Molecular Mechanisms of FAN1 in CAG Repeat Instability: Implications for HD and SCAs

Berdien Maring
01/12/2024
Department of Genetics
Supervisors: H.Westers, C. van Diemen

Abstract

CAG repeat instability is a hallmark of Huntington's disease (HD) and spinocerebellar ataxias (SCAs), contributing to disease onset and progression through somatic expansions in affected tissues. Genome-wide association studies (GWAS) have identified FAN1 as a key genetic modifier of the age of onset in both disorders. This review synthesizes current research on FAN1's role in maintaining genomic stability, particularly through its dual nuclease activities and interaction with DNA repair pathways, such as the mismatch repair (MMR) system.

This review highlights findings that FAN1 counteracts MMR-induced instability, preventing repeat expansions. We explore the molecular mechanisms by which FAN1 stabilizes CAG repeats, including its structural domains, nuclease functions, and competition with MMR proteins. Specific FAN1 variants, particularly those in 3' UTR regions and TPR-domain, are associated with delayed disease onset and slower progression. These insights suggest therapeutic potential for targeting FAN1 in HD and SCAs, with approaches such as antisense oligonucleotides under development to enhance its protective role.

Although FAN1's role in HD is well investigated, more research is needed to confirm its role in SCAs. While the mechanisms may be similar, the specific interactions and effects in SCAs require further elucidation.

Contents

1. Introduction

Trinucleotide repeat disorders are a group of genetic conditions caused by the pathological expansion of specific trinucleotide sequences within certain genes. Among these, Huntington's disease (HD) and spinocerebellar ataxias (SCAs) are notable examples of autosomal dominant neurodegenerative disorders specifically caused by the expansion of CAG trinucleotide repeats. In HD, the *HTT* gene is affected, while in SCA, the affected gene depends on the type of SCA (**Table 1**) [1]. These expansions lead to the production of mutant proteins with elongated polyglutamine tracts, which are toxic to neurons [1-4]. The length of the CAG repeat is a critical factor in determining the severity and onset of these diseases, with longer repeats generally correlating with earlier onset and more rapid progression [5].

Table 1: Characteristics of CAG repeat diseases. *Table adapted from Bettencourt et al. [6]*

Disorder	Affected Gene	Normal CAG repeats	Pathological CAG repeats
HD	*HTT*	6-35	40-121
SCA1	*ATXN1*	6-38	45-83
SCA2	*ATXN2*	15-31	33-500
SCA3	*ATXN3*	12-44	52-87
SCA6	*CACNA1A*	4-18	20-33
SCA7	*ATXN7*	3-19	37-460
SCA17	*TBP*	25-40	49-66

The biological consequences of these CAG expansions are profound. The repeated CAG sequences are translated into elongated polyglutamine (polyQ) tracts within the mutant proteins [3]. These polyQ proteins disrupt cellular functions, leading to neuronal dysfunction and death [3]. In HD, this primarily affects the basal ganglia and cortex, resulting in motor dysfunction, cognitive decline, and psychiatric symptoms [2]. In SCAs, the cerebellum and spinal cord are predominantly affected, causing progressive ataxia and coordination problems [2].

A key feature of these disorders is CAG repeat instability, characterized by the progressive expansion of repeats in somatic tissues. This instability exacerbates neurodegeneration by increasing the toxic effects of mutant proteins and accelerating disease onset and severity [2-4]. Slipped-DNA structures formed during replication, repair, or transcription are thought to contribute to this instability by promoting repeat expansions. The presence of somatic expansions in tissues such as the brain and striatum underscores their critical role in disease progression. However, while CAG repeat length is a major determinant of age of onset (AOO), it cannot fully explain the variability observed among patients, suggesting that additional disease-modifying factors are involved.

Genome-wide association studies (GWAS) have identified several genetic modifiers that influence the stability of CAG repeats and the progression of HD and SCAs [6, 7]. FAN1 (Fanconi anemia-associated nuclease 1) has emerged as a significant modifier among these [6]. FAN1 is a structure-specific nuclease involved in DNA repair, playing a crucial role in maintaining genomic stability [8]. Its interaction with components of the mismatch repair (MMR) pathway, which is essential for correcting DNA replication errors, further underscores its significance [8]. Understanding FAN1's role is crucial for advancing our knowledge of these disorders and developing effective therapeutic strategies.

Recent studies have highlighted the role of FAN1 in modulating the stability of CAG repeats [8-10]. FAN1 knockouts in HD mouse models have demonstrated increased somatic expansions [9], underscoring its critical role in limiting repeat instability. Understanding the precise

mechanisms by which FAN1 influences repeat stability, particularly in neuronal tissues where replication is minimal and transcription-associated instability may dominate, could provide valuable insights into the pathogenesis of these disorders and inform potential therapeutic strategies.

This essay explores the mechanisms by which FAN1 contributes to the stability of CAG repeats in Huntington's disease and spinocerebellar ataxias. By examining FAN1's molecular functions, interactions with DNA repair pathways, and the impact of genetic variants, we aim to elucidate its role in disease progression and investigate potential therapeutic interventions.

2. Molecular mechanism of FAN1

This chapter begins with a summary of the comprehensive review by Deshmukh et al. to establish a foundational understanding of the role of FAN1 in CAG trinucleotide diseases [8]. This review synthesizes existing research up to 2021, emphasizing FAN1's role as a structure-specific nuclease with critical functions in maintaining genomic stability. It describes FAN1's dual nuclease activities and its ability to bind complex DNA structures, such as flapped DNA and interstrand crosslinks (ICLs).

