Technical Note

CleanPlex® Technology: Ultrafast, simple, and scalable targeted sequencing solution for Illumina® and Ion Torrent™

Introduction

Next-generation sequencing (NGS) has accelerated research efforts in diverse biological fields such as molecular biology, oncology, drug development, infectious diseases, reproductive and genetic health, agrigenomics and environmental metagenomics. Scientists routinely use NGS methods with different library preparation strategies to address their scientific questions and to decode the genomic information in their samples. NGS library preparation is therefore a fundamental and critical part of sequencing studies, as any information lost or bias introduced during the process will not be restored. The sequencing data is only as good as the quality of the NGS libraries generated[1].

While whole genome sequencing is preferred or required in many studies, such as in population genetics[2], a targeted sequencing approach is often more suitable, efficient and cost-effective in many scenarios[3]. Sequencing large number of complex genomes in their entirety can be prohibitive in terms of laboratory time and funding and can also place a substantial burden on the informatics infrastructure for storing and processing immense amount of data. A targeted sequencing approach selectively captures genomic regions of interest through an enrichment process for sequencing[3,4]. Target enriched NGS libraries are cheaper and faster to sequence, and resulting data that is less cumbersome to process and analyze.

There are two main enrichment strategies for targeted sequencing: amplicon- or PCR-based and hybridization capture-based enrichment (hybrid capture)[4]. Each strategy has its own advantages and disadvantages. Amplicon-based methods feature a simplified workflow and require smaller amounts of input DNA. This can be beneficial in a laboratory testing setting for streamlining operations and ensuring reproducibility, especially for processing of clinically-relevant samples, such as FFPE tissues and liquid biopsies which often have limited or low DNA amounts. However, amplicon-based assays are difficult to optimize, limited to small panels of a few hundred amplicons, and tend to have high PCR background and poor uniformity of coverage. Hybrid capture-based methods are effective for targeting large genomic regions, such as the whole exome, and perform better with respect to uniformity of coverage. On the downside, hybrid capture-based methods require large DNA inputs, as well as highly trained operators and specialized equipment. The protocols can be time-consuming, laborious to perform, and the results can often be highly variable due to the number of steps involved.

To overcome shortcomings in existing techniques and to bridge the gap between amplicon- and hybrid capture-based methods, scientists at Paragon Genomics have developed the CleanPlex® technology to provide a better tool to prepare NGS libraries for targeted sequencing. This technology is an ultra-high multiplex PCR-based target enrichment method that features a proprietary background cleaning chemistry, which allows tens of thousands of amplicons to be pooled in one reaction to target genomic regions as large as a few megabases in size. CleanPlex technology has a simple workflow, with a low input requirements and yet produces high performance NGS libraries that allow for efficient use of sequencing reads with accurate detection of variants. In this technical note, we present data to demonstrate the performance of CleanPlex technology.

Fast, streamlined target enrichment workflow

This technology has a streamlined single-tube workflow that is performed in 3 steps and in just 3 hours (Figure A):

(1) **Multiplex PCR (mPCR):** Targets of interest are amplified in a multiplex PCR reaction which uses primer pool(s) designed using our proprietary ParagonDesigner™ algorithm. The mPCR mix is formulated to minimize GC bias and improve uniformity.

(2) **Background cleaning:** Primer-dimers, non-specific PCR products, and complex molecular-debris are biochemically removed in a digestion reaction.

(3) **Indexing PCR:** Library is barcoded with platform specific sample indexes and amplified in a PCR reaction. The resulting libraries are ready for sequencing on either Illumina® or Ion Torrent™ sequencers.

