Review

Mechanisms of the *FMR1* Repeat Instability: How Does the CGG Sequence Expand?

Elisabetta Tabolacci [1] , Veronica Nobile [1] , Cecilia Pucci [1] and Pietro Chiurazzi [1,2,*]

1 Dipartimento Scienze della Vita e Sanità Pubblica, Sezione di Medicina Genomica, Università Cattolica del Sacro Cuore, Fondazione Policlinico Universitario "A. Gemelli" IRCCS, 00168 Rome, Italy; elisabetta.tabolacci@unicatt.it (E.T.); veronicanobile88@gmail.com (V.N.); c.pucci91@yahoo.it (C.P.)
2 UOC Genetica Medica, Fondazione Policlinico Universitario "A. Gemelli" IRCCS, 00168 Rome, Italy
* Correspondence: pietro.chiurazzi@unicatt.it; Tel.: +39-06-3015-4606

Abstract: A dynamic mutation in exon 1 of the *FMR1* gene causes Fragile X-related Disorders (FXDs), due to the expansion of an unstable CGG repeat sequence. Based on the CGG sequence size, two types of *FMR1* alleles are possible: "premutation" (PM, with 56–200 CGGs) and "full mutation" (FM, with >200 triplets). Premutated females are at risk of transmitting a FM allele that, when methylated, epigenetically silences *FMR1* and causes Fragile X syndrome (FXS), a very common form of inherited intellectual disability (ID). Expansions events of the CGG sequence are predominant over contractions and are responsible for meiotic and mitotic instability. The CGG repeat usually includes one or more AGG interspersed triplets that influence allele stability and the risk of transmitting FM to children through maternal meiosis. A unique mechanism responsible for repeat instability has not been identified, but several processes are under investigations using cellular and animal models. The formation of unusual secondary DNA structures at the expanded repeats are likely to occur and contribute to the CGG expansion. This review will focus on the current knowledge about CGG repeat instability addressing the CGG sequence expands.

1. Introduction

Instability of short repeated sequences underlies the so-called dynamic mutations responsible for more than 40 mostly neurological clinical conditions also known as repeat expansion disorders or REDs [1,2]. REDs are intrinsically dynamic, i.e., repeated sequences (tri-, tetra-, penta-nucleotides and so forth) which continue to mutate across generations and within tissues of an individual, while "normal" mutations (e.g., single nucleotide variants or copy number variants) are transmitted to offspring and are retained in somatic tissues without further changes. REDs also often display genetic anticipation, a phenomenon by which disease severity increases and age of onset decreases from one generation to the other, in parallel with expansion of the repeat sequence. Most REDs are caused by a trinucleotide repeat expansion, as happens in Fragile X syndrome (FXS; OMIM #300624), the first such condition, identified in 1991. FXS is a very frequent cause of monogenic intellectual disability (ID), due to the expansion of a CGG tract in exon 1 of the *FMR1* gene, located in Xq27.3. When expansion exceeds 200 repeats (FM, full mutation), cytosines in the CGG repeat and in the neighboring CpG island become methylated (MFM, methylated FM) and the *FMR1* gene is silenced [3]. FXS is eventually caused by a loss-of-function effect produced by the dynamic mutation, causing local epigenetic modifications and absence of the *FMR1* protein, FMRP. The estimated frequency of FM in the general population is 1:7143 males (who will be affected by FXS) and 1:11,111 females, who present a variable phenotype ranging from learning disability or mild ID to a full-blown phenotype like that of affected males [4,5]. To perform a high sensitivity diagnostic examination for FXS, the

gold standard should include CGG sizing and methylation analysis, the latter is particularly indicated during prenatal diagnosis [6].

The CGG repeat is highly polymorphic in length, with a mode of 30 triplets in the general population, and it is usually interrupted by an AGG every 9-10 CGG triplets [7,8]. AGG interruptions have no effect on *FMR1* transcription efficiency [9]. Based on size, the CGG sequence may be classified in four allele categories, each with a different degree of instability through generations and within tissues. Normal alleles (5–44 repeats) are stably transmitted from parents to children, although rarely very small repeat size changes are detected particularly during paternal transmission. Intermediate or gray zone alleles (45–55 repeats) sporadically undergo length instability in some families, again with a greater instability in paternal than maternal transmissions [10,11]. The greatest instability is well documented for premutation alleles (PM, 56–200 repeats), that expand to full mutation (FM, >200 repeats) during maternal meiosis, causing FXS to manifest in the offspring. Until now, the smallest PM known to have expanded to FM in a single maternal meiosis had 56 repeats [12]. The risk of passing FM to the offspring increases with the length of maternal PM and it is nearly 100% for mothers with >90 CGGs [13]; it also seems that increasing maternal age directly correlates with the increase in PM-to-FM expansion risk [14].

This increase in risk of having more FXS children in parallel with the increase in length of the CGG repeat as generations pass, is known as the Sherman paradox [15]. Hence, the number of FXS affected individuals tends to increase over generations. An actual genetic anticipation phenomenon, i.e., a younger age of onset with passing generations, as described in myotonic dystrophy, is not described in FXS families. The so-called "all or nothing" effect consists of having the classical FXS phenotype when an MFM is transmitted, no matter how large it is, since the phenotype depends on the epigenetic changes occurring when the CGG expansion exceeds 200 units.

Paternal PM alleles are typically transmitted to all their daughters without expanding to FM, although extremely rare cases of father to daughter FM transmission have been reported [16,17]. Furthermore, although FXS males typically do not have offspring, it is expected that all their daughters would inherit a PM not a FM, since surprisingly it appears that FM contracts to PM in the germline of affected males [18]. On the contrary, females with a FM have 50% risk of transmitting it to their offspring, whether male or female, possibly with a larger CGG expansion.

It is important to point out that PM alleles are not only more meiotically unstable but also associated with a late-onset neurodegenerative condition known as Fragile X Tremor/Ataxia Syndrome (FXTAS, OMIM #300623), that affect 33% of carrier males and approximately 5–10% of carrier females [19,20]. Premutated females are also at risk of developing a premature menopause, called Fragile X-associated primary ovarian insufficiency (FXPOI, OMIM # 311360) [21]. The estimated prevalence in the general population of males carrying a PM allele is 1:855, while that of carrier females is 1:291 [5]. It is also worth of note that FXTAS and FXPOI are never observed in FXS patients, since the underlying mechanism is different: FM alleles are silenced and both *FMR1* mRNA and FMRP are lacking (loss-of-function), while PM alleles are transcribed and the expanded CGG repeat in the 5′ untranslated region of the mRNA causes neurodegeneration with at least two gain-of-function mechanisms, namely, RNA toxicity and production of abnormal polypeptides due to repeat-associated non-AUG (RAN)-dependent translation [22].

The presence of AGG interruptions inversely correlates with the risk of transmitting a FM to the offspring if a mother carries a PM: more interspersed AGGs in a PM allele correlate with a lower risk of FM transmission, i.e., of PM-to-FM expansion in maternal meiosis. For example, a PM allele with 75 repeats with no interspersed AGGs has a risk of expansion to FM of 77%, but the risk is lowered to 12% for PM of the same length with two AGG interruptions [23]. Interestingly, loss of AGG interruptions in the CGG repeat seems to occur after contraction events of a maternal PM [11].

In fact, although the instability of the *FMR1* repeat mostly consists of expansions of the CGG sequence, contractions to smaller length are observed in all classes of alleles [11].

Expansion and contraction events both contribute to the repeat size mosaicism frequently displayed by PM and FM alleles [24]. Contractions occur more frequently in paternal than maternal transmissions, and the frequency of paternal contractions increases linearly with repeat size, being more frequent in PM range. Most paternal PMs contract to a slightly shorter PM, but maternal PM alleles may also be transmitted as PM, intermediate, or even normal alleles [25,26].

Finally, it is important to remember that mitotic instability of FM alleles very often results in somatic mosaicism with cells of the same individual carrying CGG repeats of different size and sometimes even different methylation status [27,28]; for example, a post-mortem study of a high-functioning male demonstrated the presence of an unmethylated mutation with a wide span of sizes (from PM to FM) in leukocytes and most brain areas, while in the cortex of the parietal lobe and in other non-brain tissues a methylated FM of a single size was found [29].