2.1. FAN1 is structured in domains

The domain organization of FAN1 (Figure 1) underscores its multifunctional role in DNA repair, with each domain contributing uniquely to its activity. As noted by Deshmukh et al., the most important domain is the catalytical domain, central to FAN1's dual nuclease activity [8]. This is a virus-type replication-repair nuclease (VRR_NUC) domain, commonly found as a standalone domain in bacteriophages and bacteria. It is uniquely present in FAN1, the only known eukaryotic nuclease containing this domain, which highlights FAN1's unique role in maintaining genomic stability. Chapter 2.1. provides a detailed description of FAN1's dual nuclease activity.

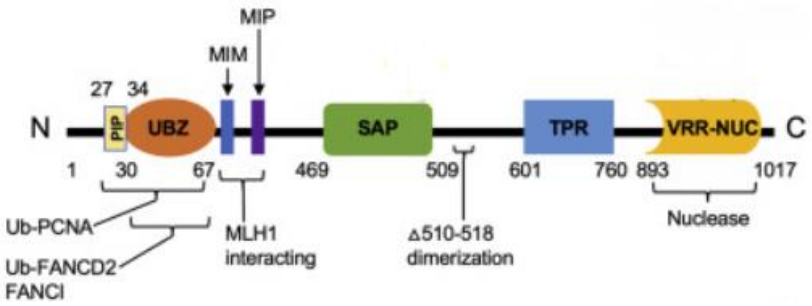

Figure 1: Schematic representation of FAN1 protein domains. *Schematic representation of the FAN1 protein domains. FAN1 comprises multiple functional domains, including the UBZ, SAP, TPR, and VRR-NUC domains. Figure adapted from Deshmukh et al. [11].*

The SAP (SAF-A/B, Acinus, and PIAS) domain is located at the center of FAN1 and is responsible for binding specific DNA structures. The TPR (tetratricopeptide repeat) domain is located between the SAP and VVR_NUC domain, is involved in interdomain and protein-protein interactions, and facilitates the dimerization of FAN1. While FAN1 is a monomer in solution in the presence of DNA, the SAP domain of one FAN1 molecule will interact with the TPR and NUC domain of the second FAN1 molecule to form a dimer.

The N-terminus contains several important domains for connecting with other proteins together with the TPR domain. The largest domain here is the ubiquitin-binding zinc (UBZ) finger domain, which facilitates FAN1's localization to sites of DNA interstrand crosslinks (ICLs) and stalled replication forks. The UBZ domain works with the proliferating cell nuclear antigen (PCNA)-interacting peptide (PIP) box motif to bind to ubiquitylated PCNA. PCNA is often loaded to DNA as part of repair pathways. In addition, the UBZ domain can bind FANCD2-FANCI, a protein that recruits nucleases for cross-link repair. FAN1 was initially identified as an MLH1 and PMS2-interacting protein. MLH1 interacts with FAN1 via two interacting sites, MLH1-interacting protein (MIP) and MLH1-interacting motif (MIF) [12]. The interaction between FAN1 and MLH1 is well-established and suggests that FAN1 may play a role in modulating or cooperating with MMR activities. However, Deshmukh et al. could only hypothesize that this interaction may play a structural role in determining DNA cleavage specificity [8].

2.2. FAN1 has dual nuclease activity to resolve various unusual DNA structures

As reviewed by Deshmukh et al., FAN1's dual nuclease activities, comprising endonuclease and exonuclease functions, are central to its role in DNA repair [8]. Its dual endo- and exonuclease activities allow it to process a variety of DNA substrates (Figure 2). Interestingly, FAN1 requires dimerization to perform its endonuclease activity, whereas its exonuclease activity operates independently of dimerization. Notably, this dimerization is unique to higher eukaryotes, suggesting an evolutionary adaptation of FAN1 to meet the demands of more complex genomes.

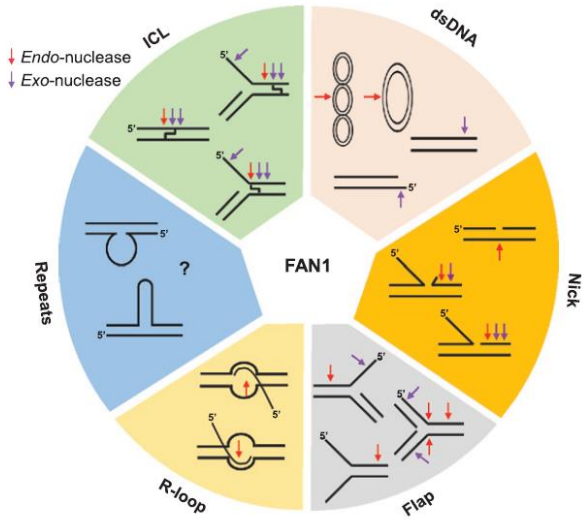

Figure 2: FAN1 can process various unusual DNA structures, *including interstrand cross-linked DNA (ICL), dsDNA (supercoiled, covalently closed circular, linear DNA), nicked DNA, flap-DNA, R-loop DNA, and repeat-containing DNA (unknown). Purple and red arrows denote exo-nuclease and endo-nucleolytic cleavage. Picture and caption are copied from Deshmukh et al 2021[8].*

FAN1 plays a critical role in DNA repair by addressing interstrand crosslinks (ICLs), which are severe forms of DNA damage that link two double-helix strands, obstructing replication and transcription. If ICLs are not repaired efficiently, they can be highly toxic to cells. FAN1 is recruited to replication forks stalled by ICLs, acting as a structure-specific nuclease. Its endonuclease activity unhooks the crosslink, while its exonuclease activity manages the resulting DNA ends, preparing them for further repair by other proteins. This dual functionality is essential for maintaining genomic integrity, particularly in rapidly dividing cells where replication stress can increase ICL formation.

In addition, FAN1 is crucial for processing nicked DNA, which can arise from base excision repair intermediates, replication stress, or incomplete lesion repair. FAN1 recognizes and binds to the damaged site via its DNA-binding domains. Subsequently, its endonuclease activity cleaves near the nick, while its exonuclease activity removes damaged or mispaired bases and trims the DNA ends, providing clean substrates for downstream repair. This function is particularly critical during replication stress, where unprocessed nicks could lead to double-strand breaks or replication fork collapse. FAN1 stabilizes the replication fork by repairing these nicks and prevents further genomic damage.

