Linked Target Capture: rapid nucleic acid target enrichment with broad applications including cancer hotspot mutation identification, minimal residual disease (MRD), immune sequencing, single cell and metagenomics

W. Lloyd Ung[1], Jana Jajarmi[1], David Broemeling[1], Andre Marziali[1,2] and **Joel Pel** (joel@borealgenomics.com)[1]

[1] Boreal Genomics Inc, Vancouver, BC, Canada, [2] University of British Columbia, Vancouver, BC, Canada

I. Linked Target Capture (LTC)

LTC combines the rapid workflow and low input mass advantages of amplicon enrichment with the UMI integration and scalability of hybridization capture. LTC does this by using capture probes that are physically linked to universal PCR primers, called Probe Dependent Primers (PDPs), to simultaneously enrich and amplify molecules matching the probe sequences*.

- **Simple workflow:** Total library prep and target capture time is less than 6 hours and is easily automated
- **Scalable:** Single-tube panels have been demonstrated from <50bp to >1Mb, with exome-sized panels possible
- **Efficient:** High uniformity and on-target regardless of panel size, leading to reduced sequencing costs
- **Low level variant detection:** SNVs, indels, CNVs and translations are detected using duplex UMIs introduced during library prep, providing detailed molecular information including start/stop coordinates and enabling detection of unknown fusion pairs
- **Broad applications:** include liquid biopsy and minimal residual disease, immune sequencing, metagenomics, single cell sequencing and gene-editing characterization

II. Workflow

Ligation

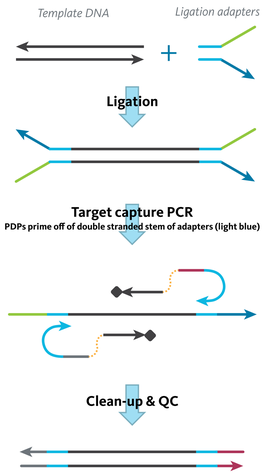
- Standard ligation with custom adapters and compatible with optional UMIs

Target Capture PCR

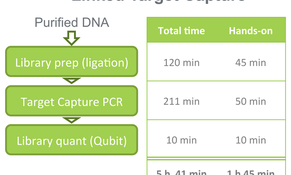
- Enrichment is achieved through proprietary Probe-Dependent Primers (PDPs) used in a PCR-like reaction
- Universal primer portion (light blue) of PDP only binds and extends if probe portion (black, blocked from extension) is bound to template
- Copies of each sense are collapsed into a single consensus read family at random to achieve low error rate (unlike Duplex Sequencing)
- Long captures and complicated pull downs of hybridization based capture are eliminated
- *Recent updates to the workflow have combined multiple PCR steps, reducing total workflow and hand-on time by an additional 30 minutes (see table below)*

Sequencing Clean-up & Quantification

- Library is cleaned up prior to quantification and then is ready for sequencing



Linked Target Capture



Total time	Hands-on
120 min	45 min
211 min	50 min
10 min	10 min
5 h 41 min	**1 h 45 min**

IDT xGen

Total time	Hands-on
125 min	50 min
80 min	50 min
285 min	50 min
105 min	95 min
150 min	40 min
10 min	10 min
12 h 35 min	**4 h 55 min**

III. Application: Oncology mutation detection and minimal residual disease (MRD)

Tumor-informed MRD testing with LTC

- A custom LTC panel was designed against all 25 variants from the SeraCare 'Complete Mutation Mix' ctDNA reference standard to simulate a patient-specific, or tumor-informed, MRD application of the technology
- LTC panels can be rapidly designed *in silico* using Boreal's proprietary machine-learning design process, minimizing turn-around time between tumor sequencing and assay implementation in clinical applications
- All four variant classes (SNVs, indels, CNVs and translocations) were covered in the assay design making LTC, to our knowledge, the only patient-specific method that has been demonstrated with all variant types

LTC assay performance

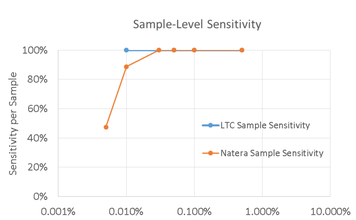
- On-target reads averaged >85% across 5kb of baits
- 100% of the panel was covered >0.2x of mean panel depth, demonstrating no dropped targets
- High on-target and uniformity combined with a small targeted panel enables efficient sequencer usage (~200k reads / target, ~5M reads / sample).

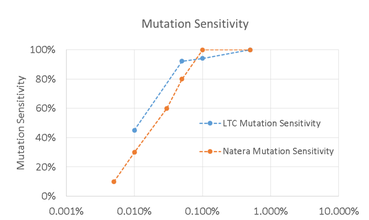
Variant Type	Specification	LTC Performance
SNVs & indels	Detection Limit	**0.01% (~1.5 molecules)** 50ng input
	Sensitivity	**98%** Detection of 0.1% variants
	PPV at 0.05%	**100%** Detection of 0.05% variants and above in 50ng WT input
	FPR at 0.05%	**0%** Detection of 0.05% variants and above in 50ng WT input
Translocations	Detection limit	**0.01% (~1.5 molecules)** 50ng input
CNVs	Detection limit	**4.5 copies/genome** 5% tumor fraction, 1ng input
All types	Sample-level sensitivity >=0.01%	**100%** 50ng input (2+ mutants detected)

LTC performance was evaluated using SeraCare and Horizon Diagnostics ctDNA standards over 24 samples and >400 variant positions as a function of variant frequency. In this study input mass was varied from 1 to 250 ng.