In summary, from the unbiased observation of FXS families, we conclude that three main factors influence the stability of the *FMR1* repeat: its size, its internal structure (AGG interruptions) and the sex of the transmitting parent. Smaller PM alleles with two (or more) AGGs are more stable than longer alleles, particularly if the latter are transmitted through oogenesis. Several hypotheses have been made to explain the apparently exclusive maternal transmission of FXS starting from a PM (Figure 1): is a maternal meiotic event occurring during gametogenesis? Or does an early embryonic mitotic event cause expansion only of a maternal PM? Do both meiotic and mitotic events contribute to the maternal PM-to-FM expansion? Which (different) mechanisms are responsible for the phenomenon of instability in the germline and in somatic cells? To answer these questions, we discuss the relevant literature, published so far, describing how errors during DNA replication, repair or recombination lead to the instability of the CGG repeat. Errors in these processes are largely caused by secondary structures formed on DNA strands with expanded repeats.
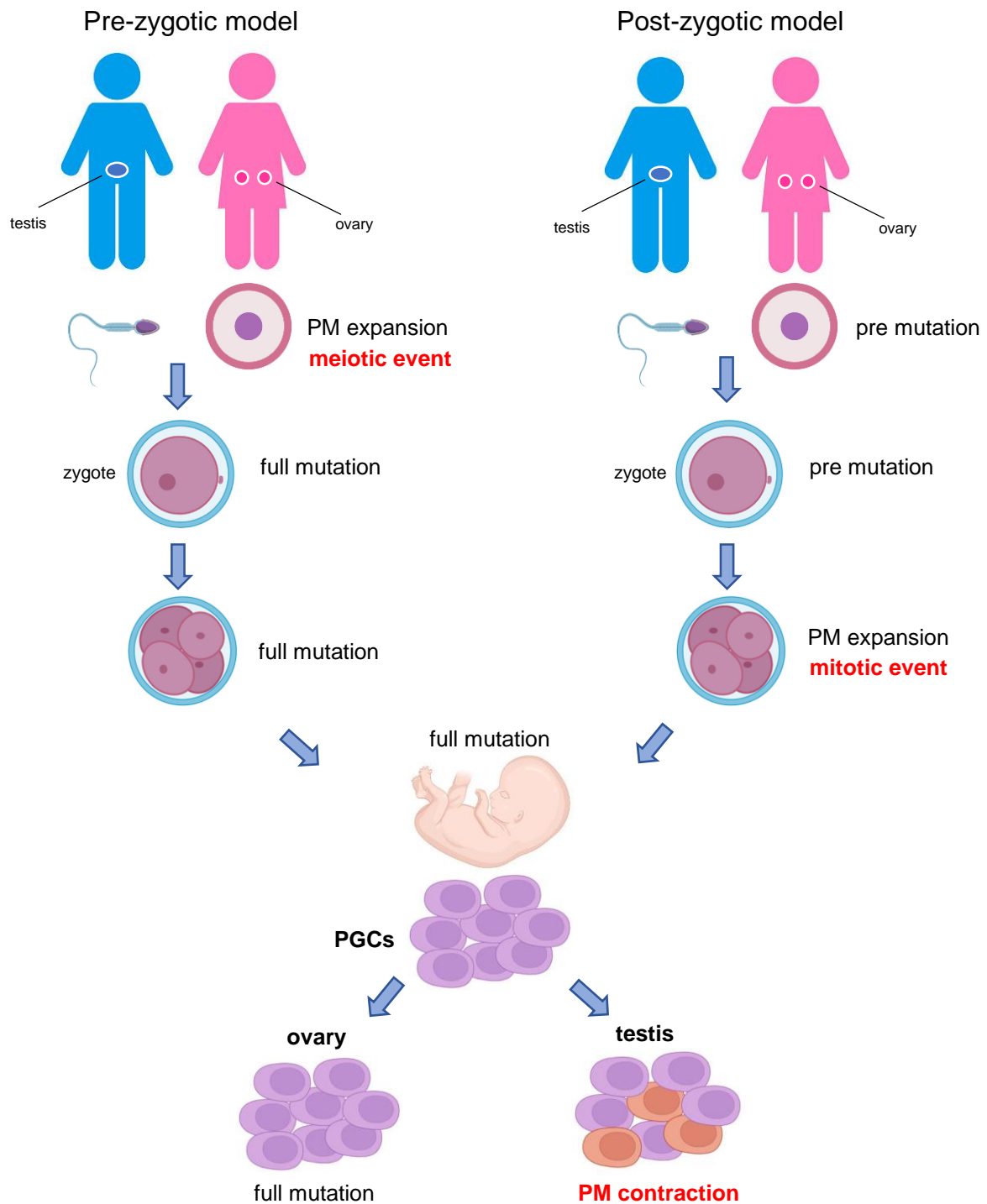
Figure 1. Two models to explain PM-to-FM expansion. The prezygotic model speculates that PM would jump to FM only during maternal germline (meiosis), possibly during recombination, and would further expand during early embryo development (on the left). The postzygotic model hypothesizes that an oocyte carrying a PM would expand to FM during the very first stages of embryogenesis (mitosis) before primordial germ cell (PGC) segregation (on the right). This model does not readily explain why a PM allele transmitted by a father to his daughters cannot expand to FM.

2. The FRAXA Fragile Site and Secondary Structures at the CGG Repeat

In 1969, Lubs wrote, "The marker X chromosome described in this paper offers an additional mechanism by which the expression of the effects of a chromosomal variant may vary from generation to generation or individual to individual and permits several basic observations about the behavior and identification of the human X chromosome" [30]. He was the first to describe a marker folate-sensitive fragile site (later denominated FRAXA) on the long arm of chromosome X (Xq27.3) in males affected by ID and their unaffected mothers of a three-generations family. Only in 1991 the *FMR1* gene with its polymorphic CGG tract was cloned in this position [3]. FRAXA belongs to the rare fragile sites, whose main feature is the late replication during S phase of the cell cycle for normal size alleles, but in the G2/M phase for FM alleles. Under folate deprivation, which further delays replication, cells enter mitosis without having completed replication of the *FMR1* region and display the FRAXA fragile site [31–33]. Delayed replication of FM alleles has been linked to the expanded CGG repeat, which may form secondary structures ("hairpins") that inhibit/retard progression of replication fork and may facilitate "replication slippage", allowing further expansion of the repeat itself [34]. Such secondary structures may be formed by virtually all repeated sequences and may be at the heart of the instability mechanisms of all REDs [35]. Actually, the expanded CGG tract of the *FMR1* gene may form a number of secondary structures, such as hairpins, triplex and quadruplex. Single-stranded DNA composed of repeated CGGs may form hairpin structures involving both Watson–Crick base pairs and mismatched base pairs. Triplex structures called R-loops may also form during transcription of long GC-rich repeats, when nascent RNA anneals with template DNA producing a Watson–Crick RNA:DNA hybrid and the non-hybridized DNA strand displaces as a single strand (ssDNA) [36]. R-loops, and thus active transcription, contribute to instability because the unbound strand in turn may itself form secondary structures. Moreover, single-stranded CGG repeats can fold into quadruplex structures stabilized by G quartets, where four guanines are linked by Hoogsteen hydrogen bonds forming thermodynamically stable four-stranded structures. Conversely, cytosines in the CCG strand can form i-motifs, another four-stranded structure stable in acidic environment. Thus, the *FMR1* coding strand containing CGG repeats forms G-quadruplex, while the non-coding strand with its CCG repeat will rather form an i-motif. Recent studies demonstrate formation of quadruplex structures in the CGG DNA (sense strand of the *FMR1* gene) and its transcript, while the corresponding antisense DNA strand (CCG) and the antisense transcript prefer hairpin structure and/or an i-motif conformation, respectively (Figure 2) [37]. In addition to these atypical structures, it should be remembered that the CGG•CCG duplex can adopt a left-handed Z-DNA conformation [38]. Not surprisingly, it has been shown that AGG interruptions modulate the structure of the expanded CGG repeat [39], diminishing the formation and stability of tetrahelical structures [40].

Hairpins and quadruplex structures formed by the repeated CGG tract during replication and/or transcription are thought to contribute to the instability of the sequence and ultimately to the expansion of the repeats; in turn, further expansion generates more secondary structures and more instability. Destabilizing such structures with specific compounds may contribute to mitigate the block of DNA replication and, possibly, could support the reactivation of *FMR1* gene transcription in FXS cells.
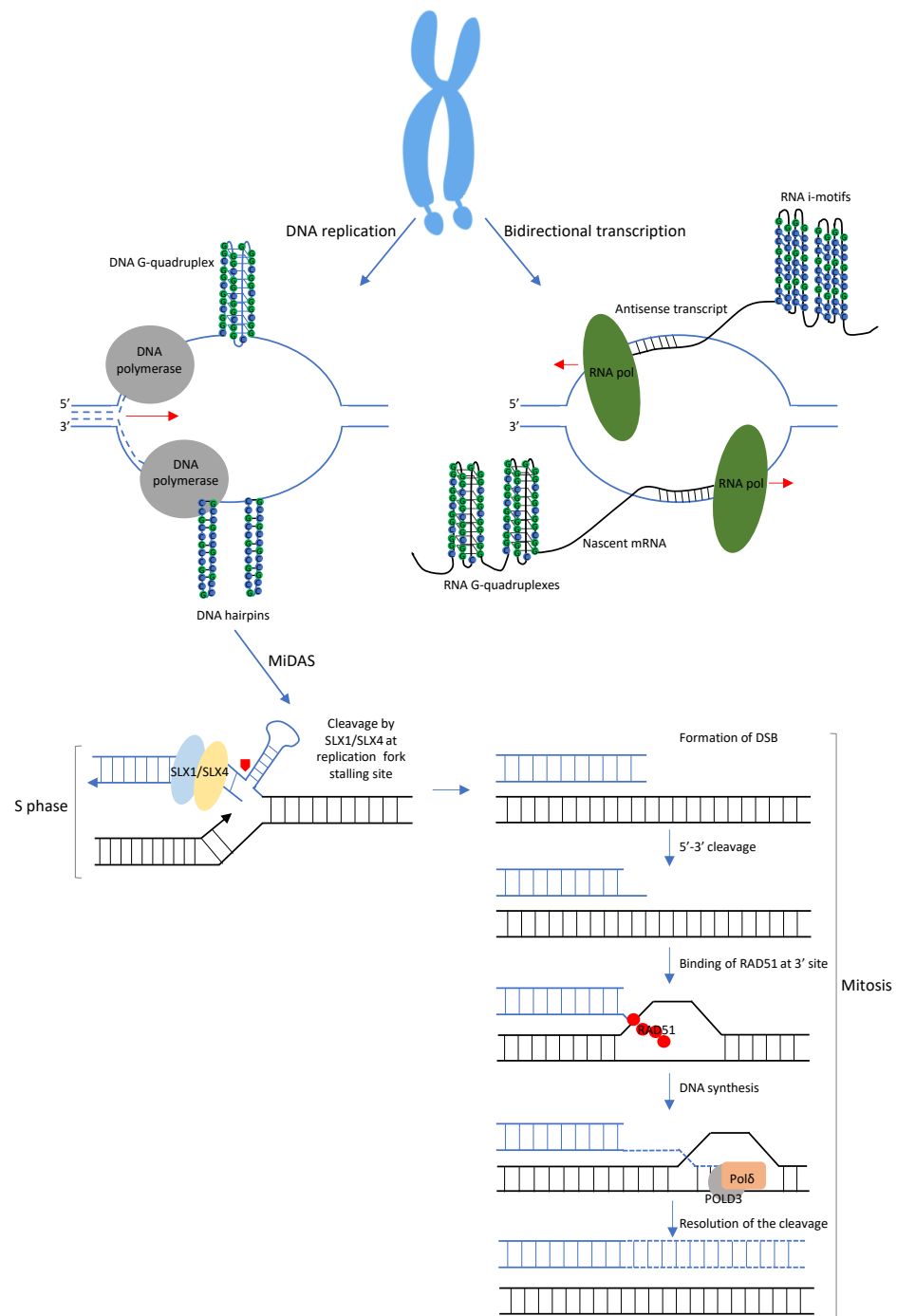
Figure 2. Secondary structures forming at the FRAXA site and their role during DNA synthesis in mitosis (MiDAS). During DNA replication at the FRAXA site, two different secondary structures may form in the CGG repeat sequence: G quadruplex structures on the leading strand and hairpins on the lagging strand. During DNA bidirectional transcription, both nascent RNAs may assume anomalous conformations: the *FMR1*-mRNA, that is G-rich, is able to form G-quadruplexes, contrariwise the antisense transcript, C-rich, prefers i-motifs. The formation of several secondary structures at the expanded CGG sequence cause failure to complete MiDAS at FRAXA locus. MiDAS starts during S phase of cell cycle with the stalling of DNA polymerase during replication due to the presence of such secondary structures. Endonucleases SLX1/SLX4 induce the formation of a DSB to remove the anomalous structures. The break will be solved during mitosis by the RAD51 recombinase, that binds 3′ end of the cleaved DNA strand and brings it near to the homologous template to complete DNA synthesis via POLD3/Polδ mechanism.