Moreover, FAN1 also resolves flap DNA structures, intermediates that arise during replication and DNA repair processes. Using its structure-specific nuclease activity, FAN1 cleaves the displaced single-stranded DNA at the junction, ensuring proper processing and preventing the accumulation of harmful DNA intermediates. This contributes to maintaining genomic stability and allowing proper ligation of the DNA fragments.

In addition, Deshmukh et al. also suggest two additional, yet unproven, nuclease activities that could be important in stabilizing CAG repeats [8]. First, they propose that FAN1 is involved in processing R-loops. R-loops, three-stranded DNA-RNA hybrids that form during transcription, are particularly prevalent in regions of repetitive DNA, such as CAG repeats. If not resolved, these structures can lead to transcriptional blockage and DNA breaks, potentially contributing to repeat expansion. Deshmukh et al. speculate that FAN1 may be involved in resolving these R-loops by cleaving either the displaced RNA or the DNA component, thus preventing the accumulation of these destabilizing structures.

In addition, hairpins and DNA extrusions that arise during DNA replication or repair are problematic in repetitive regions like CAG repeats. These secondary structures can stall replication forks and trigger DNA damage, which could ultimately lead to repeat instability. FAN1's nuclease activities, as proposed by Deshmukh et al., may play a crucial role in resolving these hairpins and extrusions, thereby ensuring the integrity of the DNA during replication and repair. Chapter 0 will further explore evidence that proves that FAN1 is involved in these mechanisms.

3. Mechanisms in repeat stability

Building on understanding FAN1's involvement in resolving secondary DNA structures, this chapter delves deeper into the specific mechanisms by which FAN1 stabilizes CAG repeats. Deshmukh et al. proposed potential nuclease activities by which FAN1 may be involved in stabilizing repeats [8]. They suggest that FAN1 contributes to repeat stability through multiple mechanisms. During replication, it is thought to help process slip-outs formed on either the template or nascent strand, which could prevent expansion or induce contraction. In non-mitotic DNA repair processes, FAN1 likely stabilizes repeat tracts by processing slip-outs in repeat-containing DNA. Furthermore, in non-dividing cells, FAN1 may resolve slip-outs that arise during transcription across expanded repeats, thus preventing repeat instability. However, given that neurons, the primary cells affected in HD and SCAs, are non-dividing, it is unlikely that FAN1's role in replication is significant in these contexts. This leads to the hypothesis that FAN1's stabilizing role is more likely associated with transcription and repair processes.

Building on this understanding, I propose a new theory: FAN1 stabilizes CAG repeats through a transcription-repair-associated mechanism. This theory posits that the repair processes triggered by transcriptional problems are central to FAN1's role in maintaining repeat stability. Specifically, FAN1 may resolve slip-outs and R-loops that form during transcription, thereby preventing repeat instability and maintaining genomic integrity.

