Technical note

A PRACTICAL GUIDE TO AMPLIFICATION-FREE PURETARGET CUSTOM PANELS

Introduction

Accurate resolution of complex and difficult-to-sequence regions of the human genome is critical for uncovering variants that cause or increase risk of disease. PacBio® HiFi sequencing offers one of the most comprehensive and accurate approaches, enabling simultaneous detection of SNPs, indels, structural variants, methylation marks, and even chromatin accessibility. For genes that short-read sequencing cannot easily capture—like repeat-rich or structurally complex loci—PacBio PureTarget™ panels provide a low-cost, Cas9-based, amplification-free solution to profile clinically relevant genes. The existing panels for carrier screening and repeat expansion targets demonstrate the power of this approach to capture challenging genomic regions with high precision.

With the release of new PureTarget products, users now have access to a set of control guide RNAs that can serve as benchmarks for designing custom panels. This flexibility extends PureTarget beyond pre-designed panels, enabling targeting of custom loci that are technically demanding yet biologically critical. Such custom panels are particularly valuable for regions with high GC content, repeat expansions, long genes requiring multiple fragments for full coverage, or loci where methylation and epigenetic context provide essential regulatory insights. This technical note provides guidance for researchers exploring custom applications and highlights considerations in guide design, panel construction, and analysis.

Types of custom panels: Removing and adding targets

Custom PureTarget panels can be broadly classified into two categories: removing targets and adding targets. Removing targets is the simpler of the two and can be implemented in two ways:

- The first is an *in-silico* panel, where no changes are made to the wet-lab workflow and unwanted targets are simply filtered out by editing the BED files in SMRT® Link.
- The second option is a "slim panel," in which a custom guide RNA (gRNA) pool is made from a set of pre-designed gRNAs available with the PureTarget collection (e-mail support@pacb.com to request designs). In this case, the corresponding BED files must also be modified to retain only the desired targets, but no further changes to analysis workflows are required.

Adding new targets to a PureTarget panel requires further considerations in guide design and workflow. These new targeted regions can generally be grouped into three types:

a. Tandem repeat expansions,

b. Non-repeat regions shorter than the length of a typical HiFi read (~15 kb),

c. Non-repeat regions longer than 15 kb.

Tandem repeats are typically captured with a single guide pair spanning a 4–5 kb region of the non-expanded allele; short non-repeat regions require a single guide pair; longer non-repeat regions need multiple guide pairs to ensure full coverage. Each modality involves slightly different strategies for library preparation, optimization, and downstream analysis, but can all be prepared with commercially available gRNA synthesis and are compatible with PureTarget kit 24 manual workflow.

NOTE: This technical note is intended as guidance for custom panel design recommendations, performance optimization, and to report example results. PacBio does not guarantee the success of custom target performance alone or when added to off-the-shelf target designs, as these panels are untested. For troubleshooting assistance, users must combine custom guides with any of our fixed panels (Control, Repeat Expansion 2.0, or Carrier) when preparing libraries (see Appendix for instructions). For best results, prepare custom panel libraries only from supported sample types (Nanobind®-extracted blood, lymphoblastoid cells, and saliva (PureTarget kit 24 and Revio system with SPRQ chemistry only), or Coriell lymphoblastoid cell DNA with GQN at 30 kb > 5) and sequence only on supported platforms (Revio® system with SPRQ™ chemistry or the Vega™ system).

PacBio

Workflow modifications	*In silico* (virtual panel) *Mask out unwanted targets*	"Slim" panel *Order and pool subset of PacBio gRNA designs*
Guide design	None	None
Reagents	None	Order guides from IDT (or Sigma)
Optimization	None	None
Library prep	None	Swap panel
Analysis	Modify BED to remove targets	Modify BED to remove targets

Table 1. Workflow modifications for removing targets.

Workflow modifications	Tandem repeat expansions	Non-repeats shorter than 15 kb	Non-repeats longer than 15 kb
Guide design	Single pair, ~5 kb fragment for non-expanded allele	Single pair, avoid SNPs	Multiple pairs to coverage larger region
Reagents	Order guides from oligo vendors	Order guides from oligo vendors	Order guides from oligo vendors
Optimization	First pair works ~80% of the time	First pair works ~80% of the time	First pair works ~80% of the time
Library prep	Spike in guides to panel	Spike in guides to panel	Split two pools at Cas9 step, pool after Cas9 heat kill
Analysis	Add repeat definitions to BED using strchive Analysis with TRGT or PureTarget repeat expansion in SMRT LInk	Add coordinate to BED and analyze with Target Enrichment CompBio tools (GitHub) for variant calling	Add coordinate to BED and analyze with Target Enrichment CompBio tools (GitHub) for variant calling

Table 2. Workflow modifications for adding targets.

Designing guide RNAs for adding new targets

A guide RNA (gRNA) is a synthetic molecule made by covalently fusing a CRISPR RNA (crRNA) with a trans-activating CRISPR RNA (tracrRNA). Across all modalities of adding new targets, a few common design principles for sgRNA apply. Guides should be oriented correctly relative to the target strand (5' for the positive strand, 3' for the negative strand), demonstrate high on-target and low off-target activity, and avoid commonly occurring SNPs in the target sequence. It is critical to review the off-target scores provided by design tools to ensure the chosen sequence does not share exact homology with unintended regions. We recommend ordering gRNAs from IDT or Millipore Sigma. IDT has web-based tooling available to measure on- and off-target scores of sgRNA designs. Additionally, gRNAs can be designed through community tools such as CHOPCHOP or Crispor. It is important to ensure that the NGG PAM motif is selected under the Cas9 enzyme options.

