Amplification-Based Nanopore Sequencing Accurately Detects *HBA* and *HBB* SNVs, Indels, and Structural Variants in Clinical Thalassemia Samples

P585

Jaime Lopes*, Marcus Cannon*, Jonathan Turner†, Monica Roberts†, Theodore Markulin†, Chris Fraher†, Bryan Killinger†, Stela Filipovic-Sadic†, and Bradley Hall†

*Cincinnati Children's Hospital Medical Center, Cincinnati OH; †Asuragen, a Bio-Techne Brand, Austin TX

SUMMARY

- Thalassemia screening currently involves a complicated workflow of MLPA, NGS, and other methods.
- The prototype PCR-nanopore assay generates accurate and comprehensive SNV/Indel and SV results for both *HBA1/2* and *HBB* across screening and affected populations in a single workflow without reflex testing.
- The algorithm correctly identified new variant combinations not previously trained on or observed during development.
- PCR-amplified and barcoded products generated with this panel can be pooled with content from other panels targeting additional carrier screening-relevant genes that are inaccessible by conventional short-read sequencing or which typically require multiple assays and reflex testing.

INTRODUCTION

Hemoglobinopathies, including alpha and beta thalassemia resulting from abnormal hemoglobin subunit structures or levels, are among the most common inherited genetic diseases. At least 5% of the global population carries a pathogenic variant. Deletions affecting the hemoglobin α subunit genes (*HBA1* and *HBA2*) cause 90% of α thalassemia, whereas *HBB* SNV/Indel variants account for 95% of β thalassemia. Multiple assays are typically required to evaluate these different variant types due to highly homologous sequences and complex recombination events that are challenging to resolve using short-read sequencing alone. We designed a single-tube assay and analysis solution to accurately resolve pathogenic SNVs, Indels, and SVs across *HBA1*, *HBA2*, and *HBB* utilizing scalable PCR enrichment with "any-length" nanopore sequencing. The prototype assay and software accurately detected pathogenic variants across all genes in a cohort of cell line, presumed normal blood, and residual clinical thalassemia samples.

METHODS

Genomic DNA from 446 samples, including 144 cell line, 247 presumed normal blood, and 55 clinical thalassemia specimens (46 blood, 5 buccal, 1 saliva, 3 unspecified source) was enriched for *HBA* and *HBB* regions in a single PCR reaction, barcoded, pooled, and prepared using a V14 ligation sequencing kit (Oxford Nanopore Technologies; ONT). Sequencing was performed using R10.4.1 flow cells and a Mk1B (ONT). Bespoke software and bioinformatic algorithms were developed to manage projects, initiate sequencing, automate data analysis, and generate reports. An SV genotyping algorithm was trained on data from 45 samples (28 cell line, 17 blood). Assay genotyping performance was evaluated for 401 samples (116 cell line, 230 presumed normal blood, 55 clinical thalassemia samples). GAP-PCR[1,2] with agarose gel electrophoresis (AGE) and Sanger sequencing were used to confirm SVs and SNV/Indels respectively for select samples. Results were compared with previous annotations obtained via NGS and MLPA when available. One cell line and some clinical samples were excluded from SV genotyping accuracy analyses either because orthogonal genotype results were unavailable or because they could not be resolved due to the assay design, which enables differentiation of the seven most common *HBA* SVs (3.7 kb, 4.2 kb, FIL, MED-I, SEA, THAI, alpha20.5) as well as the anti-3.7 and 4.2 duplications.

RESULTS

Table 1. *HBA* SV genotype identification is 99.1% accurate in cell line samples. PCR-Nanopore genotype predictions were compared with agarose gel electrophoresis of GAP-PCR products (SV identity) and/or PCR melt curve assay (*HBA1/2* copy number) in 116 cell line samples after algorithm training. Comparator method results were confirmed by PCR-Nanopore results for 115/116 (99.1% agreement) samples.

† The discrepant sample was observed as a heterozygous THAI deletion by PCR-Nanopore (Figure 3b) and melt curve analysis[3] suggested 1 copy each of *HBA1* and *HBA2* typical of an α^0 deletion. Corresponding GAP-PCR product with THAI specific primers was not observed on AGE though. Subsequent investigation suggested this sample (HG03862) is THAI-like with a large deletion affecting much of the same area as a THAI deletion (including both *HBA1* and *HBA2*) but with different breakpoints.

