Clinical Chemistry 62:10
1401–1408 (2016)

Group Testing Approach for Trinucleotide Repeat Expansion Disorder Screening

Kristjan Eerik Kaseniit,[1] Mark R. Theilmann,[1] Alexander Robertson,[1] Eric A. Evans,[1] and Imran S. Haque[1*]

BACKGROUND: Fragile X syndrome (FXS, OMIM #300624) is an X-linked condition caused by trinucleotide repeat expansions in the 5′ UTR (untranslated region) of the fragile X mental retardation 1 (*FMR1*) gene. FXS testing is commonly performed in expanded carrier screening and has been proposed for inclusion in newborn screening. However, because pathogenic alleles are long and have low complexity (>200 CGG repeats), FXS is currently tested by a single-plex electrophoresis-resolved PCR assay rather than multiplexed approaches like next-generation sequencing or mass spectrometry. In this work, we sought an experimental design based on nonadaptive group testing that could accurately and reliably identify the size of abnormally expanded *FMR1* alleles of males and females.

METHODS: We developed a new group testing scheme named StairCase (SC) that was designed to the constraints of the FXS testing problem, and compared its performance to existing group testing schemes by simulation. We experimentally evaluated SC's performance on 210 samples from the Coriell Institute biorepositories using pooled PCR followed by capillary electrophoresis on 3 replicates of each of 3 pooling layouts differing by the mapping of samples to pools.

RESULTS: The SC pooled PCR approach demonstrated perfect classification of samples by clinical category (normal, intermediate, premutation, or full mutation) for 90 positives and 1800 negatives, with a batch of 210 samples requiring only 21 assays.

CONCLUSIONS: Group testing based on SC is an implementable approach to trinucleotide repeat expansion disorder testing that offers ≥10-fold reduction in assay costs over current single-plex methods.

Fragile X syndrome [FXS; Online Mendelian Inheritance in Man (OMIM) #300624][2] is a genetic disorder causing a wide range of developmental and behavioral problems and is the leading cause of inherited mental disability *(1–4)*. The syndrome is caused by the expansion of a CGG microsatellite repeat sequence in the 5′ untranslated region (UTR) of the fragile X mental retardation 1 (*FMR1*) gene (HGNC 3775), ultimately leading to methylation and silencing of *FMR1 (1)*.

Alleles are categorized by their repeat count *(1, 4)*, for which "normal" alleles under 45 repeats in length and INT ("intermediate" or "gray zone" alleles; 45–54 repeats) are generally benign and stable across generations. PM (premutation alleles; 55–200 repeats) may expand on maternal transmission to FM (full mutation status; >200 repeats). Full mutation alleles are diagnostic for FXS in males, whereas females are affected with <50% penetrance and show milder symptoms. Men and women with the premutation are at risk for a number of related symptoms such as ataxia (FXTAS, OMIM #300623) or premature ovarian failure (FXPOI, OMIM #311360).

Testing for FXS and similar trinucleotide repeat disorders (e.g., myotonic dystrophy, Huntington disease, spinocerebellar ataxias, spinal and bulbar muscular atrophy, and Friedreich ataxia) is typically performed using Southern blot *(5)* or, more recently, by PCR and capillary electrophoresis *(6, 7)*. The large size and low complexity of the pathogenic alleles are not amenable to typical next-generation sequencing (NGS) approaches (e.g., Illumina sequencing with a typical read length under 200 bp). Long-read NGS approaches have been described *(8)* but are substantially more expensive than the PCR fragment sizing methods. Thus, although NGS has markedly reduced the cost of many genetic tests *(9)*, FXS testing has remained static in cost since the introduction of PCR-based methods.

Here we propose a multiplexed method for determining the presence and size of abnormal microsatellite

[1] Counsyl, South San Francisco, CA.
* Address correspondence to this author at: Counsyl, 180 Kimball Way, South San Francisco, CA, 94080. E-mail research@counsyl.com.
Received April 27, 2016; accepted July 22, 2016.
Previously published online at DOI: 10.1373/clinchem.2016.259796

[2] Nonstandard abbreviations: FXS, fragile X syndrome; OMIM, Online Mendelian Inheritance in Man; UTR, untranslated region; SC, StairCase; STD, Shifted Transversal Design; NGS, next-generation sequencing; NBS, newborn screening.