In practice, most gRNAs fall into a few design categories that reflect the number and quality of candidate gRNA sequences available. The ease of design depends on how many candidates meet the criteria of high on-target activity and low off-target potential. On- and off-target metrics are always assessed relative to the genome of interest, following principles discovered from off-target profiling experiments in the literature[1-4]. Designs are classified into four groups as follows:

Good: Many candidate gRNAs are available that meet design criteria, providing multiple options to test experimentally.

Limited: Only 1–3 viable candidate gRNAs are available. Additional strand-specific filtering will reduce this pool further, often leaving only a single candidate to test experimentally.

Sparse: No guides fully meet the high on-target/low off-target criteria. In these cases, a workaround is to select a guide with high on-target activity while minimizing off-targets as much as possible. A modest increase in gRNA-specific off-target activity has not been shown to impact sample multiplexing capacity or mean target coverage in internal experiments. Moreover, off-target regions may be informatically filtered because HiFi reads are long and accurate.

Recalcitrant: The design tool fails to return any candidates. This situation can occur in low-complexity regions of the genome. Splitting the target design locus into shorter genomic segments provides a path forward, but these designs may often carry higher off-target scores.

Using this above classification, we find that approximately 80% of gRNA designs fall into the "good" category (Figure 1, statistics based on ~70 *CFTR*-specific guides designed with IDT tools.). When tested experimentally, the first design iteration succeeds in delivering sufficient coverage about 85% of the time (Figure 1b). A minority of sgRNA designs require two or more iterations to achieve the desired coverage performance. The control guides included in PureTarget are used as a benchmark for expected coverage levels and include three genes: X-linked *UBL4A*, and autosomal *ACTB* and *GAPDH*.

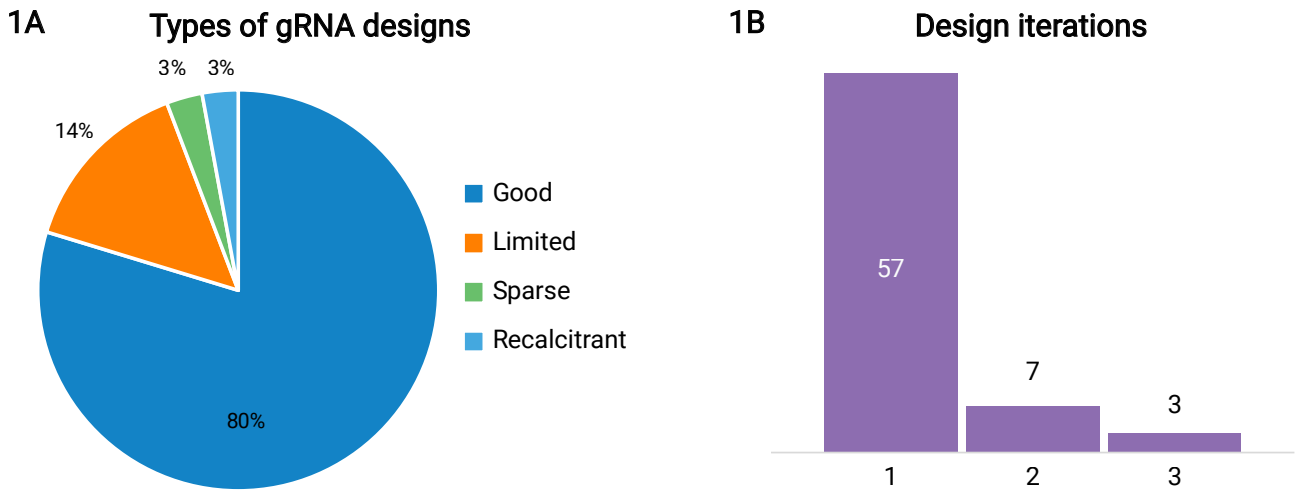

Figure 1. Summary of gRNA designs. a) Distribution of gRNA designs across four categories of availability, based on ~70 *CFTR*-specific guides designed with IDT tools. b) Number of iterations required to achieve sufficient coverage; most designs succeeded on the first attempt.

For tandem repeat expansions, the most reliable approach is to flank the non-expanded allele with a ~5 kb window. These coordinates can be annotated using strchive, a database that catalogs known tandem repeat loci and their sequence contexts, enabling consistent BED file definitions. For non-repeat regions shorter than 15 kb (the average length of a HiFi read) single gRNA pairs flanking the target region can be designed. Internal studies on the *MLH1* locus show that shorter target regions consistently achieve higher coverage. Specifically, coverage is highest with ~5 kb target regions, intermediate with ~7.5 kb, and lowest when targeting 15 kb regions (Figure 2). Targeting shorter regions may be the most practical approach in a production setting, if high coverage is needed, or when DNA samples are low quality (GQN at 30 kb < 5). If targeting a region with high sequence homology or if phasing across a region is required, targeting longer fragments may be preferred.