Cell line sample PCR-Nanopore assay *HBA* SV results		
Genotype	Number of cell line samples	Confirmed by reference method(s)
αα/αα	91	91/91 = 100%
$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$	3	3/3 = 100%
$-\alpha^{3.7}/\alpha\alpha$	9	9/9 = 100%
$-\alpha^{4.2}/\alpha\alpha$	2	2/2 = 100%
$-\alpha^{3.7}/-\alpha^{3.7}$	3	3/3 = 100%
$\alpha^{0(SEA)}/\alpha\alpha$	5	5/5 = 100%
$\alpha^{0(THAI)}/\alpha\alpha$	1	0/1 = 0%†
$\alpha^{0(SEA)}/\alpha\alpha\alpha^{anti3.7}$	1	1/1 = 100%
$\alpha^{0(FIL)}/\alpha^{0(SEA)}$	1	1/1 = 100%

Table 2. *HBA* SV genotype identification is 99.6% accurate in presumed normal blood samples. GAP-PCR AGE and/or melt curve assays were used to evaluate PCR-Nanopore genotype call accuracy in a screen of 230 presumed normal whole blood samples. Overall agreement was 229/230 = 99.6%.

† The single discordant call observed (anti-3.7) resulted from low coverage in one of the three endogenous control amplicons used for normalization rather than a true false positive. QC features will be incorporated into the analysis software to identify and flag this failure mode.

Presumed normal blood	PCR-Nanopore assay *HBA* SV result			
Sample Level Accuracy	Allele1/ Allele2	αα/αα	$-\alpha^{3.7}/\alpha\alpha$	$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$
Reference method result	αα/αα	220		1†
	$-\alpha^{3.7}/\alpha\alpha$		3	
	$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$			6

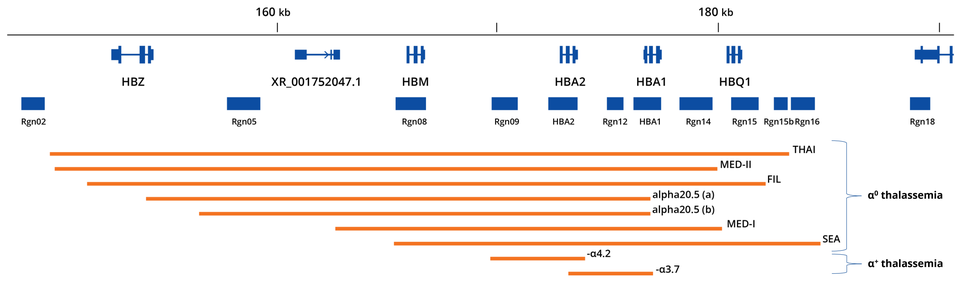


Figure 1. PCR-Nanopore assay design with the alpha cluster. The scale at the top indicates position on chromosome 16 (hg38). Genes are shown in blue on the second row. Blue rectangles in the third row indicate the positions of the *HBA* amplicons included in the PCR-Nanopore assay. Rgn01 (upstream of Rgn02) and Rgn19 (downstream of Rgn18) are not shown. Orange bars at the bottom indicate deleted portions of large SVs the assay can differentiate based on coverage pattern changes. The alpha20.5 breakpoints from two different references are included.

Table 3. *HBA* SV genotype identification performance in clinical thalassemia samples. PCR-Nanopore genotype results agreed with clinical testing results for 44/45 samples (97.8%). Within the $\alpha^+/\alpha\alpha$, $\alpha^0/\alpha\alpha$, and α^+/α^+ categories, the PCR-Nanopore assay correctly differentiated each anticipated SV, including $-\alpha^{3.7}$, $-\alpha^{4.2}$, $\alpha^{0(SEA)}$, and $\alpha^{0(FIL)}$, by name.

Clinical samples	PCR-Nanopore assay *HBA* SV result							
Sample Level Accuracy	Allele1/ Allele2	αα/αα	$\alpha^+/\alpha\alpha$	$\alpha^0/\alpha\alpha$	α^+/α^+	$\alpha^{0(SEA)}/-\alpha^{3.7}$	$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$	$\alpha\alpha\alpha^{anti4.2}/-\alpha^{3.7}$
Clinical *HBA* SV Result	αα/αα	19						
	$\alpha^+/\alpha\alpha$		10					
	$\alpha^0/\alpha\alpha$			8				
	α^+/α^+				4			
	$\alpha^{0(SEA)}/-\alpha^{3.7}$					1		
	$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$						2	
	$\alpha\alpha\alpha^{anti4.2}/-\alpha^{3.7}$	1						

