Long read sequencing: High-fidelity amplification of disease-associated complex genomic regions using DeCodi-Fi polymerase.

Poster No. #B-224

J. Hsieh, J. Gimpel, J. Sáez, J. Cáceres, C. Valdivia, I. Arriagada, T. Akentjew, S. Kamkar, I. Marx, M. Rozas
Kura Biotech Spa, Puerto Varas, Región de Los Lagos, Chile

Background

- Repeat expansion disorders such as Amyotrophic Lateral Sclerosis (**ALS**), Friedreich's Ataxia (**FrA**), Fragile X (**FrX**) and Huntington's Disease (**HD**) are notoriously difficult to detect with short-read sequencing (fragments shorter than 300 bp)
- Amplifying GC-rich, long tandem repeats often results in polymerase slippage or dropout, which can compromise diagnostic accuracy when an amplification step is required.
- Long-read sequencing platforms (e.g., Nanopore) provide a more direct and reliable approach for resolving repeat expansions, but their success depends heavily on the performance of the high-fidelity polymerase, particularly when amplifying challenging templates such as long, GC-rich regions (e.g. 6.5 kb with 100% GC content).

Objective

To evaluate the performance of the high-fidelity polymerase DeCodi-Fi compared to other high-fidelity enzymes in amplifying complex, long, and extremely GC-rich repeat regions.

Methods

Repeat Expansion Samples: DNA from unaffected control and affected individuals (Coriell Institute) were used to evaluate amplification performance across **expanded repeat loci** for each disease by different polymerases. **Table 1** summarizes target genes, repeat motifs, template DNA sources, repeat counts and expected product sizes for each disease.

Long-Range PCR Samples: To evaluate amplification performance on large genomic targets, we tested BRCA1 (8,4 kb, 23,6 kb and 36,4 kb) and SMN1 (28.2 kb), assessing DeCodi-Fi's ability to amplify long human DNA fragments compared to other high-fidelity polymerases.

Polymerases tested: DeCodi-Fi (Kura Biotech), **Platinum SuperFi II** (Thermo Fisher Scientific) and **Q5** (NEB). Relevant reaction buffers for each enzyme were used according to manufacturer's instructions. Where additional enhancers to PCR reactions were used, the same additions were made to competitor reactions.

Cycling conditions: Cycling protocols were customized and optimized for specific primer pairs and targets, using techniques such as touchdown PCR, controlled ramp rates and "wobble" temperatures. The same amplification protocol was applied across all polymerases for each target.

Detection method: Results were visualized by agarose gel electrophoresis. Long-read sequencing was performed on the PromethION (Oxford Nanopore Technologies).

Table 1: Summary of genes, repeat motifs, template DNA, repeat counts, and expected amplicon sizes for each disease studied.

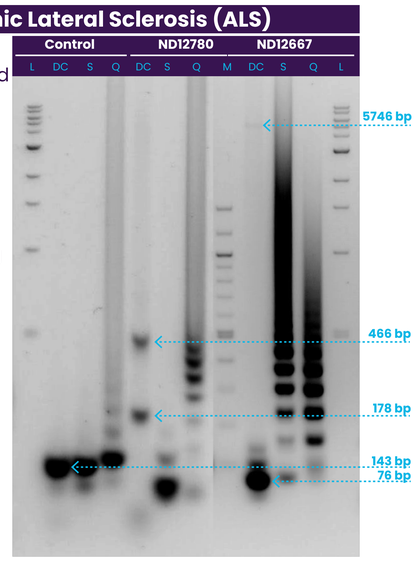
Disease	Gene	Repeat Motif	Template DNA (Coriell Sample)	Repeat Count (Allele 1/Allele 2)	Expected Size (Allele 1/Allele 2)
Amyotrophic Lateral Sclerosis (ALS)	*C9orf72*	GGGGCC	Unaffected control	< 20	143 bp
			ALS (ND12780)	22/70	178/466 bp
			ALS (ND12667)	5/~950	76/5746 bp
Friedreich's Ataxia (FrA)	*X25*	GAA	Unaffected control	< 33	520 bp
			FrA (NA16243)	670/1117	2510/4010 bp
Fragile X Syndrome (FrX)	*FMR1*	CGG	Unaffected control	<44	350 bp
			FrX (NA7862)	> 200-230	2339bp/ >3239bp
Huntington's Disease (HD)	*IT15*	CAG	Unaffected control	< 36	105 bp
			HD (NA20253)	22/100	119/353 bp
			HD (NA13509)	15/70	98/263 bp