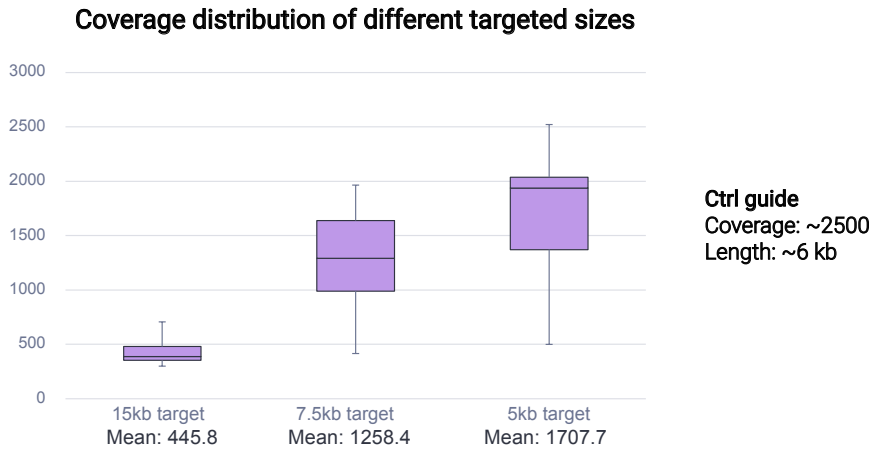

Figure 2. Coverage distribution at the *MLH1* locus using different target sizes. Coverage was compared across 15 kb, 7.5 kb, and 5 kb targets. Libraries were prepared with 4 µg of Nanobind whole blood DNA (2-plex) and sequenced on Revio with SPRQ chemistry. Shorter targets yielded higher mean coverage.

Larger non-repeat regions (≥15 kb) require strategies tailored to the research question, which can broadly be grouped into two subtypes. The first subtype tiles across the entire locus to enable allele phasing and full-length coverage but requires overlapping tiles delimited by gRNA pairs. In this approach, alternating gRNA pairs are organized into "odd" and "even" pools and must be processed separately with PureTarget library preparation during the Cas9 cleavage step. Both approaches were evaluated in the *CFTR* case study described below.

The second subtype is to cut the entire region into non-overlapping fragments, capturing experimentally relevant hotspots such as stretches containing pathogenic variants within a gene. This approach is compatible with the standard PureTarget library preparation workflow and requires only modification or augmentation of the panel.

Case study 1: Profiling *CFTR* with PureTarget

The *CFTR* gene is one of the most clinically significant loci for inherited disease, with variants that cause cystic fibrosis (CF) as well as *CFTR*-related diseases (CFTR-RDs). This gene spans 188.9 kb, contains 27 exons, and harbors thousands of reported variants in the population. The CFTR2 database, which aggregates genetic data from >120,000 CF & CFTR-RD patients, highlights the dual character of this locus from a clinical diagnostic perspective. A single mutation, F508Δ, accounts for nearly two-thirds of all cystic fibrosis cases, yet the remainder arises from a long tail of diverse pathogenic variants spread across the length of the gene (Figure 3). With an ever-growing spectrum of CFTR-related disorders (CFTR-RDs), the clinical goal is to look beyond this single mutation, and emerging sequencing technologies are enabling that shift[5,6]

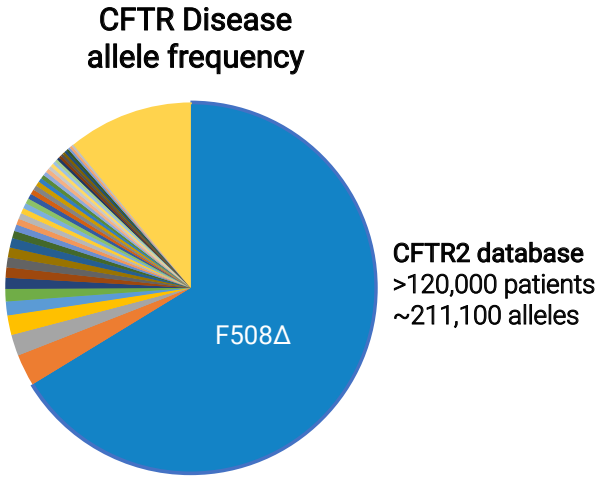

Figure 3. *CFTR* disease allele frequencies from the CFTR2 database (>120,000 patients, ~211,100 alleles). The predominant allele is F508Δ, accounting for ~67% of cases, while the remaining ~33% represent other pathogenic alleles. The larger yellow category aggregates alleles individually observed fewer than 400 times in the CFTR2 database. This distribution underscores the importance of full-gene sequencing to capture the diversity of *CFTR* variants beyond the single common mutation.

Reflecting this trend, the American College of Medical Genetics (ACMG) updated its guidance in 2023 to recommend a panel of 100 *CFTR* variants for clinical testing[5], expanding the scope of variants under formal clinical scrutiny as sequencing capabilities advance. Commercial panels, such as Asuragen's AmplideX PCR/CE CFTR Kit and Illumina's TruSight Cystic Fibrosis panel, track dozens to over a hundred pathogenic variants (e.g., Illumina's assay targets 139 variants) but neither capture all reported *CFTR* pathogenic variants. Any technology that can capture the full length of the gene, including noncoding, intronic, and regulatory variants, offers the potential to future proof an assay against potential updated guidelines. To demonstrate this, we evaluated two custom PureTarget panel designs for *CFTR*: (i) **full-gene tiling**, using overlapping guide RNAs to achieve contiguous gene coverage, support methylation profiling, and enable phasing; and (ii) **hotspot-focused panels**, using non-overlapping regions (~6 – 10 kb) placed across the gene with discrete start and stop sites to capture clinically pathogenic regions.

