# DNA Hypermethylation and Unstable Repeat Diseases: A Paradigm of Transcriptional Silencing to Decipher the Basis of Pathogenic Mechanisms

**Loredana Poeta [1], Denise Drongitis [1], Lucia Verrillo [1,2] and Maria Giuseppina Miano [1,\***

[1] Institute of Genetics and Biophysics "Adriano Buzzati-Traverso", CNR, 80131 Naples, Italy; poetaloredana@libero.it (L.P.); denise.drongitis@igb.cnr.it (D.D.); lucia.verrillo@igb.cnr.it (L.V.)

[2] Department of Environmental, Biological and Pharmaceutical Sciences and Technologies, University of Campania "Luigi Vanvitelli", 81100 Caserta, Italy

[*] Correspondence: mariag.miano@igb.cnr.it; Tel.: +39-0816132261-445

**Abstract:** Unstable repeat disorders comprise a variable group of incurable human neurological and neuromuscular diseases caused by an increase in the copy number of tandem repeats located in various regions of their resident genes. It has become clear that dense DNA methylation in hyperexpanded non-coding repeats induces transcriptional silencing and, subsequently, insufficient protein synthesis. However, the ramifications of this paradigm reveal a far more profound role in disease pathogenesis. This review will summarize the significant progress made in a subset of non-coding repeat diseases demonstrating the role of dense landscapes of 5-methylcytosine (5mC) as a common disease modifier. However, the emerging findings suggest context-dependent models of 5mC-mediated silencing with distinct effects of excessive DNA methylation. An in-depth understanding of the molecular mechanisms underlying this peculiar group of human diseases constitutes a prerequisite that could help to discover novel pathogenic repeat loci, as well as to determine potential therapeutic targets. In this regard, we report on a brief description of advanced strategies in DNA methylation profiling for the identification of unstable Guanine-Cytosine (GC)-rich regions and on promising examples of molecular targeted therapies for Fragile X disease (FXS) and Friedrich ataxia (FRDA) that could pave the way for the application of this technique in other hypermethylated expansion disorders.

**Keywords:** hypermethylated expansion disorders; DNA hypermethylation-induced transcriptional silencing; neurological and neuromuscular diseases; 5-methylcytosine; CpG site; diagnostic methods; molecular targeted therapy

## 1. Introduction

Microsatellite repeat expansions are implicated in nearly 40 human genetic disorders, including various neurodevelopmental diseases such as Fragile X syndrome (FXS; Mendelian Inheritance in Man/MIM 309550) [1] and Early Infantile Epileptic Encephalopathy (EIEE; MIM 308350) [2], and several neuromuscular disorders such as myotonic dystrophy type 1 (DM1; MIM 160900) [3]. The vast majority of these unstable repeat disorders is caused by expansions of trinucleotide repeat sequences or other repeat tracts, including tetra-, penta-, and hexanucleotide repeats that typically are highly polymorphic and collectively account for about 3% of the human genome [4–6]. This heterogeneous group can be broadly classified in three sub-groups depending on the location of the unstable repeat: i. in 5′ or 3′ regulatory elements, as found in FXS or DM1 diseases, respectively [3,5]; ii. in triplet sequences encoding homo-amino acid tracts, as established in EIEE and Huntington's disease [2,7–9]; and finally, iii. in intronic regions, as detected in amyotrophic lateral sclerosis (ALS; MIM 105400), and in Friedreich's ataxia (FRDA; MIM 229300) [10,11].

Several non-coding repeat diseases present DNA hypermethylation and gene silencing (DNA hypermethylation-induced transcriptional silencing) as a common molecular disease mechanism. Specifically, when the expansion repeats are located within or close to regulatory regions enriched with CpG repeats, an extended wave of methylation of cytosines, forming 5-methylcytosine (5mC), may occur. In the case of dense DNA methylation, this excessive epigenetic modification can repress gene transcription. Additionally, once DNA hypermethylation has taken place, a number of other functional damages may occur, altering, for example, the activity of epigenetic CCCTC binding factor protein (CTCF) or the formation of RNA/DNA hybrids (R-loop) [3,12,13].

Compared to the last review focused on epigenetic mechanisms in triplet repeat expansions, published in 2015 [14], our report looks specifically at the implications of recent findings on the relationship between DNA methylation and repeat diseases, the discovery of new pathogenic mechanisms, the application of innovative diagnostic tools and the development of promising molecular targeted-therapy approaches. Indeed, here, we discuss general features of the paradigm of DNA hypermethylation–transcription silencing in eleven repeat disease loci, some of them recently identified, emphasizing how the dynamic nature of microsatellites reveals a common epigenetic property that, similar to loss-of-function or partial-loss-of function mutations, abrogates or reduces gene expression. Depending on the combination of gene silencing degree with the type of inheritance transmission, as X-linked (XL), autosomal recessive (AR) and autosomal dominant (AD), we underlined how the outcome of this long-term phenomenon can lead to the lack of protein synthesis (null mutation) or insufficient protein synthesis (haploinsufficiency).

We also report on 5-hydroxymethylcytosine (5hmC), whose role as demethylation marker in unstable repeat disorders has strongly emerged in recent years [15–17]. Functionally, 5hmC is formed by the oxidation of 5mC catalyzed by Ten–Eleven Translocation (TET) proteins, a class of modification enzymes erasing DNA methylation in response to internal and external stimuli [18]. Particularly abundant in central nervous system, 5hmC marks promoters and bodies of transcriptionally active genes as well as enhancer elements [19,20]. In recent years, accumulating studies have proved that the conversion of 5mC to 5hmC and vice versa might have a physiological or pathological significance in relation to the degree of demethylation/methylation [15–18].

In addition, we underline the importance of investigating excessive DNA methylation to identify missing heritability, providing further insight into genetic and molecular mechanisms, and to define new diagnostic tools for the accurate detection of new hypermethylated repeat loci. Finally, we describe the application of innovative technologies to target directly the *FMR1*-expanded region in FXS by Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 (CRISPR/Cas9) methodologies, or indirectly through the correction of protein insufficiency in FRDA by synthetic long non-coding RNAs named Short interspersed nuclear element-containing translation UP-regulators (SINEUPs). Intriguingly, the versatility of these methods paves the road for the development of new therapeutic approaches in other hypermethylated expansion disorders.

## 2. Unstable GC-Rich Repeats in 5′ Regulatory Regions

### 2.1. GGC Repeat Expansion with Hypermethylation of XYLT1 Exon 1 in Baratela–Scott Syndrome (BSS)

Baratela–Scott syndrome (BSS; MIM 300881) is a rare autosomal recessive disorder characterized by cognition impairment, short stature, facial dysmorphisms and skeletal dysplasia. It is caused by mutations in Xylosyltransferase 1 (*XYLT1*; MIM 608124), a gene located in 16p12 and involved in biosynthesis of glycosaminoglycan and extracellular matrix formation [21]. Alongside various point mutations and small deletions, a GGC repeat detected in *XYLT1* promoter, close to a small CpG island, was recently found to be expanded and hypermethylated in BSS patients [22]. By applying methylation-specific PCR (MS-PCR) and Southern blot analysis, the authors found dense methylation of the region within 5′ of the *XYLT1* exon 1 and the presence of a larger allele (from 216 to 360 bp) compared to the wild type (WT) allele (ranging from 27 to 60). Thus, by using STRetch software

designed to detect short tandem repeat (STR), a 5′-GGC expansion repeat was mapped in trans with a *XYLT1* sequence variant or a gene deletion. As consequences of hypermethylation, the expression level of the expanded *XYLT1* allele was found being severely reduced in BSS fibroblasts. Notably, *XYLT1* is a gene related to vocal and facial anatomy and it has been recently demonstrated that changes in DNA methylation in this gene correlate with the evolution of vocal tract in humans [23]. Remarkably, most patients with BSS have variations in pitch and voice quality associated to mild or moderate cognition delay [22]. These findings point out the value of investigating DNA hypermethylation to identify missing genetic determinants and also to investigate on potential new epigenetic mechanisms underlining BSS disease.

### 2.2. Unstable Dodecamer Repeat in Progressive Myoclonus Epilepsy of Unverricht–Lundborg Type 1A (EPM1A)

Progressive myoclonic epilepsy of Unverricht–Lundborg type 1A (EPM1A; MIM 254800) is an autosomal recessive disorder characterized by myoclonus seizures, cerebellar ataxia and early onset neurodegeneration (between 6 and 13 years of age). Generally, the disease stabilizes in early adulthood when myoclonus and ataxia improve and only a minimal cognitive impairment remains [24]. EPM1A is caused by mutations in cystatin B (*CSTB*; MIM 601145), a gene located in 21q22 that encodes a protein inhibitor of lysosomal cathepsins present in cortical synapses [25,26]. Detected in the *CSTB* promoter, the expansions of the unstable dodecamer repeat CCCCG-CCCCG-CG were found to be the most frequent disease alleles in EPM1A patients (Figures 1 and 2). Specifically, the *CSTB* dodecamer repeat is normally replicated two or three times in the gene promoter [27] and maps a large CpG island (868-bp long) with a 22.8% of CpG density (Figure 2). In EPM1A patients, expansions of the *CSTB* dodecamer repeat (between 30 and 75 copies) were found to be in association with a severe decrease in *CSTB* transcription and insufficient protein synthesis [28,29]. To date, the molecular mechanisms underlying *CSTB* silencing are largely unknown. Moreover, the presence of an unstable GC-rich tract suggests that, like other repeat diseases located in promoter or 5′untranslated regions (5′UTR), the surrounding CpG island could be hypermethylated, inducing transcriptional repression. However, in the peripheral blood of EPM1A patients, the CpG sites were found not to be subjected to de novo methylation in comparison to individual controls [30,31]. Noteworthily, since DNA methylation has tissue-specific landscapes, up to now, we cannot exclude that *CSTB*-expanded alleles could be specifically marked by hypermethylation in disease-affected tissues. We then conclude that further studies could elucidate whether and how *CSTB*-expanded alleles are differently methylated in the central nervous system of EPM1A patients.
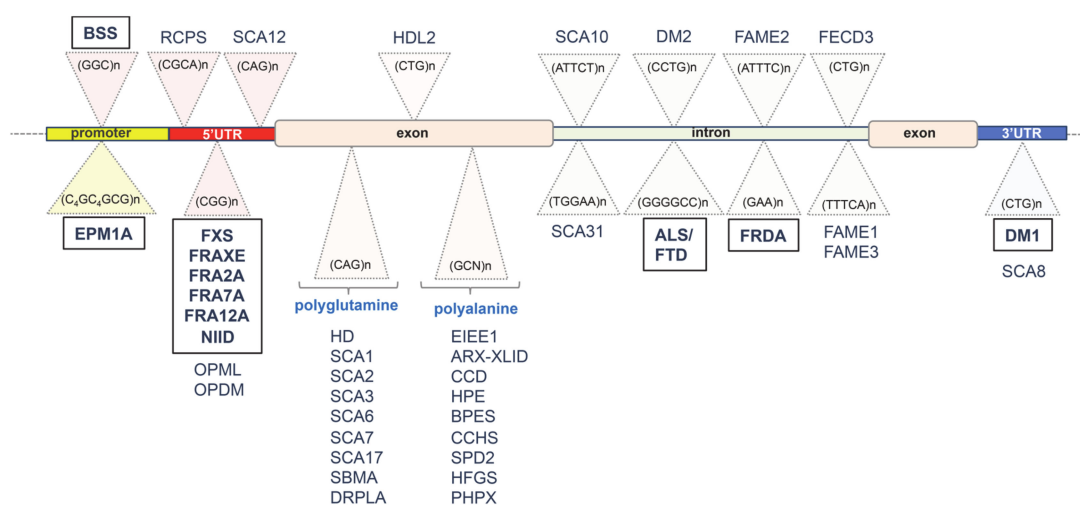