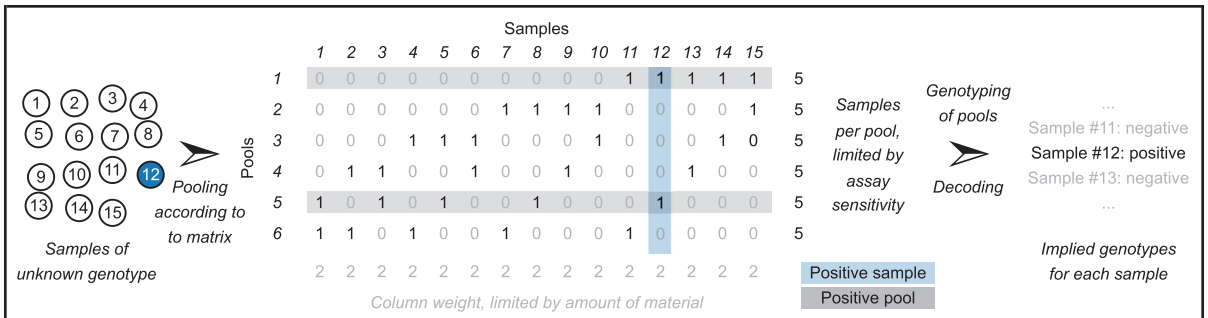

Fig. 1. Overview of the pooling and decoding processes using SC (R_{max} = 5) with sample 12 being positive.

A set of samples, mostly negative for a particular rare genotype, were pooled according to a pooling scheme; a genotyping assay was performed on each pool; the sample overlap between pools was used to decode the results and to report the implied genotype of each sample.

expansions in the context of FXS and related conditions, potentially enabling carrier screening at scales comparable to those enabled by NGS or affordable testing of every newborn for the CGG repeat expansion.

Methods

GROUP TESTING

Our proposed method was designed to take advantage of the rarity of abnormal expansions and the analytical sensitivity of PCR-based sizing *(6, 7)*. These factors allowed for multiplexed measurements using a strategy generally called group testing *(10–13)*, in which a small number of measurements of combined sets of samples (pools) were integrated to identify positive individuals.

A pooling scheme describes how to carry out a group testing experiment. In adaptive schemes, samples are first pooled in nonoverlapping (distinct) groups. Next, samples in positive pools are retested either individually or using the adaptive approach again recursively, and samples in nonpositive pools do not have to be investigated further. Nonadaptive schemes combine samples into partially overlapping pools such that each positive sample creates a known pattern of positive pools. These patterns of pools are decoded to identify the positive samples with fewer tests than when testing each sample individually.

At the core of a pooling scheme was an indicator matrix describing the assignment of N samples (columns) to T pools (rows) (Fig. 1; also see Material File 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol62/issue10 for more examples of different schemes). The column weight was the number of pools a sample was present in, and the row weight R_i was the number of samples in a pool i. A pooling scheme had decoding capability d if 1 round of pooling could determine the identity of any $p \leq d$ positive samples. If, by chance, there were more than d positive samples in the batch of N samples (a "collision"), decoding could fail and additional testing of a small number of "ambiguous" samples might be required.

The fold savings in assay costs enabled by a pooling scheme was evaluated using its amortized samples-to-tests ratio, which quantifies the mean reduction in necessary assays relative to individual testing, and is at most N/T if there are no collisions.

PRACTICAL CONSTRAINTS ON POOLING SCHEMES

In the FXS testing context, each size of allele could generally be distinguished from others by capillary electrophoresis, so collisions would occur when alleles of the same size occurred in more than d specimens in a batch. Although bigger batches generally enabled larger maximal cost savings, they also made collisions more likely.

The design of a pooling scheme must take into account practical limitations. The column weight was constrained by the quantity of input material available, because 1 sample needed to be present in this many pools. Although this constraint has been extensively studied [e.g., *(12)*], only a small amount of DNA was required for the assay used here (2 ng/sample per pool) compared to the availability of DNA from clinical samples [hundreds of nanograms from dried blood spots *(14)*], and thus the column weight was not a true limiting factor. The row weight was constrained by the analytical sensitivity of the assay. We defined R_{max} as the largest number of samples in a pool such that the signal of a single positive sample could be identified reliably. To fully utilize resources and not bias samples, we would like $R_i = R_{max}$ for all pools i.

THE STAIRCASE POOLING SCHEME

We devised a $d = 1$ nonadaptive scheme called StairCase (SC) for which each pool was fully used up to a given R_{max} limit and each sample was present in exactly 2 pools. The theoretical cost savings ratio of this scheme, achieved

when there were 0 or 1 positives in a batch, was $R_{max}/2$, which was the best theoretical ratio for a $d = 1$ matrix in which no 2 samples shared more than 1 pool (see online Supplemental Material file 2). The only parameter for creating a SC scheme was the R_{max} value, resulting in a $(R_{max} + 1)$-by-$R_{max}(R_{max} + 1)/2$ matrix. The SC(R_{max}) matrix was based on SC($R_{max} - 1$). To this, a new pool with R_{max} new samples was added (no overlap with other pools). Finally, each of the new samples were added to each of the other R_{max} pools. The base case was $R_{max} = 1$, in which 2 pools both contained the single sample. An example of the scheme with $R_{max} = 5$ is shown in Fig. 1, in which the staircase-like recursive structure is visible. Python code to generate SC(R_{max}) is given in online Supplemental Material File 3.