Human DNA helicases of the RecQ family, such as the Werner syndrome helicase/exonuclease (WRN), and heterogeneous nuclear ribonucleoproteins (hnRNPs), mainly CArG-binding factor A (CBF-A), were specifically demonstrated to unwind and destabilize quadruplex structures of CGG repeats [41–43]. It should be remembered that chromosomal fragility, i.e., fragile site FRAXA expression, is observed under replicative stress conditions (e.g., folate deprivation), when molecular mechanisms helping the replication fork advance along the expanded CGG repeat fail. Indeed, human DNA helicase B (HDHB) is enriched in S-phase and found localized at replication forks stalled in repetitive sequences of common fragile sites as well as the *FMR1* CGG repeat under non-stressful conditions, apparently preventing replication forks stalling [44]. How the "lesions" at fragile sites are repaired is poorly understood. Evidence suggests a role for homologous repair (HR) [45], although no protein involved in HR has been directly identified in the maintenance of fragile site stability. Delayed FRAXA locus replication under folate deprivation leads to failure of sister chromatids segregation at anaphase, generating further instability on entire X chromosome [33]. A higher sensitivity to chromosome breakage along the X chromosome was supported by the accumulation of micronuclei (biomarkers of DNA damage) in FXS patients compared with controls [46]. Delayed replication should also lead to activation of cell-cycle checkpoints related to DNA damage response, namely Ataxia-Telangiectasia and Rad3-related (ATR) and Ataxia-Telangiectasia Mutated (ATM) proteins that protect the integrity of replicating chromosomes and respond primarily to DNA double-strand breaks (DSBs), respectively. ATR-deficient cells were very sensitive to aphidicolin (a DNA polymerase inhibitor) and showed a very significant increase in gaps and breaks at fragile sites [47]. Conversely, ATM-deficient cells did not display increased fragile site expression, suggesting that DSBs are not their primary cause. Further evidence suggests that aphidicolin does not induce FRAXA and other fragile sites containing CGG repeats, but treatments that negatively impact thymidylate synthase, and thus the size and composition of nucleotide pools available for replication, may do. Fluorodeoxyuridine (FdU, an inhibitor of thymidylate synthase) increases the presence of FRAXA sites to such extent that prolonged treatments result in high frequency loss of the entire X chromosome carrying the FM allele, linking replication problems, fragile site expression, and aneuploidy [33]. FdU increases both the number of γ-H2AX foci (a marker of DNA damage) in normal and FXS patient cells and the frequency of *FMR1* gene co-localization with these foci in FXS cells. Combining FdU with KU55933, an ATM inhibitor, a reduction in chromosome fragility is noticed, suggesting that ATM contributes to FdU-induced chromosome fragility. Overall, these data support the hypothesis that FRAXA displays an alternative form of chromosome fragility in absence of FdU, which is normally prevented by an ATM-dependent process [48]. ATM protein and intracellular calcium signaling have been found increased in brain of mouse model of PM, suggesting a greater vulnerability to apoptotic activation and a possible link with neurodegeneration observed in FXTAS [49].

Chakraborty et al. [50] reported an increased accumulation of DSBs in FXS cell lines compared to unaffected control and DSBs appeared to co-localize with R-loop forming sequences. The exogenous re-expression of FMRP in FXS fibroblasts reduces R-loop-induced DSBs, suggesting that FMRP also plays a role in genome maintenance, preventing accumulation of R-loops.

In FXS fibroblasts and lymphoblastoid cells under folate stress, the FRAXA locus undergoes to DNA synthesis in mitosis (MiDAS), a process occurring through break-induced DNA replication and that requires the SLX1/SLX4 endonuclease complex, RAD51 recombinase and POLD3, a subunit of polymerase delta [51]. These authors reported failure to complete MiDAS at FRAXA site leading to severe locus instability and anomalous segregation during mitosis. Inhibition of MiDAS prevents chromosome fragility but increases the frequency of abnormal chromosome segregation. The fragility is the result of the delayed chromosome condensation that occurs when MiDAS fails to complete in time (Figure 2).

The interplay between instability, fragile site formation and late replication is more apparent in cells with high rate of cell division and faster replication, but less obvious for female gametes and neurons.

3. Mouse Models of CGG Instability

In the last two decades, a large body of literature has been published exploring factors and mechanisms which affect repeat (in)stability, although it is difficult to model some events in animal models. Transgenic mice have been developed to study the repeat expansion mechanism and FXS pathology caused by CGG expansions, and, in particular, *Fmr1* knock-in (KI) mice have been intensively studied [52–58]. The *Fmr1* KI mice may recapitulate some features of repeat instability since humans and mice share some molecular basis. In fact, it seems that both species show expansions and contractions events, even if expansions are more frequent than contractions in the PM range. Furthermore, expansion only occurs when the PM allele is on the active X chromosome; thus, expansion is found in approximately 50% of cells in females with normal X chromosome inactivation, confirming that DNA methylation has a protective effect on instability [57]. In addition, in mice and humans, expansions occur during male and female gametogenesis, even if expansions in male mice are observed in the spermatogonia stem cells, cells with high proliferation rate, while in female mice expansions occur in non-dividing oocytes [59]. As in humans, maternal age has also been reported to affect PM expansion in *Fmr1* KI mice, since more expansions were found in older oocytes than younger ones. However, since human reproductive lifespan is longer than that of mice, human oocytes may accumulate more CGG repeats, resulting in very large FM alleles [59]. Moreover, *Fmr1* KI mouse models are subject to other genetic modifiers that affect repeat instability in other human REDs [60,61]. The expansion mechanisms probably involve DNA damage repair or recombination events rather than chromosomal replication failure, suggesting cell proliferation is not exclusively required for expansion. In fact, it seems that different cell types have different propensity for expansion in *Fmr1* KI mice. Organs, such as heart, show very little expansion compared to testes and liver, where expansions continue to accumulate [58,59]. The high level of expansion in the brain of *Fmr1* KI mice compared to little expansion in blood suggests that in some FXS patients, the repeat size expansion found in blood may not correspond to what is found in cells related to the pathology, i.e., neurons [58]. Several proteins involved in DNA repair pathways have been identified to promote or protect against expansions in *Fmr1* KI models. For example, mutations in mismatch repair (MMR) proteins, including MSH2, MSH3, MSH6 and MLH3, affect the extent of expansion by reducing expansions altogether [57,62–64]. MSH3 (the MSH2-binding partner in the MutSβ complex) is required for 98% of germ line expansions and all somatic expansions in mice increase the stability of CCG-hairpins. Two different contraction mechanisms operate in the mouse model, a MutSβ-independent one that generates small contractions and a MutSβ-dependent one that generates larger ones [65]. It seems that two exonucleases, EXO1 and FAN1, play a protective role: if they are mutated, the expansion rate increases. EXO1 affects expansions in the germ line but not in the brain [63], contrary to FAN1 [66]. Recently, it was found that the expansion mechanism competes with non-homologous end-joining (NHEJ) repair for the processing of a DSBs intermediate, while a protective role against expansion is played by LIG4, a ligase essential for NHEJ [67]. Finally, in this mouse model, the same genetic factors required for expansion in dividing cells are required for expansion in non-dividing ones.