**Figure 1.** To date, more than 40 repeat loci have been identified: diseases and expanded repeats locations (in 5′ or 3′ regulatory elements, triplet tracts encoding homo-amino acid tracts and intronic

regions) are shown. In the box: those discussed in this review. Abbreviations: Amyotrophic lateral sclerosis/Frontotemporal dementia (ALS/FTD). ARX-related X-linked Intellectual disability (ARX-XLID). Blepharophimosis syndrome (BPES). Baratela–Scott syndrome (BSS). Cleidocranial dysplasia (CCD). Congenital central hypoventilation syndrome (CCHS). Myotonic dystrophy type 1 (DM1). Myotonic dystrophy type 2 (DM2). Dentatorubral–pallidoluysian atrophy (DRPLA). Early infantile epileptic encephalopathy type 1 (EIEE1). Progressive myoclonus epilepsy of Unverricht–Lundborg type 1A (EPM1A). Familial adult myoclonic epilepsy type 1 to type 3 (FAME1–3). Fuchs endothelial corneal dystrophy type 3 (FECD3). Fragile site 12A (FRA12A). Fragile site 2A (FRA2A). Fragile site 7A (FRA7A). Fragile site FRAXE (FRAXE). Friedreich's ataxia (FRDA). Fragile X syndrome (FXS). Huntington's disease (HD). Huntington's disease-like-2 (HDL2). Hand–foot–genital syndrome (HFGS). Holoprosencephaly (HPE). Familial neuronal intranuclear inclusion disease (NIID). Oculopharyngeal muscular dystrophy (OPMD). Oculopharyngeal myopathy with leukoencephalopathy 1 (OPML). Panhypopituitarism X-linked (PHPX). Richieri–Costa–Pereira Syndrome (RCPS). Spinal and bulbar muscular atrophy (SBMA). Spinocereberral ataxia type 1 to type 8 (SCA1–8). Spinocereberral ataxia type 10 (SCA10). Spinocereberral ataxia type 12 (SCA12). Spinocereberral ataxia type 17 (SCA17). Spinocereberral ataxia type 31 (SCA31). Synpolydactyly type 2 (SPD2). 5′ untranslated region (5′UTR). 3′ untranslated region (3′UTR).
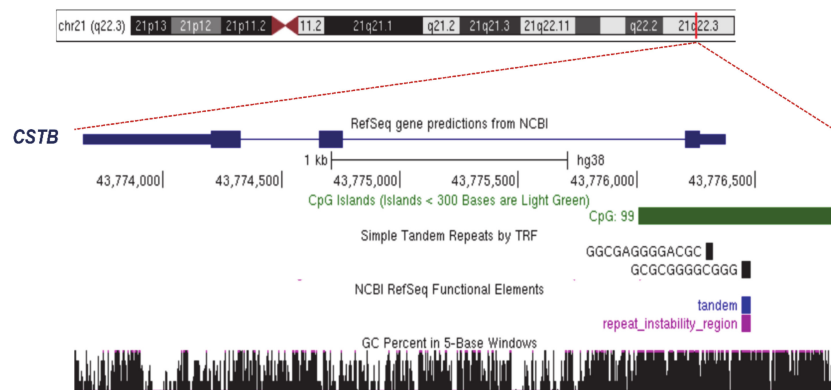


**Figure 2.** The physical map of *CSTB* locus reveals a large CpG island. A region 3321-bp long is shown. chr21:43,773,550-43,776,870 UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly.

### 2.3. Hypermethylated CGG Repeats in Fragile X Syndrome (FXS)

Fragile X syndrome (FXS; MIM 300624) is a well-known neurodevelopmental disorder (NDD) characterized by intellectual disability (ID), behavioural alterations, poor language development and seizures. Usually, males are more severely affected by this disorder than females [1]. FXS is one of the folate-sensitive fragile sites (FSFS), traditionally considered as the archetypal dynamic locus with expanding trinucleotide repeats [1,32]. It is caused by CGG repeat expansions (>200 repeats, full mutation) in the 5′ UTR of the Xq28 gene Fragile mental retardation 1 (*FMR1*; MIM 309550; Figures 1 and 3A; Table 1). This gene encodes the Fragile X mental retardation protein (FMRP), which is an RNA-binding protein involved in synaptic development and plasticity [33]. Beyond the full mutations, *FMR1* premutation expansions (from 55 to 200 CGG repeats) have been found in children with Autism spectrum disorder (ASD; Figure 3B) [33] and in two adult-onset disorders that are Fragile X-associated tremor/ataxia syndrome (FXTAS; MIM 309550) [34] and Fragile X-associated primary ovarian insufficiency (FXPOI; MIM 311360; Figure 3B) [35]. As a consequence of the full mutation, CpG sites upstream of the CGG repeat become methylated (5mC) and *FMR1* transcription is inhibited leading to the absence of its protein product [36]. These events are caused by a wave of DNA hypermethylation that spreads out from the 5′ flanking sequence (~650 bp upstream to CGG triplets) to intron 1 of *FMR1*, and thus crosses over the repeats (Figure 3A) [36]. On the contrary, *FMR1* premutation alleles are generally unmethylated [37] and give rise to ASD or other allelic diseases

through a different molecular process that involves a non- adenine-uracil-guanine (AUG)-initiated translation mechanism (RAN) and causes a RNA toxicity effect [38].

More interestingly, a recent study has demonstrated the incomplete silencing of full *FMR1* mutations in FXS male patients with autistic features [39]. This finding suggests that ID and ASD are clinical features that may vary in FXS in relation to the demethylation level of full mutations.

In the normal *FMR1* allele, 5mC is restricted to a region upstream of the CpG island and is delimited by a DNA methylation boundary. In FXS, this boundary is lost and DNA methylation spreads downstream to the *FMR1* promoter and CGG repeat region, as found in embryonic stem cells (ESCs) and blastocysts [36,40,41]. It has been hypothesized that the resulting chromatin alterations caused by the transition from the unmethylated to the fully methylated region could be responsible for the loss of the methylation barrier, leading to the progression of the methylation into the *FMR1* promoter, which, in turn, becomes silenced [36]. Additionally, the DNA methylation boundary is characterized by the presence of an insulator element that contains four binding sites for the Zinc finger protein CCCTC-binding factor (CTCF), a regulatory protein of gene expression involved in genomic imprinting, chromatin remodeling and DNA conformation change [12]. Noteworthily, the hypermethylation of FXS alleles abrogates the binding to CTCF. However, since drug-induced demethylation of CpGs does not restore the CTCF binding, it has been concluded that other epigenetic factors, such as histone configuration, could be involved [42]. In some cases, inter-individual degree variations in 5mC's conversion into its demethylation product, 5hmC, have been detected within the *FMR1* gene body as well as in proximal flanking regions, both in FXS patient blood-derived cells (PBMC) and brains [16,43]. Further studies analysing the spatial–temporal changes in 5mC/5hmC profiles will provide more information about the pathogenetic significance of the hypermethylation/demethylation correlation in *FXS*.
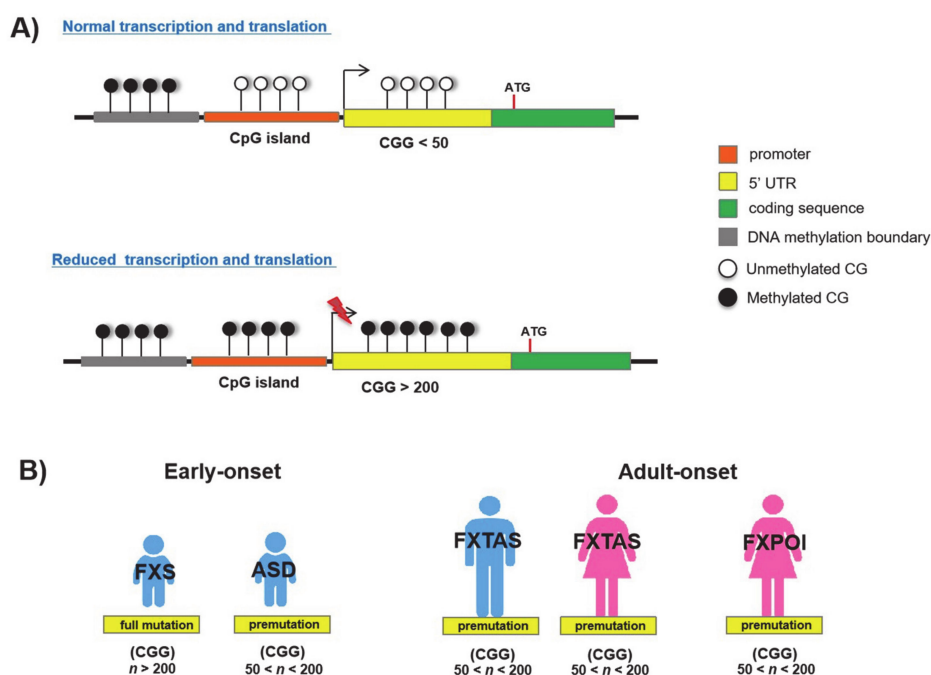


**Figure 3.** CpG island and CGG repeat in *FMR1* locus: (**A**) Normal and hypermethylated promoter regions; (**B**) CGG-expanded repeats in FXS-allelic disorders.