Results: Amplifying the Toughest Targets (6.5 kb GC Repeats)

Figure 1. Amplification of disease-associated repeat expansions using three high-fidelity polymerases. Agarose gel electrophoresis showing PCR amplification of repeat-containing loci from **unaffected control** and **disease-affected samples** across four disorders: ALS (**A**), FrA (**B**), FrX (**C**), and HD (**D**). Three high-fidelity polymerases; DeCodi-Fi **(DC)**, Platinum SuperFi II **(S)**, and Q5 **(Q)** were tested under equivalent conditions. For unaffected control amplification, DeCodi-Fi and Q5 were used with their respective GC buffer/enhancer, and SuperFi with its reaction buffer. Additional enhancers were included to amplify disease-expanded alleles. Light blue arrows indicate detected PCR products for each allele. Some gels show primer dimers or non-specific bands due to pre-cleanup loading, intended to enable raw product comparison across conditions. **M:** 100 bp ladder; **L:** 1 kb Ladder.

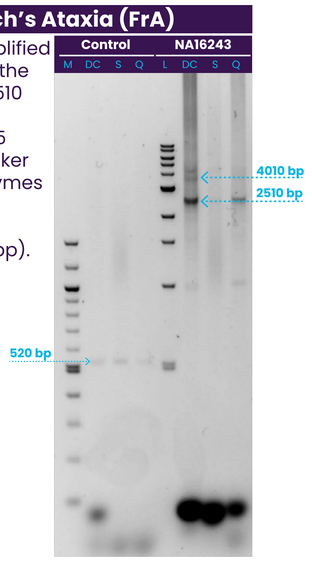
A. Amyotrophic Lateral Sclerosis (ALS)

DeCodi-Fi uniquely amplified both alleles in ND12780 (178 bp / 466 bp) and ND12667 (75 bp / 5746 bp). SuperFi and Q5 failed to amplify specific bands, producing only smears or ladders. All three enzymes amplified the control (143 bp), though Q5 showed a band that was larger than expected.



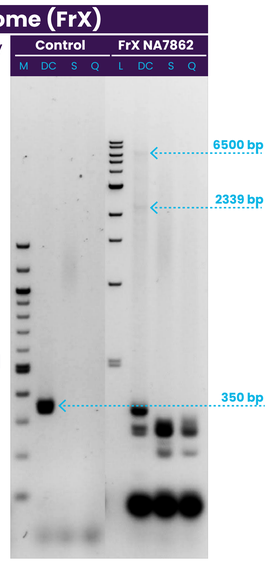
B. Friedreich's Ataxia (FrA)

DeCodi-Fi amplified both alleles in the FrA sample (2510 bp / 4010 bp). SuperFi and Q5 produced weaker bands. All enzymes successfully amplified the control (~520 bp).



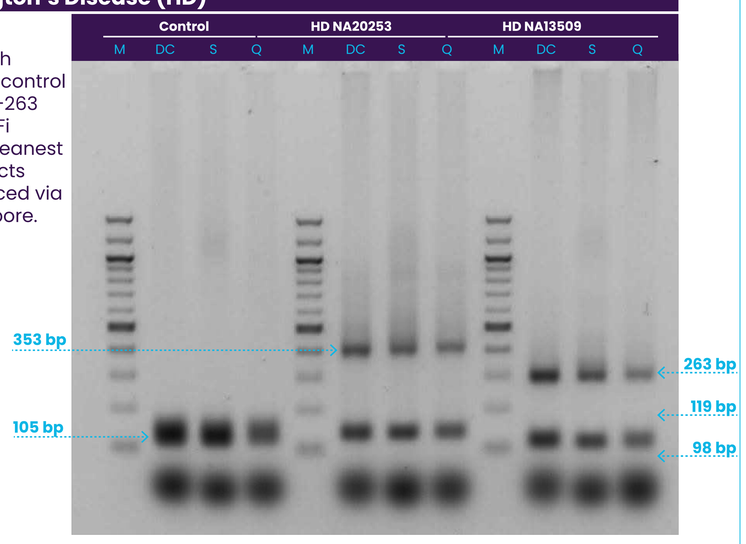
C. Fragile X Syndrome (FrX)

DeCodi-Fi was the only enzyme to amplify the expanded 6500 bp mosaic allele (100% GC), which is often missed and present only in a subpopulation of cells in this FrX affected sample. The 2339 bp allele and the unaffected control (~350 bp) were also detected with DeCodi-Fi. SuperFi and Q5 failed to amplify the control allele, which had 79.26% GC.



