P602

FXN Repeat-Primed PCR and Long-Read Sequencing Reveal Non-Canonical Repeat Expansions

Benjamin Obadia[1], Alison Nguyen[1], Tony Tran Yonce[1], Luis Jazo[1], Jhett Bordwell[1], Jack Roach[1], John Alexander[1]

1 Myriad Genetics, Inc., Salt Lake City, UT

Introduction

- Friedreich ataxia (FRDA) is an autosomal recessive, neurodegenerative condition caused by pathogenic variants in the *FXN* gene. The majority of cases (>95%) are caused by GAA trinucleotide repeat expansions in intron 1
  - The estimated incidence is 1 in 15,000 to 1 in 40,000 live births
  - The size of the GAA trinucleotide repeat correlates with the expected phenotype (**Table 1**)
- Here we demonstrate the implementation of a repeat-primed PCR (RP-PCR) assay followed by capillary electrophoresis (CE) to detect GAA repeat expansion in *FXN* (**Figure 1A**)
  - The assay detects full length PCR products from GAA alleles ≤ 100 repeats
  - Larger alleles with >100 GAA repeats generate a characteristic stutter pattern, indicative of a pathogenic expansion
- The assay was validated using known positive controls and single-molecule long-read sequencing for orthogonal confirmation (**Figure 1B**)
- Long-read sequencing (LR-seq) was leveraged to elucidate the sequence composition of alleles ≥36 repeats identified from the initial cohort analyzed by the CE-based assay
  - Triplet repeat consensus in *FXN* is GAA, however alternative patterns (GAG, GGA) have been reported

Table 1. *FXN* repeat expansion characterization, acceptable deviation, and carrier status used in this study

ndividuals with one allele above the normal range (>33 repeats) are reported as carriers because those alleles can become full expansions (≥66) in future generations.[a] Full expansion alleles with more than 100 triplet repeats are not specifically sized on the CE assay and reported as ">100".[b]

FXN allele type	Number of triplet repeats	Acceptable deviation	Result type
Normal	5-33	± 2	Negative
Premutation[a]	34-43	± 2	Carrier
Borderline	44-65	± 2	Carrier
Full expansion[b]	66-100	± 5	Carrier
N/A	Biallelic ≥44	N/A	Individual at risk

Methods

Assay (Foresight; Myriad Genetics, Inc.) validation

- 187 unique samples (internal and Coriell repository)
- Sample set:
  - 30 previously known positive samples with repeat expansions (carrier and individual at risk)
  - 8 newly identified positive samples
  - 149 newly identified negative samples
- Individual *FXN* allele sizes from the CE were compared to their sizing from LR-seq to assess sensitivity, specificity, concordance, and validity (**Table 2, Figure 2**)

Post-validation pilot study

- 1,736 and 1,839 random patient samples were tested using CE-based and LR-seq assays
- 235 unique samples were run on both assays, calls for triplet expansion were compared, and repeat composition was analyzed from LR-seq (**Table 3, Figure 3**)

Results

CE-based assay validation

- High sensitivity and specificity for detecting GAA expansions
- 100% concordance based on FRDA clinical classification
- 100% of sample alleles within acceptable validity ranges

Figure 1. Overview of the PCR–CE and long-read sequencing workflows

For the PCR–CE assay **(A)**, the repeat-primed PCR uses a set of 3 primers (arrows) to amplify a ≥173-bp region from the GAA repeat locus (green box). For the long-read sequencing assay **(B)**, Cas9 associated with guide RNAs (gRNA1 and gRNA2, purple blob) cut a 3947-bp region surrounding the GAA repeat locus (green box).

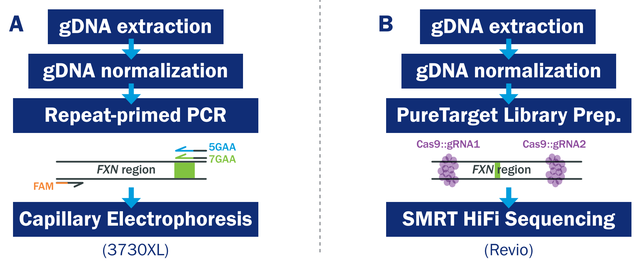


Table 2. Confusion matrix of patient status between CE and LR-seq assays (validation study)

N=187		CE		
		Non-carrier	Carrier	Individual at risk
LR-seq	Non-carrier	151	0	0
	Carrier	0	18	0
	Individual at risk	0	0	18

Figure 2. Correlation between triplet repeat calls from CE and LR-seq (validation study)

Pairwise (Pearson) correlation (red linear fit) between long-read sequencing (LR-seq) and capillary electrophoresis (CE) triplet repeat calls. The density ellipse (green) includes 99% of the datapoints (N=374 alleles from 187 samples).

