure factors have been obtained from the Protein Data Bank (PDB). DOI: [10.2210/pdb5AOR/pdb](https://doi.org/10.2210/pdb5AOR/pdb) NDB: [NA3382](https://www.rcsb.org/entry/NA3382). This figure was generated using PyMOL Molecular Graphics System.

The potential role of *DHX30* in neuronal development was initially suggested by the observation that biallelic loss of the *DHX30* ortholog in mice, *HelG*, during murine embryogenesis led to early developmental defects in the central nervous system (CNS) and lethality during early embryonic development in mice (Zheng et al., 2015). Eldomery et al. (2017) confirmed this hypothesis in an international large-scale sequencing study by classifying *DHX30* as a candidate gene for human NDDs. Ultimately, the identification of 6 *de novo* missense mutations in *DHX30* in 12 unrelated individuals severely affected by GDD and ID (Lessel et al., 2017), has provided even stronger evidence for assessing the indispensability of this gene in the development of the CNS. Each of the identified mutations was found to lead to the substitution of a conserved residue within the highly conserved helicase core region. Specifically, p.R493H identified in two individuals affected motif Ia, p.H562R, identified in a single individual, affected motif II. p.G781D, identified in two individuals, p.R782W, identified in three individuals, p.R785C also identified in three individuals and p.R785H, unique, all resulted in amino acid alterations residing in motif VI. When mutations in these highly conserved motifs were introduced in Prp43, they were found to either interfere with the splicing activity of this protein, to result in a cold sensitive growth retardation (Martin et al., 2002; Tanaka & Schwer, 2006) or to be lethal (Martin et al., 2002). All *DHX30* mutations identified by Lessel et al. (2017) were exceedingly rare: their absence from different public catalogues of human variation and genotype data (dbSNP, 1000 Genomes, the ExAC or the gnomAD browser) indicates that they represent rare variants. Additionally, *DHX30* has been identified as a gene in which missense mutations are strongly selected against, as it is predicted to be very intolerant to missense mutations based on the ExAC (Exome Aggregation Consortium) sequencing data of about 64.000 healthy individuals. Here *DHX30* was ranked at position 31 out of 18.000 analyzed genes by its missense Z

score of 6.82 (Lek et al., 2016), which is even higher than the average Z score for genes involved in developmental disorders (Samocha et al., 2014). Thus, the genetic data by themselves, together with published structural and functional findings on other RHs, already provided evidence that the patient mutations in *DHX30* were very likely to be deleterious. These findings were then corroborated by several cellular and functional assays. Specifically, all six variants were found to interfere with either RNA binding or ATPase activity in *in vitro* assays. The observation that each of the identified mutations led to aberrant formation of SGs and concomitant impairment of translation highlighted the importance of proper translation in neurodevelopment and began to uncover a previously unknown role for DHX30 in SG assembly and global translation control. However, at this point, the exact physiological function of DHX30 in cellular RNA metabolism was still mostly unknown. So far, our understanding of the cellular function of DHX30 relied on a single proteomic analysis performed by Antonicka and Shoubridge (Antonicka & Shoubridge, 2015). In an attempt to characterize the proteome of mitochondrial RNA granules, they uncovered a prominent role for DHX30 in the assembly of mitochondrial ribosomes. These findings supported a previous observation that a specific isoform of DHX30 localized to RNA granules in the vicinity of mitochondrial DNA nucleoids (Wang & Bogenhagen, 2006). According to the Mitop2 software, this specific isoform (XM_011533492.2) was predicted to have a mitochondrial signal peptide of 43 amino acids with a 99% probability of mitochondrial import. However, none of the patients presented by Lessel et al. (2017) exhibited a typical mitochondrial phenotype suggesting that the pathological phenotype of these patients was likely to be determined by other, non-mitochondrial isoforms of DHX30.

With this thesis, and the associated publication by Mannucci et al. (2021), we have aimed to provide further understanding of the DHX30-related NDDs and novel insights into the molecular and cellular function of DHX30. In this work, we expanded the clinical and molecular spectrum of DHX30-related NDDs by identifying 25 previously unreported individuals carrying mutations in *DHX30*. Interestingly, although some of the new patients were recruited through Gene-Matcher and our network of collaborators, the majority of new patients was enrolled through a social media-based family support group, an observation that highlights the usefulness of social media in the definition of novel Mendelian disorders.

Similar to the 12 patients previously reported by Lessel and colleagues (Lessel et al., 2017), 19 of the newly identified ones carry heterozygous missense variants affecting highly conserved residues within helicase core motifs (HCMs) and present with a severe NDD. Within the same cohort, we identified 2 individuals carrying mutations that affect residues outside HCMs. Specifically, the p.R725H variant affects the linker region between Motif IV and Motif V and the p.R908Q variant affects a highly conserved residue within the RL domain, outside the helicase core region. The clinical presentation of these two individuals clearly differs from that of the patients carrying HCM missense mutations, thereby suggesting a different type of pathomechanism. In contrast to all other cases, the homozygous R725H mutation was inherited from two heterozygous, healthy parents. The patient was severely affected and died in early infancy. The p.R908Q mutant was associated with initially normal development, followed by development of ataxia and cognitive problems at the age of 8 years. Due to the singular nature of these mutations, it was not immediately possible to confirm their causality. Therefore, future studies will benefit from the identification of similar variants in similarly affected individuals. This observation further highlights the importance of the recurrence of mutations as a prerequisite to establish their pathogenicity as in the case of the HCM mutations.

Among the 25 newly identified individuals, 4 carry heterozygous variants resulting in either haploinsufficiency or synthesis of truncated proteins and present with a relatively milder clinical course in comparison to the patients carrying HCM missense mutations. This observation supports the previous hypothesis that missense variants affecting a HCM have a more severe effect than the loss of one copy of the gene (Lessel et al., 2017).

Clinical and genetic findings were then corroborated by a broad spectrum of *in vitro* and *in vivo* assays. In line with the observations from the previous study (Lessel et al., 2017), all novel missense mutations affecting one of the highly conserved ATP binding/hydrolysis motifs within the helicase core region (p.G462E, p.A734D, p.S737F, p.T739A and p.R782Q) resulted in reduced ATPase activity. Sequence similarities to other RNA helicases led to the identification of DHX30 as an RNA helicase, but so far this had not been experimentally verified. As RNA helicases unwind RNA secondary structures and remodel RNP complexes by using the energy of ATP hydrolysis, ATPase activity is a prerequisite for helicase activity. Based on the previous observation that ATPase activity of DHX30 is strongly stimulated by the presence of exogenous RNA (Lessel et al., 2017), we assessed here whether DHX30 can also unwind partially double-stranded RNA duplex substrates by using the free energy provided by ATP hydrolysis. Incubation of a radiolabeled RNA duplex with increasing amounts of purified WT protein revealed that DHX30 unwinds the RNA duplex in a dose dependent manner, as outlined by the appearance of a single-stranded RNA and the concomitant disappearance of the double-stranded RNA substrate. Interestingly, 10 ng of purified DHX30 were sufficient to unwind the RNA-duplex thereby confirming that DHX30 indeed possesses the RNA unwinding activity. Importantly, the omission of ATP entirely disrupted the helicase activity, showing that DHX30 acts as an ATP dependent RNA helicase.

In agreement with these observations, it is not surprising that mutations affecting ATPase activity result in a loss of helicase activity as well. In addition, the p.R493H variant, which was shown to interfere with RNA binding but not ATPase activity (Lessel et al., 2017) appears here to disrupt the ability to unwind RNA, similar to the ATPase activity-deficient mutants. Obviously, both ATPase activity and RNA binding are necessary for proper helicase activity.

All missense mutations in HCMs (from this study, and from the previous study by Lessel et al., 2017) showed an increased propensity to trigger SG formation and concomitantly inhibited global translation. ATP hydrolysis was suggested to trigger the release of RNA targets from DEAD-box ATPases thus resulting in the disassembly of RNA-containing membrane-less organelles (Hondele et al., 2019). Given that, one might speculate that the increased propensity to trigger SG formation observed in the ATPase activity-deficient mutants may be a consequence of their inability to catalyze the ATP-dependent release of their RNA client. However, the fact that overexpression of p.R493H also leads to hyper-assembly of SGs suggests that the lack of RNA unwinding activity might also be responsible for a longer dwelling time of the mutant on the mRNA target resulting in aberrant formation of SGs. The hypothesis that mutations interfering with the RNA unwinding activity of a protein slow down its ability to release the target was previously proposed for missense mutations identified in *AGO2* (Lessel et al., 2020), as discussed above.

