# COMPREHENSIVE CHARACTERIZATION OF REPEAT EXPANSIONS WITH PURETARGET

## Shine a light on dark regions of the genome

Expansions of repetitive DNA sequences have been linked to over 50 monogenic disorders and cancers where the repeat length, sequence context and methylation can be factors in disease severity and/or age of onset. Once notoriously difficult to characterize, these regions can now be comprehensively genotyped at scale with the PacBio<sup>®</sup> PureTarget<sup>™</sup> repeat expansion panel. Offering a gene panel for 20 of the most important repeat expansions for human health, PureTarget now grants a clear view of these once-dark regions of the genome.

### What can PureTarget do?

Powered by the *Tandem Repeat Genotyping Tool* (TRGT)<sup>1</sup> and deep coverage of long and accurate HiFi reads, the PureTarget panel provides [A] single-base resolution of repeats, [B] precise sizing of long repeats, [C] methylation profiles, and [D] deep coverage to characterize mosaicism.



Resolve repeat sequence

A. Single-base resolution of *FMR1* repeat sequence detects AGG interruptions in female carrier (NA06905). Consensus alleles are 23 and 79 repeats long, respectively.<sup>2</sup> Axis unit is base pairs.





Unmethylated

C. Profile at *FMR1* shows consistent methylation of expanded allele in female carrier (NA07537). Axis unit is number of 3 bp repeat motifs.

### Size long repeat expansions

**PureTarget**<sup>™</sup>



B. Double expansion of pathogenic AAGGG repeat in *RFC1* with long allele consensus length of 1300 repeats ( $\sim$ 6.5 kb).<sup>3</sup> Axis unit is number of repeat motifs.

Profile mosaicism



D. Waterfall plot shows deep coverage and distribution of repeat lengths within a sample with *C9orf72* expansion. Sample was prepared with 2 µg of DNA and sequenced in 16-plex on Sequel<sup>®</sup> IIe. A total of 458 reads span the *C90RF72* repeat, 66 of which span the long allele. Axis unit is base pairs.



### Stay on target

PureTarget offers an easy and scalable workflow to capture repeat expansions and brings you from sample to answer in three days. Compatible with the Revio<sup>™</sup> and Sequel II systems, PureTarget libraries can be multiplexed with up to 48 samples, generating deep coverage across the 20-gene target panel.

		2 hours 8 hours	Spec	Metric	
Nanobind® DNA extraction PureTarget library			DNA input <sup>1</sup>	2 µg/sample	
			DNA quality <sup>2</sup>	GQN at 30 kb >5	
			Mean target coverage <sup>3</sup>	>200-fold	
			Minimum target coverage	50-fold	
			Sample multiplexing <sup>4</sup>	48 —Revio system 24 —Sequel II systems	
HiFi sequencing		1 hour prep 24 hour movie	Library size <sup>5</sup>	4-5 kb	
			<b>Methylation</b> <sup>6</sup>	Detected	
SMRT <sup>™</sup> Analysis		15 minutes	<ol> <li>1-4 µg supported</li> <li>50% of mass of DNA molecules longer th (Agilent). Official product support for Na cell line.</li> <li>Mean and minimum target coverage is fi types (Nanobind-extracted human blood</li> <li>Kit supports smaller batches in multiples</li> <li>Inserts with expanded alleles will be long</li> <li>Methylation probabilities for CpG sites et</li> </ol>	μg supported of mass of DNA molecules longer than 30 kb as measured on Femto Pulse lent). Official product support for Nanobind-extracted DNA from human blood and line. In and minimum target coverage is for 2 μg of input DNA from supported samples is (Nanobind-extracted human blood and cell line) for unexpanded alleles. Supports smaller batches in multiples of 8 samples. Ints with expanded alleles will be longer. hylation probabilities for CpG sites encoded in BAM.	
Disease			Targets		
Spinocerebellar ataxia			ATN1, ATXN1, ATXN2, ATXN10, CACNA1A, Pł	ATN1, ATXN1, ATXN2, ATXN3, ATXN7, ATXN8, ATXN10, CACNA1A, PPP2R2B, TBP	
Friedrich's ataxia			FXN	FXN	
CANVAS disease			RFC1	RFC1	
Fragile-X disease (FXS	5)		FMR1	FMR1	
Huntington's disease			HTT	НТТ	
Myotonic dystrophy			DMPK, CNBP	DMPK, CNBP	
Amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD)			(FTD) C9orf72	C9orf72	
Fuchs endothelial corr	neal dystrophy	,	TCF4	TCF4	
Spinal bulbar muscula	r atrophy / Ke	nnedy's disease	AR	AR	
Oculopharyngeal mus	cular dystroph	Ŋ	PABPN1	PABPN1	

#### **KEY REFERENCES**

- 1. Dolzhenko E, English A, Dashnow H, et al., (2024). Characterization and visualization of tandem repeats at genome scale. Nature Biotechnology, 1–9.
- 2. https://downloads.pacbcloud.com/public/dataset/PureTargetRE/Coriell/
- 3. https://downloads.pacbcloud.com/public/dataset/PureTargetRE/RFC1/ (sample NG24)



#### Learn more: pacb.com/target

Research use only. Not for use in diagnostic procedures. © 2024 Pacific Biosciences of California, Inc. ("PacBio"). All rights reserved. Information in this document is subject to change without notice. PacBio assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of PacBio products and/or third-party products. Refer to the applicable PacBio terms and conditions of sale and to the applicable license terms at pacb.com/license. Pacific Biosciences, the PacBio logo, PacBio, Circulomics, Omniome, SMRT, SMRTbell, Iso-Seq, Sequel, Nanobind, SBB, Revio, Onso, Apton, Kinnex, and PureTarget are trademarks of PacBio.