4. CGG Instability during Gametogenesis and in Tissues with Low Rate of Cell Division

However, not just how, but also when and where PM expansion exactly occurs it is still being debated. It has been proposed that PM-to-FM expansion does not occur during meiosis but in an early post-zygotic stage [68,69]. Two main models emerge from scientific literature that explain PM-to-FM expansion: a prezygotic model, in which PM expansion would occur during recombination events in maternal germline and later expands during early embryo development; a postzygotic model, in which an oocyte carrying a PM would expand to FM during the very first stages of embryogenesis before primordial germ cell (PGC) segregation (Figure 1) [70]. Females carriers of a PM allele, in fact, show the presence of the expansion already in pre-implantation embryos [71,72]. Not even studies on embry-

onic stem cells (ESCs) isolated by FXS fetuses that preserve FM expansion are diriment to solve this question [73–75]. Assuming the postzygotic model, an embryo carrying a FM derived from maternal PM should exhibit PM expansions in gametes and not in other tissues, since the sex of embryo is not determined before PGC segregation. From the analysis of intact ovaries and testes derived from FM fetuses of PM mothers, CGG expansions in the range of PM were not detected in gametes, but only FM were found as in other somatic fetal tissues (i.e., brain, skin and lung) [76]. Only a 17-week testis showed a faint PM signal with very low levels of FMRP expression. Together these data disproved the mitotic model of CGG expansion in the early embryonic stages sparing the germline. The PM detected in the sperm of FXS affected males may derive from a positive selection mediated by mitotic contraction events, taking place in spermatogonia. Furthermore, in the paper by Malter and co-authors [76] a relevant issue has been addressed: the FM expansion precedes its methylation in both males and females. In fact, the authors detected an unmethylated FM in fetal oocytes and concluded that the PM-to-FM expansion event may occur in the maternal germline or very early in embryogenesis, prior to *de novo* DNA methylation. Since instability of the repeated tract involves both maternally derived expansions and contractions in gametes of adult FM males, it has been proposed that the absence of aberrant CpG methylation, such as in PGCs, may enhance repeat deletions through an undefined process. Two articles of the same group addressed this matter in 2015, using an SV40 primate replication system containing a CGG expansion and bidirectional ORIs (origin of replication sites), which were methylated using *Sss*I methylase, when necessary [77,78]. With this model they demonstrated that CGG contractions occur during replication and depend on the length of the CGG repeat and CpG methylation. Precisely, the proximity of ORI to the CGG tract and the positioning of the repeat on the lagging strand turned out to be the principal determinants of CGG contractions, while the role of DNA methylation is to stabilize the repeated sequence. Taken together, these results make parental gender the major risk factor for the transmission of FXS, although not the only one.

Since expansion seems to occur in non-dividing cells such as oocytes and neurons, DNA repair systems may be involved in the pathogenic mechanism. One model suggests a role for base excision repair (BER) of 7,8-dihydro-8-oxoguanine (8-oxoG), the most common oxidation product of DNA (Figure 3A). During repair events, strand slippage or displacement may generate hairpins that are bound by the MutS proteins [79]. Formation of secondary structures such as hairpins may trigger multiple rounds of BER, since guanines included in the hairpins are susceptible to DNA damage and are difficult to repair. The role of BER in repeat expansions is supported by the observation that mutations in *OGG1* and *NEIL1*, encoding DNA glycosylases involved in 8-oxoG repair, reduce expansions in a mouse model of Huntington disease [80,81]. The BER mechanism may also play a role during transcription, when single-stranded DNA regions form R-loops that are more susceptible to oxidative damage, generating 8-oxoG. Expansion of the *FMR1* gene located on the active X chromosome may be due to the BER mechanism [57].

On the other hand, contraction events seem to follow different pathways (Figure 3B). Contractions of the CGG repeat are apparently very common during spermatogenesis, which may explain why PM males do not transmit a FM to their daughters and why FXS males have only PM alleles in their sperm. Methylated FM alleles do contract [82], thus some contractions are transcription independent, but small contractions seem to be related to DNA synthesis. It has been proposed that hairpin formation on the repeated template strand cause strand-slippage during replication, thus the nascent strand includes fewer repeats than the template one. A second round of DNA replication would generate a small, contracted allele. Conversely, larger deletions characterized in normal, PM and FM carriers present with breakpoints associated with microhomologies (MHs) of 2–9 nucleotides [83,84], which may be responsible of a process known as microhomology mediated end joining (MMEJ). MMEJ represents a repair mechanism of a DSB and starts from an end-resection that reveals MHs at both sides of the break. Non-homologous flaps will be removed

allowing the annealing of MHs strands, while gaps will be filled by Polθ or Polβ and the ends will be ligated by either ligase 3 (Lig3) or ligase 1 (Lig1) [85].
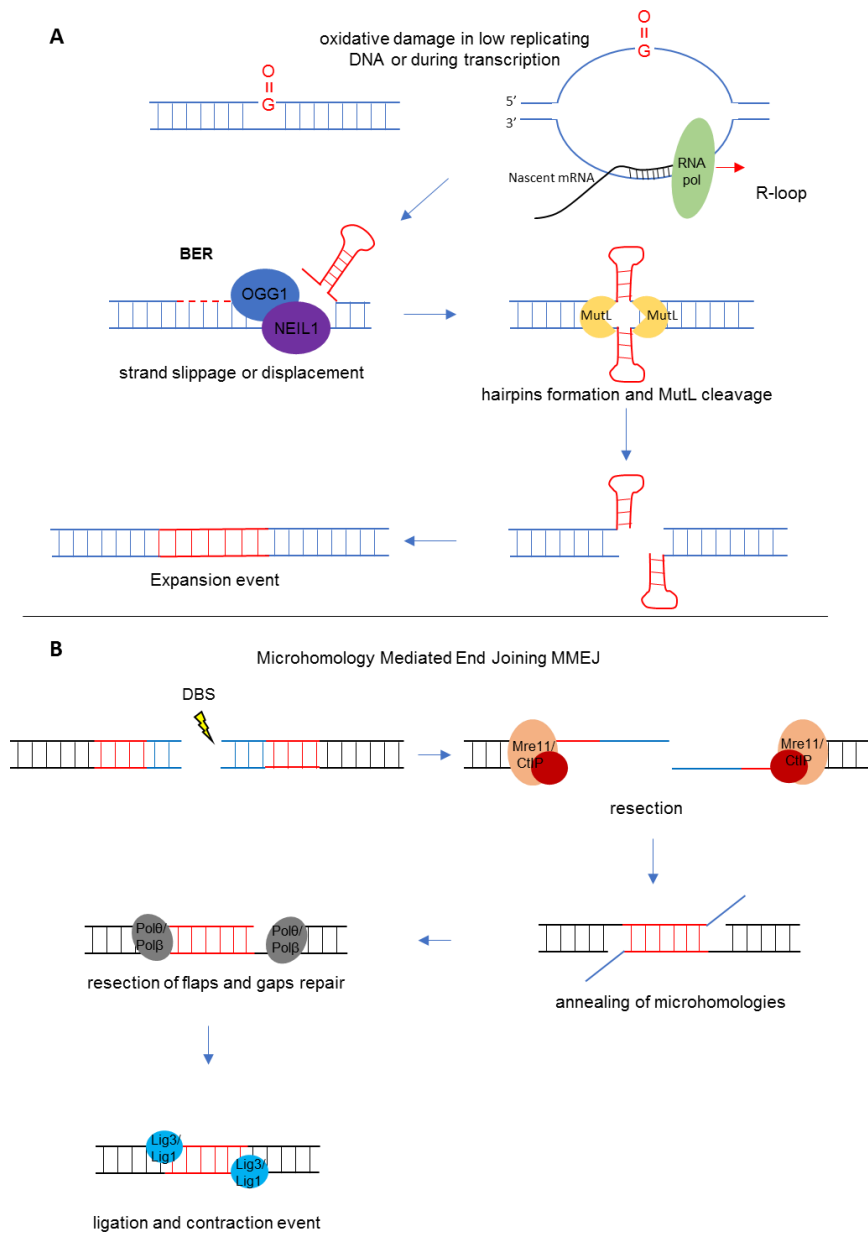

Figure 3. *Proposed mechanisms of CGG instability in tissues with low rate of cell division.* (**A**) Oxidation of guanine in DNA strands or during transcription in the displaced DNA strand of the R loop activates BER mechanisms through the recruitment of OGG1 and NEIL1 enzymes, that remove the anomalous base. This damage generates a displacement of the DNA strand favoring hairpin formation. MutL enzymes bind the region containing hairpins in an attempt to remove them; the resulting cleavage produces a DSB which will be repaired, causing the repeat expansion. (**B**) Large contraction events could be addressed by MMEJ repair mechanism. A DSB may expose sequences of microhomology (MH, in blue) at both sides of the break, which may anneal through the removal of nonhomologous flaps at both ends. Polθ/Polβ complex will fill the gaps, which will be ligated by Lig3 and Lig1, resulting in contraction.

Mechanisms other than repair, such as those related to DNA synthesis and active transcription (see below), cannot be completely excluded also in cells with low rate of division.

5. CGG Instability in Tissues with High Rate of Cell Division

Although CGG expansions display their highest instability in the germline, also the analysis of cultured single lymphocytes exhibited modest mitotic instability, confirming that in FXS the CGG repeat is unstable both through germline transmission and in somatic cells during the lifetime of the individual [11,86]. In other words, some post-zygotic mitotic instability occurs, and specific tissues show heterogeneous degrees of CGG size mosaicism [87]. The length of the expanded repeat remains the main determinant of FXS, but both in cis elements (localized in the expansion itself) and in trans elements (physically distant from the CGG repeat) may contribute to instability, as summarized in Figure 4.
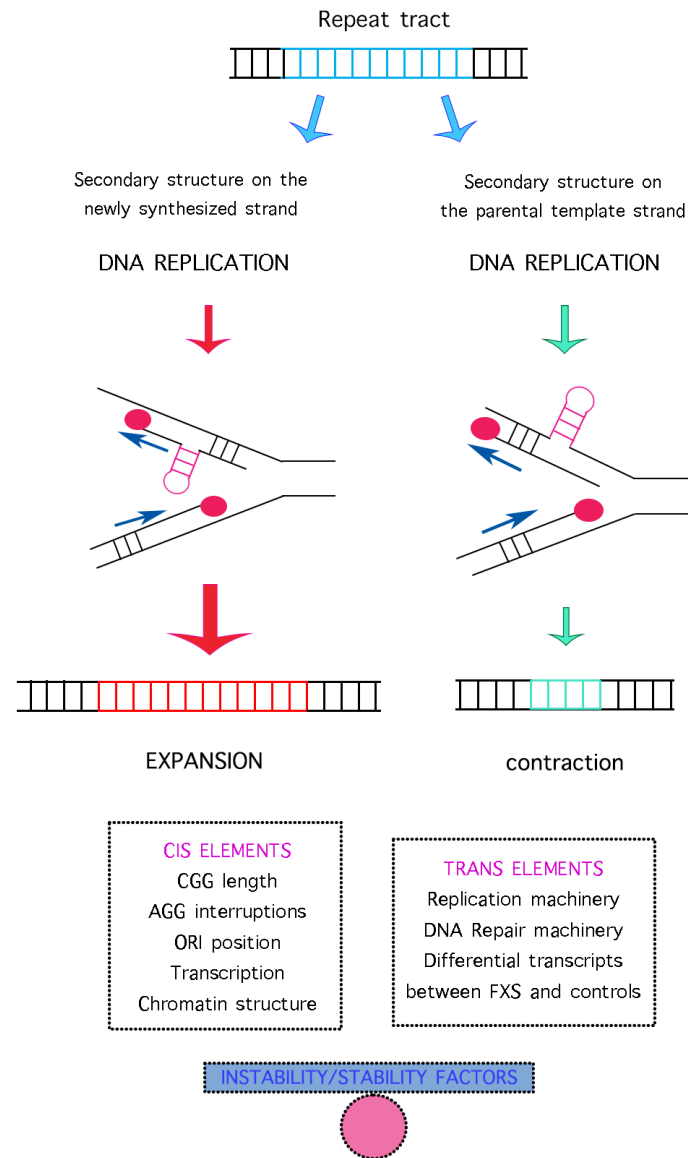

