A Deep Learning Method for High-Throughput *FMR1* Triplet Repeat Screening

Lando Ringel, Daniel Borders, Elliot Hallmark, Brian C Haynes, Jessica L Larson
Asuragen, Inc., Austin, TX

Summary

- PCR/Capillary Electrophoresis (CE) trace interpretation requires trained manual operators or customized heuristic methods.
- We present *FMR1* DeepNet*, a deep learning based prototype for automated genotyping short tandem repeat (STR) variants in *FMR1*.
- In a validation set of residual clinical specimens, *FMR1* DeepNet classified STR alleles with ≥96% sensitivity and ≥97% PPV. The categorical accuracy of the model on these specimens was ≥97%.

Introduction

Approximately 1 in 200 females and 1 in 450 males are carriers of fragile X syndrome (FXS) in the US[1]. High-throughput carrier screening and diagnostic testing requires an accurate and robust method for *FMR1* genotyping, which has been historically problematic due to the difficult-to-amplify CGG repeat that causes FXS in >99% of cases. Advances in PCR/CE technologies have enabled the amplification and sizing of these triplet repeats within the clinically-relevant range. However, existing CE trace interpretation requires trained manual operators or computational heuristic methods tailored to signal idiosyncrasies and peak morphologies. To address this issue, we developed a deep learning algorithm that can automate reliable determination of the *FMR1* STR genotype.

Methods

Blood and cell-line specimens were collected and processed with the AmplideX® PCR/CE FMR1 Kit*,†. Multi-channel CE relative fluorescence unit values (RFU) were pre-processed, sized, and scaled; candidate peaks above sample- and location-specific background thresholds were extracted for consideration. A multi-layered, multi-output convolutional neural network (CNN) classifier was developed to differentiate *FMR1* STR peaks from background noise (Figure 1). Repeat length was used to classify specimens into normal/intermediate (<55 repeat), premutation (≥55 repeat and ≤200 repeat), and full-mutation (>200 repeat) categories. The trained CNN was evaluated on two independent clinical validation cohorts with reference *FMR1* genotypes assessed using a secondary PCR/CE technique (dual-PCR) and/or Southern blot analysis.

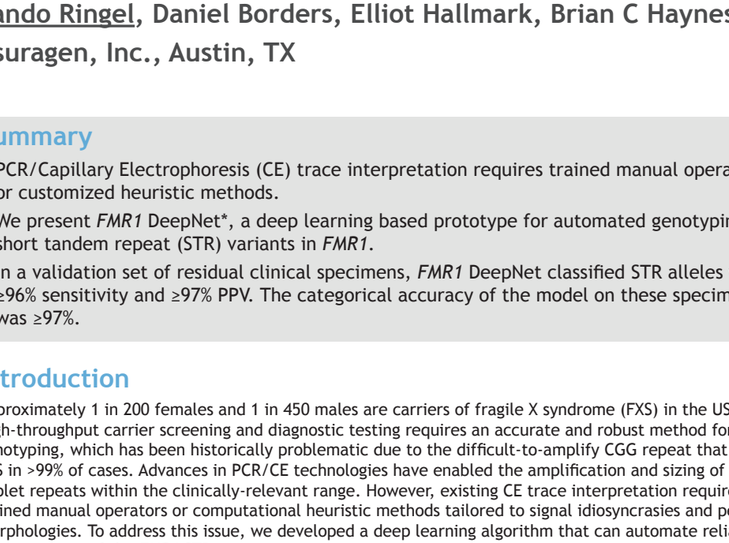

Figure 1. A) Primer design for targeting *FMR1* STR pattern. **B)** Fluorescence signal output from Applied Biosystems 3500 Genetic Analyzer and Asuragen's AmplideX PCR/CE *FMR1* Kit depicted in electropherogram format. This example contains a heterozygous female full-mutation sample with a 31|>200 genotype. Background noise and PCR artifacts resembling peaks are labelled "Not Peak." **C)** CE peaks above sample-specific background-noise thresholds were extracted and annotated into repeat-size based categories. The signals were treated as 1-dimensional images and passed through a series of convolutional and fully connected layers. The final layer returned a probability score for both "Peak" and "Not Peak" categories. Peaks were categorized into "Normal," "Intermediate," "pre-mutation," and "full mutation" classes based upon x-axis location in electropherogram.

*Proof-of-concept data only. Future availability and performance cannot be ensured. **CE-IVD for US export only.
†Research Use Only. Not for use in diagnostic procedures
Presented at AMP 2019

Results

Genotype Category # CGG	Normal/Intermediate <55	Premutation 55-200	Expanded >200	Total
FMR1 DeepNet (training cohort)	1428	1704	579	3711

Table 1. The training cohort contained 3,711 internally validated samples, consisting of 38.5% normal/intermediate, 45.9% premutation, and 15.6% full mutation classifications. Peaks were annotated by manual review process and confirmed by automated heuristic techniques. Based upon maximum STR size, each sample was assigned clinically relevant categories. The pipeline took 130.5 minutes to execute peak extraction, signal pre-processing, and CNN training.