Additionally, in order to provide novel insights into the impact of DHX30 missense variants *in vivo*, the effect of selected HCM missense variants, namely p.R493H and p.R785C, was analyzed by our collaborators in a zebrafish model. Both mutations resulted in embryonic developmental defects in 75-90% of embryos suggesting that these mutations interfere with normal embryonic development.

Given the differences in terms of both the localization within the structure and the clinical presentation of the respective patients when compared to the HCM mutations, the homozygous missense mutation p.R725H and the heterozygous missense mutation p.R908Q necessitate a separate discussion. Specifically, the p.R725H variant, was found to impair ATPase activity without triggering SG hyper-assembly. Similar to DHX30-WT, its distribution is clearly cytoplasmic diffuse. As it was not possible to purify this variant from bacteria, its impact on helicase activity could not be tested. However, the unsuccessful attempts to purify it might provide an indication of the nature of this variant. Indeed, likely due to misfolding of the protein and consequent deposition of the insoluble product in inclusion bodies or its direct degradation, this biallelic variant might lead to a loss of function effect. The p.R908Q variant impairs ATPase but not helicase activity

of DHX30, a finding that uncovers a potential role for the RL domain in coupling helicase activity to ATP hydrolysis. A similar effect has previously been described for two mutations in IGHMBP2, a ribosome-associated helicase. Interestingly, mutations in IGHMBP2 have been shown to disrupt its enzymatic activity thereby causing degeneration of α -motoneurons in distal spinal muscular atrophy type 2. Specifically, these two mutations in IGHMBP2 impair ATP hydrolysis but not RNA unwinding and yet result in a pathological phenotype similar to the remaining nine mutations found in this gene (Guenther et al., 2009). This suggests that the loss of either both ATPase and helicase activity or ATPase activity alone, as in the case of p.R908Q, might equally contribute to the pathogenicity of DHX30 missense mutations. However, the observations that overexpression of GFP-tagged R908Q resulted in localization to ATXN2-positive clusters in only 50% of transfected cells challenge the pathogenicity of this variant. The developmental defects observed in zebrafish embryos upon expression of both p.R725H and p.R908Q variants actually support the pathogenicity of these two mutations. Moreover, the inclusion of two common non-synonymous DHX30 variants found in gnomAD, namely p.V556I and p.E948K, in the functional analysis not only highlights the robustness of the finding in respect to the pathogenicity of the HCM variants but also provides some clarity into the conflicting results observed for p.R725H and p.R908Q variants. Noteworthy, p.V556I and p.E948K, despite having a similar location as p.R725H and p.R908Q (within the helicase core region but outside HCMs, the first one and within the RL domain, the second one) display ATPase activity similar to DHX30-WT, remain mostly diffuse in the cytoplasm and have little or no impact on zebrafish embryonic development. These findings, together with the functional characterization of the two variants, point to a potential pathogenic role and might establish a genotype-phenotype correlation. However, the lack of similarly affected individuals carrying similar variants, together with the possibility that these two individuals carry additional variants previously undetected by trio-whole exome sequencing, do not allow for a definitive establishment of their causality.

Whereas HCM mutations spontaneously stimulate aberrant formation of SGs and concomitant impairment of translation resulting in a severe phenotype, mutations affecting residues outside HCMs only partially trigger assembly of SGs leading to a relatively milder phenotypic outcome. This observation indicates that depending on the location of mutations within the structure, DHX30 variants behave differently with respect to SG assembly. This suggests a clear correlation between the hyper-ability of mutations to induce SG formation and the severity of the phenotype. Given the central role of SG hyper-assembly in determining the pathogenicity of DHX30 missense mutations, I further investigated the potential function of DHX30-WT in this aspect of RNA metabolism. I therefore generated a CRISPR/Cas9 based knockout (KO) of DHX30 in HEK293T cells. DHX30 KO led to a significant reduction of SG-positive cells upon incubation at 43.5°C, a condition upon which SGs are routinely observed and endogenous DHX30 is recruited to these SGs in normal cells (Lessel et al., 2017). Similarly, our collaborators confirmed this result by showing that homozygous *dhx30* mutant zebrafish (-/-) display significantly lower SG formation upon heat shock at 42°C when compared to *dhx30*-WT animals (+/+). These findings uncover a previously unknown, essential and evolutionary conserved role of DHX30 in SG assembly. Delineating the exact function of DHX30 in SG assembly/disassembly will be the focus of future extensive work. Careful dissection of multiple signalling pathways leading to SG formation in DHX30-deficient cells will help to understand at which step DHX30 is crucial for the regulation of this rather delicate aspect of mRNA metabolism.

Importantly, these findings shed light on the conflicting results obtained in our attempt to clarify the nature of the missense mutations, and to allow for a distinction between a “dominant negative” and a “loss-of-function” pathomechanism. Two results point to a dominant negative effect of HCM mutations: (1) the coexpression of RFP-tagged DHX30-WT and several GFP-tagged DHX30 mutants resulted in each case in the

recruitment of the WT into Ataxin-2 positive cytoplasmic clusters. (2) Co-injection of DHX30-WT with p.R493H and p.R785C in zebrafish only partially rescued the embryonic developmental defects induced by the mutants alone. On the other hand, as observed by our collaborators, the addition of DHX30-WT to increasing amounts of p.H562R and p.R785C mutant forms of DHX30 completely rescued their inability to unwind the RNA duplex (although only partially when added to increasing amounts of p.R493H). This finding suggests that HCM mutations cause a loss of function with respect to helicase activity. The observation that KO of DHX30 impairs SG formation and the fact that HCM missense mutations induce SG assembly even without endogenous or exogenous stressors, actually point to a detrimental gain of function effect. Thus, the functional role played by DHX30-WT in SG formation appears to be excessively active in HCM mutant forms of the protein.

In order to bridge the gap between the cellular defect and the brain phenotype and to provide further understanding of the genotype-phenotype correlation observed for both the HCM and the loss of function variants, our collaborators performed behavioral modeling of *dhx30*-deficient zebrafish. When compared to WT (+/+) and heterozygous (+/-) siblings, the homozygous *dhx30*-deficient (-/-) animals displayed social behavioral deficits and compromised sleep/wake behaviors partially resembling the clinical features of the DHX30-related NDD. The developmental and brain-specific changes observed in zebrafish might be useful to carry forward in a mouse model for future analyses. Specifically, the generation of a DHX30 knock-in mouse will allow to analyse the impact of *DHX30* missense mutations in brain development and function and to infer molecular and functional properties of DHX30 in the development of the central nervous system.

This PhD thesis and the attached publication by Mannucci et al. (2021) has expanded the clinical and genetic spectrum of the DHX30-related NDD and has provided substantial improvements in understanding the nature of the pathogenic variants. Clinical and genetic findings, corroborated by several functional *in vitro* and *in vivo* assays, contributed to the identification of two distinct clinical subtypes defined by location and nature of the pathogenic variants. Missense variants affecting HCMs cause a severe phenotype by disrupting ATPase and/or helicase activity and by leading to a detrimental gain of function in respect to SG formation. By contrast, DHX30 loss of function variants cause a milder phenotype thereby confirming that the loss of one copy of the gene is less deleterious than the occurrence of missense variants affecting a helicase core motif. Importantly, this study has provided further insights into the role of DHX30 in cellular RNA metabolism by formally establishing this protein as an ATP-dependent RNA helicase and an evolutionary conserved factor in SG assembly. However, further extensive studies will be required to elucidate the roles of DHX30 in different aspects of RNA metabolism. Specifically, given that our cellular and zebrafish data point to a role of DHX30 in the regulation of mRNA metabolism, a further emerging question is whether transcriptome and proteome changes will be observed upon loss of DHX30 in neuronal as well as non-neuronal cells. Elucidating whether loss of DHX30 affects the translational status of specific mRNAs and identifying which RNA targets directly bind to DHX30 will be the focus of future work. Moreover, investigating whether pathogenic variants of DHX30 result in different RNA binding sites and whether there are changes in the translational status of iPSC-derived neurons obtained from patients will contribute to further clarify the molecular and cellular dysfunction associated with *DHX30* missense mutations. Generating iPSC-derived neurons from patient fibroblasts will also offer the opportunity for potential therapeutic intervention. Targeting SGs by using small molecules known to affect SG dynamics might represent a possible strategy to treat the disorders associated with DHX30 dysfunction.