EVALUATING COST REDUCTION ENABLED BY POOLING

To evaluate the performance of various pooling schemes, we performed simulations by analyzing randomized batches of samples; the amortized cost savings was calculated by dividing the batch size by the mean number of tests required for complete decoding.

The prevalence of CGG expansion sizes was estimated from a previously published dataset *(2)*. A single count was added to each bin in the histogram of CGG expansions (45–150) to account for undersampling. Processing of at least 5 million samples in total were simulated per scheme variant, with a random mix of males and females.

The Shifted Transversal Design (STD) *(10)* scheme, the SC scheme (this work), and a recursive adaptive scheme were simulated. When a collision occurred, ambiguous samples were simulated as being retested individually, except in the recursive adaptive scheme. Batch sizes that were optimal for the schemes under consideration or prevalent in practice (96 and 384 samples) were simulated. To create SC schemes for suboptimal batch sizes, a smaller optimal scheme was used on nonoverlapping subsets of the batch.

PCR FOR *FMR1* CGG REPEAT COUNT DETERMINATION

PCR followed by capillary electrophoresis for detecting and sizing expansions in the *FMR1* 5′ UTR, as recommended by the American College of Medical Genetics and Genomics *(1)*, has been previously described *(15)*. In this study, we performed pooled PCR amplification on samples from the Coriell Institute biorepositories using previously described unlabeled forward and FAM (fluorescein amidite)-labeled reverse primers *(15)*. Each PCR was performed in a volume of 15 μL with a total of 40 ng input genomic DNA. To increase analytical sensitivity, PCR was run with extra cycles and the following thermocycling protocol: 5 min at 95 °C; 31 cycles of 35 s at 97 °C, 35 s at 62 °C, 4 min at 68 °C; 10 min at 72 °C. An aliquot of the completed PCR (diluted 7.5×) was added to highly deionized formamide (Applied Biosystems), along with the MapMaker 1000 ROX-labeled DNA size ladder (BioVentures). The mix was denatured for 2 min at 95 °C. Capillary electrophoresis was run on an Applied Biosystems 3730xL Genetic Analyzer (50 cm capillary) with an injection voltage of 3500 V and an injection time of 20 s. The increased number of PCR cycles could cause the fluorescence measurements to saturate in the normal CGG ranges, as well as lead to bleedthrough into the fluorescence channel of the size ladder. Because this assay was meant for determining abnormal expansion sizes and the effects were limited to the normal size range, the bleedthrough and saturation could be ignored. The time axis returned by the capillary electrophoresis instrument was converted to base pair units by using a known-size DNA ladder mixed into each injection. We report the observed CGG repeat counts by semiautomatically identifying peaks in the fluorescence intensity trace (Fig. 2A). Briefly, the fluorescence intensity area was integrated over a small window (1 CGG repeat) for all assayed pools. If the maximum area of any pool was above 800 (indicating the presence of a large signal; arbitrary units) and the median of the window-normalized area was smaller than 0.078 (indicating a large signal-to-noise ratio), then all pools with a normalized area >0.2 in the window were called as positive for the CGG repeat size in the given window. The first 2 cutoffs were determined by maximizing the harmonic mean of precision and recall in an independent set of experiments, whereas the latter was determined heuristically (see online Supplemental Material File 4). Finally, a curator assessed the calls and made the final determinations.

POOLED TRINUCLEOTIDE REPEAT EXPANSION SIZE DETERMINATION

Three variants of the SC($R_{max} = 20$) scheme were experimentally tested. Each variant differed only by the order of samples in the pooling matrix, and hence the sample composition of each pool. Each scheme used 21 assays to process 210 samples, 10 of which had distinct abnormal CGG expansion sizes (45 or more), with the layouts detailed in online Supplemental Material File 5. In each variant, all pools contained DNA from 20 different samples with an average X chromosome count of 30 (approximately equal mix of males and females). For each variant, PCRs were set up in triplicate, with a single capillary electrophoresis run for each triplicate.