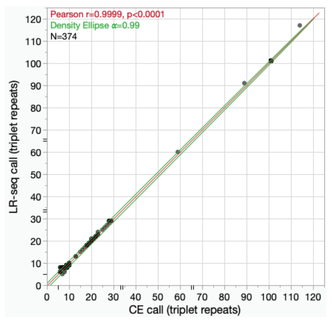


Table 3. Confusion matrix of patient status between CE and LR-seq assays (post-validation study)

N=235		CE		
		Non-carrier	Carrier	Individual at risk
LR-seq	Non-carrier	230	0	0
	Carrier	0	5	0
	Individual at risk	0	0	0

Figure 3. Correlation between triplet repeat calls from CE and LR-seq (post-validation study).

Pairwise (Pearson) correlation (red linear fit) between long-read sequencing (LR-seq) and capillary electrophoresis (CE) triplet repeat calls. The density ellipse (green) represents 99% of the datapoints (N=470 alleles from 235 samples).

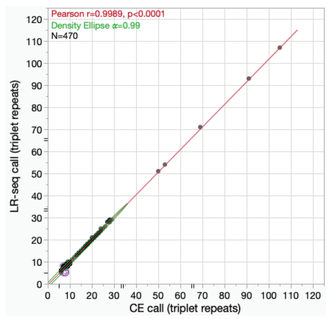


Table 4. Summary of carrier samples screened on LR-seq assay

Sample ID	Coverage allele 1	Coverage allele 2	Motif repeats allele 1				Motif repeats allele 2				LR-seq call	Classification allele 1	Classification allele 2	GAA interruption
			A	GAA	GGA	GAG	A	GAA	GGA	GAG				
043-040	256	262	23	8	0	0	25	36	0	0	(8,36)	Normal	Premutation	NO
055-024	228	210	23	8	0	0	26	49	0	1	(8,50)	Normal	Borderline	YES
055-066	102	98	25	9	0	0	24	46	0	8	(9,54)	Normal	Borderline	YES
009-024	70	70	23	19	0	0	25	50	0	10	(19,60)	Normal	Borderline	YES
055-087	58	54	25	9	0	0	25	54	0	16	(9,70)	Normal	Full Expansion	YES
026-012	22	26	23	8	0	0	29	44	33	0	(8,77)	Normal	Full Expansion	YES
039-010	457	400	24	13	0	0	27	25	54	1	(13,80)	Normal	Full Expansion	YES
044-042	176	151	25	9	0	0	25	70	0	17	(9,87)	Normal	Full Expansion	YES
042-092	73	77	25	9	0	0	25	72	0	21	(9,93)	Normal	Full Expansion	YES
055-079	101	78	25	9	0	0	26	53	40	0	(9,93)	Normal	Full Expansion	YES
024-029	198	185	23	8	0	0	25	74	0	20	(8,94)	Normal	Full Expansion	YES
046-022	170	130	25	9	0	0	34	94	1	3	(9,98)	Normal	Full Expansion	YES
040-082	275	232	24	19	0	0	19	46	31	25	(19,102)	Normal	Full Expansion	YES
048-046	188	155	25	9	0	0	26	57	46	0	(9,103)	Normal	Full Expansion	YES
055-076	83	104	25	9	0	0	25	60	49	0	(9,109)	Normal	Full Expansion	YES
032-027	529	543	26	9	0	0	31	104	17	1	(9,122)	Normal	Full Expansion	YES
018-013	140	153	24	9	0	0	24	105	0	19	(9,124)	Normal	Full Expansion	YES
015-076	27	25	23	15	0	0	24	83	76	0	(15,159)	Normal	Full Expansion	YES
049-007	128	47	22	8	0	0	101	381	2	30	(8,413)	Normal	Full Expansion	YES
055-044	484	115	25	9	0	0	188	585	8	47	(9,640)	Normal	Full Expansion	YES
036-055	76	39	25	9	0	0	219	700	8	19	(9,727)	Normal	Full Expansion	YES
055-048	53	27	24	8	0	0	207	697	8	61	(8,766)	Normal	Full Expansion	YES
050-029	484	64	23	8	0	0	285	666	15	96	(8,777)	Normal	Full Expansion	YES
021-062	181	66	25	9	0	0	246	787	7	21	(9,815)	Normal	Full Expansion	YES
040-039	413	81	23	8	0	0	264	867	13	17	(8,897)	Normal	Full Expansion	YES
024-079	161	71	24	16	0	0	242	842	10	68	(16,920)	Normal	Full Expansion	YES
040-004	452	49	25	9	0	0	262	826	17	83	(9,926)	Normal	Full Expansion	YES

