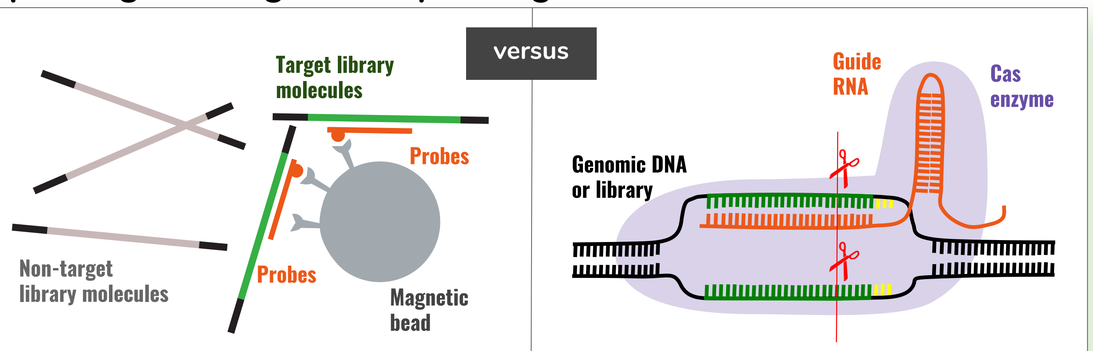
The CRISPR/Cas Toolkit for Targeted High-Throughput Sequencing:

Enriching complex regions and depleting unwanted elements

Jacob Enk[1], Brian Brunelle[1], Alison Devault[1], Benjamin Steil[1], Jean-Marie Rouillard[1,2]
1 Arbor Biosciences, Ann Arbor, MI, USA | 2 University of Michigan, Ann Arbor, MI, USA

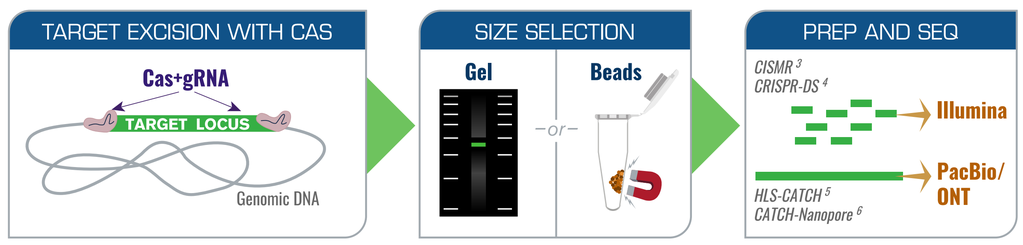
Established target enrichment technologies, like locus-specific PCR amplification and hybridization-capture, have transformed genomics research by reducing the cost of sequencing genomic targets by orders of magnitude. However, they are of limited use for resolving certain types of genomic loci, and can still be relatively expensive. The CRISPR/Cas system overcomes many of these challenges and has opened new dimensions in targeted sequencing capabilities. As a programmable, highly-specific nuclease system, **CRISPR/Cas can be used both to excise target regions from high-molecular weight genomic fragments, or to sever unwanted molecules in sequencing libraries and render them non-sequenceable.** Here we outline several applications of CRISPR/Cas-driven targeted sequencing, and highlight their versatility, affordability, and tremendous potential for highly customizable target enrichment and/or depletion prior to high-throughput sequencing on any platform. With Arbor Biosciences' scalable and efficient technology for producing large pools of guide RNAs, enriching or depleting thousands of target regions is both rapid and cost-effective.

Expanding the Targeted Sequencing Toolkit

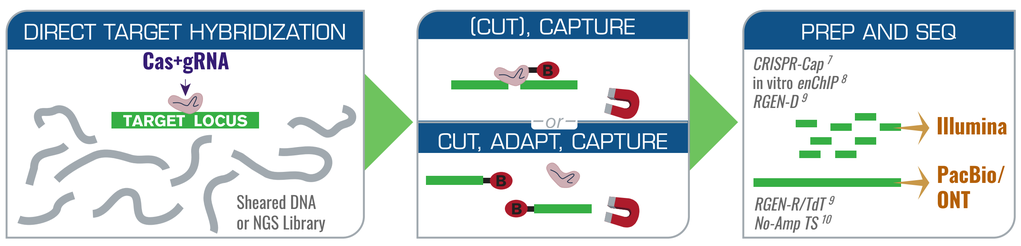


In-solution hybridization capture
Biotinylated oligo probes are hybridized to NGS library molecules, and then the hybrids are sequestered on streptavidin-coated magnetic beads. Target molecules must be denatured before probe hybridization, and probes must be tiled extensively across a target region since they are usually coupled with short-insert (<700 bp) NGS libraries.