In the first variant of the scheme, each pool had at most 1 abnormal sample. In the second variant, 1 of the 2 pools for each abnormal sample was the same, whereas the other was unique, so that there were 10 positive samples in a single pool. The third variant exhibited pools with 0–4 abnormal samples. In batches of 210 samples one would expect to see up to 4 positives 80% of the time

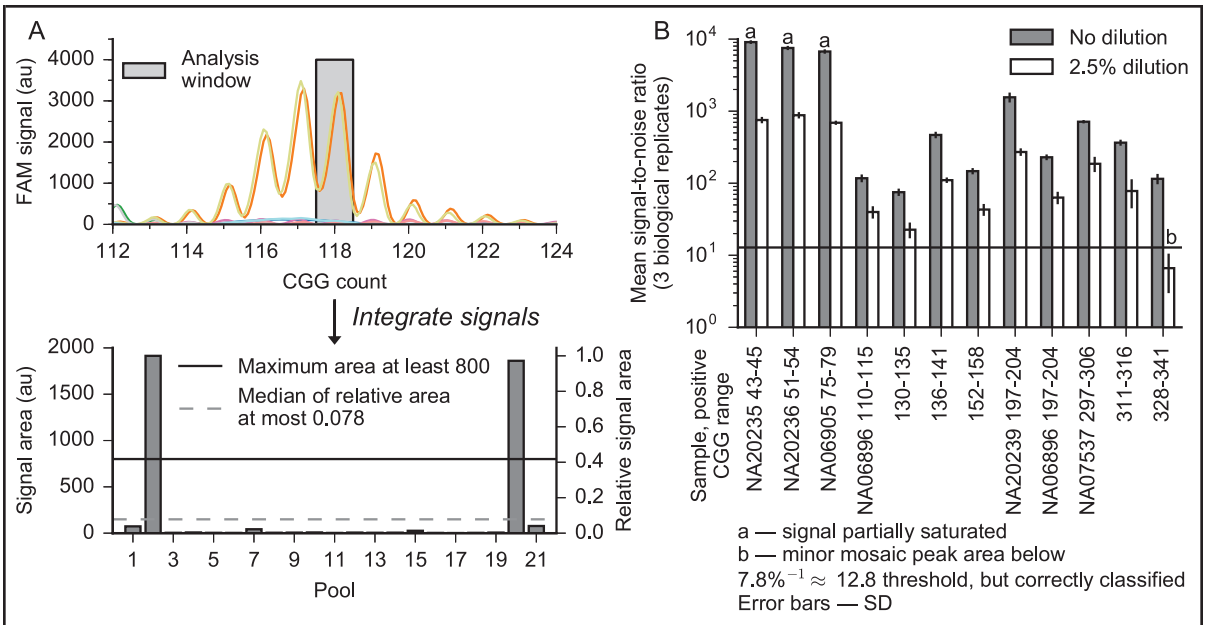

Fig. 2. Signal analysis.

(A), Signals in a window are assessed for strength (maximum area cutoff) and quality (median relative area cutoff). (B), Positive samples were assayed independently ("No dilution") or in a background of 39 normal alleles (2.5% dilution). The signal-to-noise ratio was calculated by dividing the median positive signal area by the median of the non-positive signal.

and more than 10 positives of any genotype less than once in 2000 batches; the probability of any 2 samples ending up in the same pool was about 9% in this scheme.

Each call was run through a decoding algorithm individually, marking only the identified pools as positive and the rest as negative for the purposes of decoding a particular CGG expansion. For SC, the decoding algorithm can be a simple lookup table, because 2 positive pools unambiguously identify a sample. If there are more than 2 positive pools, the results are ambiguous, and the *m*-choose-2 combinations of pairs of *m* positive pools identify a potential positive sample. For example, 2 true-positive samples can cause 4 pools to show up positive, causing 6 samples to be identified as having an ambiguous genotype, necessitating retests of those 6 samples.

Results

FX SYNDROME ASSAY ANALYTICAL SENSITIVITY

As previously reported, *FMR1*-specific PCR methods are highly analytically sensitive and can detect mosaicism down to 1% *(15)*. In concordance with these results, we were able to detect positive CGG repeats at the 2.5% dilution level in a background of normal samples (Fig. 2B). This dilution level represented a single X chromosome in a pool of 40 chromosomes, which was the worst-case example of an all-female 20-member pool, whereas a mean pool with 20 individuals was expected to contain only 30 X chromosomes, each 3.3% of the mixture. Although information about low-level mosaicism may be lost in a pooled setting because of the increased dilution, the major variants of the CGG expansion could be picked up with high confidence.

POOLING SCHEMES FOR FX SYNDROME

The frequency of abnormal *FMR1* alleles was used to compare 3 different pooling schemes in simulations (Fig. 3). Online Supplemental Material File 1 shows example pooling matrices for the 3 alternatives under consideration. The SC scheme described here fully used all pools using a recursive pattern of pooling. The STD used number theory to design a pooling layout, but was not always optimal. An adaptive scheme pools samples without overlap and recursively tested any positive pools by splitting them in half. The first 2 schemes could detect the identity of up to 1 "positive" per batch in 1 iteration, whereas the adaptive approach required several rounds of pooling.