Figure 4. Factors that influence expansion and contraction events during DNA replication at the CGG repeat. The presence of the repeated CGG sequence (blue bars) on the nascent strand induces the formation of hairpins or other secondary structures. At the end of a second round of replication, an expansion (red bars) will result (on the left), whereas the presence of similar secondary structures on the template strand results in a stall with a late restart of the DNA polymerase that causes the exclusion of the repeated tract included in the secondary structure, thus leading to contraction events (green bars) after the second round of replication (on the right). The balance between listed cis and trans elements will hold the repeat sequence stable. An imbalance of these factors will favor instability (contractions or expansions).

The first in cis elements are the length of the CGG repeats and the presence of interspersed AGGs. The ability of the unstable DNA to expand in somatic cells is again likely due to the secondary structures formed by the repeated DNA: if the CGG sequence is longer (possibly without interrupting AGGs), more hairpins will form on the nascent strand and the repeat may increase more in size. The main expansion model is associated with replication fork stalling and restarting with polymerase slippage (Figure 4). The presence of unwound DNA during replication allows secondary structures formation (most likely guanine quadruplexes), which, if not correctly recognized by the repair mechanisms, would favor contractions or expansions, depending on which strand (template or nascent) they will be formed [88]. Slipped-strand DNA mispairing may be favored also during transcription through the same mechanisms [89]. During DNA replication, the probability of expansions or contractions are related with the CGG repeat orientation relative to the ORIs. In the normal CGG tract of the *FMR1* gene, two ORIs on both strands are active, hence expansions and contractions happen with the same probability and repeat size remains stable. If only one ORI works, as it has been reported in FM alleles [90,91], replication leads to secondary structures formation on the nascent strand containing CGG repeats and consequent expansion of the tract. According to the "origin switch" model, the location of the closest ORI to the repeat changes with a consequent switch in the replication fork direction. Consequently, secondary structures will form in the lagging strand and not in the leading one, promoting CGG instability. In the "fork shift" model, instead, altered replication fork progression might shift the location of the Okazaki initiation zone relative to the repeat tract, allowing the formation of a mutagenic secondary DNA structure and ultimately leading to repeat instability [1,89]. DNA slippage and secondary structure formation on the newly synthesized strand, followed by reannealing on the repeated sequence, will generate an expansion in the repeat tract that will not be compensated by contraction events [1,35,89,92,93]. Leading- and lagging-strand behave differently during replication; a possible explanation could be that Polε and Polδ respond differently to replication stress [94]. As McMurray reviewed in 1995, DNA stands at a crossroad in presence of secondary structures: maintaining or repairing the heteroduplex molecule. Failure in repairing the secondary structure results in non-Mendelian segregation of the repeat sequence at the first mitotic division. Consequently, the ratio of the insertion allele to the original allele will be higher in the daughter cells during mitosis [93]. Hairpin and G-quadruplex must be considered as barriers during replication, and triggers of genome instability, although the precise mechanism through which the CGG repeat causes fork stalling remains yet unclear.

Another cis-acting factor shown to influence repeat instability is transcription: bidirectional transcription, known to occur at the *FMR1* locus [95], increases instability compared to unidirectional transcription [96]. The transcription complex, passing through the expansion, leave the non-template strand uncoupled, promoting the hairpin and quadruplex formation and consequently increasing instability [24].

Chromatin structure is a further cis-element: histone modifications, DNA methylation, chromatin binding proteins and chromatin looping may all affect repeat instability by influencing transcription levels, facilitating the formation of secondary structures, or promoting access to DNA by repair/recombination machinery [24,97].

To summarize, several cis-modifiers contribute to increasing the instability of the CGG expanded repeat, such as AGG interruptions, secondary structure formation, ORI position, transcriptional activity (both sense and antisense) and chromatin organization.

To complicate the picture, it should be considered that secondary structures are closely related to trans-modifiers and that DNA replication and repair mechanisms can play together to increase instability [98]. Recently, it was observed that proteins involved in DNA repair, such as MSH2, MLH3, MSH6, PCNA, RPA1, FEN1, LIG1, HMGB1, Polβ, influence trinucleotide repeat instability [24]. For example, Polβ, a protein involved in BER, was proved to be involved in CGG repeat expansion in the FXS murine model [99]. BER could facilitate CGG instability by repairing the DNA oxidative damage. Furthermore, differential

expression of transcripts encoding components of DNA repair and DNA replication has been detected in FXS patients relative to control individuals [100].

Recently, Kononenko et al. [101] proposed a role for break-induced replication (BIR) in the expansion of the *FMR1* CGG repeat: replication forks stalling cause SMARCAL1 recruitment to promote forks resolution; after isomerization and enzymatic resolution of the Holliday junction, fork regression is converted into an one-ended DSB, with only one free end that will be repaired by break-induced replication (BIR). This single strand can invade the sister chromatid with the help of RAD52 and RAD51 and the newly formed loop is extended by Polδ; given that the invasion of the repetitive strand is out of register, it can easily result in expansion, contraction or mutagenesis [101]. Furthermore, it has been proposed that unstable DNA repeats dramatically increase mutation rates in surrounding DNA segments and that these mutations can occur up to kilobases away from the repetitive tract itself: such a process has been called RIM, i.e., repeat-induced mutagenesis [102].

Finally, three other mechanisms should be mentioned that may promote repeat expansions: failure to remove displaced 5′ DNA flaps and their subsequent incorporation into DNA, misalignment during recombinational repair and synthesis during DSB repair.

6. Concluding Remarks

Despite the amount of experimental data produced so far, indicating that the determinants of *FMR1* gene instability reside mainly in the CGG repeat itself, more research is needed to better understand all the mechanisms responsible for expansion and contraction events and their timing. In fact, only precise knowledge of the pathogenetic pathways will potentially suggest tailored therapeutic intervention. For example, recent works targeted small compounds to CGG secondary structures in order to interfere with their deleterious effects; indeed, antisense oligonucleotides and small molecules targeting CGG-RNA hairpins have been shown to reduce R-loop formation with beneficial effects on the *Fmr1* KI mouse models [103–106]. Although an in vivo reduction in FM alleles in FXS patients seems unrealistic, detailed understanding of the *FMR1* CGG repeat instability and consequent gene silencing may disclose new therapeutic approaches to treat Fragile X-related disorders.

Author Contributions: E.T., V.N. and C.P. wrote and edited the review; P.C. edited and supervised the review. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pearson, C.E.; Edamura, K.N.; Cleary, J.D. Repeat instability: Mechanisms of dynamic mutations. *Nat. Rev. Genet.* **2005**, *6*, 729–742. [CrossRef]
2. Depienne, C.; Mandel, J.L. 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? *Am. J. Hum. Genet.* **2021**, *108*, 764–785. [CrossRef]
3. Verkerk, A.J.; Pieretti, M.; Sutcliffe, J.S.; Fu, Y.-H.; Kuhl, D.P.; Pizzuti, A.; Reiner, O.; Richards, S.; Victoria, M.F.; Zhang, F.; et al. Identification of a gene (FMR1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **1991**, *65*, 905–914. [CrossRef]
4. Rousseau, F.; Heitz, D.; Oberle, I.; Mandel, J.L. Selection in blood cells from female carriers of the fragile X syndrome: Inverse correlation between age and proportion of active X chromosomes carrying the full mutation. *J. Med Genet.* **1991**, *28*, 830–836. [CrossRef]
5. Hunter, J.; Rivero-Arias, O.; Angelov, A.; Kim, E.; Fotheringham, I.; Leal, J. Epidemiology of fragile X syndrome: A systematic review and meta-analysis. *Am. J. Med Genet. Part A* **2014**, *164*, 1648–1658. [CrossRef]

6. Monaghan, K.G.; Lyon, E.; Spector, E.B. ACMG Standards and Guidelines for fragile X testing: A revision to the disease-specific supplements to the Standards and Guidelines for Clinical Genetics Laboratories of the American College of Medical Genetics and Ge-nomics. *Genet. Med.* **2013**, *15*, 575–586. [CrossRef]

7. Fu, Y.-H.; Kuhl, D.P.; Pizzuti, A.; Pieretti, M.; Sutcliffe, J.S.; Richards, S.; Verkert, A.J.; Holden, J.J.; Fenwick, R.G.; Warren, S.T.; et al. Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. *Cell* **1991**, *67*, 1047–1058. [CrossRef]