4.3 | Do AGO2 and DHX30 operate in similar pathways during neuronal development?

Most of the functions of RNA in eukaryotic cells are carried out together with proteins in the form of ribonucleoprotein (RNP) complexes (Cech, 2012). Indeed, a previous proteomic analysis of RNP complexes formed by Argonaute proteins revealed that DHX30 is specifically found in these RNPs (Hock et al., 2007). The results of this initial affinity purification experiment followed by mass spectrometry analysis suggested a role for DHX30 as a potential interaction partner of AGO2. In order to confirm this finding, I performed co-immunoprecipitation assays in cells transfected with GFP-tagged DHX30 WT and found that AGO2 specifically co-precipitates with overexpressed DHX30-WT.

Given the established role of AGO2 in RNA interference, these findings prompted me to investigate the potential involvement of DHX30 in shRNA-mediated gene silencing. This possibility was tested using the CRISPR/Cas9-based DHX30 deficient cell line (Mannucci et al., 2021). I investigated whether in the absence of DHX30, an shRNA vector delivered into cells was still able to suppress the expression of its respective target gene. It is worth noting that the efficiency of this vector in reducing the expression of the target gene was previously established in the lab (Lessel et al., 2020). Comparisons between the magnitude of the target gene repression in the presence versus the absence of DHX30 revealed a silencing efficiency of 77% and 71% respectively, thus not supporting the hypothesis that DHX30 might have a role in RNAi.

Although these data exclude the possibility that DHX30 and AGO2 cooperate in the context of RNAi to regulate post-transcriptional gene expression, both our works by Lessel et al. (2020) and Mannucci et al. (2021) highlight the importance of these two RBPs in contributing to mRNA compartmentalization during mRNP trafficking. Whereas it is quite well established that AGO2 carries out its repression function within GW-bodies/P-bodies due to its interaction with GW182 (Yao et al., 2013), our work by Mannucci et al. (2021) has just started to address a possible role for DHX30 in SG assembly. However, the exact step of the SG life cycle in which DHX30 is involved is yet to be determined. Careful dissection of the SG assembly/disassembly process along with an extensive investigation of the signalling pathways leading to SG formation will be necessary to gain further insights into the role of DHX30 within this aspect of mRNA metabolism. Despite the differences in terms of cellular function and localization, the common point of convergence between *AGO2* and *DHX30* is represented by their essential role as neurodevelopmental-associated genes. Biallelic disruption of both *Ago2* and *HelG/DHX30* during murine embryogenesis results in early developmental defects in the central nervous system (CNS) and lethality during early embryonic development in mice (Liu et al., 2004; Zheng et al., 2015). Based on different public catalogues of human variation and genotype data, *AGO2* and *DHX30* are predicted to be very intolerant to missense mutations. According to the genome aggregation consortium (gnomAD), both *AGO2* and *DHX30* are extremely loss of function intolerant, as predicted by their probability of being loss-of function intolerant (pLI) score of 1. Therefore, mutations occurring in these two genes are very likely to be deleterious. None of the *AGO2* and *DHX30* alterations analyzed in our two publications was present in publicly available datasets, indicating that they represent extremely rare variants, most likely associated to disease. Missense mutations identified in *AGO2* and *DHX30* were found to lead to the substitution of conserved amino acid residues within the respective catalytic core. Interestingly, both the structural domains of *AGO2* (Swarts et al., 2014) and the helicase core motifs of *DHX30* (Lessel et al., 2017) are evolutionary conserved. However, whereas the molecular function of *AGO2* structural domains (namely N-terminal, PAZ, MID and PIWI domains) has been extensively characterized (Wu et al., 2020), the role of the highly conserved motifs within the *DHX30* helicase core region has been so far determined only based on structural similarities and published structures of other SF2 RNA helicases (Lessel et al., 2017). The functional analysis of missense mutations in *DHX30* performed here and in the previous publication by Lessel et al. (2017) have contributed to infer putative molecular and functional properties of

DHX30 helicase core motifs in respect to ATP binding/hydrolysis and RNA binding/unwinding activities. Taken together, the results of both the genetic and functional analyses of *AGO2* and *DHX30* missense mutations indicate that substitutions of highly conserved amino acid residues within the catalytic core of RNA binding proteins result in impaired enzymatic activity.

The phenotypic consequences of *AGO2* and *DHX30* dysfunction reflect the importance of proper post-transcriptional regulation of gene expression in the development of the CNS. Despite the differences in respect to the pathomechanism underlying the *AGO2* and *DHX30*-associated NDDs, the clinical features of the affected individuals are often overlapping. Indeed, all individuals presented in these two studies display intellectual disability, motor developmental delay and impaired speech ability. Additional common features were muscular hypotonia, gait abnormalities, structural brain anomalies, features of autism spectrum disorder, strabismus, feeding difficulties and seizures. The identification of further affected individuals will contribute not only to establish novel genotype-phenotype correlations and eventually different clinical subtypes but also to clarify the causality of variants classified as “likely pathogenic”.

Conclusion

Regulation of gene expression at the post-transcriptional level has recently emerged as a key factor for development and maintenance of complex neural circuits (Swanger & Bassell, 2011). Therefore, it is not surprising that altered functions of RNA-binding proteins (RBPs), as critical effectors of RNA metabolism, underlie the origin of many human neurological disorders (Lukong et al., 2008). The introduction of several high-throughput approaches to identify RBPs and their RNA binding sites has led to the discovery of more than 4000 candidate RBPs, many of which are mutated in genetic diseases. Subsequent Gene Ontology analysis showed that the majority of these potentially disease-linked RBPs are associated with metabolic processes and development of the nervous system, further highlighting the close relationship between RNA regulation and neuropathologies (Gebauer et al., 2021). However, the lack of a systematic experimental investigation of the phenotypic consequences of mutations in RBPs does not allow for a mechanistic understanding of the pathology and even less, for therapy opportunities. Among the very few exceptions, *AGO2* and *DHX30* are two examples of disease-associated RBPs for which the genetic and clinical findings were extensively corroborated by an experimental approach. Although in both cases the genetic data *per se* provided strong evidence for the pathogenicity of the mutations identified in these two genes, the functional analysis has significantly contributed to provide deeper understanding of the underlying pathomechanism and to clarify the genotype-phenotype correlations. Moreover, given that missense mutations identified in both *AGO2* and *DHX30* result in aberrant formation of P-bodies and SGs, respectively, targeting these mRNP granules could represent a potential therapeutic strategy for the treatment of the respectively associated diseases.