In the assay described here, each distinct CGG expansion size could be considered as a "positive" on its own for the purposes of decoding the pooled experiment. Hence, although the carrier rate of having an abnormal CGG expansion was too high (about 1.5%) *(2)* to enable large cost savings due to frequent collisions," each partic-

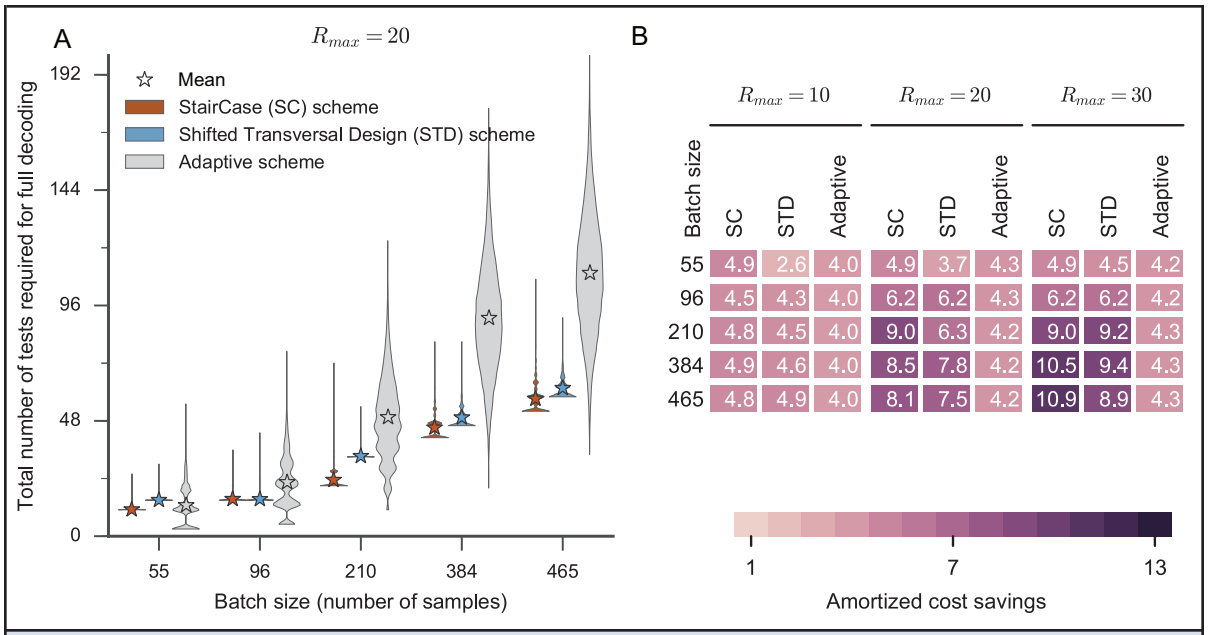

Fig. 3. Simulations showing a potential 10-fold improvement with SC.

Genotypes were drawn at random according to expected allele frequencies. (A), Smoothed violin plots showing the number of tests required. (B), The amortized cost savings, calculated by dividing the batch size by the mean number of tests required to identify the genotype of all input samples.

ular expansion size had a frequency of <0.2%, making pooling cost-effective with a $d = 1$ scheme.

In general we found that an adaptive approach ($d = 0$) had poor performance because it required retesting of pools if any nonnormal allele was observed. The nonadaptive schemes (STD and SC, both with $d = 1$) did not require samples to be retested very often because it was relatively unlikely that 2 alleles of the same size would be found in a batch. Even when there was such a collision, a limited number of samples had to be retested (e.g., at most 6 samples in the SC designs), compared to $\geq R_{max}$ samples in the adaptive scheme. Although the STD design was seemingly the most flexible in terms of freedom to set up the scheme with different parameters (batch size, d, and R_{max}), it was nearly always outperformed by the SC scheme described here, which achieved amortized cost savings of 9.0 with 210 samples ($R_{max} = 20$) compared to 6.3 in the same conditions for STD (Fig. 3; also see online Supplemental Material File 6).

Although a larger batch size and increasing R_{max} allowed for better theoretical cost savings ratios [e.g., 10 for SC($R_{max} = 20$) with 210 samples compared to 15 for SC($R_{max} = 30$) with 465 samples], the chance of collision also increased and the expected cost savings ratios decreased to 9.0 and 10.9 for the 2 examples (Fig. 3B). The SC($R_{max} = 20$) scheme with a batch size of 210 was thus chosen for experimental verification because it offered a good balance for performance and was compatible with the analytical sensitivity of the assay.

POOLED FX SYNDROME ASSAY

Three variants of the 210-sample pooling SC($R_{max} = 20$) scheme were tested, differing by the assignment of samples to pools (see online Supplemental Material File 5) and thus the number of positive samples that ended up in the same pool (ranging from 0–10).