*2.4. Beyond the FMR1 Locus: Hypermethylated GC-Rich Repeats in Rare Folate-Sensitive Fragile Sites (FSFS)*

Apart from the most extensively studied FXS site, excessive DNA methylation has been found to be causative of gene silencing at four rare folate-sensitive fragile sites that are FRAXE, FRA2A, FRA7A and FRA12A, associated with various forms of NDDs with XL or AD inheritance (Figure 1 and

Table 1) [1,44,45]. They are all characterized by CGG/CCG repeat expansions located within the 5′ UTR region of a specific resident disease gene. In a similar manner to the FXS locus, the expanded repeat and any adjacent CpG islands become hypermethylated and the related gene is silenced.

The FRAXE locus (MIM 309548) maps the Xq28 gene AF4/FMR2 family member 2 (*AFF2*), in which 5′UTR CCG-expanded alleles were found in male patients with mild ID, learning defects, attention deficit hyperactivity disorder (ADHD) and ASD [46]. In FRAXE patients, CGG triplets have >200 repeats and the surrounding CpG island becomes hypermethylated, causing the transcriptional repression of *AFF2* [47]. Functionally, *AFF2* is one of the lymphoid nuclear protein related to AF4/FMR2 (ALF) transcription factors involved in neurogenesis, learning and memory processes in mice [48]. It has also been involved in RNA toxicity induced by expanded repeats in *C9ORF72* [49].

In the FRA2A locus (MIM 601464), which maps in 2q11, a CGG repeat in a conserved brain-active alternative promoter located in intron 2 of *AFF3* has been found, an autosomal paralog gene of *AFF2*. In patients with ID, this nucleotide triplet was found to be repeated and the surrounding CpG island was hypermethylated. As expected, in FRA2A carriers, the transcription of *AFF3* was repressed, causing a strong reduction in protein synthesis and therefore determining a haploinsufficiency condition [44].

The FRA7A locus (MIM 616181), maps the 7p11.2 gene Zinc finger protein 713 (*ZNF713*; MIM 61618), which is highly expressed in the brain and encodes a Kruppel-type Zinc finger protein with an unknown function. Detected in 5′ intron of *ZNF713*, a CGG repeat tract (approximately 450 bp) was found expanded in ASD patients [45]. As a consequence of this instability, the adjacent CpG islands turn out to be hypermethylated, the transcription of *ZNF713* is silenced and the protein synthesis is reduced [45].

Located in 12q13.1, the FRA12A locus (MIM 136630) is characterized by a CGG repeat tract (650 to 860 bp) detected in the CpG island in the promotor region of Disco-interacting protein 2 drosophila homolog (*DIP2B; MIM 611379).* This gene encodes a DMAP1-binding domain with a crucial role in DNA methylation machinery [50]. Like other folate-sensitive fragile sites, FRA12A-CGG-expanded alleles (1050 to 1150 bp) were found to be hypermethylated in ID patients. The consequences of this process were the decrease in *DIP2B* expression and haploinsufficiency [50].

Collectively, FXS and FSFS studies are consistent with the concept that exists in this group of NDD disorders—a firm molecular link between the hypermethylation of GC-rich repeats and surrounding CpG islands and gene silencing.

## 2.5. CGG Repeat Instability in Familial Neuronal Intranuclear Inclusion Disease (NIID)

Familial neuronal intranuclear inclusion disease (NIID; MIM 603472) is a rare autosomal dominant disease with progressive neurodegeneration, characterized by neuronal and glial intranuclear inclusions present in the central and peripheral nervous systems. Several independent studies report on the identification of, as causal disease mutations, expansions of the CGG repeat detected in 5′ UTR of NOTCH2 N-Terminal-like C gene (*NOTCH2NLC*; MIM 618025), which maps on chromosome 1q21 [51–53]. Located upstream to a CpG island, this CGG repeat, which normally contains fewer than 43 repeats, may expand and generate a wide range of NIID repeat units (66 to 517). Although very rare, the NOTCH2NLC repeat expansions have also been found in patients with Alzheimer's disease (AD) or frontotemporal dementia-like phenotypes, as well as in patients with leukoencephalopathy [52,54]. In both NIID and allelic disorders, it is not clear whether hypermethylation may repress *NOTCH2NLC* [51–53]. Based on the *FXS* study and other FSFS studies, it has been hypothesized that loci with more than 200 CGG triplets, located close to CpG islands, could be recognized by DNA methyltransferases and thus become hypermethylated. However, an analysis of the DNA methylation of *NOTCH2NLC*-CGG repeats and the surrounding CpG islands in peripheral blood has revealed that not all expanded alleles detected in NIID patients were hypermethylated [51–53]. These ambiguous results suggest that, like other unclear cases, further studies are required to establish whether tissue-specific rounds of DNA hypermethylation may occur in affected tissues.

### 3. Unstable Hypermethylated Repeats in Intronic Regions

*3.1. GGGGCC Expanded Repeat in Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD)*

Amyotrophic lateral sclerosis (ALS; MIM 105550) and frontotemporal dementia (FTD; MIM 600274) are two autosomal dominant neurodegenerative disorders characterized by the adult onset of one or both of these clinical manifestations with high intra-familial variability. Specifically, ALS patients show progressive muscle weakness and atrophy caused by the degeneration of motor neurons, while FTD patients present alterations in behavior and cognition due to the loss of cortical neurons [10]. The most common cause of ALS and FTD is a heterozygous hexanucleotide repeat expansion (GGGGCC) located in intron 1 of the chromosome 9 open reading frame 72 gene (*C9orf72*; MIM 614260) on chromosome 9p21.2 [10]. The hexanucleotide repeat mutation accounts for 20–50% of familial and 5–15% of sporadic ALS and FTD cases [10,55]. In ALS/FTD patients, the number of GGGGCC repeats abnormally expands to more than 30 copies and becomes increasingly unstable (over 2000 repeats) [56]. To date, multiple pathogenic mechanisms for ALS/FTD have been suggested, including haploinsufficiency, abnormal translation (RAN translation) and RNA toxicity [10,55,57]. However, the accumulated data strongly suggest that the hypermethylation of *C9orf72*-expanded alleles contributes to ALS/FTD pathogenesis [58,59]. The GGGGCC repeat sequence is flanked by two CpG islands spanning a region of 1 kb, from the promoter to *C9orf72* intron 1. In control individuals with 5–20 hexanucleotide repeats, this region typically remains unmethylated. On the contrary, in patients with larger expansions, the formation of one large CpG island with a high degree of methylation was detected [58]. Like other hypermethylated repeat expansions, this aberrant epigenetic modification was found to be associated with a decrease in the *C9orf72* mRNA level that, in turn, may trigger haploinsufficiency [58]. Very interestingly, in ALS iPSC-derived neurons and ALS post-mortem brains, the *C9orf72* promoter was found enriched in the 5hmc levels, which suggests a more complex picture of the regulatory framework of *C9orf72* promoter methylation [60]. The functional impact of these processes on disease outcome and progression is still unknown. Noteworthily, *C9orf72* promoter hypermethylation was theorized to also have a protective function because, as direct consequences of *C9orf72* transcriptional silencing, a reduced number of toxic products were observed in patient cells, leading to a reduced loss of neurons in the brain. Further investigation regarding the contribution and timing of *C9orf72* hypermethylation/demethylation in ALS/FTD pathogenesis is required to clarify this dual role [61,62].

*3.2. Intronic GAA Repeat Expansion in Friedreich Ataxia (FRDA)*

Friedreich's ataxia (FRDA) is a lethal autosomal recessive neurodegenerative disorder characterized by gait and limb ataxia, slurred speech, muscle weakness, sensory loss and cardiomyopathy [11]. The primary cause of FDRA is the hyper-expansion of the GAA trinucleotide repeat detected in intron 1 of *FXN*, a gene located in 9q21.1. Generally, normal individuals have five to 30 FXN-GAA repeat expansions, whereas FRDA individuals have from 70 to more than 1000 GAA triplets [11]. As a consequence of the repeated instability, *FXN* transcription is silenced and a marked reduction in the steady-state level of mature frataxin protein is observed (from 4% to 29% compared to that of normal levels) [63]. Additionally, in line with the role of *FXN* in mitochondrial iron homeostasis, FRDA patients present a defective iron–sulphur (Fe–S) biosynthesis [64] and an increased susceptibility to oxidative stress [65]. The proposed mechanism by which *FXN*-GAA repeat expansion leads to the insufficient synthesis of frataxin involves a gene silencing process mediated by rounds of methylation and hydroxymethylation [17,66,67]. Indeed, studies using FRDA patient blood and lymphoblastoid cell lines have revealed that CpG sites located upstream of the GAA repeat are hypermethylated [17,66]. Moreover, an analysis of FRDA human and transgenic mouse brain uncovered a peculiar DNA methylation profile with hypermethylated sites located upstream of the GAA repeat and hypomethylated sites located downstream of the GAA repeat [67]. The methylation centre was located in an Alu sequence that led to a bi-directional spread of DNA methylation.