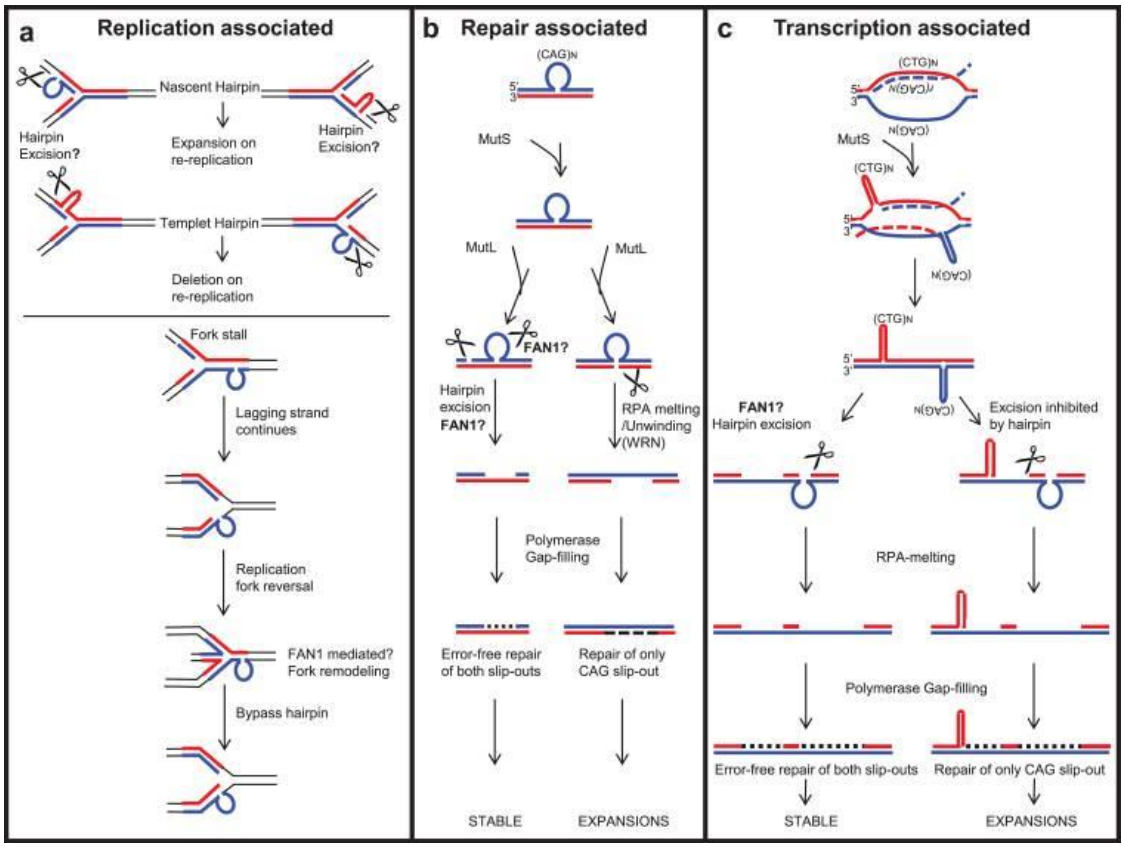

Figure 3: Proposed mechanism of FAN1 in CAG repeat stability by Deshmukh et al. ***(a)*** *Replication-associated repeat instability; FAN1 may process slip-outs formed on a template or nascent strand during replication and may stabilize repeat length or induce contraction.* ***(b)*** *A slip-out formed by repeat-containing DNA can be processed by a non-mitotic DNA repair-associated mechanism. FAN1 may act as one of the critical nucleases to process repeat slip-outs to stabilize the repeat tract against length variations.* ***(c)*** *In non-replicating cells, repeat instability is associated with active transcription across the expanded repeat. FAN1 may process the slip-outs formed following unidirectional or bidirectional transcription across the expanded repeat, to stabilize repeat tract against length variations. Figure and caption retrieved from Deshmukh et al. [8]*

3.1. FAN1 stabilizes trinucleotide repeats

The essential role of FAN1's nuclease activity in stabilizing CAG repeats in neuronal-like cells has been demonstrated in a cellular HD model [13]. These findings support the transcription-repair-associated mechanism by showing that FAN1's enzymatic function is crucial for maintaining repeat stability even when no transcription exists. The observation that a nuclease-dead FAN1 variant (p.D960) leads to similar expansion rates as FAN1-deficient cells underscores the necessity of FAN1's nuclease activity in this process.

Further insights into FAN1's mechanism were provided using biochemically characterized DNA substrates and purified FAN1 protein [11]. It was shown that FAN1 binds specifically to slip-out DNA structures containing CAG/CTG repeats, which form during transcription but not to fully duplex DNA. This indicates that FAN1 requires these slip-out structures to function. Additionally, FAN1 exhibits different exonuclease activity on repeat versus non-repeat DNA, cutting every third nucleotide in non-repeat DNA and almost every nucleotide in repeat DNA, demonstrating a specific role in processing trinucleotide repeats. Importantly, FAN1's exonuclease activity occurs in cycles of cleavage, release, and reassociation, with pauses that may allow other repair enzymes to engage, supporting the idea that FAN1 collaborates with other repair pathways to stabilize repetitive sequences.

There is also evidence suggesting that FAN1 functions during replication, requiring PCNA and RFC [14]. This does not exclude FAN1's role in transcription but highlights its versatility in maintaining genomic stability. FAN1's involvement in replication and transcription processes indicates its comprehensive function in genomic maintenance. The transcription-repair mechanism may be more prominent in non-dividing cells, such as neurons, whereas FAN1's replication role is critical in dividing cells. This dual functionality suggests that FAN1 resolves DNA structures arising from replication and transcription, preventing repeat expansions.

3.2. FAN1 competes with the MMR pathway

The mismatch repair (MMR) pathway is an essential DNA repair mechanism that maintains genomic stability by correcting mismatched bases and slipped DNA structures. This pathway involves key proteins such as MutSβ (MSH2/MSH3) and MutLα (MLH1/PMS2), which detect and repair these structural abnormalities. Interestingly, while MMR is crucial for genomic stability, it paradoxically contributes to the instability of repeat DNA sequences. For instance, in Huntington's disease (HD) knock-in mice, the absence of MLH1 prevents somatic expansion of CAG repeats, indicating MLH1's role in destabilizing these repeats [9]. Given the protein-protein interactions between FAN1 and MLH1, this raises an important question: What is the precise relationship between FAN1 and the MMR pathway, and how does this interaction influence repeat stability?

Evidence shows that FAN1 competes with the MMR pathway for binding to MLH1 [14, 15]. This competition is crucial for understanding FAN1's role in stabilizing CAG repeats. By competing with MMR proteins, particularly MSH3, FAN1 prevents the MMR pathway from exacerbating repeat expansions, underscoring FAN1's protective role in maintaining genomic stability.

The competition between FAN1 and the MMR pathway highlights the complex interplay between FAN1 and DNA repair processes. By preventing the formation of the MutSβ-MutLα complex and limiting excision tract lengths, FAN1 plays a pivotal role in stabilizing CAG repeats (**Figure 4**). This is vital for reducing somatic expansions and delaying disease onset in HD and SCAs. The proposed transcription-repair-associated mechanism, supported by FAN1's competitive interaction with the MMR pathway, provides a comprehensive framework for understanding FAN1's function in non-dividing cells.

Porro et al. found that FAN1 is not essential for the MMR pathway, suggesting that FAN1 operates through a distinct mechanism. Goold et al. and Phadte et al. reveal that FAN1 limits the extension of CAG repeats mediated by MutSβ (MSH2/MSH3) during the repair of slipped-DNA structures (). This protective role is facilitated by FAN1's binding with MLH1, which effectively prevents MSH3 from binding and, consequently, prevents the activation of the MMR pathway. When FAN1 levels are higher than those of MutSβ, MLH1 predominantly binds to FAN1, inhibiting MutSβ activation.