In every pooling variant and PCR triplicate we were able to correctly determine the diagnostic classification (normal, intermediate, premutation, full mutation) and in each case determine the major expansion size, defined as the largest peak when assayed individually (Table 1). Fig. 4 shows the calls (CGG repeat size ranges) across all experiments and online Supplemental Material File 7 shows examples of the obtained signals and decoding process. The analytical sensitivity of detecting positive expansions was 0.984 (95% CI, 0.944–0.998) and analytical specificity 1.000 (95% CI, 0.998–1.000; Table 1).

In 2 cases across 3 layouts and their 3 replicates (126 possible allele observations), a minor expansion size of a mosaic sample was observed, but not called because of not meeting one of the thresholds for calling. Online Supplemental Material File 8 discusses the data in question, showing how the maximum signal area was slightly

Table 1. Confusion matrix across 3 different pooling layouts, with each PCR performed in triplicate. Layouts differed in assignment of samples to pools.[a]

	Actual			
Sample classifications	Intermediate	Premutation	Full mutation	Normal
Identified				
Intermediate	18	0	0	0
Premutation	0	54	0	0
Full mutation	0	0	18	0
Normal	0	0	0	1800

	Actual	
Genotype calls	Positive	Negative
Identified		
Positive	124	0
Negative	2	1800+

[a] A genotype was correct if the known and called genotypes overlapped. Pooling does not give information on the number of X chromosomes, so the true negative count is underestimated. The 2 false negatives are discussed in online Supplemental Material File 8.

below the threshold, whereas the median signal area met the threshold for calling. The major genotype of this mosaic was called in each iteration and triplicate and the sample was always correctly classified.

Mosaicism in cell line samples with expansions such as those used here is very common, but is also observed in patient samples and can vary across tissues *(16)*. Mosaicism is clinically relevant, as the presence of smaller expansions in full mutation individuals can slightly rescue the phenotype *(16)*. If any indication of mosaicism exists (presence of high quality but low-amplitude peaks below the calling threshold as observed here), samples should be retested individually.

Discussion

ASSAY SENSITIVITY

Prior reports differ in their estimates of the analytical sensitivity of PCR to detect large *FMR1* expansions; some report sensitivity down to 1% mosaicism *(15)*, others have reported difficulty in amplifying large alleles *(17–18)* or amplifying premutation alleles in a background of full mutations *(19)*. In this work we adapted our PCR protocol to the desired application of detecting abnormal repeat expansions. By adding additional cycles of PCR [31 cycles here vs 25 in *(15)*], we sacrifice performance on small (normal) alleles, which are amplified outside the linear range of the assay, while improving amplification of large alleles.

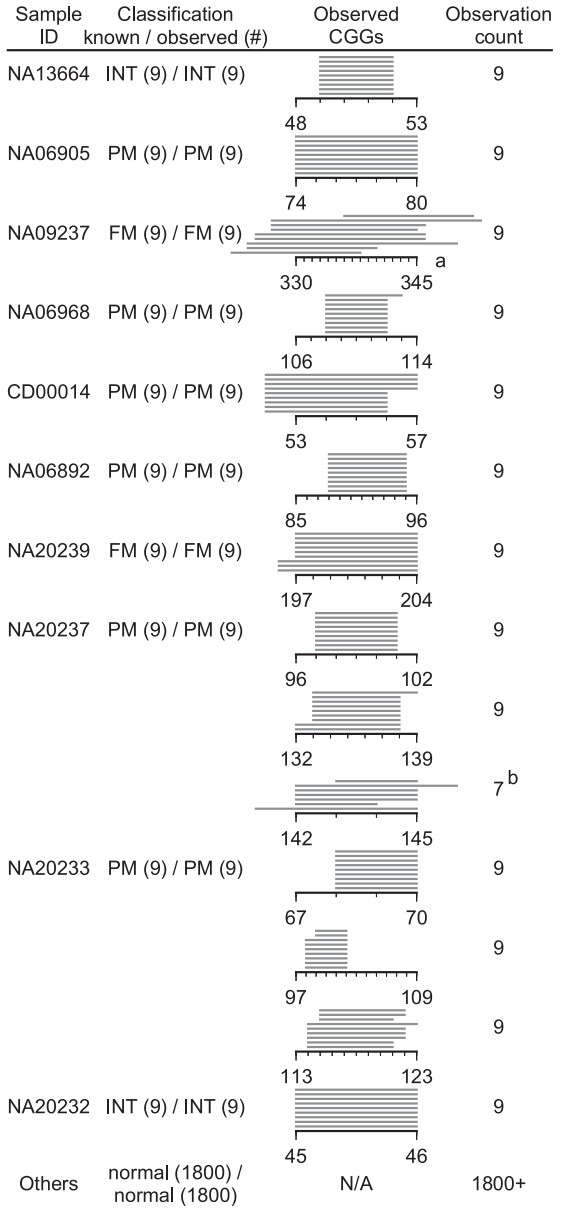

Fig. 4. Summary of all decodings indicates consistent and accurate calling in a pooled setting.