Interestingly, variations in 5hmC/5mC profiles and decreased CTCF occupancy have been found in the tissues of FRDA patients [17]. Remarkably, the oxidative stress caused by disruptions in the iron metabolism detected in FRDA tissues may in itself prompt modifications to the 5hmC and 5mC landscapes [68]. However, further studies are required to more precisely define the correlation between *FXN*-GAA expansions and levels of DNA methylation.

## 4. Unstable GC-Rich Repeats in 3′ Untranslated Region (3′UTR)

*CTG Triplet Expansion in Myotonic Dystrophy Type 1 (DM1)*

Myotonic dystrophy type 1 (DM1; MIM 160900) is an autosomal dominant multisystem disorder presenting myotonia, muscular dystrophy, cataracts, hypogonadism and heart defects [3]. The disease is caused by a CTG repeat expansion embedded in a large 3.5-kb CpG island detected in the 3′ UTR of the myotonic dystrophia protein kinase (*DMPK; MIM 605377*), a gene located on 19q13.32. Generally, DM1 patients with a larger CTG expansion present a more severe phenotype than those with a smaller one. Based on this feature, separate congenital, infantile, juvenile, adult and late-onset forms of DM1 are described, with respective mean repeat lengths of ~1000, 800, 600, 400, and 200 according to clinical records (Figure 4) [3,69]. Noteworthily, DM1 patients display high levels of CTG instability between different tissues of the same individual that correlate with variations in hypermethylation and clinical variability (Figure 4) [3,69,70]. In DM1-affected human embryonic stem cell (hESC) lines, it has been observed that hypermethylation varies in relation to expansion size [71]. Remarkably, when the CTG repeat tract expands, a reduction in the expression of the neighboring gene six homeobox 5 (*SIX5*; MIM 600963) is commonly reported. This secondary genetic defect causes the peculiar Christmas tree cataract, a rare type of lens opacification typically found in DM1 patients [72]. Additionally, methylation was found to be highly polarized because it was only present at each CpG site located upstream of the CTG repeat (Figure 4). On the contrary, methylation was absent in the downstream CpG site, suggesting that the CTG repeats block methylation progression [73,74]. More interestingly, the hypermethylation of *DM1* alleles was found to functionally impair other epigenetic factors [73,74]. Several studies showed that the hypermethylation of CTCF sites, flanking the CTG triplet, disrupts the binding of the CTCF protein [3,12,75]. Moreover, it has been proposed that the methylation-sensitive CTCF binding sites, located between the CTG repeat tract and replication origin, may affect the replication process and thereby have a role in tissue instability [3].
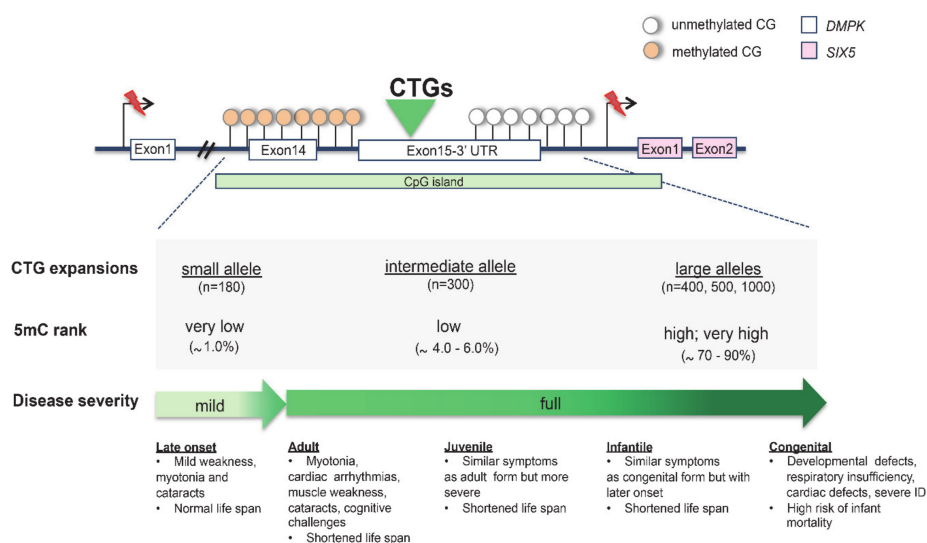


**Figure 4.** Hypermethylated CpG island in expanded CTG repeat of *DMPK*. The correlation between (CTG)n alleles, 5mC levels and clinical phenotypes is shown. The percentage of DNA methylation was analysed in DM1 human embryonic stem cells (hESCs) [71].

**Table 1.** Unstable disease repeats and methylation.

| Location | Repeat | Gene | Locus/ Disease | Inheritance | Hyper- Methylation | Changes in 5hmC | Reference |
|---|---|---|---|---|---|---|---|
| **Promoter** | | | | | | | |
| | (GGC)n | *XYLT1* | BSS | AR | Yes | (?) | [22] |
| | (C₄GC₄GCG)n | *CSTB* | EPM1A | AR | (?) | (?) | [28,31] |
| **5′ UTR** | (CGG)n | *FMR1* | FXS | XL | Yes | Yes | [16,33,36,37,43] |
| | | *AFF2* | FRAXE | XL | Yes | (?) | [46] |
| | | *AFF3* | FRA2A | AD | Yes | (?) | [44] |
| | | *ZNF713* | FRA7A | AD | Yes | (?) | [45] |
| | | *DIP2B* | FRA12A | AD | Yes | (?) | [50] |
| | | *NOTCH2NLC* | NIID | AD | Yes | (?) | [51] |
| **intron** | (GGGGCC)n | *C9orf72* | ALS/FTD | AD | Yes | Yes | [10,60] |
| | (GAA)n | *FXN* | FRDA | AR | Yes | Yes | [11,17,68] |
| **3′ UTR** | (CTG)n | *DMPK* | DM1 | AD | Yes | (?) | [74] |

Abbreviations: Progressive myoclonus epilepsy of Unverricht–Lundborg type 1A (EPM1A). Familial adult myoclonic epilepsy type 1 to type 3 (FAME1–3). Fuchs endothelial corneal dystrophy type 3 (FECD3). Fragile site 12A (FRA12A). Fragile site 2A (FRA2A). Fragile site 7A (FRA7A). Fragile site FRAXE (FRAXE). Friedreich's ataxia (FRDA). Fragile X syndrome (FXS). Huntington's disease (HD). Huntington's disease-like-2 (HDL2). AF4/FMR2 family member 2 (*AFF2*). AF4/FMR2 family member 3 (*AFF3*). Chromosome 9 open reading frame 72 gene (*C9orf72*). Cystatin B (*CSTB*). Disco-interacting protein 2 drosophila homolog (*DIP2B*). Myotonic dystrophy type 1 (DM1). Dystrophia myotonica protein kinase (*DMPK*). Fragile mental retardation 1 (*FMR1*). Frataxin (FXN). N-Terminal-like C NOTCH2 (*NOTCH2NLC*). Xylosyltransferase 1 (*XYLT1*). Zinc finger protein 713 (*ZNF713*). Autosomal recessive (AR). X-linked (XL). Autosomal dominant (AD). Tissue-specific assay to be determined (?).

## 5. Diagnostic Tools

Unstable expanded alleles and hypermethylated expanded alleles are both difficult detect by Next-Generation Short-Read Sequencing (SRS) approaches, since the length of pathogenic expanded repeats (>100 bp) is commonly bigger than the SRS reads. Repeat regions are generally analysed through ad hoc polymerase chain reaction-based assays, even if technical difficulties due to repeat size, consecutive GC regions and DNA methylation or combinations of these factors often interfere with the conventional PCR methods. However, emerging long-read sequencing (LRS) technologies, typified by PacBio single-molecule and real-time (SMRT) sequencing, offer complementary strengths to those of the SRS methods and are well suited to determine expanded tract length and variation in repeat numbers [76]. Applied to Fragile-X diagnosis, LRS technology enabled researchers to completely sequence expanded full mutation *FMR1* alleles, with up to 750 CGG repeats, which translates to >2 kb of 100% CGG-repeat DNA [77,78]. In DM1 patients, LRS sequencing detected de novo repeat interruptions at the DMPK locus, associated with the reduced somatic instability of the repeat expansion [79]. In patients with familial cortical myoclonic tremor with epilepsy (FCMTE; MIM 615400), LRS sequencing identified, as the causative mutations, intronic repeat expansions in *SAMD12* [80]. Recently, various applications of massively parallel sequencing coupled with innovative algorithms have also contributed to identifying new repeat loci, highlighting the importance of considering the genetic contribution of unstable repeats, where these lack a clear genetic cause [51–53,81,82]. In addition, a genome-wide DNA methylation assay has been successfully applied to the molecular diagnosis of genetically unresolved cases, uncovering new hypervariable trinucleotide repeat loci [72,83]. Collectively, emerging Next-Generation Sequencing (NGS) technologies offer efficient strategies in the characterization of expandable regions that are difficult to assess with NGS short-read approaches and, thus, we believe that novel hypermethylated expansion disorders could be identified in the future [72,76,78]. Fascinatingly, these new technologies could lead a new field of studies to trace DNA methylation in repeat ancestors by comparing ancient and modern human genomes and highlighting epigenetic differences at disease repeat loci. This aspect is well integrated with the study of ancient repeat-derived regions, an emerging research area whose development will allow us to deepen our knowledge about the history of regions where disease repeats reside [84].

## 6. Therapeutic Approaches

Currently, there is no cure or treatment for unstable repeat diseases, likely because of the lack of a mechanistic understanding of the specific pathophysiology at the molecular and cellular level. However, DNA hypermethylation-induced transcriptional silencing may turn out to be a potential druggable target in specific unstable repeat diseases. CGG hypermethylated *FMR1* alleles were tested as direct targets of chemicals that inhibit DNA hypermethylation and thus trigger the pharmacological reactivation of *FMR1* gene expression. One of these drugs is 5-azadeoxycytidine (5-aza-dC), which is able to rescue *FMR1* silencing in patient-derived fibroblast cells [85]. As the DNA helix extends, this compound is inserted in place of cytidine and irreversibly and covalently binds to DNA methyltransferase 1 (DNMT1), inhibiting DNA hypermethylation and enhancing the level of 5hmC. Other epigenetic compounds, such as FDA-approved histone deacetylase (HDAC) inhibitors, have been shown to reactivate, even if weakly, *FMR1* expression in Fragile X cell lines [86]. Since these epi-compounds have effects on global DNA methylation, causing several adverse effects, their applicability in clinical trials needs to be accurately investigated [87]. However, the development of new methodologies that allow us to precisely target the molecular defect leads the way to overcoming these limits.