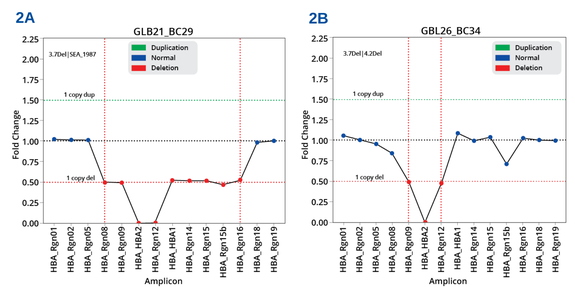
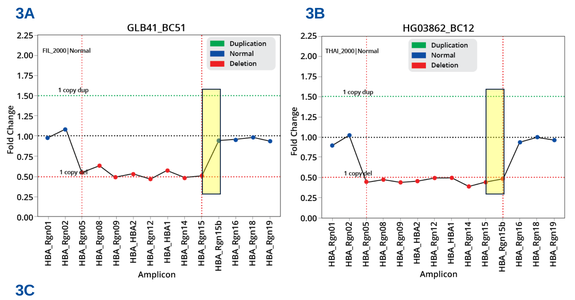


Figure 2. PCR-Nanopore correctly identifies genotypes of clinical samples with structural variant combinations not previously observed together. The assay was designed to identify these complex variants and the algorithm successfully resolved them though it had not previously been trained on samples. **A)** A 50% reduction in coverage of all amplicons from Rgn08 through Rgn16 relative to wild-type levels indicates an SEA deletion on one allele; further reduction of *HBA2* and Rgn12 coverage to near zero indicates a 3.7 kb deletion on the other allele. With one copy of *HBA1* and no copies of *HBA2*, this sample is affected by HbH disease. The genotype was called correctly by the algorithm. **B)** A 4.2 kb deletion affects Rgn09 and *HBA2* coverage, and the 3.7 kb deletion affects Rgn12 and *HBA2* coverage. This was observed and genotyped correctly by the algorithm.



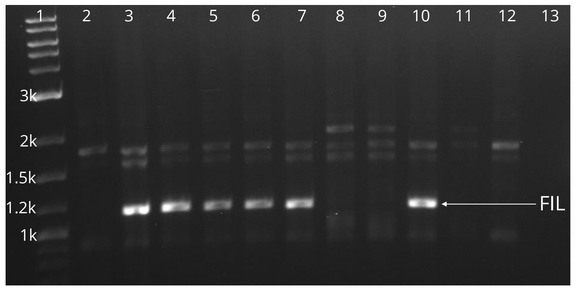


Figure 3. PCR-Nanopore assay distinguishes between THAI and FIL deletions. A) Representative algorithm data from one of six clinical samples that could be either FIL or THAI by MLPA orthogonal method was distinguished as a FIL deletion by the PCR-Nanopore assay. **B)** Representative result showing coverage pattern for a presumed heterozygous THAI-like sample. **C)** Comparator assay (GAP-PCR on AGE) confirmed PCR-Nanopore "FIL" call in all clinical samples (five samples represented in lanes 3-7; the sixth $\alpha^{0(FIL)}/\alpha\alpha$ sample was run on a separate gel and is not pictured here) compared to a cell line sample with heterozygous FIL deletion (lane 10). The FIL product is absent in samples with other *HBA* SVs (lanes 8, 9, 11, 12), a sample without any *HBA* SV PCR products (lane 2), and the NTC (lane 13).