Stacked lines indicate the width of each peak determined to be associated with the given sample over 3 assay replicates of 3 pooling scheme iterations, with the true CGG range given by the axis limits. [a], outside the linear range of electrophoresis measurement, [b], 2 false negatives.

RELATION TO PREVIOUS POOLING SCHEMES

The row weight (R_{max}) limitation rules out highly compressed schemes such as those in which the number of required pools scales as the logarithm of the number of

samples *(11)*. Less compressed schemes such as those with T being approximately $\sqrt{N}$ [e.g., *(13)*] are suitable, but are not balanced and can thus be improved. Few studies have explicitly considered the analytical sensitivity of assays as a limiting factor in pooling design and have not provided desirable cost reduction. For example, Kainkaryam *(20)* proposed a scheme in which $R_i \leq 10$, but achieved a cost-savings ratio of only 2.5, whereas a ratio of 5 could be achieved with SC under the same constraint.

CLINICAL APPLICATIONS

Our method, by reducing analytical costs nearly 10-fold, may make newborn screening (NBS) for FX syndrome feasible. Such ubiquitous testing would reduce the time-to-diagnosis from years to weeks, and enable early intervention services to be used as soon as possible *(21)*. Early intervention can mitigate the symptoms of FXS by improving development through various therapies *(22)*. In addition, parents of affected children have more time to find support and prepare for raising a child with FXS, and can make informed reproductive planning decisions going forward *(21)*.

Carrier screening requires high multiplexing to be able to report results to a large number of patients with a fast turnaround time. This multiplexing has been enabled by NGS techniques for most conditions tested in (expanded) carrier screening. Our method brings carrier testing for abnormal *FMR1* alleles on par with these conditions with a 10-fold improvement over current approaches.

Coffee et al. described an adaptively pooled methylation-sensitive PCR assay suitable for determining *FMR1* hypermethylation in males *(23)*. The method did not indicate the size of the CGG expansion, which has clinical value and is of interest in carrier screening *(1, 23)*. Furthermore, analytical sensitivity in female samples dropped significantly because of normal methylation of the inactivated X chromosome. Compared to this approach, our PCR-based assay is applicable to males as well as females, can determine the size of the CGG expansion and can detect mosaicism in many cases.

A limitation of our approach is that, in females, our method is unable to distinguish between high levels of mosaicism because of 1 expanded copy and the extremely rare case of both X chromosomes having an abnormal *FMR1* expansion. Additionally, all experiments were performed on DNA from cell lines. Before clinical implementation, other sample types must be validated, potentially including blood, saliva, or buccal swabs for carrier screening and dried blood spots for NBS.

APPLICATIONS TO OTHER DISEASES

Other microsatellite-expansion diseases are candidates for our approach as well, including myotonic dystrophy and Huntington disease. For a successful adoption of our approach to such diseases, the abnormal genotypes should be sufficiently rare, with FXS being comparatively more common.

PCR-based assays have remained the standard for measuring microsatellite expansion as the expansions in these diseases are generally much longer than the short reads produced by NGS, and single-molecule long-read sequencing approaches are too expensive and low throughput *(8)*. Because of this, cost reductions have lagged for microsatellite expansion testing, and an order of magnitude decrease as described here makes a substantial step toward catching up with genetic tests accessible by short read sequencing.

Conclusion

In conclusion, the SC approach to pooled PCR testing for the *FMR1* 5′ UTR repeat expansion can reduce the assay cost for FX testing by over 10-fold, enabling potential applications in population carrier or NBS. Widespread carrier screening and NBS for *FMR1* CGG repeat expansions could reduce the diagnostic odyssey faced by many parents of affected children. The general approach of pooling samples, performing PCR, resolving the amplicon sizes using capillary electrophoresis, followed by mathematical decoding of the pooled results could also be applied to other hard to sequence trinucleotide expansion alleles in several other severe diseases.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: K.E. Kaseniit, Counsyl; M.R. Theilmann, Counsyl; A. Robertson, Counsyl; E.A. Evans, Counsyl; I.S. Haque, Counsyl.
Consultant or Advisory Role: None declared.
Stock Ownership: K.E. Kaseniit, Counsyl; M.R. Theilmann, Counsyl; A. Robertson, Counsyl; E.A. Evans, Counsyl; I.S. Haque, Counsyl.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: K.E. Kaseniit, serial no. 62/323,441; M.R. Theilmann, serial no. 62/323,441; A. Robertson, serial no. 62/323,441; E.A. Evans, serial no. 62/323,441; I.S. Haque, serial no. 62/323,441.

Role of Sponsor: No sponsor was declared.