Recently, gene editing using Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) methodologies have been applied to efficiently and directly demethylate the *FMR1* triplet expansions (Figure 5A) [88]. This result was obtained through a new DNA demethylation editing tool named dCas9-Tet1 that combines a catalytically inactivate Cas9 with the DNA methylation modification enzyme Tet1 belonging to TET family (Figure 5A) [89]. Noteworthily, this strategy allows for the targeting of a DNA methylation erasure in a locus specific manner because it combines a locus-specific gene editing with the catalytic activity of TET proteins [90]. In recent years, applications of dCas9-Tet1 have been emerging as powerful tools for the correction of FXS-associated hypermethylation in patient-derived induced pluripotent stem cells (iPSCs) [88,90]. As a consequence of the demethylation editing, the CpG island became hypomethylated, the *FMR1* silencing was unlocked, and the FMRP expression was restored. Furthermore, the injection of FXS-edited iPSCs into murine brains showed that *FMR1* reactivation was also sustained in vivo [88]. However, since reversing DNA methylation was not able to remove all FXS-associated defects, more recently, an antisense oligonucleotide (ASO) strategy has been proposed to selectively block CGG RAN translation and thus reduce RNA toxicity [38]. In combination with the reactivation strategy, ASOs could inhibit the translation of upstream open reading frames (uORFs) and block the toxicity of large demethylated repeat alleles [38]. The CRISPR-based methodology turned out to be a promising and multipurpose strategy to correct *FMR1* silencing. In fact, other groups have used the CRISPR/Cas9 approach to shorten the CGG trinucleotide repeats, rescuing FMRP expression and normalizing its physiological function in FXS patient-derived cells [91,92]. Similarly, other repeat expansion disorders are being targeted with CRISPR-based methodologies through the deletion of the expanded coding allele, as in Huntington disease (HTT) and DM1 disease models [93]. Overall, this evidence supports the wide scope of developing CRISPR-designed strategies to edit DNA hypermethylation in repeat sequences or to shorten the expansions that could be used to reactivate disease genes and thus to reverse phenotypic abnormalities in patient-derived cells.

Recently, an innovative RNA therapeutic technology emerged to balance the final impact caused by DNA methylation-induced transcriptional silencing. Designated as SINE element-containing translation UP-regulators (SINEUPs), these new functional classes of long non-coding RNAs (lncRNAs) are able to enhance the translation process and thus to correct the decrease in protein dosage and haploinsufficiency [94]. These synthetic lncRNAs target the antisense sequence to a selected mRNA and, thus, through a still unknown mechanism, enhance translation and increase protein quantity [94]. Very recently, this therapeutic methodology was successfully applied to increase the endogenous quantities of mature frataxin in FRDA-patient derived cells (Figure 5B) [95]. Specifically, FXN-SINEUP molecules are characterized by i. the binding domain (BD), which is a short sequence that overlaps, in

antisense orientation, the sense frataxin-coding mRNA and ii. the SINEUP effector domain containing the inverted SINEB2 element that confers the activation of protein synthesis [95]. The application of the FXN-SINEUP strategy rescues defective mitochondrial aconitase activity, which is one of the disease hallmarks of FRDA [95]. Previous approaches point to restoring sufficient frataxin levels in FRDA-derived cells and in FRDA animal models by using non-specific drugs, such as the epi-drug Vorinostat, which is an HDAC inhibitor capable of forcing gene transcription [96,97]. Compared to the previous pharmacological strategies proposed, FXN-SINEUP presents the unique advantage of being the first scalable method that allows us to specifically FXN translation, and thus to finely modulate frataxin dosage in FRDA [95]. Given these results, we strongly believe that a large group of repeat diseases associated to haploinsufficiency would benefit from the application of SINEUPs, aiming to restore protein dosage of a specific target when low gene expression is pathogenic.

Although the applications of the targeted methodologies CRISPR/Cas9 and SINEUP in hypermethylated expansion disorders are particularly innovative, it should be emphasized that, at the moment, none of them have been used in pre-clinical and clinical trials. Obviously, every effort in this direction will require important information for the potential transfer of this method to a clinical setting.
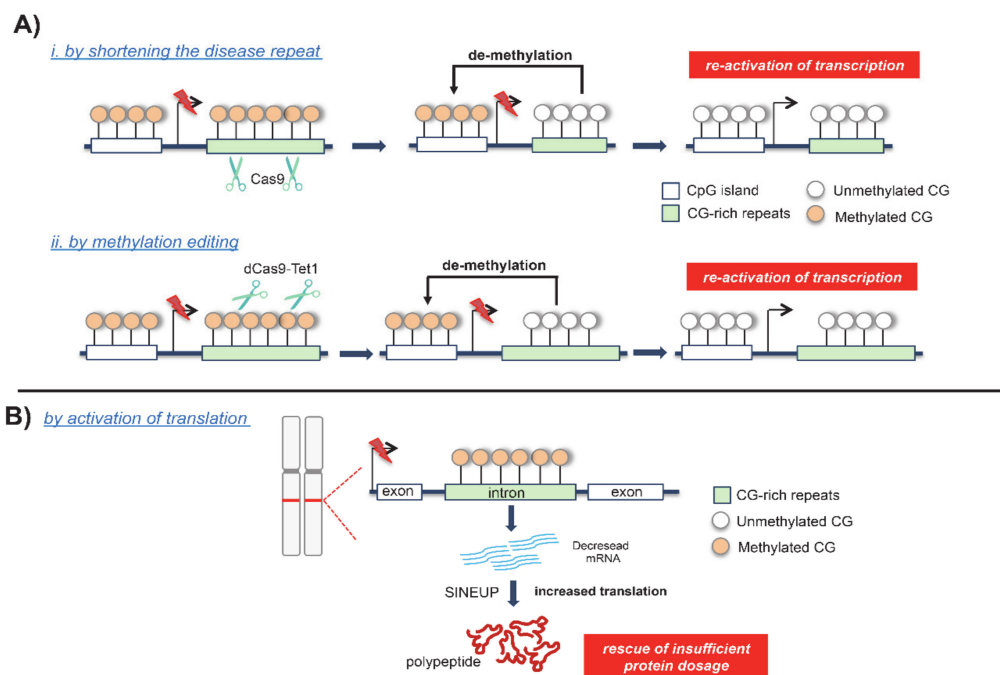


**Figure 5.** Innovative strategies applied to correct hypermethylated-expanded disorders: (**A**) Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 (CRISPR/Cas9) methodologies to rescue transcription in FXS by shortening the disease repeat (i) and by DNA methylation editing (ii); (**B**) Short interspersed nuclear element -containing translation UP-regulator (SINEUP) strategy to rescue protein insufficiency in FRDA.

## 7. Conclusions

Hypermethylation is the common denominator for several expanded repeat disorders, but its contribution as pathogenic modifier is still not fully understood, given the complexity of each genetic condition. Remarkable efforts have been made in understanding the molecular correlation between repeat expansion disorders and the dynamics of DNA methylation as a diagnostic bridge between patient genotype and disease phenotype. More interestingly, behind the 5mC profile, the analysis of the 5hmC profile represents a complementary molecular tool to track epigenetic perturbations in unstable repeat disorders. Therefore, a comprehensive understanding of the physiological and

pathological significance of hypermethylation/demethylation processes will provide insights into novel disease mechanisms. Discoveries over the last few years have yielded promising molecular-targeted therapies, including CRISPR/Cas9-based methodologies (Cas9 and dCas9/Tet1) and the SINEUP strategy. Moreover, the application of one or more of these strategies to other hypermethylated expansion disorders could represent a new challenge in the coming years.

## References

1. Kumar, R.; Nagpal, G.; Kumar, V.; Usmani, S.S.; Agrawal, P.; Raghava, G.P.S. HumCFS: A database of fragile sites in human chromosomes. *BMC Genom.* **2019**, *19*, 985. [CrossRef] [PubMed]

2. Poeta, L.; Fusco, F.; Drongitis, D.; Shoubridge, C.; Manganelli, G.; Filosa, S.; Paciolla, M.; Courtney, M.; Collombat, P.; Lioi, M.B.; et al. A Regulatory Path Associated with X-Linked Intellectual Disability and Epilepsy Links KDM5C to the Polyalanine Expansions in ARX. *Am. J. Hum. Genet.* **2013**, *92*, 114–125. [CrossRef] [PubMed]

3. Lanni, S.; Pearson, C.E. Molecular genetics of congenital myotonic dystrophy. *Neurobiol. Dis.* **2019**, *132*, 104533. [CrossRef] [PubMed]

4. Mirkin, S.M. Expandable DNA repeats and human disease. *Nature* **2007**, *447*, 932–940. [CrossRef]

5. Nelson, D.L.; Orr, H.T.; Warren, S. The Unstable Repeats—Three Evolving Faces of Neurological Disease. *Neuron* **2013**, *77*, 825–843. [CrossRef]

6. The 1000 Genomes Project Consortium; 1000 Genomes Project Consortium; Abecasis, G.R.; Altshuler, D.; Auton, A.; Brooks, L.D.; Durbin, R.; Gibbs, R.A.; Hurles, M.E.; McVean, G.A.; et al. A map of human genome variation from population-scale sequencing. *Nature* **2010**, *467*, 1061–1073. [CrossRef]

7. Shoubridge, C.; Fullston, T.; Gécz, J. ARX spectrum disorders: Making inroads into the molecular pathology. *Hum. Mutat.* **2010**, *31*, 889–900. [CrossRef]