8. Eichler, E.E.; Holden, J.J.; Popovich, B.W.; Reiss, A.L.; Snow, K.; Thibodeau, S.N.; Richards, C.S.; Ward, P.A.; Nelson, D.L. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat. Genet.* **1994**, *8*, 88–94. [CrossRef]

9. Yrigollen, C.; Tassone, F.; Durbin-Johnson, B.; Tassone, F. The Role of AGG Interruptions in the Transcription of FMR1 Premutation Alleles. *PLoS ONE* **2011**, *6*, e21728. [CrossRef]

10. Sullivan, A.K.; Crawford, D.C.; Scott, E.H.; Leslie, M.L.; Sherman, S.L. Paternally Transmitted FMR1 Alleles Are Less Stable than Maternally Transmitted Alleles in the Common and Intermediate Size Range. *Am. J. Hum. Genet.* **2002**, *70*, 1532–1544. [CrossRef]

11. Nolin, S.L.; Glicksman, A.; Tortora, N.; Allen, E.; Macpherson, J.; Mila, M.; Vianna-Morgante, A.M.; Sherman, S.L.; Dobkin, C.; Latham, G.J.; et al. Expansions and contractions of the FMR1 CGG repeat in 5508 transmissions of normal, intermediate, and premutation alleles. *Am. J. Med. Genet.* **2019**, *179A*, 1148–1156. [CrossRef] [PubMed]

12. Fernandez-Carvajal, I.; Posadas, B.L.; Pan, R.; Raske, C.; Hagerman, P.J.; Tassone, F. Expansion of an FMR1 Grey-Zone Allele to a Full Mutation in Two Generations. *J. Mol. Diagn.* **2009**, *11*, 306–310. [CrossRef] [PubMed]

13. Nolin, S.L.; Brown, W.T.; Glicksman, A.; Houck, J.G.E.; Gargano, A.D.; Sullivan, A.; Biancalana, V.; Bröndum-Nielsen, K.; Hjalgrim, H.; Holinski-Feder, E.; et al. Expansion of the Fragile X CGG Repeat in Females with Premutation or Intermediate Alleles. *Am. J. Hum. Genet.* **2003**, *72*, 454–464. [CrossRef] [PubMed]

14. Yrigollen, C.; Martorell, L.; Durbin-Johnson, B.; Naudo, M.; Genoves, J.; Murgia, A.; Polli, R.; Zhou, L.; Barbouth, D.; Rupchock, A.; et al. AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. *J. Neurodev. Disord.* **2014**, *6*, 24. [CrossRef]

15. Sherman, S.L.; Jacobs, P.A.; Morton, N.E.; Froster-Iskenius, U.; Howard-Peebles, P.N.; Nielsen, K.B.; Partington, M.W.; Sutherland, G.R.; Turner, G.; Watson, M. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. *Qual. Life Res.* **1985**, *69*, 289–299. [CrossRef]

16. Zeesman, S.; Zwaigenbaum, L.; Whelan, D.T.; Hagerman, R.J.; Tassone, F.; Taylor, S.A. Paternal transmission of fragile X syndrome. *Am. J. Med. Genet.* **2004**, *129A*, 184–189. [CrossRef]

17. Alvarez-Mora, M.I.; Guitart, M.; Rodriguez-Revenga, L.; Madrigal, I.; Gabau, E.; Milà, M. Paternal transmission of a *FMR1* full mutation allele. *Am. J. Med Genet. Part A* **2017**, *173*, 2795–2797. [CrossRef]

18. Reyniers, E.; Vits, L.; De Boulle, K.; Van Roy, B.; Van Velzen, D.; De Graaff, E.; Verkerk, A.J.; Jorens, H.Z.; Darby, J.K.; Oostra, B.; et al. The full mutation in the FMR–1 gene of male fragile X patients is absent in their sperm. *Nat. Genet.* **1993**, *4*, 143–146. [CrossRef]

19. Hagerman, P.J.; Hagerman, R.J. The Fragile-X Premutation: A Maturing Perspective. *Am. J. Hum. Genet.* **2004**, *74*, 805–816. [CrossRef]

20. Greco, C.M.; Tassone, F.; Garcia-Arocena, D.; Tartaglia, N.; Coffey, S.M.; Vartanian, T.K.; Brunberg, J.A.; Hagerman, P.J.; Hagerman, R.J. Clinical and Neuropathologic Findings in a Woman With the FMR1 Premutation and Multiple Sclerosis. *Arch. Neurol.* **2008**, *65*, 1114–1116. [CrossRef]

21. Allingham-Hawkins, D.J.; Babul-Hirji, R.; Chitayat, D.; Holden, J.J.; Yang, K.T.; Lee, C.; Hudson, R.; Gorwill, H.; Nolin, S.L.; Glicksman, A.; et al. Fragile X premutation is a significant risk factor for premature ovarian failure: The Interna-tional Collaborative POF in Fragile X study—Preliminary data. *Am. J. Med. Genet.* **1999**, *83*, 322–325. [CrossRef]

22. Glineburg, M.R.; Todd, P.K.; Charlet-Berguerand, N.; Sellier, C. Repeat-associated non-AUG (RAN) translation and other molecular mechanisms in Fragile X Tremor Ataxia Syndrome. *Brain Res.* **2018**, *1693*, 43–54. [CrossRef] [PubMed]

23. Yrigollen, C.M.; Durbin-Johnson, B.; Gane, L.; Nelson, D.L.; Hagerman, R.; Hagerman, P.J.; Tassone, F. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syn-drome. *Genet. Med.* **2012**, *14*, 729–736. [CrossRef]

24. Zhao, X.-N.; Usdin, K. Ups and Downs: Mechanisms of Repeat Instability in the Fragile X-Related Disorders. *Genes* **2016**, *7*, 70. [CrossRef]

25. Maia, N.; Loureiro, J.R.; Oliveira, B.; Marques, I.; Santos, R.; Jorge, P.; Martins, S. Contraction of fully expanded FMR1 alleles to the normal range: Predisposing haplotype or rare events? *J. Hum. Genet.* **2017**, *62*, 269–275. [CrossRef]

26. Tabolacci, E.; Pietrobono, R.; Maneri, G.; Remondini, L.; Nobile, V.; Della Monica, M.; Pomponi, M.G.; Genuardi, M.; Neri, G.; Chiurazzi, P. Reversion to Normal of FMR1 Expanded Alleles: A Rare Event in Two Independent Fragile X Syndrome Families. *Genes* **2020**, *11*, 248. [CrossRef]

27. Wöhrle, D.; Hirst, M.C.; Kennerknecht, I.; Davies, K.E.; Steinbach, P. Genotype mosaicism in fragile X fetal tissues. *Hum. Genet.* **1992**, *89*, 114–116. [CrossRef]

28. Pretto, D.I.; Mendoza-Morales, G.; Lo, J.; Cao, R.; Hadd, A.; Latham, G.J.; Durbin-Johnson, B.; Hagerman, R.; Tassone, F. CGG allele size somatic mosaicism and methylation inFMR1premutation alleles. *J. Med Genet.* **2014**, *51*, 309–318. [CrossRef]

29. Taylor, A.K.; Tassone, F.; Dyer, P.N.; Hersch, S.M.; Harris, J.B.; Greenough, W.T.; Hagerman, R.J. Tissue heterogeneity of the FMR1 mutation in a high-functioning male with fragile X syndrome. *Am. J. Med. Genet.* **1999**, *84*, 233–239. [CrossRef]

30. Lubs, H.A. A marker X chromosome. *Am. J. Hum. Genet.* **1969**, *21*, 231–244.

31. Webb, T. Delayed replication of Xq27 in individuals with the fragile X syndrome. *Am. J. Med Genet.* **1992**, *43*, 1057–1062. [CrossRef]

32. Hansen, R.S.; Canfield, T.K.; Lamb, M.M.; Gartler, S.M.; Laird, C.D. Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* **1993**, *73*, 1403–1409. [CrossRef]

33. Bjerregaard, V.A.; Garribba, L.; McMurray, C.T.; Hickson, I.D.; Liu, Y. Folate deficiency drives mitotic missegregation of the human *FRAXA* locus. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 13003–13008. [CrossRef] [PubMed]

34. Chen, X.; Mariappan, S.V.; Catasti, P.; Ratliff, R.; Moyzis, R.K.; Laayoun, A.; Smith, S.S.; Bradbury, E.M.; Gupta, G. Hairpins are formed by the single DNA strands of the fragile X triplet repeats: Structure and biological im-plications. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5199–5203. [CrossRef]

35. McMurray, C.T. DNA secondary structure: A common and causative factor for expansion in human disease. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1823–1825. [CrossRef]

36. Loomis, E.; Sanz, L.A.; Chedin, F.; Hagerman, P.J. Transcription-Associated R-Loop Formation across the Human FMR1 CGG-Repeat Region. *PLoS Genet.* **2014**, *10*, e1004294. [CrossRef]

37. Ajjugal, Y.; Kolimi, N.; Rathinavelan, T. Secondary structural choice of DNA and RNA associated with CGG/CCG trinucleo-tide repeat expansion rationalizes the RNA misprocessing in FXTAS. *Sci. Rep.* **2021**, *11*, 8163. [CrossRef]

38. Renčiuk, D.; Kypr, J.; Vorlíčková, M. CGG repeats associated with fragile X chromosome form left-handed Z-DNA structure. *Biopolymers* **2010**, *95*, 174–181. [CrossRef]

39. Jarem, D.A.; Huckaby, L.V.; Delaney, S. AGG interruptions in (CGG)(n) DNA repeat tracts modulate the structure and ther-modynamics of non-B conformations in vitro. *Biochemistry* **2010**, *49*, 6826–6837. [CrossRef]