CE-based assay post-validation pilot study

- 24 carriers (1.4%) identified (none at risk)
- Distinct alleles for sizes up to 153 repeats

Long-read sequencing assay post-validation pilot study

- 27 carriers (1.5%) identified (none at risk)
- Repeat range: 5 to 926 repeats
  - 96.3% (26/27) of expanded alleles exhibited GAA interruptions, characterized by GGA (19.2%), GAG (30.8%), or both motifs (50%) (**Table 4**)
  - 33.3% (9/27) of expanded alleles exhibited an expanded adenine (A) tract, and all (9/9) alleles with expanded A-tract had a cumulative GAA, GGA, or GAG beyond 400 repeats (**Table 4, Figures 4-5**)
- Interrupted samples tested on both assays were concordant, despite the presence of alternative triplet expansions (**Table 3**)

Figure 4. Overview of triplet repeat interruptions for a subset of carrier samples

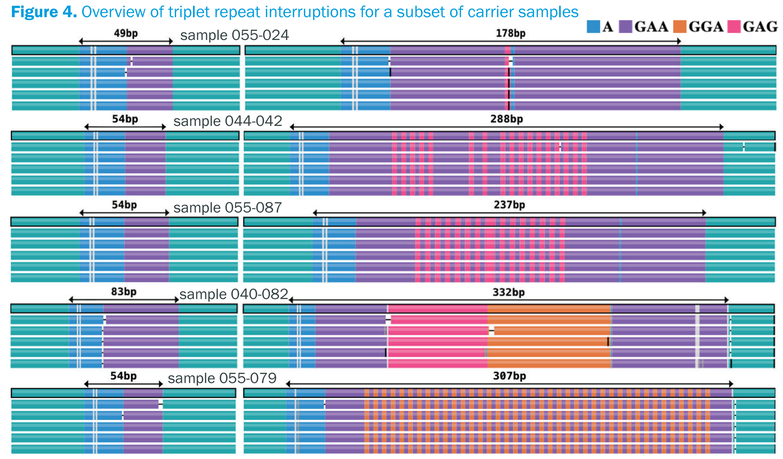
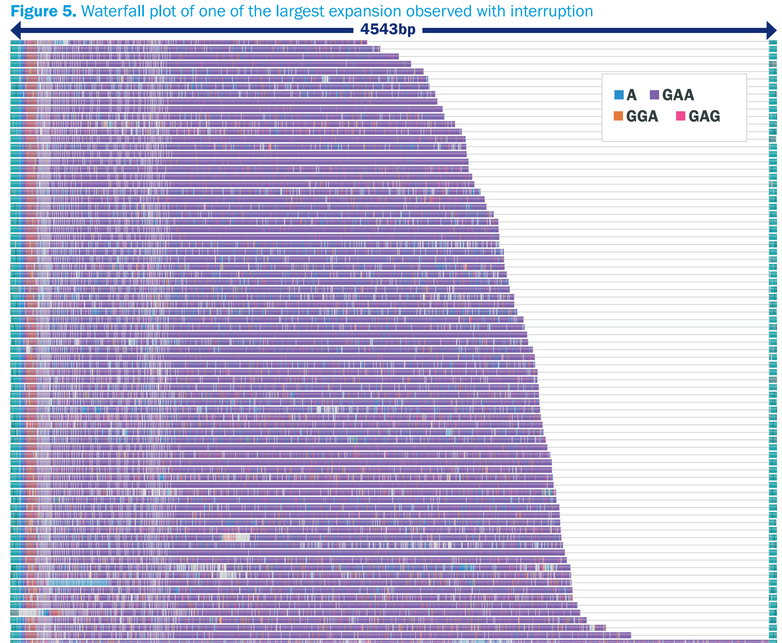


Figure 5. Waterfall plot of one of the largest expansion observed with interruption



Conclusions

- CE-based assay was validated as an efficient and accurate screening method for measuring repeat expansions in *FXN*
- LR-seq revealed that GAA interruptions were present in most expanded alleles from the study cohort
- Identifying interruptions may have significant clinical relevance; however, additional larger scale studies and an interpretation paradigm will be needed prior to incorporating interruption data into routine testing

Disclosures: All authors were employees of Myriad Genetics, Inc. at the time of this study and received salary and stock as compensation.

Email for questions: benjamin.obadia@myriad.com