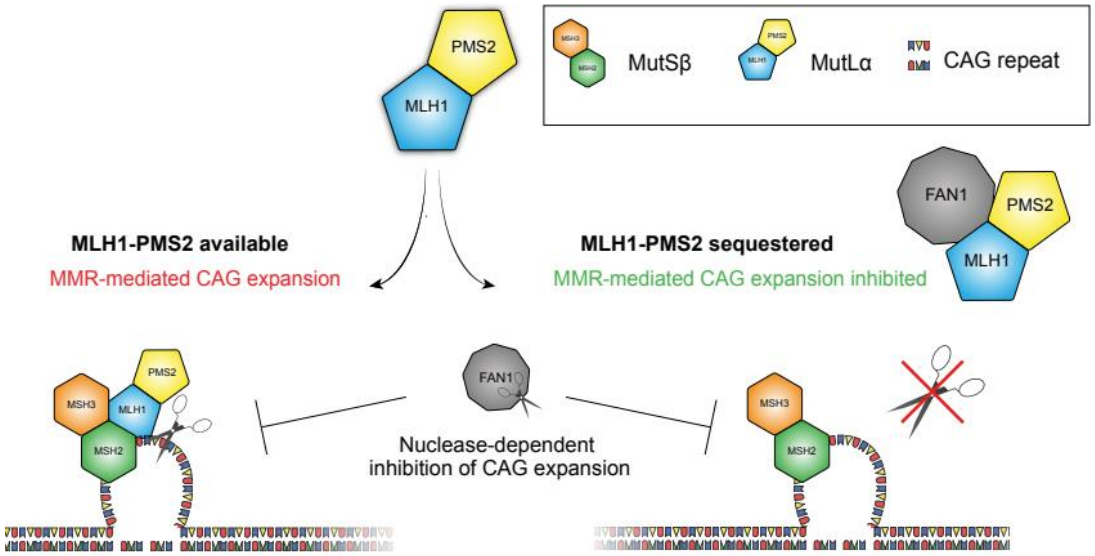

Figure 4: Proposed mechanism of FAN1 in limiting somatic expansion in CAG repeats. *Picture adapted from Goold et al.[15]*

Combined, these findings show that FAN1 is a competitor of MMR and does not take part in the MMR pathway in regions with repetitive DNA. By preventing the formation of the MutSβ-MutLα complex and limiting excision tract lengths, FAN1 plays a pivotal role in stabilizing CAG repeats, vital for reducing somatic expansions and delaying disease onset in HD and SCAs.

4. FAN1 variants and their effect on HD and SCA

As stated above, GWAs have identified FAN variants that affect the stability of CAG repeats and thereby act as a significant modifier of disease onset and progression in SCAs. Several studies have been performed to identify genetic modifiers of polyQ diseases. The first study by the GeM-HD consortium analyzed 4,082 HD patients of European ancestry to identify modifiers that could influence the age of onset [7]. They found over 200 significant SNPs in chromosome 15 and suggest that *FAN1* is the strongest candidate on this chromosome to influence AOO. In addition, a smaller study by Bettencourt et al used 445 HD patients, 177 SCA1, 294 SCA2, 397 SCA3, 69 SCA6, 73 SCA7 and 7 SCA17 patients [6]. They found rs3512 in FAN1 to be the most significant modifier of AOO in all HD and SCA patients combined. In addition, rs3512 was the only significant modifier of AOO in all SCA patients combined. To identify rare variants associated with clinical effects, McAllister et al. used exosome sequencing on 683 HD patients with extremes of onset or phenotype relative to CAG length [13]. The patients were selected from European and United States-based cohorts.

The variants related to AOO in HD and SCA often affect FAN1's functional domains, such as its nuclease domain or DNA binding regions, which are essential for stabilizing repetitive DNA and preventing somatic repeat expansion. By exploring how specific variants alter FAN1's structure and activity, we can gain deeper insights into the mechanisms driving CAG repeat instability and identify potential therapeutic targets to delay disease onset and progression in HD and SCAs. An overview of the known FAN1 variants and how they influence HD and SCAs can be found in Table 2.

4.1. Reduced nuclease activity causes earlier disease onset

Having established the significance of FAN1 variants in HD and SCAs, we now delve into the specific domains affected by these variants and their impact on disease progression. The nuclease domain is crucial for FAN1's endonuclease and exonuclease activities. This domain stabilizes CAG repeats, which is essential in preventing somatic repeat expansion [11, 13, 14]. Reduced nuclease activity impairs FAN1's ability to stabilize CAG repeats, leading to their expansion and exacerbating disease symptoms [13]. The pivotal role of FAN1's nuclease domain in stabilizing CAG repeats is highlighted by McAllister et al., who linked damaging variants in this domain to earlier disease onset and progression in HD [13]. Exome sequencing was used in a cohort of 683 individuals with HD with extreme phenotypes compared to the CAG repeat length. They identified common and rare damaging variants in the nuclease domain of FAN1. In addition, they performed in vitro analyses of mutated FAN1 proteins to investigate the functional consequences. The results demonstrated that these variants reduced nuclease activity to less than 50% of the wild-type protein's capacity. Functional analysis using CRISPR-edited patient-derived induced pluripotent stem cells (iPSC) confirmed that FAN1 knockout and nuclease-inactive variants, like D960A, significantly increased somatic CAG repeat expansions. These findings directly associate impaired nuclease activity with reduced repeat stability, highlighting the critical role of FAN1 in stabilizing CAG repeats in HD patients.

Table 2: Overview of FAN1 variants categorized by functional domain, associated diseases, and their impact on FAN1 activity and disease progression. Variants are linked to Huntington's disease (HD) and spinocerebellar ataxias (SCA), with functional and phenotypic consequences described.