References

- Alberti, S., Gladfelter, A., & Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell*, *176*(3), 419-434. doi:10.1016/j.cell.2018.12.035
- Alrahbeni, T., Sartor, F., Anderson, J., Miedzybrodzka, Z., McCaig, C., & Muller, B. (2015). Full UPF3B function is critical for neuronal differentiation of neural stem cells. *Mol Brain*, *8*, 33. doi:10.1186/s13041-015-0122-1
- Anderson, P., & Kedersha, N. (2002). Stressful initiations. *J Cell Sci*, *115*(Pt 16), 3227-3234.
- Anderson, P., & Kedersha, N. (2008). Stress granules: the Tao of RNA triage. *Trends Biochem Sci*, *33*(3), 141-150. doi:10.1016/j.tibs.2007.12.003
- Andrei, M. A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., & Luhrmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA*, *11*(5), 717-727. doi:10.1261/rna.2340405
- Antonicka, H., & Shoubridge, E. A. (2015). Mitochondrial RNA Granules Are Centers for Posttranscriptional RNA Processing and Ribosome Biogenesis. *Cell Rep*, *10*(6), 920-932. doi:10.1016/j.celrep.2015.01.030
- Aulas, A., Fay, M. M., Lyons, S. M., Achorn, C. A., Kedersha, N., Anderson, P., & Ivanov, P. (2017). Stress-specific differences in assembly and composition of stress granules and related foci. *J Cell Sci*, *130*(5), 927-937. doi:10.1242/jcs.199240
- Aylett, C. H., & Ban, N. (2017). Eukaryotic aspects of translation initiation brought into focus. *Philos Trans R Soc Lond B Biol Sci*, *372*(1716). doi:10.1098/rstb.2016.0186
- Bain, J. M., Cho, M. T., Telegrafi, A., Wilson, A., Brooks, S., Botti, C....Chung, W. K. (2016). Variants in HNRNPH2 on the X Chromosome Are Associated with a Neurodevelopmental Disorder in Females. *Am J Hum Genet*, *99*(3), 728-734. doi:10.1016/j.ajhg.2016.06.028
- Balak, C., Benard, M., Schaefer, E., Iqbal, S., Ramsey, K., Ernoult-Lange, M....Piton, A. (2019). Rare De Novo Missense Variants in RNA Helicase DDX6 Cause Intellectual Disability and Dysmorphic Features and Lead to P-Body Defects and RNA Dysregulation. *Am J Hum Genet*, *105*(3), 509-525. doi:10.1016/j.ajhg.2019.07.010
- Bardoni, B., Abekhouk, S., Zongaro, S., & Melko, M. (2012). Intellectual disabilities, neuronal posttranscriptional RNA metabolism, and RNA-binding proteins: three actors for a complex scenario. *Prog Brain Res*, *197*, 29-51. doi:10.1016/B978-0-444-54299-1.00003-0
- Bicknell, A. A., & Ricci, E. P. (2017). When mRNA translation meets decay. *Biochem Soc Trans*, *45*(2), 339-351. doi:10.1042/BST20160243
- Bregues, M., Teixeira, D., & Parker, R. (2005). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science*, *310*(5747), 486-489. doi:10.1126/science.1115791
- Buchan, J. R. (2014). mRNP granules. Assembly, function, and connections with disease. *RNA Biol*, *11*(8), 1019-1030. doi:10.4161/15476286.2014.972208
- Buchan, J. R., Kolaitis, R. M., Taylor, J. P., & Parker, R. (2013). Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell*, *153*(7), 1461-1474. doi:10.1016/j.cell.2013.05.037
- Burnett, J. C., Rossi, J. J., & Tiemann, K. (2011). Current progress of siRNA/shRNA therapeutics in clinical trials. *Biotechnol J*, *6*(9), 1130-1146. doi:10.1002/biot.201100054
- Buttner, K., Nehring, S., & Hopfner, K. P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol*, *14*(7), 647-652. doi:10.1038/nsmb1246
- Caruthers, J. M., & McKay, D. B. (2002). Helicase structure and mechanism. *Curr Opin Struct Biol*, *12*(1), 123-133. doi:10.1016/s0959-440x(02)00298-1

- Cech, T. R. (1986). A model for the RNA-catalyzed replication of RNA. *Proc Natl Acad Sci U S A*, 83(12), 4360-4363. doi:10.1073/pnas.83.12.4360
- Cech, T. R. (2012). The RNA worlds in context. *Cold Spring Harb Perspect Biol*, 4(7), a006742. doi:10.1101/cshperspect.a006742
- Chiu, S. Y., Lejeune, F., Ranganathan, A. C., & Maquat, L. E. (2004). The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex. *Genes Dev*, 18(7), 745-754. doi:10.1101/gad.1170204
- Chu, C. Y., & Rana, T. M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol*, 4(7), e210. doi:10.1371/journal.pbio.0040210
- Chung, W. J., Okamura, K., Martin, R., & Lai, E. C. (2008). Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr Biol*, 18(11), 795-802. doi:10.1016/j.cub.2008.05.006
- Cooke, C., & Alwine, J. C. (1996). The cap and the 3' splice site similarly affect polyadenylation efficiency. *Mol Cell Biol*, 16(6), 2579-2584. doi:10.1128/mcb.16.6.2579
- de Nadal, E., Ammerer, G., & Posas, F. (2011). Controlling gene expression in response to stress. *Nat Rev Genet*, 12(12), 833-845. doi:10.1038/nrg3055
- Decker, C. J., & Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb Perspect Biol*, 4(9), a012286. doi:10.1101/cshperspect.a012286
- Decker, C. J., Teixeira, D., & Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol*, 179(3), 437-449. doi:10.1083/jcb.200704147
- Detzer, A., Engel, C., Wunsche, W., & Sczakiel, G. (2011). Cell stress is related to re-localization of Argonaute 2 and to decreased RNA interference in human cells. *Nucleic Acids Res*, 39(7), 2727-2741. doi:10.1093/nar/gkq1216
- Dever, T. E., Dinman, J. D., & Green, R. (2018). Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb Perspect Biol*, 10(8). doi:10.1101/cshperspect.a032649
- Dever, T. E., & Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb Perspect Biol*, 4(7), a013706. doi:10.1101/cshperspect.a013706
- Ding, S. W., & Voinnet, O. (2007). Antiviral immunity directed by small RNAs. *Cell*, 130(3), 413-426. doi:10.1016/j.cell.2007.07.039
- Eldomery, M. K., Coban-Akdemir, Z., Harel, T., Rosenfeld, J. A., Gambin, T., Stray-Pedersen, A....Lupski, J. R. (2017). Lessons learned from additional research analyses of unsolved clinical exome cases. *Genome Med*, 9(1), 26. doi:10.1186/s13073-017-0412-6
- Emara, M. M., Fujimura, K., Sciaranghella, D., Ivanova, V., Ivanov, P., & Anderson, P. (2012). Hydrogen peroxide induces stress granule formation independent of eIF2alpha phosphorylation. *Biochem Biophys Res Commun*, 423(4), 763-769. doi:10.1016/j.bbrc.2012.06.033
- Eystathiou, T., Chan, E. K., Tenenbaum, S. A., Keene, J. D., Griffith, K., & Fritzler, M. J. (2002). A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. *Mol Biol Cell*, 13(4), 1338-1351. doi:10.1091/mbc.01-11-0544
- Eystathiou, T., Jakymiw, A., Chan, E. K., Seraphin, B., Cougot, N., & Fritzler, M. J. (2003). The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA*, 9(10), 1171-1173. doi:10.1261/rna.5810203
- Fairman-Williams, M. E., Guenther, U. P., & Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol*, 20(3), 313-324. doi:10.1016/j.sbi.2010.03.011

- Farny, N. G., Kedersha, N. L., & Silver, P. A. (2009). Metazoan stress granule assembly is mediated by P-eIF2alpha-dependent and -independent mechanisms. *RNA*, *15*(10), 1814-1821. doi:10.1261/rna.1684009
- Filipowicz, W. (1978). Functions of the 5'-terminal m7G cap in eukaryotic mRNA. *FEBS Lett*, *96*(1), 1-11. doi:10.1016/0014-5793(78)81049-7
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, *391*(6669), 806-811. doi:10.1038/35888
- Friedman, R. C., Farh, K. K., Burge, C. B., & Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res*, *19*(1), 92-105. doi:10.1101/gr.082701.108
- Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J....Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. *Mol Cell*, *56*(1), 79-89. doi:10.1016/j.molcel.2014.09.005
- Ganassi, M., Mateju, D., Bigi, I., Mediani, L., Poser, I., Lee, H. O....Carra, S. (2016). A Surveillance Function of the HSPB8-BAG3-HSP70 Chaperone Complex Ensures Stress Granule Integrity and Dynamism. *Mol Cell*, *63*(5), 796-810. doi:10.1016/j.molcel.2016.07.021
- Garneau, N. L., Wilusz, J., & Wilusz, C. J. (2007). The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol*, *8*(2), 113-126. doi:10.1038/nrm2104
- Gebauer, F., & Hentze, M. W. (2004). Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol*, *5*(10), 827-835. doi:10.1038/nrm1488
- Gebauer, F., Schwarzl, T., Valcarcel, J., & Hentze, M. W. (2021). RNA-binding proteins in human genetic disease. *Nat Rev Genet*, *22*(3), 185-198. doi:10.1038/s41576-020-00302-y
- Gennarino, V. A., Palmer, E. E., McDonnell, L. M., Wang, L., Adamski, C. J., Koire, A....Zoghbi, H. Y. (2018). A Mild PUM1 Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures. *Cell*, *172*(5), 924-936 e911. doi:10.1016/j.cell.2018.02.006
- Gerstberger, S., Hafner, M., & Tuschl, T. (2014). A census of human RNA-binding proteins. *Nat Rev Genet*, *15*(12), 829-845. doi:10.1038/nrg3813
- Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet*, *10*(2), 94-108. doi:10.1038/nrg2504
- Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L. M., & Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell*, *15*(12), 5383-5398. doi:10.1091/mbc.e04-08-0715
- Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett*, *582*(14), 1977-1986. doi:10.1016/j.febslet.2008.03.004
- Golden, R. J., Chen, B., Li, T., Braun, J., Manjunath, H., Chen, X....Mendell, J. T. (2017). An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. *Nature*, *542*(7640), 197-202. doi:10.1038/nature21025
- Grudzien-Nogalska, E., & Kiledjian, M. (2017). New insights into decapping enzymes and selective mRNA decay. *Wiley Interdiscip Rev RNA*, *8*(1). doi:10.1002/wrna.1379
- Guenther, U. P., Handoko, L., Laggenbauer, B., Jablonka, S., Chari, A., Alzheimer, M....Fischer, U. (2009). IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1). *Hum Mol Genet*, *18*(7), 1288-1300. doi:10.1093/hmg/ddp028
- Gunderson, S. I., Vagner, S., Polycarpou-Schwarz, M., & Mattaj, I. W. (1997). Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes Dev*, *11*(6), 761-773. doi:10.1101/gad.11.6.761

- Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, *466*(7308), 835-840. doi:10.1038/nature09267
- Gwon, Y., Maxwell, B. A., Kolaitis, R.-M., Zhang, P., Kim, H. J., & Taylor, J. P. (2021). Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner. *Science*, *372*(6549). doi:10.1126/science.abf6548
- Hagerman, R. J., Berry-Kravis, E., Hazlett, H. C., Bailey, D. B., Jr., Moine, H., Kooy, R. F....Hagerman, P. J. (2017). Fragile X syndrome. *Nat Rev Dis Primers*, *3*, 17065. doi:10.1038/nrdp.2017.65
- Hamilton, A. J., & Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, *286*(5441), 950-952. doi:10.1126/science.286.5441.950
- Hamm, J., & Mattaj, I. W. (1990). Monomethylated cap structures facilitate RNA export from the nucleus. *Cell*, *63*(1), 109-118. doi:10.1016/0092-8674(90)90292-m
- He, Y., Andersen, G. R., & Nielsen, K. H. (2010). Structural basis for the function of DEAH helicases. *EMBO Rep*, *11*(3), 180-186. doi:10.1038/embor.2010.11
- Hellen, C. U. T. (2018). Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harb Perspect Biol*, *10*(10). doi:10.1101/cshperspect.a032656
- Hendrickson, D. G., Hogan, D. J., McCullough, H. L., Myers, J. W., Herschlag, D., Ferrell, J. E., & Brown, P. O. (2009). Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol*, *7*(11), e1000238. doi:10.1371/journal.pbio.1000238
- Hock, J., Weinmann, L., Ender, C., Rudel, S., Kremmer, E., Raabe, M....Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep*, *8*(11), 1052-1060. doi:10.1038/sj.embor.7401088
- Holt, C. E., Martin, K. C., & Schuman, E. M. (2019). Local translation in neurons: visualization and function. *Nat Struct Mol Biol*, *26*(7), 557-566. doi:10.1038/s41594-019-0263-5
- Hondele, M., Sachdev, R., Heinrich, S., Wang, J., Vallotton, P., Fontoura, B. M. A., & Weis, K. (2019). DEAD-box ATPases are global regulators of phase-separated organelles. *Nature*, *573*(7772), 144-148. doi:10.1038/s41586-019-1502-y
- Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Bálint, E., Tuschl, T., & Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science*, *293*(5531), 834-838. doi:10.1126/science.1062961
- Ivanov, P., Kedersha, N., & Anderson, P. (2019). Stress Granules and Processing Bodies in Translational Control. *Cold Spring Harb Perspect Biol*, *11*(5). doi:10.1101/cshperspect.a032813
- Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol*, *11*(2), 113-127. doi:10.1038/nrm2838
- Jain, S., Wheeler, J. R., Walters, R. W., Agrawal, A., Barsic, A., & Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell*, *164*(3), 487-498. doi:10.1016/j.cell.2015.12.038
- Jinek, M., & Doudna, J. A. (2009). A three-dimensional view of the molecular machinery of RNA interference. *Nature*, *457*(7228), 405-412. doi:10.1038/nature07755
- Jonas, S., & Izaurralde, E. (2013). The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev*, *27*(24), 2628-2641. doi:10.1101/gad.227843.113
- Kaida, D. (2016). The reciprocal regulation between splicing and 3'-end processing. *Wiley Interdiscip Rev RNA*, *7*(4), 499-511. doi:10.1002/wrna.1348

- Karim, M. E., Tha, K. K., Othman, I., Borhan Uddin, M., & Chowdhury, E. H. (2018). Therapeutic Potency of Nanoformulations of siRNAs and shRNAs in Animal Models of Cancers. *Pharmaceutics*, *10*(2). doi:10.3390/pharmaceutics10020065
- Kedersha, & Anderson, P. (2007). Mammalian Stress Granules and Processing Bodies. In *Translation Initiation: Cell Biology, High-Throughput Methods, and Chemical-Based Approaches* (pp. 61-81).
- Kedersha, Chen, S., Gilks, N., Li, W., Miller, I. J., Stahl, J., & Anderson, P. (2002). Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol Biol Cell*, *13*(1), 195-210. doi:10.1091/mbc.01-05-0221
- Kedersha, Cho, M. R., Li, W., Yacono, P. W., Chen, S., Gilks, N....Anderson, P. (2000). Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J Cell Biol*, *151*(6), 1257-1268. doi:10.1083/jcb.151.6.1257
- Kedersha, Panas, M. D., Achorn, C. A., Lyons, S., Tisdale, S., Hickman, T....Anderson, P. (2016). G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *J Cell Biol*, *212*(7), 845-860. doi:10.1083/jcb.201508028
- Kedersha, Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M. J....Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol*, *169*(6), 871-884. doi:10.1083/jcb.200502088
- Kim, J. H., Shinde, D. N., Reijnders, M. R. F., Hauser, N. S., Belmonte, R. L., Wilson, G. R....Ahn, E. Y. E. (2016). De Novo Mutations in SON Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectual-Disability Syndrome. *Am J Hum Genet*, *99*(3), 711-719. doi:10.1016/j.ajhg.2016.06.029
- Kimball, S. R., Horetsky, R. L., Ron, D., Jefferson, L. S., & Harding, H. P. (2003). Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. *Am J Physiol Cell Physiol*, *284*(2), C273-284. doi:10.1152/ajpcell.00314.2002
- Klum, S. M., Chandradoss, S. D., Schirle, N. T., Joo, C., & MacRae, I. J. (2018). Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition. *Embo j*, *37*(1), 75-88. doi:10.15252/emboj.201796474
- Kole, R., Krainer, A. R., & Altman, S. (2012). RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov*, *11*(2), 125-140. doi:10.1038/nrd3625
- Konarska, M. M., Padgett, R. A., & Sharp, P. A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell*, *38*(3), 731-736. doi:10.1016/0092-8674(84)90268-x
- Kremer, E. J., Yu, S., Pritchard, M., Nagaraja, R., Heitz, D., Lynch, M....et al. (1991). Isolation of a human DNA sequence which spans the fragile X. *Am J Hum Genet*, *49*(3), 656-661.
- Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, *31*(1), 147-157. doi:10.1016/0092-8674(82)90414-7
- Kyburz, A., Friedlein, A., Langen, H., & Keller, W. (2006). Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Mol Cell*, *23*(2), 195-205. doi:10.1016/j.molcel.2006.05.037
- Łabno, A., Tomecki, R., & Dziembowski, A. (2016). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochim Biophys Acta*, *1863*(12), 3125-3147. doi:10.1016/j.bbamcr.2016.09.023
- Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *Embo j*, *23*(20), 4051-4060. doi:10.1038/sj.emboj.7600385

- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T....Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616), 285-291. doi:10.1038/nature19057
- Leppek, K., Das, R., & Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol*, 19(3), 158-174. doi:10.1038/nrm.2017.103
- Lessel, D., Schob, C., Kury, S., Reijnders, M. R. F., Harel, T., Eldomery, M. K....Kreienkamp, H. J. (2017). De Novo Missense Mutations in DHX30 Impair Global Translation and Cause a Neurodevelopmental Disorder. *Am J Hum Genet*, 101(5), 716-724. doi:10.1016/j.ajhg.2017.09.014
- Lessel, D., Zeitler, D. M., Reijnders, M. R. F., Kazantsev, A., Hassani Nia, F., Bartholomaeus, A....Kreienkamp, H. J. (2020). Germline AGO2 mutations impair RNA interference and human neurological development. *Nat Commun*, 11(1), 5797. doi:10.1038/s41467-020-19572-5
- Leung, A. K., Calabrese, J. M., & Sharp, P. A. (2006). Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc Natl Acad Sci U S A*, 103(48), 18125-18130. doi:10.1073/pnas.0608845103
- Leung, A. K., & Sharp, P. A. (2006). Function and localization of microRNAs in mammalian cells. *Cold Spring Harb Symp Quant Biol*, 71, 29-38. doi:10.1101/sqb.2006.71.049
- Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J....Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, 305(5689), 1437-1441. doi:10.1126/science.1102513
- Liu, J., Valencia-Sanchez, M. A., Hannon, G. J., & Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol*, 7(7), 719-723. doi:10.1038/ncb1274
- Loschi, M., Leishman, C. C., Berardone, N., & Boccaccio, G. L. (2009). Dynein and kinesin regulate stress-granule and P-body dynamics. *J Cell Sci*, 122(Pt 21), 3973-3982. doi:10.1242/jcs.051383
- Lukong, K. E., Chang, K. W., Khandjian, E. W., & Richard, S. (2008). RNA-binding proteins in human genetic disease. *Trends Genet*, 24(8), 416-425. doi:10.1016/j.tig.2008.05.004
- Lund, E., & Dahlberg, J. E. (2006). Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb Symp Quant Biol*, 71, 59-66. doi:10.1101/sqb.2006.71.050
- Luo, Y., Na, Z., & Slavoff, S. A. (2018). P-Bodies: Composition, Properties, and Functions. *Biochemistry*, 57(17), 2424-2431. doi:10.1021/acs.biochem.7b01162
- Mahboubi, H., & Stochaj, U. (2017). Cytoplasmic stress granules: Dynamic modulators of cell signaling and disease. *Biochim Biophys Acta Mol Basis Dis*, 1863(4), 884-895. doi:10.1016/j.bbadis.2016.12.022
- Mannucci, I., Dang, N. D. P., Huber, H., Murry, J. B., Abramson, J., Althoff, T....Lessel, D. (2021). Genotype-phenotype correlations and novel molecular insights into the DHX30-associated neurodevelopmental disorders. *Genome Med*, 13(1), 90. doi:10.1186/s13073-021-00900-3
- Martin, A., Schneider, S., & Schwer, B. (2002). Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem*, 277(20), 17743-17750. doi:10.1074/jbc.M200762200
- Mazroui, R., Huot, M. E., Tremblay, S., Filion, C., Labelle, Y., & Khandjian, E. W. (2002). Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Hum Mol Genet*, 11(24), 3007-3017. doi:10.1093/hmg/11.24.3007
- Meijer, H. A., Smith, E. M., & Bushell, M. (2014). Regulation of miRNA strand selection: follow the leader? *Biochem Soc Trans*, 42(4), 1135-1140. doi:10.1042/BST20140142

- Muller-McNicoll, M., & Neugebauer, K. M. (2013). How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat Rev Genet*, *14*(4), 275-287. doi:10.1038/nrg3434
- Ng, C. K., Shboul, M., Taverniti, V., Bonnard, C., Lee, H., Eskin, A....Reversade, B. (2015). Loss of the scavenger mRNA decapping enzyme DCPS causes syndromic intellectual disability with neuromuscular defects. *Hum Mol Genet*, *24*(11), 3163-3171. doi:10.1093/hmg/ddv067
- Niemi, M. E. K., Martin, H. C., Rice, D. L., Gallone, G., Gordon, S., Kelemen, M....Barrett, J. C. (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*, *562*(7726), 268-271. doi:10.1038/s41586-018-0566-4
- Nojima, T., Hirose, T., Kimura, H., & Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *J Biol Chem*, *282*(21), 15645-15651. doi:10.1074/jbc.M700629200
- Nonhoff, U., Ralser, M., Welzel, F., Piccini, I., Balzereit, D., Yaspo, M. L....Krobitsch, S. (2007). Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. *Mol Biol Cell*, *18*(4), 1385-1396. doi:10.1091/mbc.e06-12-1120
- Olina, A. V., Kulbachinskiy, A. V., Aravin, A. A., & Esyunina, D. M. (2018). Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry (Mosc)*, *83*(5), 483-497. doi:10.1134/S0006297918050024
- Pabis, M., Neufeld, N., Steiner, M. C., Bojic, T., Shav-Tal, Y., & Neugebauer, K. M. (2013). The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA*, *19*(8), 1054-1063. doi:10.1261/rna.037069.112
- Paine, I., Posey, J. E., Grochowski, C. M., Jhangiani, S. N., Rosenheck, S., Kleyner, R....Lupski, J. R. (2019). Paralog Studies Augment Gene Discovery: DDX and DHX Genes. *Am J Hum Genet*, *105*(2), 302-316. doi:10.1016/j.ajhg.2019.06.001
- Pakos-Zebrucka, K., Koryga, I., Mnich, K., Lujic, M., Samali, A., & Gorman, A. M. (2016). The integrated stress response. *EMBO Rep*, *17*(10), 1374-1395. doi:10.15252/embr.201642195
- Panas, M. D., Ivanov, P., & Anderson, P. (2016). Mechanistic insights into mammalian stress granule dynamics. *J Cell Biol*, *215*(3), 313-323. doi:10.1083/jcb.201609081
- Parker, R., & Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. *Mol Cell*, *25*(5), 635-646. doi:10.1016/j.molcel.2007.02.011
- Pestova, T. V., & Hellen, C. U. (2001). Functions of eukaryotic factors in initiation of translation. *Cold Spring Harb Symp Quant Biol*, *66*, 389-396. doi:10.1101/sqb.2001.66.389
- Prabu, J. R., Muller, M., Thomae, A. W., Schussler, S., Bonneau, F., Becker, P. B., & Conti, E. (2015). Structure of the RNA Helicase MLE Reveals the Molecular Mechanisms for Uridine Specificity and RNA-ATP Coupling. *Mol Cell*, *60*(3), 487-499. doi:10.1016/j.molcel.2015.10.011
- Quévillon Huberdeau, M., Zeitler, D. M., Hauptmann, J., Bruckmann, A., Fressigné, L., Danner, J....Meister, G. (2017). Phosphorylation of Argonaute proteins affects mRNA binding and is essential for microRNA-guided gene silencing in vivo. *Embo j*, *36*(14), 2088-2106. doi:10.15252/embj.201696386
- Rajman, M., & Schratt, G. (2017). MicroRNAs in neural development: from master regulators to fine-tuners. *Development*, *144*(13), 2310-2322. doi:10.1242/dev.144337
- Ramanathan, A., Robb, G. B., & Chan, S. H. (2016). mRNA capping: biological functions and applications. *Nucleic Acids Res*, *44*(16), 7511-7526. doi:10.1093/nar/gkw551
- Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., & Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA*, *11*(11), 1640-1647. doi:10.1261/rna.2191905