	Southern Blot/Dual-PCR Comparison Cohort	Dual-PCR Comparison Cohort	Total
Number of specimens	207	207	414
Number of Extracted Peaks	1,972	2,490	4,462
True Positive	322	400	722
False Negative	12	5	17
False Positive	8	11	19
Sensitivity	96.4%	98.8%	97.7%
PPV	97.6%	97.3%	97.4%

Table 2. Our trained *FMR1* DeepNet CNN model was validated upon 4,462 peaks extracted from two independent clinically derived cohorts. Each sample genotype was confirmed by either dual-PCR or Southern blot and dual-PCR comparison method. The algorithm correctly identified 722 out of 739 (97.7%) expected positive *FMR1* peaks. The algorithm misclassified 19 out of 3265 (0.5%) expected negative peaks. When combining both independently validated clinical cohorts, *FMR1* DeepNet achieves 97.7% sensitivity and 97.4% PPV.

		Southern Blot/Dual-PCR Comparison Cohort			
FMR1 DeepNet Prediction	Genotype Category # CGG	Normal/Intermediate <55	Premutation 55-200	Expanded >200	Total
	Normal/Intermediate <55	67	0	1	68
	Premutation 55-200	2	67	1	70
	Expanded >200	0	0	67	67
	No Peaks	0	2	0	2
	Total	69	69	69	207

Table 3. The CNN achieved 97.1% categorical agreement with a Southern blot/Dual-PCR reference method for diagnostic-relevant full, premutation, and normal/intermediate genotypes. The CNN failed to detect two low-intensity expanded peaks below the 100 RFU background noise threshold. The CNN identified two "Normal" peaks as 54 repeats but sized them incorrectly as "Premutation" (55 repeats). This is a boundary case that reflects a failure on the peak sizing algorithm. The CNN failed to detect two low RFU (<5,000) premutation peaks in two male samples. We attribute these false negatives to their peak morphology not being represented in the training cohort.

		Dual-PCR Comparison Cohort			
FMR1 DeepNet Prediction	Genotype Category # CGG	Normal <45	Intermediate 45-54	Full/premutation >54	Total
	Normal <45	68	0	0	68
	Intermediate 45-54	0	70	0	70
	Full/premutation >54	1	0	68	69
	Total	69	70	68	207

Table 4. *FMR1* DeepNet achieved >98% categorical agreement with the dual-PCR reference method for carrier screening-relevant full/premutation, intermediate, and normal genotypes. Given 99.5% concordance with the manually annotated dual-PCR cohort, our CNN based genotyping method can be considered a clinically viable technique for PCR/CE peak calling. The "Full/premutation" misclassification was due to an air-bubble artifact that evaded detection during preprocessing.

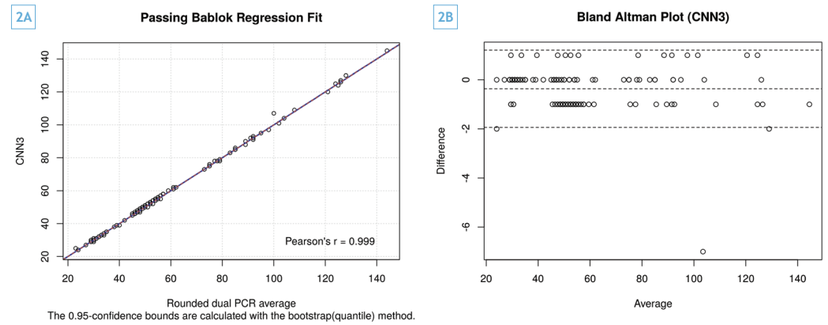

Figure 2. There is no statistical difference between the repeat length reported by the DeepNet CNN and the independent dual-PCR reference method (Passing-Bablok regression 95% bootstrap CIs for slope (1.00, 1.00) and intercept (0.00, 0.00)). **A)** Estimated CGG repeat number for the reference method (x-axis) versus the CNN (y-axis) for all alleles <200 CGGs is plotted. The non-parametric Passing-Bablok regression line is plotted. **B)** Bland-Altman bias plot shows minimal bias present in the CNN output (average between the two methods on the x-axis, difference on the y-axis). The extreme outlier is the normal that was miscalled as a full/premutation as noted in Table 4.

Conclusions

- *FMR1* DeepNet is a robust and highly accurate deep learning system to detect, size, and categorize *FMR1* repeat lengths from PCR/CE data.
- *FMR1* DeepNet is well suited for high volume settings, offering advantages over manual interpretation in terms of speed, convenience, and consistency.
- Our approach has the potential to expand FXS screening and may also improve PCR/CE analysis of other repetitive, and non-repetitive, genetic markers.

References

1. Saldarriaga W, Tassone F, González-Teshima LY, Forero-Forero JV, Ayala-Zapata S, Hagerman R. Fragile X syndrome. Colomb Med (Cali). 2014;45(4):190-198. Published 2014 Dec 30.

Asuragen®
www.asuragen.com