Domain	SNP	Variant	Rare variant	Disease	Functional impact	Impact on disease	Source
3'-UTR	rs3512	-	No	HD, SCA	Reduced sensitivity to miR-124-3p, leading to elevated FAN1 levels	Late onset	[6, 7, 16]
DNA binding domain	rs151322829	R377W	No	HD	Reduced nuclease activity	Early/more severe disease	[7, 11, 13]
	rs150393409	R507H	No	HD	Deleterious to FAN1 function, reduced DNA binding, and reduced nuclease activity	Early/more severe disease	[6, 7, 11, 13]
Nuclease domain	-	R982C	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
	-	C1004G	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
	-	D960A	No	HD	Nuclease death variant	Early/more severe disease	[7, 11, 13, 14]
	-	A949V	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
	-	V963VL	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
	-	R969L	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
	-	R982C	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
	-	C1004G	Yes	HD	Reduced nuclease activity	Early/more severe disease	[13]
TPR domain	-	R685W	Yes	HD	Likely alters interaction with MLH1, preventing MMR from increasing CAG repeats	Late/less severe disease	[13]
	-	D702E	Yes	HD	Likely alters interaction with MLH1, preventing MMR from increasing CAG repeats	Late/less severe disease	[13]
	-	L713I	Yes	HD	Likely alters interaction with MLH1, preventing MMR from increasing CAG repeats	Late/less severe disease	[13]
	-	K794R	Yes	HD	Likely alters interaction with MLH1, preventing MMR from increasing CAG repeats	Late/less severe disease	[13]

4.2. Reduced DNA binding leads to earlier disease onset

The DNA binding domain binds slip-out DNA so the nuclease domain can cut out the CAG repeats [11]. Alternation of the DNA binding domain is also associated with earlier onset and a more severe phenotype of HD [13]. Most frequent variants, like R977W and R507H, alter the ability of FAN1 to bind to slipped DNA [11]. When FAN1 can't bind to DNA, it can't perform its exonuclease activity to stabilize the CAG repeats. Therefore, it causes earlier disease onset and a severe phenotype.

4.3. Alterations in the TPR domain delay disease onset

The TPR domain is involved in interdomain and protein-protein interactions and facilitates the dimerization of FAN1 [8]. McAllister et al. showed that damaging variants in this domain are associated with later disease onset and less severe disease in HD [8]. It is argued that this is due to changes in the protein-protein interaction with MLH1. Variants of FAN1 are likely to bind MLH1 more strongly, preventing it from activating MutSβ. However, further studies are needed to prove whether FAN1 variants in the TPR domain exhibit stronger binding to MLH1. For example, binding affinity assays can be used to determine the binding kinetics of the variants and compare these with the variants to show that the variants bind MLH1 stronger compared to the wildtype.

4.4. Reduced post-transcriptional regulation delays disease onset

The 3'UTR is a non-coding region located downstream of a gene and is crucial for post-transcriptional regulation of gene expression. This region does not code for proteins but contains elements that regulate mRNA's stability, localization, and translation. Variants in the 3' UTR can affect gene function, particularly by modulating interactions with microRNAs (miRNAs) or RNA-binding proteins. The strongest modifier of HD and SCA is rs3512, located in the 3'UTR region [6, 7].

Recent studies revealed that FAN1 mRNA levels are post-transcriptionally regulated by miR-124-3p, a microRNA that normally destabilizes FAN1 mRNA, reducing FAN1 protein levels [16]. Lower FAN1 levels tip the balance toward the mismatch repair (MMR) pathway, which promotes CAG repeat expansion in the HTT gene, exacerbating repeat instability. However, in patients carrying the alternative rs3512 allele, miR-124-3p binds less efficiently to the FAN1 3'UTR. This reduced binding results in greater mRNA stability, increasing FAN1 protein levels. The elevated FAN1 levels counteract MMR activity, stabilizing CAG repeats and delaying disease onset in HD.

While this mechanism has been confirmed for HD, the same process may apply to SCA, as rs3512 is also the strongest modifier in these disorders [6]. However, direct evidence linking the rs3512 variant to repeat stability in SCA is currently lacking. Future studies should investigate whether this variant affects CAG repeat stability in genes associated with SCA, as in HTT for HD. This research could elucidate whether similar therapeutic strategies targeting the 3'UTR region or miR-124-3p interactions could benefit patients with SCA.

4.5. Combined effects of FAN1 variants

While individual FAN1 variants significantly influence disease phenotype, their combined effects may vary. For instance, R377W or R507H variants in the nuclease domain are typically associated with earlier disease onset and more severe phenotypes. However, six patients with one of these variants had later disease onset instead of earlier onset [13]. These patients had either an extra FAN1 variant or a genetic variation in one of the other DNA repair genes. This shows

that the consequences of having a FAN1 variant are not black and white but depend on additional genetic modifiers.

To summarize, the effect of FAN1 variants depends on the functional domain in which they are located. Alterations in the DNA binding and nuclease domains result in earlier onset and more severe phenotypes due to reduced FAN1 activity and increased CAG instability. In contrast, variants that promote FAN1 protein levels result in later onset and less severe phenotypes. However, most of these variants have only been recorded in HD patients.

Further research into the combined effects of FAN1 variants and their interaction with other genetic modifiers is necessary to fully understand their role in disease progression and identify therapeutic opportunities.

5. Therapeutic potential of FAN1

FAN1's to stabilize repetitive DNA makes it a promising therapeutic target for addressing CAG repeat instability in HD and SCAs. Combined with its role as a genetic modifier of AOO, linking specific FAN1 variants, such as rs3512, to delayed onset [6, 7]. These discoveries provide a strong rationale for exploring FAN1 modulation to mitigate somatic expansions, critical drivers of disease progression. While DNA repair pathways have been increasingly recognized as therapeutic targets in neurodegenerative diseases, FAN1's unique position in regulating repeat stability offers distinct opportunities for intervention.

5.1. Potential therapeutic approaches

One potential approach involves upregulating FAN1 expression in neurons, where somatic expansions are most detrimental. Gene therapy using viral vectors, such as adeno-associated viruses (AAVs), could deliver functional FAN1 genes directly to affected tissues [17]. Additionally, targeting post-transcriptional regulators like miR-124-3p, which suppresses FAN1 mRNA stability, presents another promising avenue. However, ensuring tissue-specific delivery and minimizing off-target effects remain significant challenges.