- Reijns, M. A., Alexander, R. D., Spiller, M. P., & Beggs, J. D. (2008). A role for Q/N-rich aggregation-prone regions in P-body localization. *J Cell Sci*, *121*(Pt 15), 2463-2472. doi:10.1242/jcs.024976
- Riggs, C. L., Kedersha, N., Ivanov, P., & Anderson, P. (2020). Mammalian stress granules and P bodies at a glance. *J Cell Sci*, *133*(16). doi:10.1242/jcs.242487
- Roy, B., & Jacobson, A. (2013). The intimate relationships of mRNA decay and translation. *Trends Genet*, *29*(12), 691-699. doi:10.1016/j.tig.2013.09.002
- Salpietro, V., Efthymiou, S., Manole, A., Maurya, B., Wiethoff, S., Ashokkumar, B....Houlden, H. (2018). A loss-of-function homozygous mutation in DDX59 implicates a conserved DEAD-box RNA helicase in nervous system development and function. *Hum Mutat*, *39*(2), 187-192. doi:10.1002/humu.23368
- Samocha, K. E., Robinson, E. B., Sanders, S. J., Stevens, C., Sabo, A., McGrath, L. M....Daly, M. J. (2014). A framework for the interpretation of de novo mutation in human disease. *Nat Genet*, *46*(9), 944-950. doi:10.1038/ng.3050
- Sartor, F., Anderson, J., McCaig, C., Miedzybrodzka, Z., & Muller, B. (2015). Mutation of genes controlling mRNA metabolism and protein synthesis predisposes to neurodevelopmental disorders. *Biochem Soc Trans*, *43*(6), 1259-1265. doi:10.1042/BST20150168
- Saxton, R. A., & Sabatini, D. M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, *168*(6), 960-976. doi:10.1016/j.cell.2017.02.004
- Schalk, A., Cousin, M. A., Challman, T. D., Wain, K. E., Powis, Z., Minks, K....Gerard, B. (2020). doi:10.1101/2020.12.23.424103
- Scheller, U., Pfisterer, K., Uebe, S., Ekici, A. B., Reis, A., Jamra, R., & Ferrazzi, F. (2018). Integrative bioinformatics analysis characterizing the role of EDC3 in mRNA decay and its association to intellectual disability. *BMC Med Genomics*, *11*(1), 41. doi:10.1186/s12920-018-0358-6
- Schwartz, D. C., & Parker, R. (1999). Mutations in translation initiation factors lead to increased rates of deadenylation and decapping of mRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol*, *19*(8), 5247-5256. doi:10.1128/mcb.19.8.5247
- Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., & Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, *115*(2), 199-208. doi:10.1016/s0092-8674(03)00759-1
- Shimotohno, K., Kodama, Y., Hashimoto, J., & Miura, K. I. (1977). Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proc Natl Acad Sci U S A*, *74*(7), 2734-2738. doi:10.1073/pnas.74.7.2734
- Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem*, *76*, 23-50. doi:10.1146/annurev.biochem.76.052305.115300
- Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R....Kleefstra, T. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am J Hum Genet*, *97*(2), 343-352. doi:10.1016/j.ajhg.2015.07.004
- Solomon, S., Xu, Y., Wang, B., David, M. D., Schubert, P., Kennedy, D., & Schrader, J. W. (2007). Distinct structural features of caprin-1 mediate its interaction with G3BP-1 and its induction of phosphorylation of eukaryotic translation initiation factor 2alpha, entry to cytoplasmic stress granules, and selective interaction with a subset of mRNAs. *Mol Cell Biol*, *27*(6), 2324-2342. doi:10.1128/MCB.02300-06
- Stewart, M. (2019). Polyadenylation and nuclear export of mRNAs. *J Biol Chem*, *294*(9), 2977-2987. doi:10.1074/jbc.REV118.005594

- Stoecklin, G., & Kedersha, N. (2013). Relationship of GW/P-bodies with stress granules. *Adv Exp Med Biol*, 768, 197-211. doi:10.1007/978-1-4614-5107-5_12
- Stoecklin, G., Stubbs, T., Kedersha, N., Wax, S., Rigby, W. F., Blackwell, T. K., & Anderson, P. (2004). MK2-induced tristetraprolin:14-3-3 complexes prevent stress granule association and ARE-mRNA decay. *Embo j*, 23(6), 1313-1324. doi:10.1038/sj.emboj.7600163
- Sudhakar, A., Ramachandran, A., Ghosh, S., Hasnain, S. E., Kaufman, R. J., & Ramaiah, K. V. (2000). Phosphorylation of serine 51 in initiation factor 2 alpha (eIF2 alpha) promotes complex formation between eIF2 alpha(P) and eIF2B and causes inhibition in the guanine nucleotide exchange activity of eIF2B. *Biochemistry*, 39(42), 12929-12938. doi:10.1021/bi0008682
- Sutton, M. A., & Schuman, E. M. (2006). Dendritic protein synthesis, synaptic plasticity, and memory. *Cell*, 127(1), 49-58. doi:10.1016/j.cell.2006.09.014
- Swanger, S. A., & Bassell, G. J. (2011). Making and breaking synapses through local mRNA regulation. *Curr Opin Genet Dev*, 21(4), 414-421. doi:10.1016/j.gde.2011.04.002
- Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V....van der Oost, J. (2014). The evolutionary journey of Argonaute proteins. *Nat Struct Mol Biol*, 21(9), 743-753. doi:10.1038/nsmb.2879
- Tanaka, N., & Schwer, B. (2006). Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. *Biochemistry*, 45(20), 6510-6521. doi:10.1021/bi052656g
- Tanner, N. K., & Linder, P. (2001). DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell*, 8(2), 251-262. doi:10.1016/s1097-2765(01)00329-x
- Tauchert, M. J., Fourmann, J. B., Luhrmann, R., & Ficner, R. (2017). Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *Elife*, 6. doi:10.7554/eLife.21510
- Tawamie, H., Martianov, I., Wohlfahrt, N., Buchert, R., Mengus, G., Uebe, S....Abou Jamra, R. (2017). Hypomorphic Pathogenic Variants in TAF13 Are Associated with Autosomal-Recessive Intellectual Disability and Microcephaly. *Am J Hum Genet*, 100(3), 555-561. doi:10.1016/j.ajhg.2017.01.032
- Tebbenkamp, A. T., Willsey, A. J., State, M. W., & Sestan, N. (2014). The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. *Curr Opin Neurol*, 27(2), 149-156. doi:10.1097/WCO.0000000000000069
- Teixeira, D., Sheth, U., Valencia-Sanchez, M. A., Brengues, M., & Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA*, 11(4), 371-382. doi:10.1261/rna.7258505
- Thoreen, C. C., Chantranupong, L., Keys, H. R., Wang, T., Gray, N. S., & Sabatini, D. M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature*, 485(7396), 109-113. doi:10.1038/nature11083
- Tokita, M. J., Chow, P. M., Mirzaa, G., Dikow, N., Maas, B., Isidor, B....Prontera, P. (2015). Five children with deletions of 1p34.3 encompassing AGO1 and AGO3. *Eur J Hum Genet*, 23(6), 761-765. doi:10.1038/ejhg.2014.202
- Tourriere, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J. M., Bertrand, E., & Tazi, J. (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol*, 160(6), 823-831. doi:10.1083/jcb.200212128
- Tseng-Rogenski, S. S., & Chang, T. H. (2004). RNA unwinding assay for DExD/H-box RNA helicases. *Methods Mol Biol*, 257, 93-102. doi:10.1385/1-59259-750-5:093
- Umate, P., Tuteja, N., & Tuteja, R. (2011). Genome-wide comprehensive analysis of human helicases. *Commun Integr Biol*, 4(1), 118-137. doi:10.4161/cib.4.1.13844