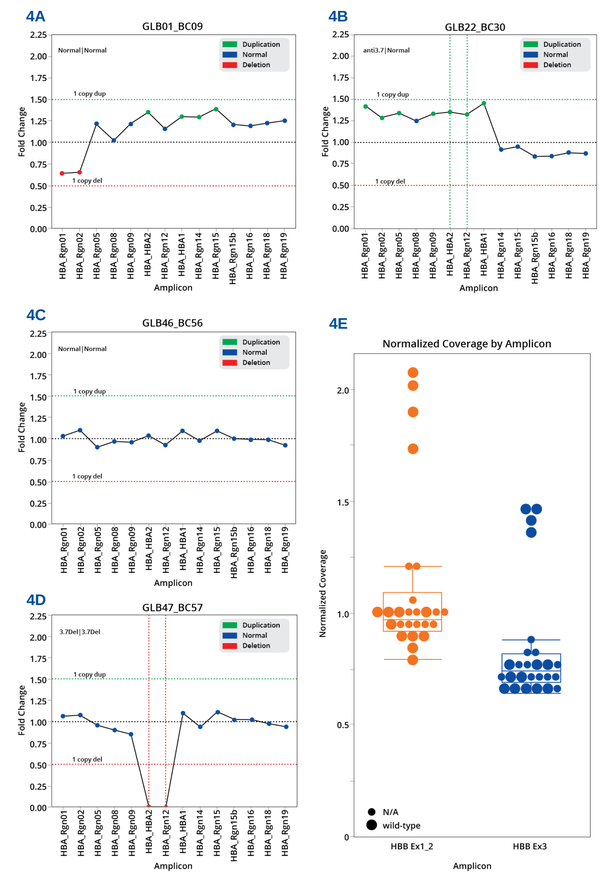


Figure 4. PCR-Nanopore results may be capable of flagging additional rare *HBA* SVs. A) Clinical MLPA results indicated a heterozygous HS-40 deletion in this sample. PCR-Nanopore assay reports a genotype of αα/αα. However, fold-change values are consistent with a heterozygous deletion affecting the alpha cluster Rgn01 and Rgn02 amplicons, overlapping typical HS-40 breakpoints. Amplification calls for the *HBA2*, *HBA1*, Rgn14, and Rgn15 amplicons are the result of reduced coverage of Rgn02, which is used for normalization, and are not associated with another SV. Future algorithm updates will enable flagging of samples with similar results to recommend further testing for HS-40. **B)** Clinical results for this sample indicated a large heterozygous duplication that does not match any previously reported *HBA1/2* SV. PCR-Nanopore results indicate duplication of several alpha cluster amplicons, including *HBA1* and *HBA2*, consistent with a large duplication of non-standard breakpoints. The genotype call of $\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$ instead of QC flagging will be corrected in future algorithm updates. **C)** Clinical results indicated a large deletion affecting the entire alpha cluster. PCR-Nanopore assay called an αα/αα. This is due to the algorithm's within-sample coverage normalization strategy. A future version of the algorithm will incorporate *HBB* read depth to identify and flag these samples. **D)** Clinical results indicated this sample as compound heterozygous for a complete alpha deletion and a 3.7 kb deletion. Because the PCR-Nanopore genotyping algorithm cannot currently detect rare full alpha deletions, and no *HBA2* coverage is observed, the algorithm reports as homozygous $-\alpha^{3.7}$. Future versions of the software will flag full alpha deletions. **E)** PCR-Nanopore *HBB* amplicon normalized coverage for the samples shown in **C** and **D** are ~2-fold higher (consistent with 4 total copies) than those of samples known or presumed to have no *HBB* SVs (2 copies). This suggests the presence of a full alpha cluster deletion in these samples because *HBB* amplicon coverage is normalized to coverage of amplicons in the alpha cluster. Box and whisker plots show quartiles when the full alpha cluster deletion samples (top four data points for both amplicons) are excluded as outliers. All samples were tested in duplicate, and results from both technical replicates are included in this graph.

Table 4. PCR-Nanopore assay *HBB* amplicon coverage is consistent with clinical SV results. Manual analysis of PCR-Nanopore coverage was consistent with clinical *HBB* SV genotype results for 34/34 (100%) thalassemia samples. The PCR-Nanopore assay can distinguish between SVs that affect only exons 1 and 2 (e.g., Hb Lepore) and those that affect the entire *HBB* gene (e.g., HPFH-1) but cannot differentiate the SV present by name.

Clinical sample *HBB* SV results		
Genotype	Number of cell-line samples	PCR-Nanopore coverage consistent with clinical results
Wild-type	24	24/24 = 100%
Sicilian Δβ 0-Thal (13.4 kb)	2	2/2 = 100%
HPFH-1	1	1/1 = 100%
HPFH-2	2	2/2 = 100%
Hb Lepore	1	1/1 = 100%
Anti-Lepore	1	1/1 = 100%
Other heterozygous deletion of full *HBB* gene	4	4/4 = 100%