Given FAN1's competition with the MMR pathway, another strategy involves modulating MMR activity. Small molecules targeting MutSβ or MutLα could reduce their influence on repeat expansions [18], allowing FAN1 to exert its protective effects more effectively. Combination therapies that enhance FAN1's activity while suppressing MMR-mediated repeat instability may offer synergistic benefits.

5.2. Ongoing therapeutic development

Currently, one drug, HRN001, an antisense oligonucleotide (ASO), is under development by Harness Therapeutics to increase FAN1 levels in the brains of patients with HD by Harness Therapeutics [19]. The company has raised funding to initiate Clinical Trial Application (CTA)-enabling studies in 2025, aiming to gather the necessary data for regulatory approval to begin clinical trials.

ASOs are short, synthetic, single-stranded DNA or RNA molecules that bind specifically to complementary sequences in target RNA [20]. While traditionally used to downregulate gene expression via mRNA degradation, ASOs can also be engineered to upregulate expression by interfering with inhibitory regulatory elements. Although the exact mechanism of HRN001 is not yet disclosed, it is plausible that the ASO targets the binding site of miR-124-3p, a microRNA that destabilizes FAN1 mRNA, reducing its translation. By preventing miR-124-3p binding, HRN001 could potentially enhance FAN1 levels, stabilizing CAG repeats and delaying disease onset.

ASOs offer several advantages as a therapeutic platform. First, their sequence specificity minimizes the risk of off-target effects, a critical consideration for neurodegenerative disorders [20]. Second, their reversible effects reduce the potential for long-term adverse consequences, although this necessitates regular administration [20]. Additionally, the rapid clinical translation of ASOs is supported by the approval of similar therapies, such as Spinraza, which treats spinal muscular atrophy by modulating SMN2 splicing [20, 21].

Interestingly, the genetic variant rs3512, which modifies miR-124-3p binding to FAN1 mRNA, is also a major modifier of AOO in SCAs. This suggests that ASO-based therapies like HRN001, primarily developed for HD, may also hold significant potential for spinocerebellar ataxias. Expanding clinical investigations to include SCAs could broaden the therapeutic

applicability of HRN001, addressing shared mechanisms of CAG repeat instability in both diseases.

Despite these promising strategies, significant hurdles remain. Brain-specific targeting is critical, as somatic expansions primarily affect neurons. However, achieving efficient and safe delivery to the central nervous system is challenging. Additionally, long-term modulation of FAN1 must be studied to ensure it does not compromise other DNA repair functions, such as ICL repair. Expanding research to include SCAs, where FAN1's role is less understood than HD, will also be necessary to broaden its therapeutic applicability.

6. Discussion and conclusion

This review investigated FAN1's role in stabilizing CAG repeats in HD and SCAs, focusing on its molecular mechanisms as a genetic modifier and potential therapeutic target. FAN1's dual nuclease activities and its capacity to bind secondary DNA structures are central to its function in maintaining genomic stability. Deshmukh et al. and McAllister et al. highlighted FAN1's specialized exonuclease activity on repetitive DNA [11, 13], which operates in cycles and is likely to coordinate with other repair enzymes. This supports a model where FAN1 resolves complex DNA structures formed during transcription or replication. FAN1's versatility is underscored by its ability to process various DNA substrates, suggesting a broad role in genomic maintenance.

FAN1's transcription-repair-associated mechanism offers a plausible explanation for its function in non-dividing cells, such as neurons, where transcriptional stress may generate secondary DNA structures. The proposed mechanism aligns with findings of FAN1's activity on slipped-DNA and R-loop structures, which are frequent intermediates in transcription-associated instability. However, further studies are needed to link FAN1's activity to transcription-induced DNA damage directly.

FAN1's competitive interaction with the MMR pathway emerged as a central theme in its role in repeat stability. While the MMR pathway generally ensures genomic integrity, it paradoxically drives repeat expansions in repetitive DNA sequences. Studies by Goold et al. and Phadte et al. demonstrate that FAN1 inhibits MMR-mediated somatic expansions by binding to MLH1, thereby preventing MSH3 from activating the MMR pathway [14, 15]. This mechanism underscores FAN1's protective role against repeat instability.

The observation that FAN1 operates independently of the MMR pathway reinforces the hypothesis that FAN1 stabilizes repeats through distinct mechanisms. However, the exact dynamics of FAN1's competition with MMR proteins, particularly in transcription-associated contexts, remain to be fully elucidated.

FAN1 variants are significant modifiers of disease onset and severity in HD and SCAs. However, only one variant is linked to SCAs. Variants affecting DNA binding and nuclease activity generally cause earlier and more severe disease onset through decreased nuclease activity. Variants in the TPR domain and 3'UTR delay disease onset. The strongest modifier of the age of onset, rs3512 in the 3'-UTR region, diminishes the binding of miR-124-3p, leading to less FAN1 protein. However, the mechanism behind the disease-delaying variants in the TPR domain remains unclear. Additionally, the variants' influence has only been investigated in HD, and while the mechanism in SCA is likely similar, this must be validated.

Identifying FAN1 as a therapeutic target opens promising avenues for HD and SCAs. Strategies like ASOs to upregulate FAN1 levels (e.g., HRN001) or gene editing to enhance its function hold the potential to mitigate somatic expansions. However, challenges such as delivery specificity and long-term effects require further investigation. Given FAN1's involvement in other repeat expansion disorders, these approaches could have broader applications in treating diseases characterized by genomic instability.

Beyond HD and SCAs, FAN1 has been associated with other brain-related conditions, including epilepsy, autism spectrum disorder (ASD), and schizophrenia [8]. While autosomal recessive mutations in FAN1 lead to karyomegalic interstitial nephritis. Although these diseases are not all directly linked to trinucleotide repeat expansions, these disorders share common themes of genomic instability and dysregulated DNA repair pathways, suggesting a potential indirect role for FAN1 in modulating neuronal integrity.