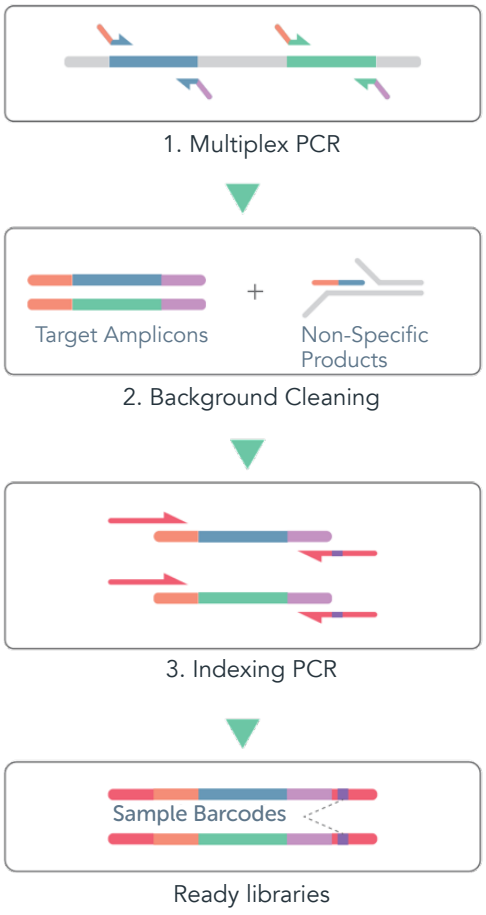


Figure A. CleanPlex technology. The chemistry of this technology features a simple 3-step workflow

PARAGON GENOMICS

High-quality NGS libraries powered by background cleaning

CleanPlex technology uses a unique background cleaning step to ensure that only sequences of interest are converted into NGS libraries, thus resulting in highly efficient use of sequencing reads. CleanPlex background cleaning chemistry is essential for eliminating undesirable by-products before creating sequencing-ready libraries (Figure B).

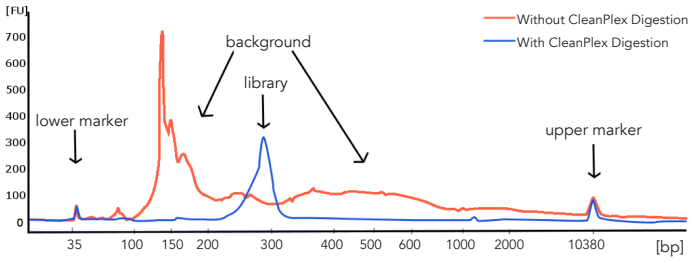


Figure B. Effective removal of PCR background. Libraries were prepared using the CleanPlex OncoZoom Cancer Hotspot Panel with (blue trace) or without (red trace) using the CleanPlex digestion reagent and examined using an Agilent® Bioanalyzer®. Without CleanPlex digestion, significant PCR background was formed, which would result in low mapping rate and poor on-target rate and require more sequencing reads to obtain adequate data. With CleanPlex digestion, nearly no background was generated and a sharp and clean library peak was produced in the Bioanalyzer trace.

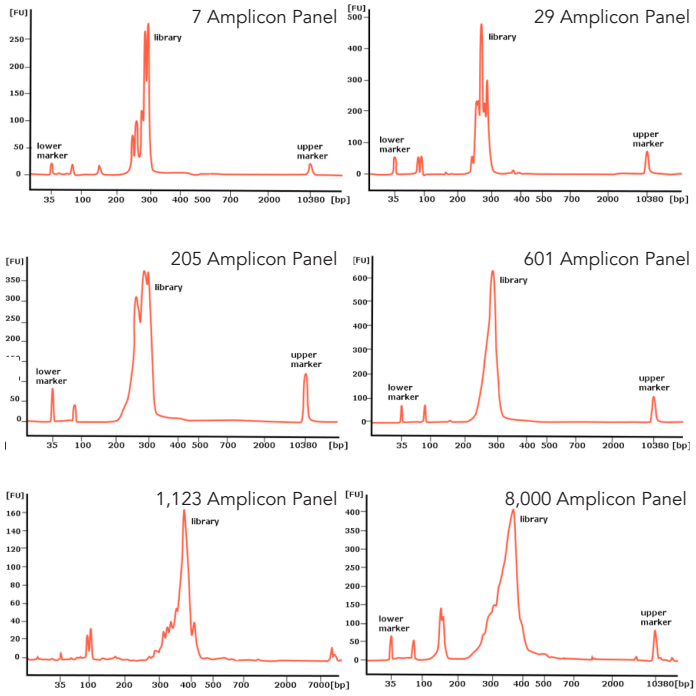
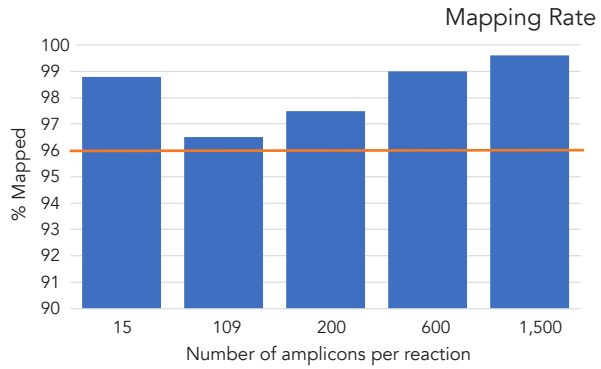
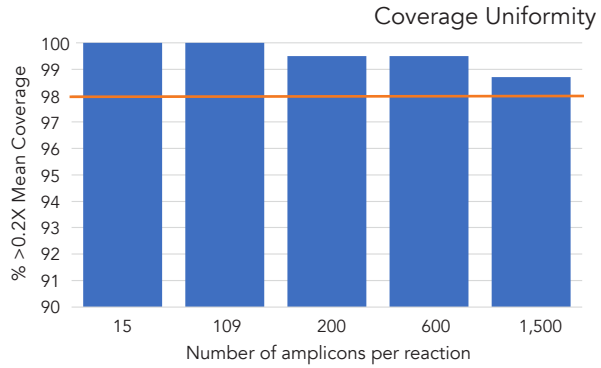