8. Laperuta, C.; Spizzichino, L.; D'Adamo, P.; Monfregola, J.; Maiorino, A.; D'Eustacchio, A.; Ventruto, V.; Neri, G.; D'Urso, M.; Chiurazzi, P.; et al. MRX87 family with Aristaless X dup24bp mutation and implication for polyAlanine expansions. *BMC Med. Genet.* **2007**, *8*, 25. [CrossRef]

9. Zuccato, C.; Cattaneo, E. Huntington's disease. *Handb. Exp. Pharmacol.* **2014**, *220*, 357–409.

10. Pattamatta, A.; Cleary, J.D.; Ranum, L.P. All in the Family: Repeats and ALS/FTD. *Trends Neurosci.* **2018**, *41*, 247–250. [CrossRef]

11. Delatycki, M.B.; Bidichandani, S.I. Friedreich ataxia- pathogenesis and implications for therapies. *Neurobiol. Dis.* **2019**, *132*, 104606. [CrossRef] [PubMed]

12. Filippova, G.N. Genetics and Epigenetics of the Multifunctional Protein CTCF. *Current. Top. Dev. Biol.* **2007**, *80*, 337–360. [CrossRef]

13. Groh, M.; Lufino, M.M.P.; Wade-Martins, R.; Gromak, N. R-loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome. *PLoS Genet.* **2014**, *10*, e1004318. [CrossRef] [PubMed]

14. Nageshwaran, S.; Festenstein, R. Epigenetics and Triplet-Repeat Neurological Diseases. *Front. Neurol.* **2015**, *6*, 1010. [CrossRef]

15. Madrid, A.; Papale, L.A.; Alisch, R.S. New hope: The emerging role of 5-hydroxymethylcytosine in mental health and disease. *Epigenomics* **2016**, *8*, 981–991. [CrossRef]

16. Brasa, S.; Mueller, A.; Jacquemont, S.; Hahne, F.; Rozenberg, I.; Peters, T.; He, Y.; McCormack, C.; Gasparini, F.; Chibout, S.-D.; et al. Reciprocal changes in DNA methylation and hydroxymethylation and a broad repressive epigenetic switch characterize FMR1 transcriptional silencing in fragile X syndrome. *Clin. Epigenetics* **2016**, *8*, 15. [CrossRef]

17. Al-Mahdawi, S.; Sandi, C.; Pinto, R.M.; Pook, M.A. Friedreich Ataxia Patient Tissues Exhibit Increased 5-Hydroxymethylcytosine Modification and Decreased CTCF Binding at the FXN Locus. *PLoS ONE* **2013**, *8*, e74956. [CrossRef]

18. Wu, X.; Zhang, Y. TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nat. Rev. Genet.* **2017**, *18*, 517–534. [CrossRef]

19. Cheng, Y.; Bernstein, A.; Chen, D.; Jin, P. 5-Hydroxymethylcytosine: A new player in brain disorders? *Exp. Neurol.* **2015**, *268*, 3–9. [CrossRef]

20. Khare, T.; Pai, S.; Koncevičius, K.; Pal, I.; Kriukienė, E.; LiutkeviČiŪtĖ, Z.; Irimia, M.; Jia, P.; Ptak, C.; Xia, M.; et al. 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nat. Struct. Mol. Boil.* **2012**, *19*, 1037–1043. [CrossRef]

21. Prydz, K.; Dalen, K.T. Synthesis and sorting of proteoglycans. *J. Cell Sci.* **2000**, *113*, 193–205. [PubMed]

22. Lacroix, A.J.; Stabley, D.; Sahraoui, R.; Adam, M.P.; Mehaffey, M.; Kernan, K.; Myers, C.T.; Fagerstrom, C.; Anadiotis, G.; Akkari, Y.M.; et al. GGC Repeat Expansion and Exon 1 Methylation of XYLT1 Is a Common Pathogenic Variant in Baratela-Scott Syndrome. *Am. J. Hum. Genet.* **2019**, *104*, 35–44. [CrossRef] [PubMed]

23. Gokhman, D.; Nissim-Rafinia, M.; Agranat-Tamir, L.; Housman, G.; García-Pérez, R.; Lizano, E.; Cheronet, O.; Mallick, S.; Nieves-Colón, M.A.; Li, H.; et al. Differential DNA methylation of vocal and facial anatomy genes in modern humans. *Nat. Commun.* **2020**, *11*, 1–21. [CrossRef] [PubMed]

24. Ramachandran, N.; Girard, J.-M.; Turnbull, J.; A Minassian, B. The autosomal recessively inherited progressive myoclonus epilepsies and their genes. *Epilepsia* **2009**, *50*, 29–36. [CrossRef]

25. Pennacchio, L.A.; Lehesjoki, A.-E.; Stone, N.E.; Willour, V.L.; Virtaneva, K.; Miao, J.; D'Amato, E.; Ramirez, L.; Faham, M.; Koskiniemi, M.; et al. Mutations in the Gene Encoding Cystatin B in Progressive Myoclonus Epilepsy (EPM1). *Science* **1996**, *271*, 1731–1734. [CrossRef]

26. Penna, E.; Cerciello, A.; Chambery, A.; Russo, R.; Cernilogar, F.M.; Pedone, E.M.; Perrone-Capano, C.; Cappello, S.; Di Giaimo, R.; Crispino, M. Cystatin B Involvement in Synapse Physiology of Rodent Brains and Human Cerebral Organoids. *Front. Mol. Neurosci.* **2019**, *12*, 195. [CrossRef]

27. Joensuu, T.; Kuronen, M.; Alakurtti, K.; Tegelberg, S.; Hakala, P.; Aalto, A.; Huopaniemi, L.; Aula, N.; Michellucci, R.; Eriksson, K.; et al. Cystatin B: Mutation detection, alternative splicing and expression in progressive myclonus epilepsy of Unverricht-Lundborg type (EPM1) patients. *Eur. J. Hum. Genet.* **2006**, *15*, 185–193. [CrossRef]

28. Alakurtti, K.; Virtaneva, K.; Joensuu, T.; Palvimo, J.J.; Lehesjoki, A.-E. Characterization of the cystatin B gene promoter harboring the dodecamer repeat expanded in progressive myoclonus epilepsy, EPM1. *Genetics* **2000**, *242*, 65–73. [CrossRef]

29. O'Brien, A.; Marshall, C.R.; Blaser, S.; Ray, P.N.; Yoon, G. Severe neurodegeneration, progressive cerebral volume loss and diffuse hypomyelination associated with a homozygous frameshift mutation in CSTB. *Eur. J. Hum. Genet.* **2017**, *25*, 775–778. [CrossRef]

30. Weinhaeusel, A.; Morris, M.A.; Antonarakis, S.E.; Haas, O.A. DNA deamination enables direct PCR amplification of the cystatin B (CSTB) gene-associated dodecamer repeat expansion in myoclonus epilepsy type Unverricht-Lundborg. *Hum. Mutat.* **2003**, *22*, 404–408. [CrossRef]

31. Horiuchi, H.; Osawa, M.; Furutani, R.; Morita, M.; Tian, W.; Awatsu, Y.; Shimazaki, H.; Umetsu, K. Polymerase Chain Reaction-Based Analysis Using Deaminated DNA of Dodecamer Expansions in CSTB, Associated with Unverricht-Lundborg Myoclonus Epilepsy. *Genet. Test.* **2005**, *9*, 328–333. [CrossRef] [PubMed]

32. Voutsinos, V.; Munk, S.H.N.; Oestergaard, V.H. Common Chromosomal Fragile Sites—Conserved Failure Stories. *Genes* **2018**, *9*, 580. [CrossRef] [PubMed]

33. Hagerman, R.; Berry-Kravis, E.; Hazlett, H.C.; Bailey, D.B.; Moine, H.; Kooy, F.R.; Tassone, F.; Gantois, I.; Sonenberg, N.; Mandel, J.-L.; et al. Fragile X syndrome. *Nat. Rev. Dis. Prim.* **2017**, *3*, 17065. [CrossRef] [PubMed]

34. Hagerman, P.J.; Hagerman, R. Fragile X-associated Tremor/Ataxia Syndrome (FXTAS). *Ment. Retard. Dev. Disabil. Res. Rev.* **2004**, *10*, 25–30. [CrossRef]

35. Miano, M.G.; Laperuta, C.; Chiurazzi, P.; D'Urso, M.; Ursini, M.V. Ovarian dysfunction and FMR1 alleles in a large Italian family with POF and FRAXA disorders: Case report. *BMC Med. Genet.* **2007**, *8*, 18. [CrossRef]

36. Naumann, A.; Hochstein, N.; Weber, S.; Fanning, E.; Doerfler, W. A Distinct DNA-Methylation Boundary in the 5′- Upstream Sequence of the FMR1 Promoter Binds Nuclear Proteins and Is Lost in Fragile X Syndrome. *Am. J. Hum. Genet.* **2009**, *85*, 606–616. [CrossRef] [PubMed]

37. Tabolacci, E.; Moscato, U.; Zalfa, F.; Bagni, C.; Chiurazzi, P.; Neri, G. Epigenetic analysis reveals a euchromatic configuration in the FMR1 unmethylated full mutations. *Eur. J. Hum. Genet.* **2008**, *16*, 1487–1498. [CrossRef]

38. Rodriguez, C.M.; Wright, S.E.; Kearse, M.G.; Haenfler, J.M.; Flores, B.N.; Liu, Y.; Ifrim, M.F.; Glineburg, M.R.; Krans, A.; Jafar-Nejad, P.; et al. A native function for RAN translation and CGG repeats in regulating fragile X protein synthesis. *Nat. Neurosci.* **2020**, *23*, 386–397. [CrossRef]

39. Baker, E.K.; Arpone, M.; Vera, S.A.; Bretherton, L.; Kraan, C.M.; Bui, M.; Slater, H.; Ling, L.; Francis, D.; Hunter, M.F.; et al. Incomplete silencing of full mutation alleles in males with fragile X syndrome is associated with autistic features. *Mol. Autism* **2019**, *10*, 21. [CrossRef]

40. Kuznitsov-Yanovsky, L.; Mayshar, Y.; Ben-Yosef, D. *Modeling FXS: Human Pluripotent Stem Cells and In Vitro Neural Differentiation*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019; Volume 1942, pp. 89–100.