Note: In addition to cystic fibrosis, pathogenic CFTR variants are linked to a spectrum of CFTR-related disorders (CFTR-RDs), including congenital bilateral absence of the vas deferens (CBAVD), chronic pancreatitis, bronchiectasis, and certain forms of sinus disease. These conditions typically result from reduced, rather than absent, CFTR function and represent an important part of the broader clinical landscape[7].

Full-gene tiling of *CFTR* (subtype 1 strategy)

For full-gene tiling, we aimed to achieve contiguous coverage of the entire *CFTR* locus across its ~189 kb span. The gene was divided into overlapping fragments ("tiles"), with an average fragment size of ~10 kb, selected to balance sequencing depth associated with DNA quality requirements with guide design and library prep complexity. Overlap was deliberately prioritized to enable phasing across long stretches of the gene, since variant-rich overlap regions provide anchor points for phasing algorithms. To guide this design, we referenced a list of *CFTR* variants from ClinVar annotated as benign or likely benign, ensuring overlap regions were aligned with common polymorphisms prevalent in global populations.

Once tile coordinates were defined, specific gRNA pairs were designed for each fragment following standard principles of guide design. Approximately 85% of gRNAs functioned effectively in the first round, yielding robust coverage, while the remainder required optimization. After design, the gRNAs were grouped by tile into odd and even sets, and the library preparation workflow was adjusted to prevent cross-Cas9 cleavage (see Library Preparation Modifications). This strategy delivered high-quality coverage across the entire gene, enabling integrated analysis of variation, methylation, and phasing.

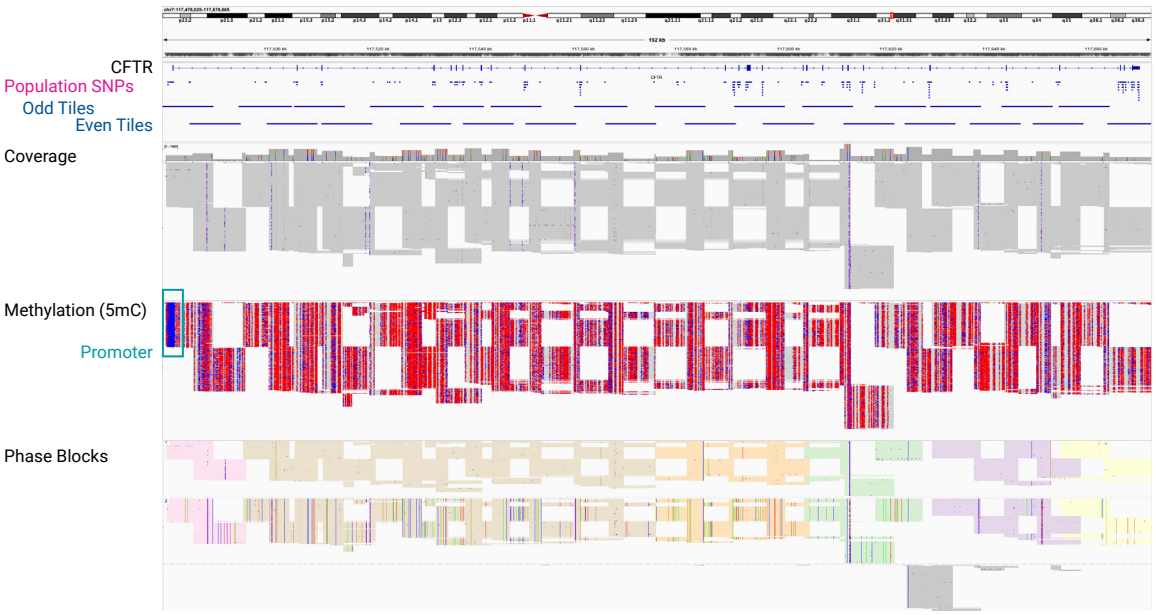

Figure 4 IGV visualization of *CFTR* full-gene tiling. Tracks 1–3: *CFTR* locus structure, population SNPs, and designed tiles. Track 4 (Coverage): Grey bars illustrate sequencing depth across all tiles, showing contiguous coverage. Track 5 (Methylation): Coverage colored by methylation status (blue = unmethylated, red = methylated). Notably, the *CFTR* promoter region is clearly unmethylated. Track 6 (Phasing): Phase blocks generated using DeepVariant variant calls and HiPhase phasing; overlap between tiles allows phase blocks to extend across long stretches of the locus. Note that DeepVariant has not been trained on PureTarget data.

Pathogenic *CFTR* hotspot-focused panels (subtype 2 strategy)

For users who prefer minimal changes to the library preparation workflow, we designed a separate panel of non-overlapping gRNAs targeting pathogenic hotspots of the *CFTR* gene highlighted in the ACMG100 panel plus regions of interest to key diagnostic providers (see Figure 5). In this approach, the standard workflow remains unchanged; the only modification is the substitution of the custom panel alongside either the custom control or the carrier panel (PN 103-633-300 or PN 103-633-200, respectively). This strategy delivered excellent coverage across the targeted regions, providing a streamlined option for hotspot-focused profiling.