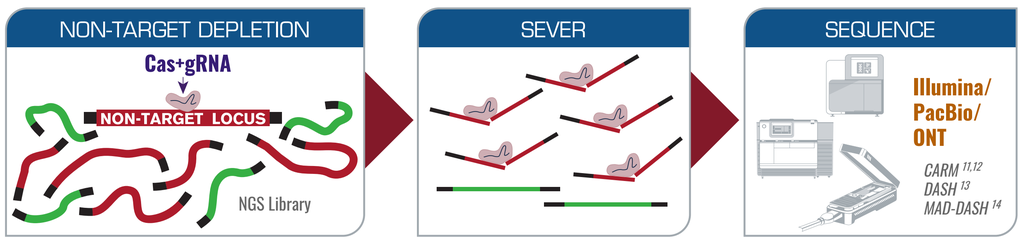
Targeted endonuclease system
Guide RNAs with a ~20 bp site-specific region are coupled with various types of Cas enzymes (forming an "RNP"), which hybridize to a target (typically dsDNA) and then usually cleave the target. The cleaved or RNP-hybridized products can then be manipulated in various ways to perform target enrichment or depletion.



Enrichment via dual-sided target excision. Beginning with high molecular weight genomic DNA, site-specific Cas-sgRNA complexes (RNPs) cut 5' and 3' of a target region or regions. Depending on the resulting fragment size, these excised molecules can then be isolated with pulsed-field gel electrophoresis (PFGE)[3,5,6] or SPRI beads[4], converted to a sequencing library, and sequenced on the desired platform.



Enrichment via hybridization to internal target regions. Following sample fragmentation (and sometimes adapter ligation), RNPs hybridize internal to a target region. The site is then either cleaved and tagged with a modified adapter[9,10], or a tagged RNP is left hybridized to the site, and the tagged complex is isolated with e.g. magnetic beads[7,8,9]. These fragments are then released, library-prepped, and sequenced.



Depletion via cutting of unwanted molecules. Site-specific RNPs hybridize with and then cleave molecules that are already converted to a sequencing library. Only the remaining non-cleaved molecules are sequenceable by downstream platforms, effectively enriching the libraries for non-cleaved molecules of interest. Extremely powerful, but particularly sensitive to off-target cleavage.

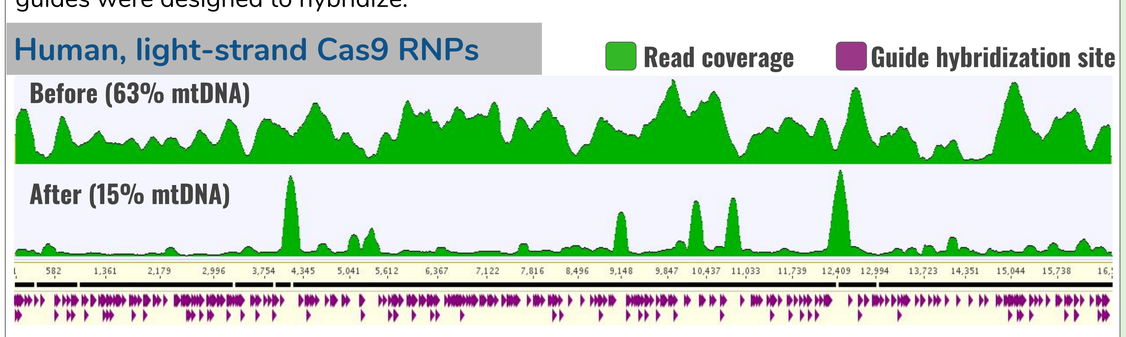
CRISPR-powered target depletion in action

In many cases, different biological nucleic acid samples can, once converted to NGS sequencing libraries, contain a high proportion of undesired molecules that consume sequencing space. For example, ATAC-Seq libraries can be dominated by mitochondrial DNA molecules, cDNA libraries with rRNAs, and gDNA libraries with repetitive DNA and/or adapter-dimers. Site-specific RNPs can be used to cleave these unwanted molecules, resulting in substantial enrichment of the remaining material with minimal ascertainment bias during sequencing.