41. Colak, D.; Zaninovic, N.; Cohen, M.S.; Rosenwaks, Z.; Yang, W.-Y.; Gerhardt, J.; Disney, M.D.; Jaffrey, S.R. Promoter-Bound Trinucleotide Repeat mRNA Drives Epigenetic Silencing in Fragile X Syndrome. *Science* **2014**, *343*, 1002–1005. [CrossRef]

42. Lanni, S.; Goracci, M.; Borrelli, L.; Mancano, G.; Chiurazzi, P.; Moscato, U.; Ferrè, F.; Helmer-Citterich, M.; Tabolacci, E.; Neri, G. Role of CTCF Protein in Regulating FMR1 Locus Transcription. *PLoS Genet.* **2013**, *9*, e1003601. [CrossRef]

43. Esanov, R.; Andrade, N.S.; Bennison, S.; Wahlestedt, C.; Zeier, Z. The FMR1 promoter is selectively hydroxymethylated in primary neurons of fragile X syndrome patients. *Hum. Mol. Genet.* **2016**, *25*, 4870–4880. [CrossRef] [PubMed]

44. Metsu, S.; Rooms, L.; Rainger, J.; Taylor, M.S.; Bengani, H.; Wilson, D.I.; Chilamakuri, C.S.R.; Morrison, H.; Vandeweyer, G.; Reyniers, E.; et al. FRA2A Is a CGG Repeat Expansion Associated with Silencing of AFF3. *PLoS Genet.* **2014**, *10*, e1004242. [CrossRef] [PubMed]

45. Metsu, S.; Rainger, J.K.; Debacker, K.; Bernhard, B.; Rooms, L.; A Grafodatskaya, D.; Weksberg, R.; Fombonne, E.; Taylor, M.S.; Scherer, S.W.; et al. A CGG-Repeat Expansion Mutation inZNF713Causes FRA7A: Association with Autistic Spectrum Disorder in two Families. *Hum. Mutat.* **2014**, *35*, 1295–1300. [CrossRef]

46. Gécz, J.; Gedeon, Á.K.; Sutherland, G.R.; Mulley, J.C. Identification of the gene FMR2, associated with FRAXE mental retardation. *Nat. Genet.* **1996**, *13*, 105–108. [CrossRef]

47. Knight, S.; Flannery, A.; Hirst, M.; Campbell, L.; Christodoulou, Z.; Phelps, S.; Pointon, J.; Middleton-Price, H.; Barnicoat, A.; Pembrey, M.; et al. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. *Cell* **1993**, *74*, 127–134. [CrossRef]

48. Gu, Y.; McIlwain, K.L.; Weeber, E.J.; Yamagata, T.; Xu, B.; Antalffy, B.A.; Reyes, C.; Yuva-Paylor, L.; Armstrong, D.; Zoghbi, H.Y.; et al. Impaired Conditioned Fear and Enhanced Long-Term Potentiation inFmr2 Knock-Out Mice. *J. Neurosci.* **2002**, *22*, 2753–2763. [CrossRef]

49. Yuva-Aydemir, Y.; Almeida, S.; Krishnan, G.; Gendron, T.F.; Gao, F.B. Transcription elongation factor AFF2/FMR2 regulates expression of expanded GGGGCC repeat-containing C9ORF72 allele in ALS/FTD. *Nat. Commun.* **2019**, *10*, 5466. [CrossRef]

50. Winnepenninckx, B.; Debacker, K.; Ramsay, J.; Smeets, M.; Smits, A.; Fitzpatrick, D.R.; Kooy, R.F. CGG-Repeat Expansion in the DIP2B Gene Is Associated with the Fragile Site FRA12A on Chromosome 12q13.1. *Am. J. Hum. Genet.* **2006**, *80*, 221–231. [CrossRef]

51. Ishiura, H.; Shibata, S.; Yoshimura, J.; Suzuki, Y.; Qu, W.; Doi, K.; Almansour, M.A.; Kikuchi, J.K.; Taira, M.; Mitsui, J.; et al. Noncoding CGG repeat expansions in neuronal intranuclear inclusion disease, oculopharyngodistal myopathy and an overlapping disease. *Nat. Genet.* **2019**, *51*, 1222–1232. [CrossRef]

52. Sone, J.; Mitsuhashi, S.; Fujita, A.; Mizuguchi, T.; Hamanaka, K.; Mori, K.; Koike, H.; Hashiguchi, A.; Takashima, H.; Sugiyama, H.; et al. Long-read sequencing identifies GGC repeat expansions in NOTCH2NLC associated with neuronal intranuclear inclusion disease. *Nat. Genet.* **2019**, *51*, 1215–1221. [CrossRef]

53. Tian, Y.; Wang, J.-L.; Huang, W.; Zeng, S.; Jiao, B.; Liu, Z.; Chen, Z.; Li, Y.; Wang, Y.; Min, H.-X.; et al. Expansion of Human-Specific GGC Repeat in Neuronal Intranuclear Inclusion Disease-Related Disorders. *Am. J. Hum. Genet.* **2019**, *105*, 166–176. [CrossRef] [PubMed]

54. Sun, Q.-Y.; Xu, Q.; Tian, Y.; Hu, Z.-M.; Qin, L.-X.; Yang, J.-X.; Huang, W.; Xue, J.; Li, J.-C.; Zeng, S.; et al. Expansion of GGC repeat in the human-specific NOTCH2NLC gene is associated with essential tremor. *Brain* **2020**, *143*, 222–233. [CrossRef] [PubMed]

55. Gendron, T.F.; Belzil, V.V.; Zhang, Y.-J.; Petrucelli, L. Mechanisms of toxicity in C9FTLD/ALS. *Acta Neuropathol.* **2014**, *127*, 359–376. [CrossRef] [PubMed]

56. Dols-Icardo, O.; García-Redondo, A.; Rojas-García, R.; Sanchez-Valle, R.; Noguera, A.; Gómez-Tortosa, E.; Pastor, P.; Hernández, I.; Esteban-Pérez, J.; Suárez-Calvet, M.; et al. Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. *Hum. Mol. Genet.* **2013**, *23*, 749–754. [CrossRef]

57. DeJesus-Hernandez, M.; MacKenzie, I.R.; Boeve, B.F.; Boxer, A.L.; Baker, M.; Rutherford, N.J.; Nicholson, A.M.; Finch, N.A.; Flynn, H.; Adamson, J.; et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72*, 245–256. [CrossRef]

58. Xi, Z.; Zinman, L.; Moreno, D.; Schymick, J.; Liang, Y.; Sato, C.; Zheng, Y.; Ghani, M.; Dib, S.; Keith, J.; et al. Hypermethylation of the CpG Island Near the G4C2 Repeat in ALS with a C9orf72 Expansion. *Am. J. Hum. Genet.* **2013**, *92*, 981–989. [CrossRef]

59. Belzil, V.V.; Katzman, R.B.; Petrucelli, L. ALS and FTD: An epigenetic perspective. *Acta Neuropathol.* **2016**, *132*, 487–502. [CrossRef]

60. Esanov, R.; Belle, K.C.; Van Blitterswijk, M.; Belzil, V.V.; Rademakers, R.; Dickson, D.W.; Petrucelli, L.; Boylan, K.B.; Dykxhoorn, D.M.; Wuu, J.; et al. C9orf72 promoter hypermethylation is reduced while hydroxymethylation is acquired during reprogramming of ALS patient cells. *Exp. Neurol.* **2016**, *277*, 171–177. [CrossRef]

61. Bauer, P.O. Methylation of C9orf72 expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. *Neurosci. Lett.* **2016**, *612*, 204–209. [CrossRef]

62. Liu, J.; Hu, J.; Ludlow, A.T.; Pham, J.T.; Shay, J.W.; Rothstein, J.D.; Corey, D.R. c9orf72 disease-related foci are each composed of one mutant expanded repeat RNA. *Cell. Chem. Biol.* **2017**, *24*, 141–148. [CrossRef] [PubMed]

63. Evans-Galea, M.V.; Carrodus, N.; Rowley, S.; Corben, L.A.; Tai, G.; Saffery, R.; Galati, J.C.; Wong, N.; Craig, J.M.; Lynch, D.R.; et al. FXN methylation predicts expression and clinical outcome in Friedreich ataxia. *Ann. Neurol.* **2012**, *71*, 487–497. [CrossRef]

64. Bradley, J.; Blake, J.; Chamberlain, S.; Thomas, P.; Cooper, J.; Schapira, A.H. Clinical, biochemical and molecular genetic correlations in Friedreich's ataxia. *Hum. Mol. Genet.* **2000**, *9*, 275–282. [CrossRef] [PubMed]

65. Wong, A.; Yang, J.; Cavadini, P.; Gellera, C.; Lonnerdal, B.; Taroni, F.; Cortopassi, G.A. The Friedreich's ataxia mutation confers cellular sensitivity to oxidant stress which is rescued by chelators of iron and calcium and inhibitors of apoptosis. *Hum. Mol. Genet.* **1999**, *8*, 425–430. [CrossRef] [PubMed]

66. Castaldo, I.; Pinelli, M.; Monticelli, A.; Acquaviva, F.; Giacchetti, M.; Filla, A.; Sacchetti, S.; Keller, S.; E Avvedimento, V.; Chiariotti, L.; et al. DNA methylation in intron 1 of the frataxin gene is related to GAA repeat length and age of onset in Friedreich ataxia patients. *J. Med Genet.* **2008**, *45*, 808–812. [CrossRef] [PubMed]

67. Al-Mahdawi, S.; Pinto, R.M.; Ismail, O.; Varshney, D.; Lymperi, S.; Sandi, C.; Trabzuni, D.; Pook, M.A. The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues. *Hum. Mol. Genet.* **2007**, *17*, 735–746. [CrossRef] [PubMed]