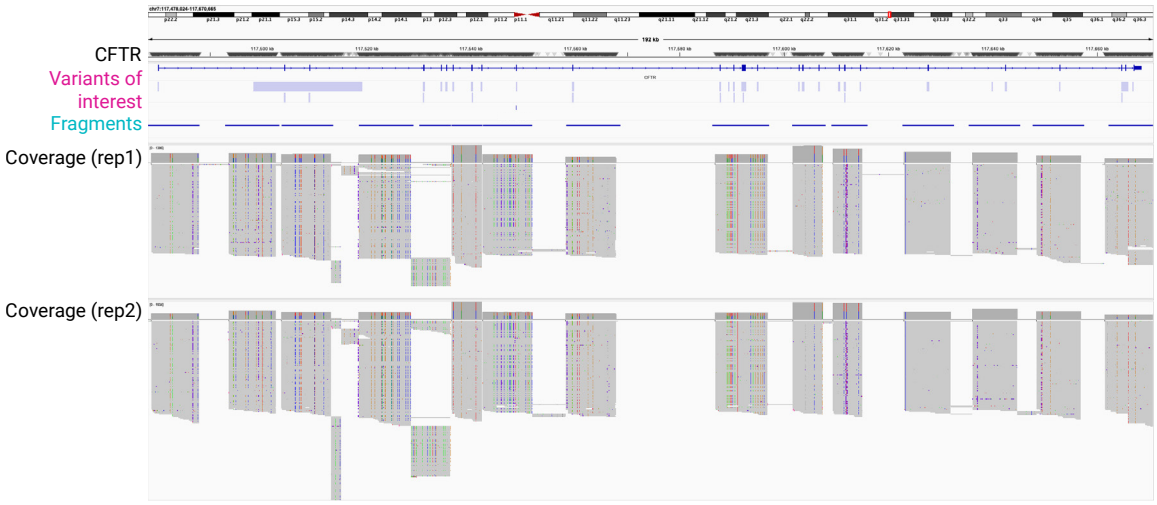

Figure 5. IGV visualization of CFTR targeting pathogenic hotspots according to ACMG100[5] and other pathogenic alleles of interest. Notably, the spread of pathogenic and likely pathogenic alleles is across the entire length of the gene.

Case study 2: Characterizing *ROSA26* insertions in humanized mice with PureTarget

Humanized mouse models are indispensable for translational research, and precise knock-ins into safe-harbor loci such as *ROSA26* form the backbone of these genetic engineering approaches. Accurate genotyping of founder animals is critical, since editing efficiency is never complete—some animals carry the desired allele, others remain unedited, and many are mosaic, making interpretation difficult. Insertions at *ROSA26* are often long, and long-range PCR approaches can be confounded by incomplete amplification, allele bias, or dropout of complex alleles. Amplification-free genomic characterization provides a path forward, as it captures all allele types directly without the biases introduced by PCR. Single-molecule HiFi sequencing is especially well suited to this application, combining long, accurate reads with the ability to resolve mosaic alleles. High per-sample coverage is essential for detecting these alleles, and the ability to multiplex many animals at once makes the approach cost-effective.

To address this challenge, Freeman et al. applied CRISPR/Cas9-based PureTarget enrichment to directly sequence engineered *ROSA26* alleles in founder and F1 animals. Custom guide RNAs flanking 12 kb on either side of the insertion site were designed and substituted into the PureTarget workflow. Multiplexing enabled sequencing of 12 animals on a single PacBio Revio SMRT® Cell, delivering per-sample coverage between 150× and 2,000×. This strategy provided complete, PCR-free characterization of insertion alleles—including correctly edited, indel, and wild-type sequences. The ability to multiplex large cohorts while maintaining high coverage lowers the barrier to comprehensive allele characterization and accelerates validation of humanized and transgenic models, reducing the time from founder generation to experimental readiness.

Case study 3: Detecting genetic mosaicism in zebrafish with PureTarget

Zebrafish are a widely used model organism for developmental biology and genome editing research, but accurate analysis of CRISPR-Cas9 outcomes has been hampered by PCR biases and short-read sequencing. Genetic mosaicism, where multiple editing events coexist within the same organism, is difficult to measure comprehensively. Founder zebrafish (F0) often harbor dozens of distinct alleles at both on-target and off-target sites, and the key challenge is determining which of these alleles are transmitted to the next generation.

To address this, Höijer et al.[8] applied an amplification-free PureTarget workflow to 32 zebrafish samples spanning both founders (F0) and offspring (F1). Four on-target loci (*ldrla*, *nbeal2*, *sh2b3*, *ywhaqa*) and three off-target sites (~5 kb each) were enriched with Cas9 and sequenced on a single Revio SMRT Cell with SPRQ chemistry. Despite partially degraded DNA that was previously collected and stored, 30 of 32 samples produced high-quality data, with an average of ~1168× target coverage and QV39 single-molecule accuracy. Critically, PureTarget provided unbiased haplotype representation (see Figure 6 below), showing that while long-range PCR (LR-PCR) produced skewed allele frequencies at heterozygous SNVs, PureTarget consistently delivered ~50:50 allele balance.

Analysis revealed extensive mosaicism in F0 zebrafish, with 7–18 distinct alleles detected at single on-target loci and up to 11 alleles at off-target sites. Several of these alleles—including a 1053 bp deletion at *sh2b3* and a 3 bp deletion at *ywhaqa* off-target 2—were transmitted to F1 offspring and segregated in juvenile fish. These findings demonstrate that mosaicism can extend into germ cells, underscoring the need to account for this when interpreting germline genome editing outcomes.