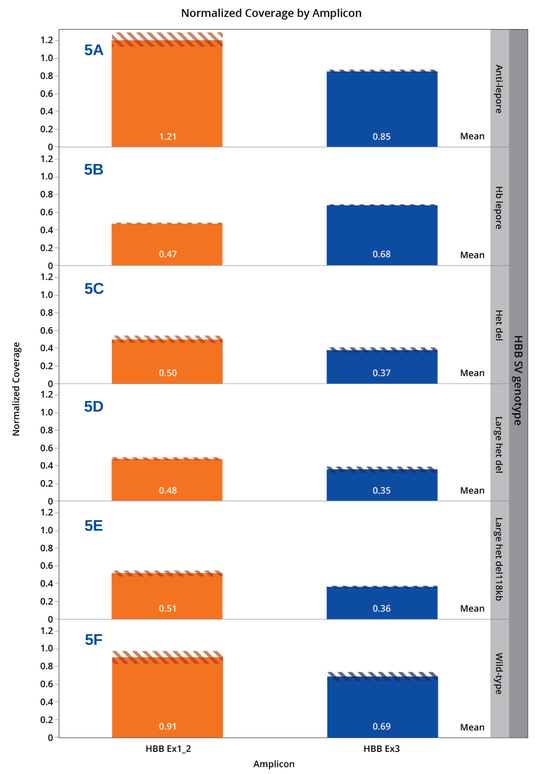


Figure 5. Normalized coverage of PCR-Nanopore *HBB* amplicons confirms the presence of SVs in clinical samples. Theoretically, heterozygous duplication would be expected to increase, and heterozygous deletion to decrease, normalized coverage of affected amplicons by 50%. **A)** *HBB* exon 1 and 2 coverage is ~33% higher in an anti-Lepore sample (duplication event associated with the Hb Lepore deletion). **B)** Exon 1 and 2 coverage is ~48% lower in an Hb Lepore sample. **C**), **D**), and **E)** Coverage of both amplicons is 43-49% lower in samples heterozygous for various deletions affecting the whole *HBB* gene. **F)** Average normalized coverage of *HBB* PCR-Nanopore amplicons in five samples unaffected by SVs. Hash bands at the tops of bars show standard deviation. Average normalized coverage values are shown at the bottom of each bar. All samples were tested in duplicate, and results from both technical replicates are included in this graph. The Het del and Large het del 118kb categories include one sample each; the large het del category includes two.

Table 5. PCR-Nanopore assay detects *HBA* and *HBB* SNV/INDELs with 100% accuracy. All variants detected in cell line and presumed normal blood samples were confirmed by Sanger sequencing. All variants reported in clinical sample results were identified by the PCR-Nanopore assay.

Gene	Sample type	Unique SNV/Indels Detected	Sample Agreement
HBA	Clinical thalassemia	c.427T>C (Hb Constant Spring), c.207C>G (Hb G-Philadelphia), c.237del, c.142G>C, c.95+2_95+6del	6/6 (100%)
	Cell line	c.427T>C (Hb Constant Spring), c.207C>G (Hb G-Philadelphia)	2/2 (100%)
HBB	Clinical thalassemia	c.20A>T (Hb S), c.27dup, c.-138C>T, c.17_18del, c.315+1G>A, c.19G>A (Hb C), c.364G>C	14/14 (100%)
	Cell line and presumed normal blood	c.19G>A (Hb C), c.79G>A (Hb E), c.20A>T (Hb S), c.126_129del, c.93-21G>A, c.-137C>G, c.-138C>T, c.316-197C>T, c.82G>T (Hb Knossos), c.-79A>G, c.404T>C, c.208G>A, c.216_217insA	16/16 (100%)

CONCLUSIONS

- PCR-Nanopore assay provides accurate *HBA* structural variant (SV) genotypes in 115/116 cell line (99.1% accuracy), 229/230 presumed normal blood (99.6%), and 44/45 clinical thalassemia samples (97.8%).
- *HBB* amplicon coverage accurately reflects the presence of SVs in 34 clinical thalassemia samples with 100% accuracy.
- SNV/Indel analysis detected 19 unique *HBB* and 5 unique *HBA1/2* variants for which orthogonal assay data was available in 38 samples with 100% accuracy.
- Our assay enables comprehensive thalassemia variant detection in an easy-to-interpret readout. Using a single PCR reaction, Nanopore sequencing, and bespoke analysis software, our streamlined workflow simultaneously identifies variants across all *HBA* and *HBB* exons with high accuracy.
- PCR-Nanopore assay results for clinical samples with rare SVs will support flagging of such samples for further evaluation upon implementation of a planned QC strategy within the analysis software.

This product is under development. Future availability and performance to be determined.

All authors except Jaime Lopes and Marcus Cannon have the financial relationship to disclose: Employment by Asuragen. Conference attendance by Jaime Lopes was sponsored by Asuragen.

References

1. Tan, ASC et al. (2001) *Blood* 98(1):250-251.
2. Oron-Karni, V et al. (1998) *Am J Hematol* 58:306-310.
3. Turner A et al. (2015) *BMC Med Genet* 16, 115.

Presented at ACMG 2024

We thank Charles T. Quinn and D. Brian Dawson from Cincinnati Children's for contribution to study design.

bio-techne® Asuragen®

Cincinnati Children's®