MRD sample-level sensitivity

- The LTC MRD panel demonstrated higher sample-level sensitivity than Natera Signatera's published data *(https://www.natera.com/wp-content/uploads/2020/11/Oncology-Clinical-Seeing-beyond-the-limit-Detect-residual-disease-and-assess-treatment-response-SGN_AV_WP.pdf)*
- Both panels were validated with SeraCare standards, considering 16 SNVs for Natera and 17 SNVs & indels for LTC (LTC performance is expected to improve further with the inclusion of translocations and CNV variants)
- LTC also demonstrated higher individual mutation sensitivity at low mutation abundance, suggesting that its higher sample-level performance would continue below 0.01% (note 0.01% represents detection of ~1.5 mutant molecules at 50 ng input mass)





IV. Application: Single cell targeted DNA sequencing

Targeted DNA sequencing

- LTC enrichment from 10x Genomics cDNA libraries has previously been demonstrated *(2020 AGBT poster, https://www.borealgenomics.com)*
- Single cell targeted DNA sequencing is more challenging than targeted cDNA sequencing and is typically limited to Multiplexed PCR, since single cell genomic DNA mass is too low for traditional target capture
- PCR methods are typically limited to a small number of amplicons and variant types providing an opportunity for LTC to demonstrate its unique advantages in capture of large target regions, recovery of all variant types and UMI integration
- To mimic targeted sequencing of a single cell, LTC was tested with decreasing amounts of gDNA to 10 pg (~1 diploid genome)

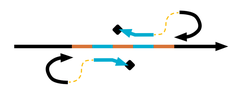
Performance Metric	1ng	100pg	10pg
On-target	95%	94%	95%
0.2x Uniformity	93%	96%	75%†
UMI conversion efficiency (including ligation loss)	40%	48%	>50%
Mutation detection sensitivity (SNVs and indels >5%)	100%	100%	–
Input genome equivalents *(No. molecules at 5% mutation)*	300 (*15*)	30 (*1.5*)	3 (*0.15*)

Performance metrics for LTC with decreasing input mass on the same 5kb panel as in MRD study. †Uniformity is low due to sampling variability from dilution, but is expected to remain high in a true single-cell workflow

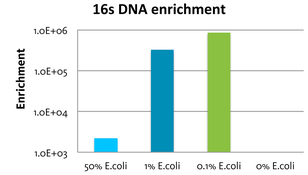
V. Application: Metagenomics

LTC for bacterial/viral sequencing & detection

- LTC enables capture of more diverse sequences than PCR as larger mismatches can be tolerated in priming regions and also allows accurate counting of targets through the use of UMIs
- As a demonstration, two LTC probes were designed against HV4 of the 16S rRNA gene, enabling broad identification of bacteria & archaea



In highly variable regions (blue and orange alternating), LTC probes are designed to bind to the blue sequence despite mismatches at 3' end (orange), while the universal adapters perfectly match the primers (black).

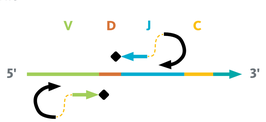


A simple demonstration shows ~1M-fold enrichment of E. coli DNA when the initial sample begins with 0.1% E. coli DNA in human DNA background. E. coli UMI reads are >99% of all reads across all samples, representing near 100% capture efficiency down to 0.1% E. coli.

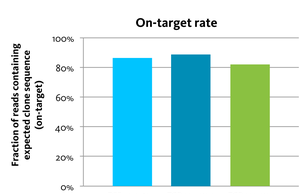
VI. Application: Immune repertoire sequencing

LTC advantages over existing V(D) J sequencing methods

- Compared to PCR workflows, LTC does not suffer from bias due to multiplexing or sample splitting, and allows the use of UMIs to evaluate library complexity while still providing a single day workflow compared to long (>2 day) capture workflows
- LTC probes were designed against VDJ target sequences and samples provided by the Pugh Lab (UHN - Toronto, ON), such that a fixed set of PDPs could be used to target all desired sequences in a single reaction
- Clones were successfully identified down to 0.007% using integrated UMIs



LTC probes are designed so that the forward PDPs target the V region and reverse PDPs target the J region



High on-target and uniformity were achieved due to PDP's high specificity. UMIs were used for molecular counting and consensus as well as evaluation of library complexity.

With reagent support from Integrated DNA Technologies Inc. (IDT)

Pel et al. PLOS ONE, Dec 2018. https://doi.org/10.1371/journal.pone.0208283