In addition to genotyping, PureTarget preserved native 5mC methylation signals, allowing simultaneous analysis of epigenetic features. No methylation differences were observed between edited and control zebrafish in this study, but the ability to pair genetic and epigenetic analysis at single-molecule resolution broadens the utility of the workflow. Overall, this case study shows how PureTarget enables deep, unbiased characterization of CRISPR editing outcomes in model organisms, capturing low-frequency alleles and inheritance patterns that would be missed by PCR-based methods.

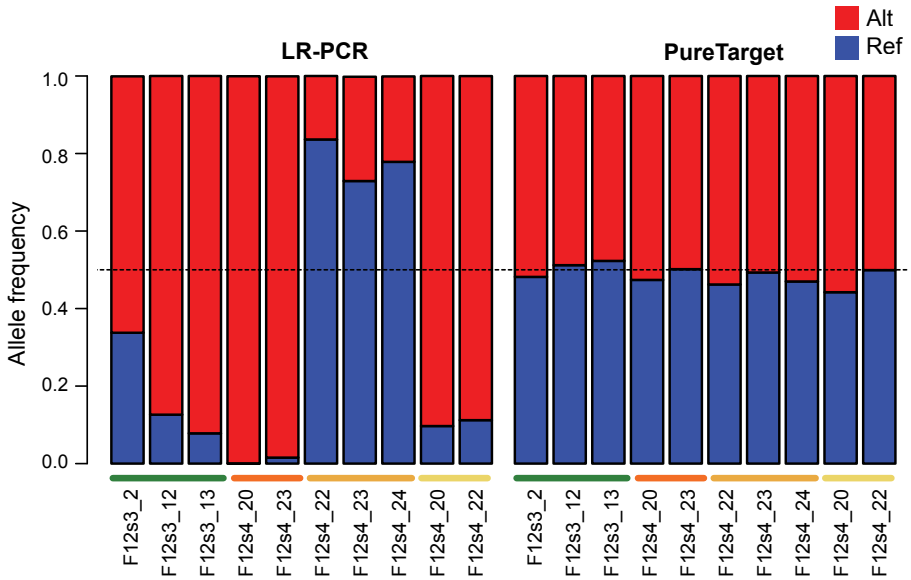

Figure 6. Frequencies of reference (blue) and alternative (red) alleles at single nucleotide variant (SNV) positions in heterozygous target regions for F1 zebrafish. The bars to the left show allele frequencies for long-range PCR (LR-PCR) while the bars to the right show allele frequencies for PureTarget in the same samples. The minor allele frequencies obtained by PureTarget are close to 50% while LR-PCR results in more uneven distribution between the two haplotypes. The colors below the bars indicate which of the target sites were examined. The target sites have been detailed below each bar. Reproduced with permission from Höijer et al. (2025), Figure 2b.

Library preparation modifications

Any customization beyond *in silico* "virtual" panels requires adjusting the composition of the guide RNA (gRNA) mixes. It is critical to include the PureTarget control panel (PN: 103-633-300) whenever custom gRNA mixes are prepared. Alternatively, the repeat expansion panel 2.0 (103-633-100) or carrier panel (103-633-200) could be used if users wish to augment one of the fixed panels. Control and fixed panel guides serve as positive control for PureTarget library prep and provide a benchmark for evaluating the performance of the new gRNAs designs and oligos. Detailed instructions for preparing equimolar gRNA mixtures are provided in the Appendix.

For removing targets to create a slim panel, the workflow modification is minimal. A new gRNA pool is created by selecting only the desired coordinates from the PureTarget 2.0 collection (email support@pacb.com to request designs), combining them with the control guides in equimolar proportion (see Appendix for mixing protocol, control guide part numbers).

For adding targets, modifications can range from simple to extensive.

- For tandem repeat expansions and non-repeat regions <15 kb, the new gRNAs (designed as described above) are simply added to the control pool, and the standard workflow can be followed without further changes.
- For non-repeat regions >15 kb, two strategies apply:
  - If targeting the entire locus with overlapping tiles (subtype 1), gRNAs must be organized into odd and even pools (two Cas9 digestion master mixes are required at Step 4.1 of the manual library prep protocol, one with odd pool gRNA and the other with even pool gRNA, each at half the volumes of all reagents). The post-dephosphorylation sample is split in half where each Cas9 digestion master mix is used for a separate Cas9 digestion to prevent formation of unwanted intermediate fragments (Step 4.3). After Cas9 digestion, heat-inactivate Cas9 at 65°C for 20 mins; then combine the odd and even post-Cas9 digestion reactions for each sample just prior to bead cleanup (Step 4.7).
  - If targeting non-overlapping fragments (subtype 2), design gRNAs to the regions of interest and incorporate them into a new gRNA pool.

A schematic overview of these workflows is provided in Figure 7. Internally, we have investigated panels containing up to 145 gRNAs (including 6 control gRNAs), which provides an upper bound for current experimental design.