68. Chia, N.; Wang, L.; Lu, X.; Senut, M.-C.; Brenner, C.A.; Ruden, D.M. Hypothesis: Environmental regulation of 5-hydroxymethylcytosine by oxidative stress. *Epigenetics* **2011**, *6*, 853–856. [CrossRef]

69. Tomé, S.; Gourdon, G. DM1 Phenotype Variability and Triplet Repeat Instability: Challenges in the Development of New Therapies. *Int. J. Mol. Sci.* **2020**, *21*, 457. [CrossRef]

70. Barbé, L.; Lanni, S.; López-Castel, A.; Franck, S.; Spits, C.; Keymolen, K.; Seneca, S.; Tomé, S.; Miron, I.; Letourneau, J.; et al. CpG Methylation, a Parent-of-Origin Effect for Maternal-Biased Transmission of Congenital Myotonic Dystrophy. *Am. J. Hum. Genet.* **2017**, *100*, 488–505. [CrossRef]

71. Yanovsky-Dagan, S.; Avitzour, M.; Altarescu, G.; Renbaum, P.; Eldar-Geva, T.; Schonberger, O.; Mitrani-Rosenbaum, S.; Levy-Lahad, E.; Birnbaum, R.Y.; Gepstein, L.; et al. Uncovering the Role of Hypermethylation by CTG Expansion in Myotonic Dystrophy Type 1 Using Mutant Human Embryonic Stem Cells. *Stem Cell Rep.* **2015**, *5*, 221–231. [CrossRef]

72. Klesert, T.R.; Cho, D.H.; Clark, J.I.; Maylie, J.; Adelman, J.; Snider, L.; Yuen, E.C.; Soriano, P.; Tapscott, S.J. Mice deficient in Six5 develop cataracts: Implications for myotonic dystrophy. *Nat. Genet.* **2000**, *25*, 105–109. [CrossRef] [PubMed]

73. Castel, A.L.; Nakamori, M.; Tomé, S.; Chitayat, D.; Gourdon, G.; Thornton, C.A.; E Pearson, C. Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum. Mol. Genet.* **2010**, *20*, 1–15. [CrossRef] [PubMed]

74. Légaré, C.; Overend, G.; Guay, S.-P.; Monckton, D.G.; Mathieu, J.; Gagnon, C.; Bouchard, L. DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DM1. *Neurol. Genet.* **2019**, *5*, e338. [CrossRef]

75. Filippova, G.N.; Thienes, C.; Penn, B.H.; Cho, D.H.; Hu, Y.J.; Moore, J.M.; Klesert, T.R.; Lobanenkov, V.V.; Tapscott, S.J. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* **2001**, *28*, 335–343. [CrossRef]

76. Mantere, T.; Kersten, S.; Hoischen, A. Long-Read Sequencing Emerging in Medical Genetics. *Front. Genet.* **2019**, *10*, 426. [CrossRef] [PubMed]

77. Hagerman, P.J.; Eid, J.; Peluso, P.; Rank, D.; Yin, J.; Hickey, L.; Tassone, F.; Loomis, E. Sequencing the Unsequenceable: Expanded CGG Repeats in the Human FMR1 Gene. *Biophys. J.* **2013**, *104*, 377a. [CrossRef]

78. Zablotskaya, A.; Van Esch, H.; Verstrepen, K.J.; Froyen, G.; Vermeesch, J.R. Mapping the landscape of tandem repeat variability by targeted long read single molecule sequencing in familial X-linked intellectual disability. *BMC Med Genom.* **2018**, *11*, 123. [CrossRef]

79. Cumming, S.; Hamilton, M.J.; Robb, Y.; Gregory, H.; McWilliam, C.; Cooper, A.; Adam, B.; McGhie, J.; Hamilton, G.; Herzyk, P.; et al. De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *Eur. J. Hum. Genet.* **2018**, *26*, 1635–1647. [CrossRef]

80. Zeng, S.; Zhang, M.-Y.; Wang, X.-J.; Hu, Z.-M.; Li, J.-C.; Li, N.; Wang, J.-L.; Liang, F.; Yang, Q.; Liu, Q.; et al. Long-read sequencing identified intronic repeat expansions in SAMD12 from Chinese pedigrees affected with familial cortical myoclonic tremor with epilepsy. *J. Med Genet.* **2018**, *56*, 265–270. [CrossRef]

81. Dashnow, H.; Lek, M.; Phipson, B.; Halman, A.; Sadedin, S.; Lonsdale, A.; Davis, M.; Lamont, P.; Clayton, J.; Laing, N.G.; et al. STRetch: Detecting and discovering pathogenic short tandem repeat expansions. *Genome Biol.* **2018**, *19*, 121. [CrossRef]

82. Genovese, L.M.; Geraci, F.; Corrado, L.; Mangano, E.; D'Aurizio, R.; Bordoni, R.; Severgnini, M.; Manzini, G.; De Bellis, G.; D'Alfonso, S.; et al. A Census of Tandemly Repeated Polymorphic Loci in Genic Regions Through the Comparative Integration of Human Genome Assemblies. *Front. Genet.* **2018**, *9*, 155. [CrossRef] [PubMed]

83. Aref-Eshghi, E.; Bend, E.G.; Colaiacovo, S.; Caudle, M.; Chakrabarti, R.; Napier, M.; Brick, L.; Brady, L.; Carere, D.A.; Levy, M.A.; et al. Diagnostic Utility of Genome-wide DNA Methylation Testing in Genetically Unsolved Individuals with Suspected Hereditary Conditions. *Am. J. Hum. Genet.* **2019**, *104*, 685–700. [CrossRef]

84. Shortt, J.A.; Ruggiero, R.P.; Cox, C.; Wacholder, A.C.; Pollock, D.D. Finding and extending ancient simple sequence repeat-derived regions in the human genome. *Mob. DNA* **2020**, *11*, 11–12. [CrossRef] [PubMed]

85. Tabolacci, E.; Mancano, G.; Lanni, S.; Palumbo, F.; Goracci, M.; Ferrè, F.; Helmer-Citterich, M.; Neri, G. Genome-wide methylation analysis demonstrates that 5-aza-2-deoxycytidine treatment does not cause random DNA demethylation in fragile X syndrome cells. *Epigenetics Chromatin* **2016**, *9*, 12. [CrossRef]

86. Dolskiy, A.; Pustylnyak, V.; Yarushkin, A.; Lemskaya, N.; Yudkin, D.V. Inhibitors of Histone Deacetylases Are Weak Activators of the FMR1 Gene in Fragile X Syndrome Cell Lines. *BioMed Res. Int.* **2017**, *2017*, 3582601. [CrossRef] [PubMed]

87. Tomaselli, D.; Lucidi, A.; Rotili, D.; Mai, A. Epigenetic polypharmacology: A new frontier for epi-drug discovery. *Med. Res. Rev.* **2019**, *40*, 190–244. [CrossRef] [PubMed]

88. Liu, X.S.; Wu, H.; Krzisch, M.; Wu, X.; Graef, J.; Muffat, J.; Hnisz, D.; Li, C.; Yuan, B.; Xu, C.; et al. Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell* **2018**, *172*, 979–992. [CrossRef]

89. Morita, S.; Horii, T.; Hatada, I. Editing of DNA Methylation Using dCas9-Peptide Repeat and scFv-TET1 Catalytic Domain Fusions. *Adv. Struct. Saf. Stud.* **2018**, 419–428. [CrossRef]

90. Liu, X.S.; Wu, H.; Ji, X.; Stelzer, Y.; Wu, X.; Czauderna, S.; Shu, J.; Dadon, D.; Young, R.A.; Jaenisch, R. Editing DNA Methylation in the Mammalian Genome. *Cell* **2016**, *167*, 233–247. [CrossRef]

91. Park, C.-Y.; Halevy, T.; Lee, D.R.; Sung, J.J.; Lee, J.S.; Yanuka, O.; Benvenisty, N.; Kim, D.-W. Reversion of FMR1 Methylation and Silencing by Editing the Triplet Repeats in Fragile X iPSC-Derived Neurons. *Cell Rep.* **2015**, *13*, 234–241. [CrossRef]

92. Xie, N.; Gong, H.; Suhl, J.A.; Chopra, P.; Wang, T.; Warren, S. Reactivation of FMR1 by CRISPR/Cas9-Mediated Deletion of the Expanded CGG-Repeat of the Fragile X Chromosome. *PLoS ONE* **2016**, *11*, e0165499. [CrossRef] [PubMed]

93. Dabrowska, M.; Olejniczak, M. Gene Therapy for Huntington's Disease Using Targeted Endonucleases. *Methods Mol. Biol.* **2020**, *2056*, 269–284. [PubMed]

94. Zucchelli, S.; Cotella, D.; Takahashi, H.; Carrieri, C.; Cimatti, L.; Fasolo, F.; Jones, M.; Sblattero, D.; Sanges, R.; Santoro, C.; et al. SINEUPs: A new class of natural and synthetic antisense long non-coding RNAs that activate translation. *RNA Boil.* **2015**, *12*, 771–779. [CrossRef] [PubMed]

95. Bon, C.; Luffarelli, R.; Russo, R.; Fortuni, S.; Pierattini, B.; Santulli, C.; Fimiani, C.; Persichetti, F.; Cotella, D.; Mallamaci, A.; et al. SINEUP non-coding RNAs rescue defective frataxin expression and activity in a cellular model of Friedreich's Ataxia. *Nucleic Acids Res.* **2019**, *47*, 10728–10743. [CrossRef]

96. Herman, D.; Jenssen, K.; Burnett, R.; Soragni, E.; Perlman, S.L.; Gottesfeld, J.M. Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. *Nat. Chem. Biol.* **2006**, *2*, 551–558. [CrossRef]

97. Poeta, L.; Padula, A.; Attianese, B.; Valentino, M.; Verrillo, L.; Filosa, S.; Shoubridge, C.; Barra, A.; E Schwartz, C.; Christensen, J.; et al. Histone demethylase KDM5C is a SAHA-sensitive central hub at the crossroads of transcriptional axes involved in multiple neurodevelopmental disorders. *Hum. Mol. Genet.* **2019**, *28*, 4089–4102. [CrossRef]