Mitochondrial sequence depletion with CRISPR/Cas

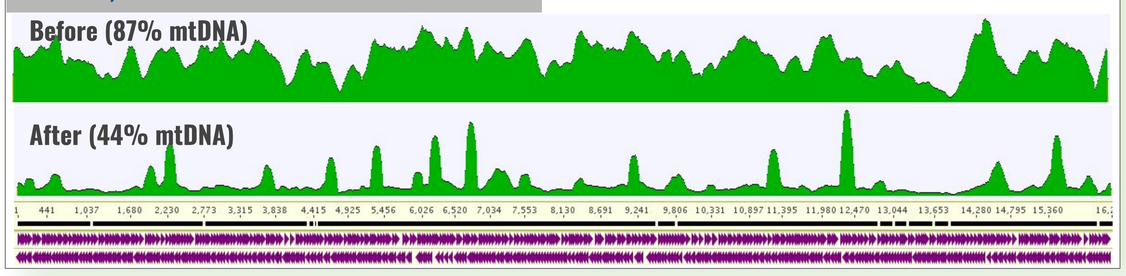
We built mtDNA-enriched human NGS libraries (64 to 92% mtDNA sequences) to simulate typical ATAC-Seq substrate. We then used 425 unique Cas9-gRNA complexes targeting only the light strand of the mitochondrial genome to sever the mtDNA sequences and render them non-sequenceable. After a few additional cycles of library PCR to swamp out the severed mtDNA remnant, we re-sequenced the libraries, which indicated **up to 85% mtDNA sequence deletion**, with coverage plots clearly indicating the lowest mtDNA coverage at regions where guides were designed to hybridize.

Human, light-strand Cas9 RNPs



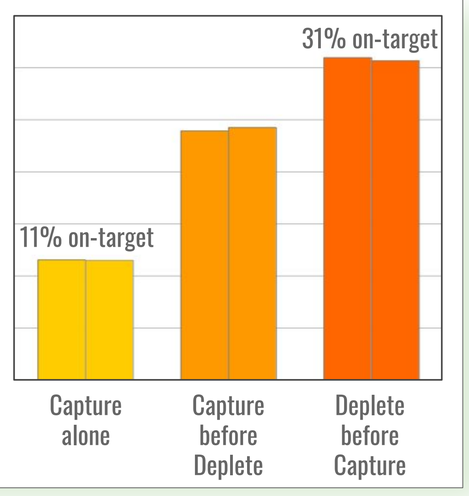
We also built mtDNA-enriched mouse libraries (73 to 97% mtDNA sequences) and then depleted the mtDNA with 741 unique Cpf1/Cas12a-gRNA RNPs targeting both strands of the mitochondrial genome. While Cas12a is known to have higher processivity than Cas9, it does require more specific guide-template hybridization to induce cleavage. Here depletion was not as thorough (**up to 49% depletion**) and areas of high remaining mtDNA coverage show no clear correspondence to guide hybridization 'gaps' across the target.

Mouse, both strands Cas12a RNPs



Combining CRISPR-powered depletion and hybridization capture

Depleting unwanted elements from a sequencing library prior to traditional target enrichment can significantly improve on-target rates. For example, we designed and manufactured a pool of more than 122,000 Cas9 guide RNAs targeting the most common sequence motifs in the transposable element repetitive portion of a large plant genome. We then combined our myNGS Guides™ depletion system with target enrichment using a myBaits® capture kit that targets various exons of that plant genome, and sequenced the result with NGS. On-target rates were more than double that of enrichment without depletion, and nearly triple when the depletion was performed prior to target enrichment.



REFERENCES

3. Bennett-Baker, P.E. & Mueller, JL. (2017). CRISPR-mediated isolation of specific megabase segments of genomic DNA. Nucleic Acid Research.
4. Nachmanson, D. et a. (2018). Targeted genome fragmentation with CRISPR/Cas9 enables fast and efficient enrichment of small genomic regions and ultra-accurate sequencing with low DNA input (CRISPR-DS). Genome Res.
5. Zhou et al., Sage Science https://nanoporetech.com/resource-centre/sagehls-catch-workflow-targeted-nanopore-sequencing.
6. Gabrieli, T. et al. (2018). Selective nanopore sequencing of human BRCA1 by Cas9-assisted targeting of chromosome segments (CATCH). Nucleic Acid Research.
7. Lee, J. et al. (2018). CRISPR-Cap: multiplexed double-stranded DNA enrichment based on the CRISPR system. NAR
8. Fujita, T. et al. (2016). Efficient sequence-specific isolation of DNA fragments and chromatin by *in vitro* enChIP technology using recombinant CRISPR ribonucleoproteins. Genes to Cells
9. Slesarev, A. et al. (2018). CRISPR/Cas9 Targeted Capture Of Mammalian Genomic Regions For Characterization By NGS. bioRxiv.
10. Tsai, YC. et al. (2017). Amplification-free, CRISPR-Cas9 Targeted Enrichment and SMRT Sequencing of Repeat-Expansion Disease Causative Genomic Regions. bioRxiv.
11. Wu, J. et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature.
12. Montefiori, L. et al. (2017). Reducing mitochondrial reads in ATAC-seq using CRISPR/Cas9. Scientific Reports.
13. Gu, W. et al. (2016). Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. Genome Biology.
14. Hardigan, A.A. et al. (2018). CRISPR/Cas9-targeted removal of unwanted sequences from small-RNA sequencing libraries. bioRxiv.

Check out Arbor's CRISPR-powered Targeted Sequencing Solutions!

BOOTH #310

myNGS Guides™