References

1. Monaghan KG, Lyon E, Spector EB. 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 Genomics. Gene Med 2013;15:575–86.
2. Tassone F, Iong KP, Tong TH, Lo J, Gane LW, Berry-Kravis E, et al. FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. Genome Med 2012;4:100.
3. Bailey DB, Raspa M, Bishop E, Holiday D. No change in the age of diagnosis for fragile X syndrome: findings from a national parent survey. Pediatrics 2009;124: 527–33.
4. Saul RA, Tarleton JC. FMR1-related disorders. In: Pagon RA, Adam MP, Ardinger HH, et al, editors. GeneReviews. Seattle (WA): University of Washington; 2016.
5. Margolis RL. Diagnosis of Huntington disease. Clin Chem 2003;49:1726–32.
6. Tassone F, Pan R, Amiri K, Taylor AK, Hagerman PJ. A rapid polymerase chain reaction-based screening method for identification of all expanded alleles of the fragile X (FMR1) gene in newborn and high-risk populations. J Mol Diagn 2008;10:43–9.
7. Chen L, Hadd A, Sah S, Filipovic-Sadic S, Krosting J, Sekinger E, et al. An information-rich CGG repeat primed PCR that detects the full range of fragile X expanded alleles and minimizes the need for southern blot analysis. J Mol Diagn 2010;12:589–600.
8. Loomis EW, Eid JS, Peluso P, Yin J, Hickey L, Rank D, McCalmon S, et al. Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. Genome Res 2013;23:121–8.
9. Azimi M, Schmaus K, Greger V, Neitzel D, Rochelle R, Dinh T. Carrier screening by next-generation sequencing: health benefits and cost effectiveness. Mol Genet Genomic Med 2016;4:292–302.
10. Thierry-Mieg N. A new pooling strategy for high-throughput screening: the Shifted Transversal Design. BMC Bioinformatics 2006;7:28.
11. Eppstein D, Goodrich MT, Hirschberg DS. Improved combinatorial group testing algorithms for real-world problem sizes. SIAM J Comput 2007;36:1360–75.
12. Erlich Y, Chang K, Gordon A, Ronen R, Navon O, Rooks M, Hannon GJ. DNA Sudoku–harnessing high-throughput sequencing for multiplexed specimen analysis. Genome Res 2009;19:1243–53.
13. Erlich Y, Gordon A, Brand M, Hannon GJ, Mitra PP. Compressed genotyping. IEEE Trans Inf Theory 2010; 56:706–23.
14. Saavedra-Matiz CA, Isabelle JT, Biski CK, Duva SJ, Sweeney ML, Parker AL, et al. Cost-effective and scalable DNA extraction method from dried blood spots. Clin Chem 2013;59:1045–51.
15. Filipovic-Sadic S, Sah S, Chen L, Krosting J, Sekinger E, Zhang W, et al. A novel FMR1 PCR method for the routine detection of low abundance expanded alleles and full mutations in fragile X syndrome. Clin Chem 2010; 56:399–408.
16. Pretto D, Yrigollen CM, Tang HT, Williamson J, Espinal G, Iwahashi CK, et al. Clinical and molecular implications of mosaicism in FMR1 full mutations. Front Genet 2014;5:318.
17. Esposito G, Ruggiero R, Savarese G, Savarese M, Tremolaterra MR, Salvatore F, Carsana A. A 15-year case- mix experience for fragile X syndrome molecular diagnosis and comparison between conventional and alternative techniques leading to a novel diagnostic procedure. Clin Chim Acta 2013;417:85–9.
18. Aliaga SM, Slater HR, Francis D, Du Sart D, Li X, Amor DJ, et al. Identification of males with cryptic fragile X alleles by methylation-specific quantitative melt analysis. Clin Chem 2016;62:343–52.
19. Seneca S, Lissens W, Endels K, Caljon B, Bonduelle M, Keymolen K, et al. Reliable and sensitive detection of fragile X (expanded) alleles in clinical prenatal DNA samples with a fast turnaround time. J Mol Diagn 2012;14:560–8.
20. Kainkaryam RM. Pooling designs for high-throughput biological experiments [Ph.D. thesis]. Ann Arbor: University of Michigan; 2010. p 34–6.
21. Tassone F. Newborn screening for fragile X syndrome. JAMA Neurol 2014;71:355–9.
22. Christie L, Wotton T, Bennetts B, Wiley V, Wilcken B, Rogers C, et al. Maternal attitudes to newborn screening for fragile X syndrome. Am J Med Genet A 2013; 161A:301–11.
23. Coffee B, Keith K, Albizua I, Malone T, Mowrey J, Sherman SL, Warren ST. Incidence of fragile X syndrome by newborn screening for methylated FMR1 DNA. Am J Hum Genet 2009;85:503–14.