- Vagner, S., Vagner, C., & Mattaj, I. W. (2000). The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev*, *14*(4), 403-413.
- Valentin-Vega, Y. A., Wang, Y. D., Parker, M., Patmore, D. M., Kanagaraj, A., Moore, J....Taylor, J. P. (2016). Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. *Sci Rep*, *6*, 25996. doi:10.1038/srep25996
- Vissers, L. E., Gilissen, C., & Veltman, J. A. (2016). Genetic studies in intellectual disability and related disorders. *Nat Rev Genet*, *17*(1), 9-18. doi:10.1038/nrg3999
- Walker, J. E., Saraste, M., Runswick, M. J., & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo j*, *1*(8), 945-951.
- Wang, & Bogenhagen, D. F. (2006). Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane. *J Biol Chem*, *281*(35), 25791-25802. doi:10.1074/jbc.M604501200
- Wang, Li, J., Fan, S., Jin, Z., & Huang, C. (2020). Targeting stress granules: A novel therapeutic strategy for human diseases. *Pharmacol Res*, *161*, 105143. doi:10.1016/j.phrs.2020.105143
- Weil, D., Piton, A., Lessel, D., & Standart, N. (2020). Mutations in genes encoding regulators of mRNA decapping and translation initiation: links to intellectual disability. *Biochem Soc Trans*, *48*(3), 1199-1211. doi:10.1042/bst20200109
- Wek, R. C. (2018). Role of eIF2alpha Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harb Perspect Biol*, *10*(7). doi:10.1101/cshperspect.a032870
- Wheeler, J. R., Matheny, T., Jain, S., Abrisch, R., & Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. *Elife*, *5*. doi:10.7554/eLife.18413
- Wickramasinghe, V. O., & Laskey, R. A. (2015). Control of mammalian gene expression by selective mRNA export. *Nat Rev Mol Cell Biol*, *16*(7), 431-442. doi:10.1038/nrm4010
- Wilczynska, A., Aigueperse, C., Kress, M., Dautry, F., & Weil, D. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. *J Cell Sci*, *118*(Pt 5), 981-992. doi:10.1242/jcs.01692
- Wolozin, B., & Ivanov, P. (2019). Stress granules and neurodegeneration. *Nature Reviews Neuroscience*, *20*(11), 649-666. doi:10.1038/s41583-019-0222-5
- Wu, J., Yang, J., Cho, W. C., & Zheng, Y. (2020). Argonaute proteins: Structural features, functions and emerging roles. *J Adv Res*, *24*, 317-324. doi:10.1016/j.jare.2020.04.017
- Yao, B., Li, S., & Chan, E. K. (2013). Function of GW182 and GW bodies in siRNA and miRNA pathways. *Adv Exp Med Biol*, *768*, 71-96. doi:10.1007/978-1-4614-5107-5_6
- Ying, S., & Khapersky, D. A. (2020). UV damage induces G3BP1-dependent stress granule formation that is not driven by mTOR inhibition-mediated translation arrest. *J Cell Sci*, *133*(20). doi:10.1242/jcs.248310
- Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., & Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. *Nat Struct Mol Biol*, *17*(1), 17-23. doi:10.1038/nsmb.1733
- Yu, J. H., Yang, W. H., Gulick, T., Bloch, K. D., & Bloch, D. B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. *Rna*, *11*(12), 1795-1802. doi:10.1261/rna.2142405
- Zamore, Tuschl, T., Sharp, P. A., & Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, *101*(1), 25-33. doi:10.1016/s0092-8674(00)80620-0
- Zheng, H. J., Tsukahara, M., Liu, E., Ye, L., Xiong, H., Noguchi, S....Ji, Z. S. (2015). The novel helicase helG (DHX30) is expressed during gastrulation in mice and has a structure similar to a human DExH box helicase. *Stem Cells Dev*, *24*(3), 372-383. doi:10.1089/scd.2014.0077

Declaration of Contribution

Chapter 1 | Introduction

This chapter was written by Ilaria Mannucci.

Chapter 2 | Germline AGO2 mutations impair RNA interference and human neurological development

Lessel D, Zeitler DM, Reijnders MRF, Kazantsev A, Hassani Nia F, Bartholomäus A, Martens V, Bruckmann A, Graus V, McConkie-Rosell A, McDonald M, Lozic B, Tan ES, Gerkes E, Johannsen J, Denecke J, Telegrafi A, Zonneveld-Huijssoon E, Lemmink HH, Cham BWM, Kovacevic T, Ramsdell L, Foss K, Le Duc D, Mitter D, Syrbe S, Merckenschlager A, Sinnema M, Panis B, Lazier J, Osmond M, Hartley T, Mortreux J, Busa T, Missirian C, Prasun P, Lüttgen S, **Mannucci I**, Lessel I, Schob C, Kindler S, Pappas J, Rabin R, Willemsen M, Gardeitchik T, Löhner K, Rump P, Dias KR, Evans CA, Andrews PI, Roscioli T, Brunner HG, Chijiwa C, Lewis MES, Jamra RA, Dymant DA, Boycott KM, Stegmann APA, Kubisch C, Tan EC, Mirzaa GM, McWalter K, Kleefstra T, Pfundt R, Ignatova Z, Meister G, Kreienkamp HJ.

Nat Commun. 2020 Nov 16;11(1):5797. doi: 10.1038/s41467-020-19572-5. PMID: 33199684; PMCID: PMC7670403.

Ilaria Mannucci contributed to establish experimental conditions for the morphological analysis of neurons, as shown in Fig. 5.

Chapter 3 | Genotype–phenotype correlations and novel molecular insights into the DHX30-associated neurodevelopmental disorders

Mannucci I, Dang NDP, Huber H, Murry JB, Abramson J, Althoff T, Banka S, Baynam G, Bearden D, Beleza-Meireles A, Benke PJ, Berland S, Bierhals T, Bilan F, Bindoff LA, Braathen GJ, Busk ØL, Chenbhanich J, Denecke J, Escobar LF, Estes C, Fleischer J, Groepper D, Haaxma CA, Hempel M, Holler-Managan Y, Houge G, Jackson A, Kellogg L, Keren B, Kiraly-Borri C, Kraus C, Kubisch C, Le Guyader G, Ljungblad UW, Brenman LM, Martinez-Agosto JA, Might M, Miller DT, Minks KQ, Moghaddam B, Nava C, Nelson SF, Parant JM, Prescott T, Rajabi F, Randrianaivo H, Reiter SF, Schuurs-Hoeijmakers J, Shieh PB, Slavotinek A, Smithson S, Stegmann APA, Tomczak K, Tveten K, Wang J, Whitlock JH, Zweier C, McWalter K, Juusola J, Quintero-Rivera F, Fischer U, Yeo NC, Kreienkamp HJ, Lessel D.

Genome Med. 2021 May 21;13(1):90. doi: 10.1186/s13073-021-00900-3. PMID: 34020708; PMCID: PMC8140440.

The following functional experiments were performed by Ilaria Mannucci: ATPase assay (Figure 2a and 2b), immunocytochemistry (Figure 3, 5, 7 and S5), generation of a CRISPR knockout cell line and stress treatment (Figure 7). Helicase assay (Figure 2c, 2d and S4) was performed by Ilaria Mannucci in collaboration with Hannes Huber. The associated figure legends and the corresponding materials and methods sections were written by Ilaria Mannucci for the first version of the manuscript. The results were analyzed by Ilaria Mannucci, Hans-Jürgen Kreienkamp, Davor Lessel, Hannes Huber and Utz Fischer. Figure 1a was designed by Davor Lessel and Hans-Jürgen Kreienkamp and modified by Ilaria Mannucci for the purpose of this manuscript. The corresponding figure legend was written by Ilaria Mannucci.

Chapter 4 | Discussion and Conclusion

This chapter was written by Ilaria Mannucci.

Prof. Dr. rer. nat Hans-Jürgen Kreienkamp

Acknowledgements

To my mentor, Prof. Hans-Jürgen Kreienkamp, thank you for your trust in me as I was a completely inexperienced Master student, having nothing but the will to learn and prove herself within an international environment. Many thanks for giving me then the opportunity to grow even further as a scientist and to develop as a person within a wonderful lab environment.

To my co-advisor, Dr. Davor Lessel, thank you for the inspiring scientific discussions and for pushing me forward when necessary.

To the past and present members of the lab, Edward and Fatemeh, Daniel and Debora, thank you for your kindness and true friendship. This journey would not have been the same all-encompassing experience without you by my side.

To Hans-Hinrich Hönck, thank you for outstanding experimental assistance. I wish my German would have been better!

To the whole Institute of Human Genetics, thank you for creating such a stimulating scientific environment and for your friendly attitude throughout these years.

To my family, thank you with all my heart for accepting my decision and letting me choose my own path, for patiently waiting for me and for your unconditional love. I would not be the person I am today without you being a living reminder of my origin and values. I am eternally grateful.

Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt.

Eidesstattliche Erklärung

Declaration of Oath

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

I hereby declare that I have written the present dissertation independently and without inadmissible third-party assistance, that I have not used any sources or aids other than those acknowledged in the thesis and that I have separately listed all of the literature I employed when producing this academic work, both literally and in content.

Furthermore, I affirm that I have not submitted the dissertation to a representative of another university or applied for the permission to take the final examination at another university.

I agree that my dissertation may be checked by the Dean's Office of the Faculty of Medicine using common plagiarism detection software.

Hamburg, den

Unterschrift