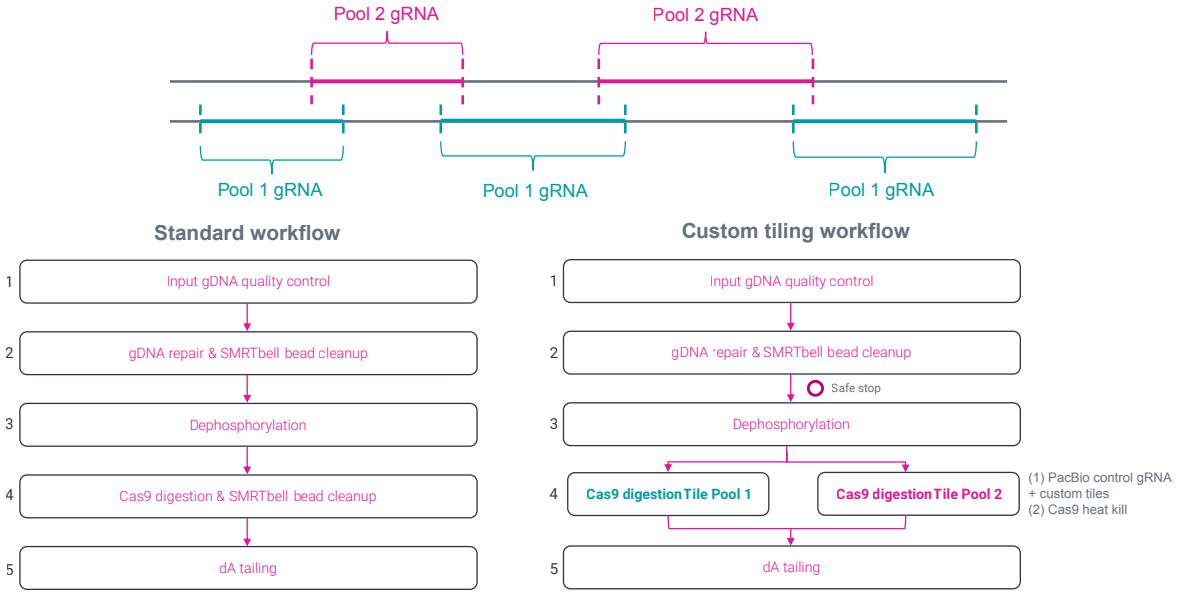

Figure 7. Schematic overview of standard and custom tiling workflow for custom panel library prep. In the custom tiling workflow, reagents must be split at the Cas9 digestion step to separate odd and even gRNA pools. Cas9 is then heat-inactivated before the samples are re-mixed for cleanup and dA tailing. This modification prevents cross-cleavage and enables efficient tiling across long loci.

Best practices for custom PureTarget library preparation

- Always include the control gRNA panel (or other off-the-shelf PureTarget panel) when preparing custom mixes.
- Prepare gRNA mixes at equimolar concentrations (see Appendix).
- For slim panels, simply swap in a new gRNA pool (picking gRNAs from the existing list) plus controls.
- For adding targets <15 kb, add new gRNAs directly to the control pool.
- For full-gene tiling, split gRNAs into odd and even pools, run Cas9 step separately, heat-inactivate Cas9, and then recombine samples before cleanup.
- For optimal performance, prepare custom panel libraries only from supported sample types (blood or lymphoblastoid cells extracted with Nanobind-extracted saliva (manual prep only), or Coriell lymphoblastoid cell DNA with $GQN_{30kb} > 5$)
- For optimal performance, sequence only on supported platform chemistry (Revio + SPRQ and Vega).

	Minimum size	Maximum size
gRNA pairs	4 guide pairs	75 guide pairs
Total panel length	20 kb	500 kb

Table 3. Recommended minimum and maximum PureTarget panel size. The minimum size is one pair of custom guides plus 3 control guides with a total length of 20 kb. The maximum recommended number of guide pairs is 75 with total length 500 kb. A feasibility study of 56 guide pairs and 322 kb total panel size was able to achieve mean target coverage of 660-fold across 8 Nanobind blood samples sequenced on Revio + SPRQ.

Conclusion

Custom PureTarget panels empower researchers to interrogate specific genomic loci with a level of depth and multiplexing efficiency that are difficult to achieve with whole genome HiFi sequencing. By extending beyond the content of the fixed PureTarget panels, this technical note outlines principles for designing custom gRNAs, benchmarking their performance, and implementing workflow modifications. The case studies presented—spanning clinically relevant human genes, humanized mouse models, and model organisms like zebrafish—illustrate the versatility of the approach for profiling both germline and somatic variation in diverse contexts. Looking forward, this framework of custom PureTarget can be adapted for emerging applications, including a deep profiling of somatic variation, expanded analyses such as methylation-based variant detection and haplotype phasing, or chromatin-accessibility by combining with the Fiber-seq assay. These possibilities are enabled by the amplification-free nature of PureTarget workflows. Advances such as SPRQ chemistry, which lower DNA input requirements and increase coverage, further expand what is possible in targeted, long-read genomic studies.

Appendix:

Preparing custom gRNA mixes

For all customizations, follow the **standard PureTarget protocol** in the Procedure & checklist - Generating PureTarget libraries with PureTarget kit 24 – manual protocol.

Special focus: Section 4 (*Cas9 digestion and SMRTbell cleanup*).
All other steps in the protocol remain unchanged.

Step 1. Prepare individual gRNAs

1. Obtain sgRNAs from IDT (2 nmol tubes supplied).
2. Dilute each sgRNA (shipped in individual tubes) to a **5 µM stock concentration** by adding 400 µL nuclease-free water to each tube.
   - This dilution makes all guides consistent and ready for pooling.

