Targeted long-read sequencing of native DNA for genetic disease diagnostic and screening research

Sarah B. Kingan, Guilherme De Sena Brandine, Jeff Zhou, Tom Mokveld, Jocelyne Bruand, Egor Dolzhenko, Michael A. Eberle, Duncan Kilburn
PacBio, 1305 O'Brien Drive, Menlo Park, CA 94025

Background

Despite advances in DNA sequencing, the **causal mutations of many human diseases remain challenging** to characterize with standard methods. The adoption of **long-read sequencing can improve and simplify** sequencing and analysis of these complex regions. The ability to detect methylation enables simultaneous generation of sequence-level and epigenetic data. We present **new panel content for the PureTarget assay**, which uses the CRISPR/Cas9 system to generate **targeted sequencing libraries of native DNA**, sequenced with **long and accurate HiFi** sequencing.

Scalable amplification-free workflow

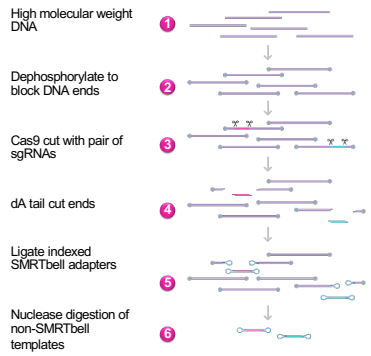

Figure 1. The PureTarget technology is a robust amplification-free approach to generate long-read HiFi sequencing libraries for panel targets. Starting with high molecular weight DNA, the workflow employs Cas9 and pairs of guide RNAs to target loci ranging in size from 3 – 14 kb.

Automated and manual workflows

	Automated	Manual
Input DNA	1 – 1.5 µg	1 – 4 µg
Sample types	Blood, saliva, cell lines	Nanobind cell line, blood (RBC-lysis)
Library prep time	16hr for 96 samples	8hr for 24 samples
Samples / SMRT cell	96 on Revio 384 per run	8 – 48 on Revio and Vega

Coverage in automated 96-plex prep

- Automation kit supports 96 samples per prep on Hamilton NGS Star
- 16 hour turn around time from extracted DNA through final pooled library (pre-ABC)

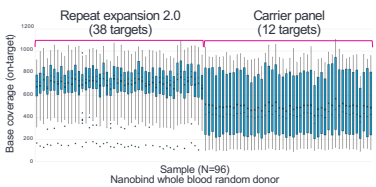

Figure 2. 96 samples of 1 µg of random donor blood extracted with Nanobind HT CBB kit from whole blood (protocol 102-573-500) was prepared on the Hamilton NGS Star. 48 samples were prepared with the repeat expansion (RE) panel 2.0, the rest with the carrier panel. Mean coverage on RE was 707-fold and on carrier is 494-fold. Fragment size on RE panel is more uniform (3.3 – 5.6kb) than carrier panel (3.9 – 14kb) which results in more variable target coverage on carrier panel.

Carrier screening panel

Target	Disease	Variant type
F8 intron 22	Hemophilia	Inversion
F8 intron 1	Hemophilia	Inversion
FXN	Friedrich's ataxia	Repeat Expansion
FMR1	Fragile-X disease (FXS)	Repeat Expansion
CYP21A2, TNXB	Congenital adrenal hyperplasia, Classical-like Ehlers-Danlos syndrome	Small variants and copy number
HBA1/2	Alpha thalassemia	Deletion
GBA	Gaucher	Small variants and copy number
SMN1/2	Spinal muscular atrophy	Small variants and copy number
ARX	Early-infantile epileptic encephalopathy (EIEE1) and Partington syndrome (PRTS)	Repeat Expansion (N=2)
HBB	Beta thalassemia	Small variants
RPGR	X-linked retinitis pigmentosa	Small variants
AFF2	Fragile X syndrome, FRAXE type	Repeat Expansion

CYP21A2 – CNV and fusions

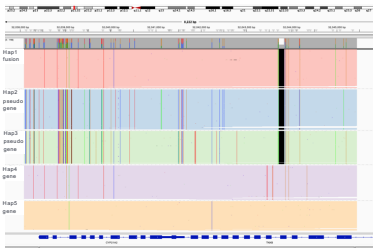

Figure 3. A pair of gRNAs capture ~7.4 kb spanning the full *CYP21A2* gene and 13 exons of *TNXB* on the 3' end. The gRNAs also cut the segmental duplication containing *CYP21A1P* and *TNXA*. HG03540 reference sample was sequenced on Revio + SPRQ and analyzed with paraphase 3.3.1 in the PureTarget carrier pipeline. Five haplotypes are captured showing 2 gene copies, 2 pseudogene copies, and a fusion allele.

HBA1/2 – large deletion

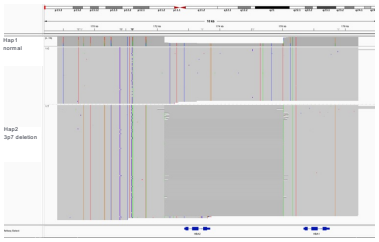

Figure 4. A pair of gRNAs captures ~9.5kb spanning *HBA1* and *HBA2*. The 3p7 deletion is neatly captured showing one copy of *HBA2* and 2 copies of *HBA1*.