The studies reviewed predominantly focus on HD, leaving significant gaps in understanding FAN1's role in SCAs. While rs3512 in the 3'UTR is the strongest modifier identified

for both HD and SCA, its specific influence on SCA genes and repeat expansions requires further validation. Larger cohorts and diverse populations are necessary to confirm FAN1 as a genetic modifier in SCA and identify additional SNPs. In addition, whether SNPs identified in HD have similar effects in SCA should be investigated. If the mechanism in which FAN1 stabilizes CAG repeats is the same in SCA as in HD, therapies developed for HD, such as ASOs like HRN001, could be repurposed for SCA.

FAN1 is a critical modulator of repeat stability, balancing transcription-repair mechanisms and its competition with the MMR pathway to prevent somatic expansions. Its dual role as a genomic stabilizer and genetic modifier positions FAN1 as a promising therapeutic target for HD, SCAs, and potentially other neurological diseases characterized by genomic instability. However, a deeper understanding of its mechanisms in diverse contexts is essential to harness its therapeutic potential fully. By bridging gaps in SCA research and extending insights from HD, FAN1-targeted therapies could benefit patients suffering from these diseases.

Total word count (Introduction – Conclusion): 5,437

The use of AI

In the process of writing this essay, I utilized AI tools, including Microsoft Copilot and Grammarly, to assist with structuring, refining, and enhancing the content. These tools provided valuable support in organizing arguments, integrating literature, and improving clarity and coherence. It is important to note that AI was not used for content generation but solely for organizing and editing the essay.

References

1. Tenchov, R., J.M. Sasso, and Q.A. Zhou, *Polyglutamine (PolyQ) Diseases: Navigating the Landscape of Neurodegeneration*. ACS Chem Neurosci, 2024. **15**(15): p. 2665-2694.
2. He, X.H., F. Lin, and Z.H. Qin, *Current understanding on the pathogenesis of polyglutamine diseases*. Neurosci Bull, 2010. **26**(3): p. 247-56.
3. Zahra, R.L.S., A. R., *Chapter 48 - Cellular and Molecular Basis of Neurodegeneration in the CAG–Polyglutamine Repeat Diseases*, in *Basic Neurochemistry*, S.T.S.G.J.A. Brady, R. W.; Price D. L., Editor. 2012, Academic Press: New York. p. 844–855.
4. Lieberman, A.P., V.G. Shakkottai, and R.L. Albin, *Polyglutamine Repeats in Neurodegenerative Diseases*. Annu Rev Pathol, 2019. **14**: p. 1-27.
5. Andrew, S.E., et al., *The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease*. Nat Genet, 1993. **4**(4): p. 398-403.
6. Bettencourt, C., et al., *DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases*. Ann Neurol, 2016. **79**(6): p. 983-90.
7. Consortium, G.M.o.H.s.D.G.-H., *Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease*. Cell, 2015. **162**(3): p. 516-26.
8. Deshmukh, A.L., et al., *FAN1, a DNA Repair Nuclease, as a Modifier of Repeat Expansion Disorders*. J Huntingtons Dis, 2021. **10**(1): p. 95-122.
9. Loupe, J.M., et al., *Promotion of somatic CAG repeat expansion by Fan1 knock-out in Huntington's disease knock-in mice is blocked by Mlh1 knock-out*. Hum Mol Genet, 2020. **29**(18): p. 3044-3053.
10. Kim, K.H., et al., *Genetic and Functional Analyses Point to FAN1 as the Source of Multiple Huntington Disease Modifier Effects*. Am J Hum Genet, 2020. **107**(1): p. 96-110.
11. Deshmukh, A.L., et al., *FAN1 exo- not endo-nuclease pausing on disease-associated slipped-DNA repeats: A mechanism of repeat instability*. Cell Rep, 2021. **37**(10): p. 110078.
12. Porro, A., et al., *FAN1-MLH1 interaction affects repair of DNA interstrand cross-links and slipped-CAG/CTG repeats*. Sci Adv, 2021. **7**(31).
13. McAllister, B., et al., *Exome sequencing of individuals with Huntington's disease implicates FAN1 nuclease activity in slowing CAG expansion and disease onset*. Nat Neurosci, 2022. **25**(4): p. 446-457.
14. Phadte, A.S., et al., *FAN1 removes triplet repeat extrusions via a PCNA- and RFC-dependent mechanism*. Proc Natl Acad Sci U S A, 2023. **120**(33): p. e2302103120.
15. Goold, R., et al., *FAN1 controls mismatch repair complex assembly via MLH1 retention to stabilize CAG repeat expansion in Huntington's disease*. Cell Rep, 2021. **36**(9): p. 109649.
16. Kim, K.H., et al., *Posttranscriptional regulation of FAN1 by miR-124-3p at rs3512 underlies onset-delaying genetic modification in Huntington's disease*. Proc Natl Acad Sci U S A, 2024. **121**(16): p. e2322924121.
17. Ye, D., et al., *Adeno-associated virus vector delivery to the brain: Technology advancements and clinical applications*. Adv Drug Deliv Rev, 2024. **211**: p. 115363.
18. Erie, D.A. and K.R. Weninger, *Combining single-molecule and structural studies reveals protein and DNA conformations and assemblies that govern DNA mismatch repair*. Curr Opin Struct Biol, 2024. **89**: p. 102917.
19. Therapeutics, H. *Huntington's Disease Programme*. 2024 [cited 2024 November 22]; Available from: https://www.harnesstx.com/our-programmes/huntingtons-disease/.
20. Rinaldi, C. and M.J.A. Wood, *Antisense oligonucleotides: the next frontier for treatment of neurological disorders*. Nat Rev Neurol, 2018. **14**(1): p. 9-21.
21. Nederland, Z. *Nusinersen - Preparaattekst*. 2024 [cited 2024 November 22]; Available from: https://www.farmacotherapeutischkompas.nl/bladeren/preparaatteksten/n/nusinersen.