40. Weisman-Shomer, P.; Cohen, E.; Fry, M. Interruption of the fragile X syndrome expanded sequence d(CGG)(n) by inter-spersed d(AGG) trinucleotides diminishes the formation and stability of d(CGG)(n) tetrahelical structures. *Nucleic Acids Res.* **2000**, *28*, 1535–1541. [CrossRef]

41. Kamath-Loeb, A.S.; Loeb, L.A.; Johansson, E.; Burgers, P.M.; Fry, M. Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facili-tate copying of tetraplex and hairpin structures of the d(CGG)n trinucleotide repeat sequence. *J. Biol. Chem.* **2001**, *276*, 16439–16446. [CrossRef]

42. Weisman-Shomer, P.; Cohen, E.; Fry, M. Distinct domains in the CArG-box binding factor A destabilize tetraplex forms of the fragile X expanded sequence d(CGG)n. *Nucleic Acids Res.* **2002**, *30*, 3672–3681. [CrossRef]

43. Khateb, S.; Weisman-Shomer, P.; Hershco, I.; Loeb, L.A.; Fry, M. Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)n is mediated by homo-log-conserved domains in three members of the hnRNP family. *Nucleic Acids Res.* **2004**, *32*, 4145–4154. [CrossRef]

44. Guler, G.D.; Rosenwaks, Z.; Gerhardt, J. Human DNA Helicase B as a Candidate for Unwinding Secondary CGG Repeat Structures at the Fragile X Mental Retardation Gene. *Front. Mol. Neurosci.* **2018**, *11*, 138. [CrossRef]

45. Sasaki, M.; Takata, M.; Sonoda, E.; Tachibana, A.; Takeda, S. Recombination repair pathway in the maintenance of chromosomal integrity against DNA interstrand crosslinks. *Cytogenet. Genome Res.* **2004**, *104*, 28–34. [CrossRef]

46. Özer, L.; Ruhi, H.I.; Bökesoy, I. Analysis of Xq27.3 fragility using the micronucleus-fluorescence In situ hybridization assay. *Genome Integr.* **2020**, *11*, 1. [CrossRef]

47. Casper, A.; Nghiem, P.; Arlt, M.F.; Glover, T.W. ATR Regulates Fragile Site Stability. *Cell* **2002**, *111*, 779–789. [CrossRef]

48. Kumari, D.; Somma, V.; Nakamura, A.J.; Bonner, W.M.; D'Ambrosio, E.; Usdin, K. The role of DNA damage response pathways in chromosome fragility in Fragile X syndrome. *Nucleic Acids Res.* **2009**, *37*, 4385–4392. [CrossRef]

49. Robin, G.; López, J.R.; Espinal, G.M.; Hulsizer, S.; Hagerman, P.J.; Pessah, I.N. Calcium dysregulation and Cdk5-ATM pathway involved in a mouse model of fragile X-associated trem-or/ataxia syndrome. *Hum. Mol. Genet.* **2017**, *26*, 2649–2666. [CrossRef]

50. Chakraborty, A.; Jenjaroenpun, P.; Li, J.; El Hilali, S.; McCulley, A.; Haarer, B.; Hoffman, E.A.; Belak, A.; Thorland, A.; Hehnly, H.; et al. Replication Stress Induces Global Chromosome Breakage in the Fragile X Genome. *Cell Rep.* **2020**, *32*, 108179. [CrossRef]

51. Garribba, L.; Bjerregaard, V.A.; Dinis, M.M.G.; Özer, Ö.; Wu, W.; Sakellariou, D.; Pena-Diaz, J.; Hickson, I.D.; Liu, Y. Folate stress induces SLX1- and RAD51-dependent mitotic DNA synthesis at the fragile X locus in human cells. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 16527–16536. [CrossRef]

52. Bontekoe, C.J.M.; Bakker, C.E.; Nieuwenhuizen, I.M.; Van Der Linde, H.; Lans, H.; De Lange, D.; Hirst, M.; Oostra, B.A. Instability of a (CGG)98 repeat in the Fmr1 promoter. *Hum. Mol. Genet.* **2001**, *10*, 1693–1699. [CrossRef]

53. Peier, A.M.; Nelson, D.L. Instability of a Premutation-Sized CGG Repeat in FMR1 YAC Transgenic Mice. *Genomics* **2002**, *80*, 423–432. [CrossRef]

54. Fleming, K.; Riser, D.; Kumari, D.; Usdin, K. Instability of the fragile X syndrome repeat in mice: The effect of age, diet and mutations in genes that affect DNA replication, recombination and repair proficiency. *Cytogenet. Genome Res.* **2003**, *100*, 140–146. [CrossRef]

55. Entezam, A.; Biacsi, R.; Orrison, B.; Saha, T.; Hoffman, G.E.; Grabczyk, E.; Nussbaum, R.L.; Usdin, K. Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. *Gene* **2007**, *395*, 125–134. [CrossRef]

56. Brouwer, J.R.; Huizer, K.; Severijnen, L.-A.; Hukema, R.K.; Berman, R.F.; Oostra, B.A.; Willemsen, R. CGG-repeat length and neuropathological and molecular correlates in a mouse model for fragile X-associated tremor/ataxia syndrome. *J. Neurochem.* **2008**, *107*, 1671–1682. [CrossRef]

57. Lokanga, R.A.; Zhao, X.-N.; Usdin, K. The Mismatch Repair Protein MSH2 is Rate Limiting for Repeat Expansion in a Fragile X Premutation Mouse Model. *Hum. Mutat.* **2013**, *35*, 129–136. [CrossRef]

58. Lokanga, R.A.; Entezam, A.; Kumari, D.; Yudkin, D.; Qin, M.; Smith, C.B.; Usdin, K. Somatic expansion in mouse and human carriers of fragile X premutation alleles. *Hum. Mutat.* **2012**, *34*, 157–166. [CrossRef]

59. Zhao, X.-N.; Usdin, K. Timing of Expansion of Fragile X Premutation Alleles During Intergenerational Transmission in a Mouse Model of the Fragile X-Related Disorders. *Front. Genet.* **2018**, *9*. [CrossRef]

60. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell* **2015**, *162*, 516–526. [CrossRef]

61. Bettencourt, C.; Moss, D.H.; Flower, M.; Wiethoff, S.; Brice, A.; Goizet, C.; Stevanin, G.; Koutsis, G.; Karadima, G.; Panas, M.; et al. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann. Neurol.* **2016**, *79*, 983–990. [CrossRef]

62. Zhao, X.-N.; Kumari, D.; Gupta, S.; Wu, D.; Evanitsky, M.; Yang, W.; Usdin, K. Mutsβ generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. *Hum. Mol. Genet.* **2015**, *24*, 7087–7096. [CrossRef]

63. Zhao, X.; Zhang, Y.; Wilkins, K.; Edelmann, W.; Usdin, K. MutLγ promotes repeat expansion in a Fragile X mouse model while EXO1 is protective. *PLoS Genet.* **2018**, *14*, e1007719. [CrossRef]

64. Miller, C.J.; Kim, G.-Y.; Zhao, X.; Usdin, K. All three mammalian MutL complexes are required for repeat expansion in a mouse cell model of the Fragile X-related disorders. *PLoS Genet.* **2020**, *16*, e1008902. [CrossRef]

65. Zhao, X.-N.; Usdin, K. Gender and Cell-Type-Specific Effects of the Transcription-Coupled Repair Protein, ERCC6/CSB, on Repeat Expansion in a Mouse Model of the Fragile X-Related Disorders. *Hum. Mutat.* **2013**, *35*, 341–349. [CrossRef]

66. Zhao, X.-N.; Usdin, K. FAN1 protects against repeat expansions in a Fragile X mouse model. *DNA Repair* **2018**, *69*, 1–5. [CrossRef]

67. Gazy, I.; Hayward, B.; Potapova, S.; Zhao, X.; Usdin, K. Double-strand break repair plays a role in repeat instability in a fragile X mouse model. *DNA Repair* **2018**, *74*, 63–69. [CrossRef]

68. Rousseau, F.; Heitz, D.; Tarleton, J.; MacPherson, J.; Malmgren, H.; Dahl, N.; Barnicoat, A.; Mathew, C.; Mornet, E.; Tejada, I. A multicenter study on genotype-phenotype correlations in the fragile X syndrome, using direct diagnosis with probe StB12.3: The first 2,253 cases. *Am. J. Hum. Genet.* **1994**, *55*, 225–237.

69. Reyniers, E.; Martin, J.-J.; Cras, P.; Van Marck, E.; Handig, I.; Jorens, H.Z.; Oostra, B.A.; Kooy, R.F.; Willems, P.J. Postmortem examination of two fragile X brothers with anFMR1 full mutation. *Am. J. Med Genet.* **1999**, *84*, 245–249. [CrossRef]

70. Moutou, C.; Vincent, M.-C.; Biancalana, V.; Mandel, J.-L. Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic. *Hum. Mol. Genet.* **1997**, *6*, 971–979. [CrossRef]

71. Sermon, K.; Seneca, S.; De Rycke, M.; Goossens, V.; Van de Velde, H.; De Vos, A.; Platteau, P.; Lissens, W.; Van Steirteghem, A.; Liebaers, I. PGD in the lab for triplet repeat diseases—Myotonic dystrophy, Huntington's disease and Fragile-X syn-drome. *Mol. Cell. Endocrinol.* **2001**, *183*, S77–S85. [CrossRef]

72. Rifé, M.; Badenas, C.; Quintó, L.; Puigoriol, E.; Rodriguez-Revenga, L.; Tazón, B.; Jiménez, L.; Sánchez, A.; Milà, M. Analysis of CGG variation through 642 meioses in Fragile X families. *Mol. Hum. Reprod.* **2004**, *10*, 773–776. [CrossRef] [PubMed]