PureTarget carrier pipeline (PTCP)

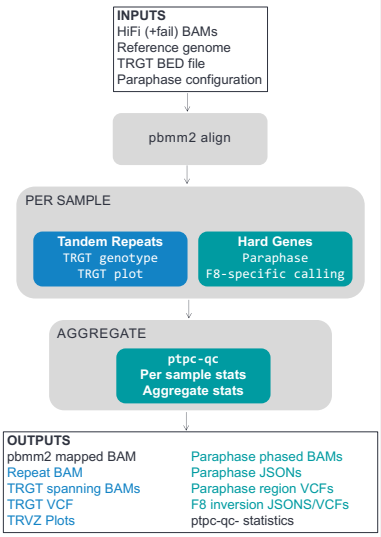

Figure 5. Schematic of PureTarget carrier pipeline for variant calling. Pipeline will be available on DNA Nexus Marketplace and for flexible deployment on cloud or on-prem server via github repository.

Repeat expansion panel 2.0

Targets	Disease
ATN1, ATXN1, ATXN2, ATXN3, ATXN7, ATXN8, ATXN10, CACNA1A, PPP2R2B, TBP ***BEAN1, DAB1, FGF14, NOP56, ZFHX3***	Spinocerebellar ataxia (SCA)
FMR1	Fragile-X disease (FXS)
AFF2	Fragile X syndrome, FRAXE type
AFF3	Intellectual disability associated with fragile site FRA2A
C9orf72	Frontotemporal dementia (FTD), amyotrophic lateral sclerosis (ALS)
FXN	Friedreich ataxia (FRDA)
RFC1	Cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS)
NOTCH2NLC	Neuronal intranuclear inclusion disease, Alzheimer disease and parkinsonism phenotype (NIID)
DMPK, CNBP	Myotonic dystrophy (DM)
HTT	Huntington disease (HD)
JPH3	Huntington's disease-like type2 (HDL2)
TCF4	Fuchs endothelial corneal dystrophy 3 (FECD3)
AR	Kennedy Disease, Spinal and bulbar muscular atrophy, (SBMA)
PABPN1	Oculopharyngeal muscular dystrophy (OPMD)
ABCD3, GIPC1, LRP12, RILPL1	Oculopharyngodistal myopathy (OPDM)
NOXD13	Syndactyly (SD5)
PHOX2B	Congenital central hypoventilation syndrome (CCHS)
PRNP	Creutzfeldt-Jakob disease (CJD)
CSTB	Progressive Myoclonic Epilepsy Type 1 (EPM1) Unverricht-Lundborg Disease (ULD)
SAMD12	Familial adult myoclonic epilepsy type 1 (FAME)

Heterozygous expansion in *FGF14*

		Observed	Expected
Repeat count	Allele 1	33	30 +/- 2
	Allele 2	283	~280
Read count	Allele 1	506	--
	Allele 2	249	--

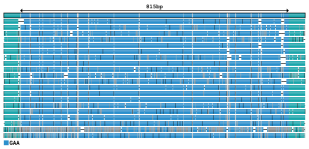

Figure 6. 2 µg of DNA, extracted from blood with QIAsymphony, was manually prepared with PureTarget repeat expansion panel 2.0, sequenced on Revio + SPRQ in an 8plex, and analyzed with TRGT v3.0.0. Expected genotype from rpPCR.

Heterozygous expansion in *FMR1*

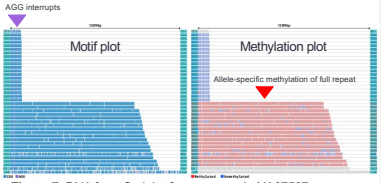

Figure 7. DNA from Coriel reference sample NA07537 was manually prepared with PureTarget repeat expansion panel 1.0, sequenced on Sequel IIe, and analyzed with TRGT v3.0.0. Expected genotype (28-29 and >200 repeats) from southern and PCR analysis. Observed genotype is 27 CGG, 2 AGG and 332 CGG and 1 AGG.

Conclusions

- PureTarget enables **scalable** targeted HiFi sequencing of **native DNA**
- Panels available for **repeat expansion** disease and carrier screening (**rare disease**)
- Manual and **automated** preps available
- **PureTarget carrier pipeline** available for genotyping with **TRGT** and **Paraphase**

Acknowledgements and References

The authors wish to thank Johannes Zschocke, Katharina Lackner and Katharina Herzog from Medical University Innsbruck and Dale Muzzey, Ben Obadia and Jack Roach from Myriad Women's Health for kindly sharing samples.

Dolzhenko et al., (2024). Nature Biotech., 1-9. https://doi.org/10.1038/s41587-023-02057-3.
Chen X, Baker D, Dolzhenko E, et al., (2025). Nature Communications. doi:10.1038/s41467-025-57505-2