D. Huntington's Disease (HD)

All enzymes amplified both affected and control samples (98–263 bp). DeCodi-Fi yielded the cleanest bands. Products were sequenced via Oxford Nanopore.



MAIN FINDING:

DeCodi-Fi robustly amplified 100% GC targets up to ~500 bp and was the only enzyme to detect larger repeats (2000–6500 bp), outperforming competitors.

Oxford Nanopore Sequencing

Huntington's disease alleles were selected for comparing sequencing results between polymerases since it was the only disease where both unaffected and affected samples could be amplified by all polymerases. DeCodi-Fi generally achieved higher coverage and in some cases more spanning reads than Platinum SuperFi and Q5.

Table 2: Motif counts observed and sequencing metrics for unaffected individual and Huntington's disease (HD) patients for Allele1(A1)/Allele2(A2). DeCodi-Fi (**DC**), Platinum SuperFi II (**S**), and Q5 (**Q**) polymerases were compared. The median motif count refers to the number of motifs detected in the sequencing results for each allele. Coverage corresponds to the total number of sequencing reads aligned over the repeat region, and spanning reads indicate the number of reads that fully span the repeat motif.

Genome	Expected motif count (A1/A2)	Enzyme	Median motif count(A1/A2)	Coverage	Spanning reads (A1/A2)
Unaffected	<36	DC	17/17	9958	2367/2067
		S	18/18	9948	2104/2254
		Q	18/18	8122	1856/1675
HD NA20253	22/100	DC	22/103	8930	2652/230
		S	22/106	8016	2526/380
		Q	22/104	5706	1766/279
HD NA13509	15/70	DC	15/71	8568	3404/511
		S	15/71	748	252/93
		Q	22/71	4720	2527/381

Amplifying the Longest Targets

DeCodi-Fi successfully amplified human DNA fragments up to 36.4 kb, while other polymerases were limited to amplifying fragments no larger than 8.4 kb.

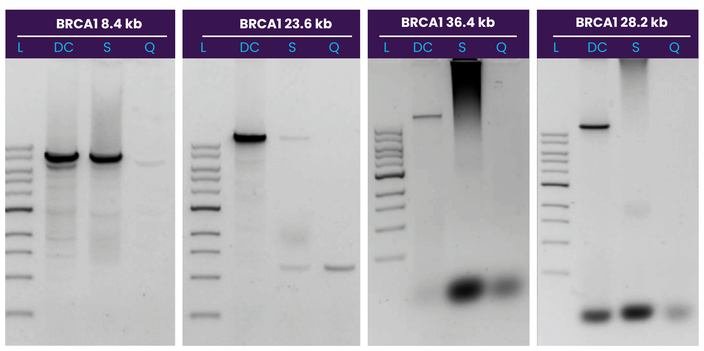


Figure 2. Long-range amplification of clinically relevant targets, including Breast Cancer 1 gene (BRCA1) and Survival of Motor Neuron 1 gene (SMN1). DeCodi-Fi (**DC**) successfully amplified all four large fragments (BRCA1 8.4 kb, 23.6 kb, 36.4 kb, and SMN1 28.2 kb), producing strong, specific bands. This capability is critical for long-read sequencing, where full-length amplification enables comprehensive analysis of structural variants and mutations. Platinum SuperFi II (**S**) amplified only the 8.4 kb fragment, while Q5 (**Q**) produced a faint band at 8.4 kb and both failed to amplify larger targets. **L:** 1 kb ladder.

Conclusion

- DeCodi-Fi reliably amplified all four disease-associated repeat expansions, including long and highly GC-rich alleles.
- It was the only polymerase to amplify the 6.5 kb, 100% GC ([CGG]>1000) repeat in Fragile X and the 5.7 kb GGGGCC expansion in ALS.
- In long-range PCR, DeCodi-Fi amplified human genomic targets up to 36.4 kb, whereas other polymerases were limited to 8.4 kb.
- Sequencing of Huntington's disease samples showed higher coverage and more spanning reads with DeCodi-Fi.
- These results support DeCodi-Fi as a robust tool for long-read sequencing of complex genomic regions.