73. Eiges, R.; Urbach, A.; Malcov, M.; Frumkin, T.; Schwartz, T.; Amit, A.; Yaron, Y.; Eden, A.; Yanuka, O.; Benvenisty, N.; et al. Developmental study of fragile X syndrome using human embryonic stem cells derived from preimplanta-tion genetically diagnosed embryos. *Cell Stem Cell* **2007**, *1*, 568–577. [CrossRef] [PubMed]

74. Avitzour, M.; Mor-Shaked, H.; Yanovsky-Dagan, S.; Aharoni, S.; Altarescu, G.; Renbaum, P.; Eldar-Geva, T.; Schonberger, O.; Levy-Lahad, E.; Epsztejn-Litman, S.; et al. FMR1 Epigenetic Silencing Commonly Occurs in Undifferentiated Fragile X-Affected Embryonic Stem Cells. *Stem Cell Rep.* **2014**, *3*, 699–706. [CrossRef]

75. Gerhardt, J.; Tomishima, M.J.; Zaninovic, N.; Colak, D.; Yan, Z.; Zhan, Q.; Rosenwaks, Z.; Jaffrey, S.; Schildkraut, C.L. The DNA Replication Program Is Altered at the FMR1 Locus in Fragile X Embryonic Stem Cells. *Mol. Cell* **2013**, *53*, 19–31. [CrossRef]

76. Malter, H.E.; Iber, J.C.; Willemsen, R.; De Graaff, E.; Tarleton, J.C.; Leisti, J.; Warren, S.T.; Oostra, B.A. Characterization of the full fragile X syndrome mutation in fetal gametes. *Nat. Genet.* **1997**, *15*, 165–169. [CrossRef]

77. Edamura, K.N.; Leonard, M.R.; Pearson, C.E. Role of Replication and CpG Methylation in Fragile X Syndrome CGG Deletions in Primate Cells. *Am. J. Hum. Genet.* **2005**, *76*, 302–311. [CrossRef]

78. Edamura, N.K.; Pearson, C.E. DNA methylation and replication: Implications for the "deletion hotspot" region of FMR1. *Hum. Genet.* **2005**, *118*, 301–304. [CrossRef]

79. Zhao, X.N.; Lokanga, R.; Allette, K.; Gazy, I.; Wu, D.; Usdin, K. A MutSβ-Dependent Contribution of MutSα to Repeat Expansions in Fragile X Premutation Mice? *PLoS Genet.* **2016**, *12*, e1006190. [CrossRef]

80. Kovtun, I.V.; Liu, Y.; Bjoras, M.; Klungland, A.; Wilson, S.; McMurray, C.T. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature* **2007**, *447*, 447–452. [CrossRef]

81. Møllersen, L.; Rowe, A.D.; Illuzzi, J.L.; Hildrestrand, G.A.; Gerhold, K.J.; Tveterås, L.; Bjølgerud, A.; Wilson, D.M., III; Bjørås, M.; Klungland, A. Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. *Hum. Mol. Genet.* **2012**, *21*, 4939–4947. [CrossRef] [PubMed]

82. Hayward, B.; Loutaev, I.; Ding, X.; Nolin, S.L.; Thurm, A.; Usdin, K.; Smith, C.B. Fragile X syndrome in a male with methylated premutation alleles and no detectable methylated full mutation alleles. *Am. J. Med Genet. Part A* **2019**, *179*, 2132–2137. [CrossRef] [PubMed]

83. De Graaff, E.; Rouillard, P.; Willems, P.J.; Smits, A.P.T.; Rousseau, F.; Oostra, B.A. Hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients. *Hum. Mol. Genet.* **1995**, *4*, 45–49. [CrossRef] [PubMed]

84. Luo, S.; Huang, W.; Chen, C.; Pan, Q.; Duan, R.; Wu, L. A novel deletion to normal size in the sperm of a fragile X full mutation male. *Clin. Genet.* **2013**, *86*, 295–297. [CrossRef] [PubMed]

85. Wyatt, D.; Feng, W.; Conlin, M.P.; Yousefzadeh, M.J.; Roberts, S.A.; Mieczkowski, P.; Wood, R.; Gupta, G.P.; Ramsden, D.A. Essential Roles for Polymerase θ-Mediated End Joining in the Repair of Chromosome Breaks. *Mol. Cell* **2016**, *63*, 662–673. [CrossRef]

86. Zhao, X.; Usdin, K. (Dys)function Follows Form: Nucleic Acid Structure, Repeat Expansion, and Disease Pathology in *FMR1* Disorders. *Int. J. Mol. Sci.* **2021**, *22*, 9167. [CrossRef]

87. Cleary, J.D.; Pearson, C.E. The contribution of cis-elements to disease-associated repeat instability: Clinical and experi-mental evidence. *Cytogenet. Genome Res.* **2003**, *100*, 25–55. [CrossRef]

88. Usdin, K.; Woodford, K.J. CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res.* **1995**, *23*, 4202–4209. [CrossRef]

89. Mirkin, E.V.; Mirkin, S. Replication Fork Stalling at Natural Impediments. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 13–35. [CrossRef]

90. Ennis, S.; Murray, A.; Brightwell, G.; Morton, N.; Jacobs, P. Closely linked cis -acting modifier of expansion of the CGG repeat in high risk FMR1 haplotypes. *Hum. Mutat.* **2007**, *28*, 1216–1224. [CrossRef]

91. Gerhardt, J.; Zaninovic, N.; Zhan, Q.; Madireddy, A.; Nolin, S.; Ersalesi, N.; Yan, Z.; Rosenwaks, Z.; Schildkraut, C.L. Cis-acting DNA sequence at a replication origin promotes repeat expansion to fragile X full mutation. *J. Cell Biol.* **2014**, *206*, 599–607. [CrossRef] [PubMed]

92. Gacy, A.M.; McMurray, C.T. Influence of Hairpins on Template Reannealing at Trinucleotide Repeat Duplexes: A Model for Slipped DNA. *Biochemistry* **1998**, *37*, 9426–9434. [CrossRef] [PubMed]

93. McMurray, C.T. Mechanisms of DNA expansion. *Chromosoma* **1995**, *104*, 2–13. [CrossRef] [PubMed]

94. Gan, H.; Yu, C.; Devbhandari, S.; Sharma, S.; Han, J.; Chabes, A.; Remus, D.; Zhang, Z. Checkpoint Kinase Rad53 Couples Leading- and Lagging-Strand DNA Synthesis under Replication Stress. *Mol. Cell* **2017**, *68*, 446–455.e3. [CrossRef]

95. Ladd, D.; Smith, L.E.; Rabaia, N.A.; Moore, J.M.; Georges, S.A.; Hansen, R.S.; Hagerman, R.J.; Tassone, F.; Tapscott, S.J.; Filippova, G.N. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum. Mol. Genet.* **2007**, *16*, 3174–3187. [CrossRef]

96. Lin, Y.; Wilson, J.H. Nucleotide Excision Repair, Mismatch Repair, and R-Loops Modulate Convergent Transcription-Induced Cell Death and Repeat Instability. *PLoS ONE* **2012**, *7*, e46807. [CrossRef]

97. Dion, V.; Wilson, J.H. Instability and chromatin structure of expanded trinucleotide repeats. *Trends Genet.* **2009**, *25*, 288–297. [CrossRef]

98. McMurray, C.T. Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* **2010**, *11*, 786–799. [CrossRef]

99. Lokanga, R.A.; Senejani, A.G.; Sweasy, J.B.; Usdin, K. Heterozygosity for a Hypomorphic Polβ Mutation Reduces the Expansion Frequency in a Mouse Model of the Fragile X-Related Disorders. *PLoS Genet.* **2015**, *11*, e1005181. [CrossRef]

100. Xu, H.; A Rosales-Reynoso, M.; Barros-Núñez, P.; Peprah, E. DNA repair/replication transcripts are down regulated in patients with Fragile X Syndrome. *BMC Res. Notes* **2013**, *6*, 90–95. [CrossRef]

101. Kononenko, A.V.; Ebersole, T.; Vasquez, K.M.; Mirkin, S.M. Mechanisms of genetic instability caused by (CGG)n repeats in an experimental mammalian system. *Nat. Struct. Mol. Biol.* **2018**, *25*, 669–676. [CrossRef] [PubMed]

102. Shah, K.A.; Mirkin, S.M. The hidden side of unstable DNA repeats: Mutagenesis at a distance. *DNA Repair* **2015**, *32*, 106–112. [CrossRef] [PubMed]

103. Disney, M.D.; Liu, B.; Yang, W.-Y.; Sellier, C.; Tran, T.; Charlet-Berguerand, N.; Childs-Disney, J.L. A Small Molecule That Targets r(CGG)^{exp} and Improves Defects in Fragile X-Associated Tremor Ataxia Syndrome. *ACS Chem. Biol.* **2012**, *7*, 1711–1718. [CrossRef] [PubMed]

104. Verma, A.K.; Khan, E.; Mishra, S.K.; Mishra, A.; Charlet-Berguerand, N.; Kumar, A. Curcumin Regulates the r(CGG)exp RNA Hairpin Structure and Ameliorate Defects in Fragile X-Associated Tremor Ataxia Syndrome. *Front. Neurosci.* **2020**, *14*, 295. [CrossRef]

105. Derbis, M.; Kul, E.; Niewiadomska, D.; Sekrecki, M.; Piasecka, A.; Taylor, K.; Hukema, R.K.; Stork, O.; Sobczak, K. Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG re-peats. *Nat. Commun.* **2021**, *12*, 1265. [CrossRef]

106. Haify, S.N.; Buijsen, R.A.; Verwegen, L.; Severijnen, L.A.W.; de Boer, H.; Boumeester, V.; Monshouwer, R.; Yang, W.Y.; Cameron, M.D.; Willemsen, R.; et al. Small molecule 1a reduces FMRpolyG-mediated toxicity in in vitro and in vivo models for FMR1 premuta-tion. *Hum. Mol. Genet.* **2021**, *30*, 1632–1648. [CrossRef]