Step 2. Pool custom guides

1. Combine **equal volumes (µL)** of each diluted sgRNA to create your **test pool**.
   - Example: For 6 guides, add 10 µL of each into a single tube (final 60 µL pool). We recommend using DNAse/RNase free DNA Lo bind tubes (Eppendorf Cat No. 022431021). These pools can be stored at -20ºC for up to 2 years.
2. Vortex mix and briefly spin down.

Step 3. Combine with Control Panel

The PureTarget control panel (103-633-300)* is comprised of 6 control target guides and must always be included for benchmarking performance. To prepare a complete customized panel including both custom and control guides, create an equimolar mix of test pool + control panel. The volume ratio between test pool and control panel should be the gRNA count ratio between the two. For example,

- **6 guides in test pool:** Add equal volume of control guides to the test pool (1:1).
- **12 guides in test pool:** Add **0.5×** volume of control guides relative to test pool.
- **30 guides in test pool:** Add **0.2×** volume of control guides relative to test pool.

* *General rule: scale the control panel so that the total molar contribution of control and test guides is balanced across all guides in the final pool.*

Step 4. Proceed with library prep

- Use the customized panel created in Step 3 in place of "PureTarget panel" for **Cas9 digestion (Section 4, step 4.1)**.
- All subsequent steps (dA tailing, adapter ligation, nuclease digestion) remain identical to the published PureTarget workflow.

Step 5. SMRT® Link run design

- Select "PureTarget custom" as the run design Application.
- Analysis Options will by default be set to run the Target Enrichment analysis workflow. A Target Region BED file and Advanced Analysis settings must be provided per guidance below in Analysis Tools.

Analysis tools

PureTarget custom panels can be analyzed using **SMRT Link Target Enrichment** workflows for panel QC and coverage assessment. These tools support benchmarking of panel performance, have been tested with GRCh38 and GRCh37 human reference genomes, but do not include variant interpretation; custom panels are **not intended for diagnostic use.**

Benchmarking guide RNA performance

Coverage at each target region can be evaluated relative to the control guides included in the assay. This provides a direct measure of enrichment efficiency and highlights underperforming sgRNAs.

Using SMRT Link target enrichment

1. **BED file preparation**
   - Append your custom targets to the PureTarget control guide BED file provided on the PacBio application kits page.
   - Ensure the file formatting matches SMRT Link requirements detailed in the User Guide Appendix E (see PureTarget control BED linked above as an example).
2. **Analysis parameters**
   - *Include fail reads:* Keep **OFF** by default. Toggle **ON** only when analyzing genomic regions with known high repeat content.
   - *Padding around regions (bp)*: Set to **0** to avoid incorrect measurement of coverage metrics.
   - *Variant calling:* Keep **OFF** for custom panels.
3. **Odd and even tiled designs (subtype 2)**
   - If using overlapping odd/even pools for full-gene tiling, run these as **two separate analyses** in SMRT Link. The workflow does not support overlapping regions in a single run.

Alternate tools

For coverage determination outside of SMRT Link, `samtools bedcov` can be used to calculate read depth over a set of target intervals. Unlike SMRT Link, which reports base-level coverage, bedcov outputs the sum of per-base read depth for each interval, producing an integer value that can be normalized to the region length if desired. Command syntax is detailed here in samtools documentation.

Genotyping and phasing tools

For downstream genotyping analyses, PacBio maintains a suite of long-read sequencing tools on GitHub. Commonly used tools include pbmm2 for mapping, sawfish for structural variant calling, TRGT for tandem repeat genotyping, paraphase for high homology genes, and hiphase for phasing. In addition, Google's DeepVariant has been used for variant calling, producing outputs that are compatible with hiphase for downstream phasing analysis. All tagged BAMs generated from these workflows can be readily visualized in IGV, enabling interactive inspection of alignments, variants, and haplotypes.

Additional Resources

PureTarget technology page

PureTarget control guide BED file provided on the PureTarget datasets page

Procedure & checklist - Generating PureTarget libraries with PureTarget kit 24 – manual protocol

References

1. Tsai SQ, Zheng Z, et al. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33(2):187–197.
2. Tsai SQ, Nguyen NT, et al. (2017) CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* 14(6):607–614.
3. Kim D, Bae S, et al. (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* 12(3):237–243.
4. Wienert B, Wyman SK, et al. (2019) Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science* 364(6437):286–289.
5. Deignan, J. L., et al. & ACMG Board of Directors. (2023). Updated recommendations for CFTR carrier screening: A position statement of the American College of Medical Genetics and Genomics (ACMG). *Genetics in Medicine*, 25(8), 100867.
6. Beauchamp, K. A., et al. (2019). Sequencing as a first-line methodology for cystic fibrosis carrier screening. *Genetics in Medicine*, 21(11), 2569-2576.
7. Castellani, C., et al. & ECFS Diagnostic Network Working Group. (2022). ECFS standards of care on CFTR-related disorders: Updated diagnostic criteria. *Journal of Cystic Fibrosis*, 21(6), 908-921.
8. Höijer, I., et al. (2025). Accurate characterization of CRISPR-Cas9 genome editing outcomes and mosaicism with near-perfect long reads. *bioRxiv*, 2025-09.