Figure C. Effective background removal for all amplicon multiplexing levels. Libraries were prepared with CleanPlex NGS Panels of varying sizes (7 to 8,000 amplicons) and examined on an Agilent Bioanalyzer. All libraries generated a clean peak indicating minimal formation of non-specific PCR products. The data shows that CleanPlex background cleaning chemistry is effective regardless of the level of multiplexing, which means new targets can be added without affecting performance.

Flexible and scalable NGS panels that can evolve to meet new challenges

CleanPlex background removal is highly effective regardless of the panel size (Figure C). As a result, high performance target-enriched NGS libraries are generated with high mapping rates, high uniformity of coverage, and high on-target rates (Figure D). CleanPlex NGS Panels can be designed to multiplex from 7 to more than 20,000 amplicons per reaction to interrogate hundreds of genes simultaneously. New gene and/or hotspot targets can be easily added without sacrificing performance, allowing the NGS assays to evolve to stay current to the latest discoveries.





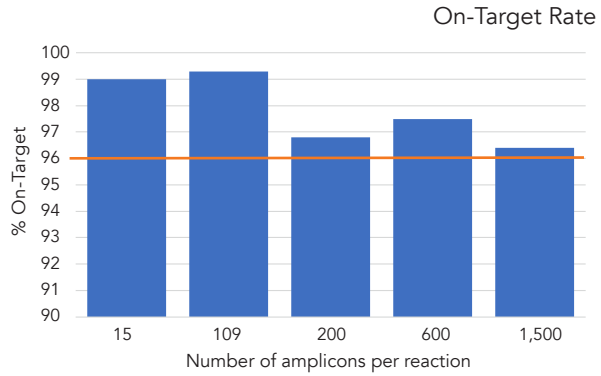


Figure D. High performance regardless of panel size. Libraries were prepared with CleanPlex NGS Panels of varying sizes (15 to 1,500 amplicons) and sequenced on an Illumina platform. **Mapping rate** was maintained above 96% and increased with the number of amplicons used per reaction, indicating that CleanPlex chemistry is even more effective on libraries prepared using larger numbers of amplicons. **Coverage uniformity**, measured as % covered by at least 0.2X mean coverage, was maintained above 96%. **On-target rate** was also higher than 96% for all panel sizes.

Uniform amplification of NGS libraries with minimal GC bias

CleanPlex technology's highly efficiently multiplex PCR-based chemistry allows hundreds to thousands of targets of varying GC content to be effectively targeted and uniformly amplified to generate high quality NGS libraries (Figure E). By minimizing GC bias, CleanPlex technology ensures high uniformity of coverage and sequencing efficiency.

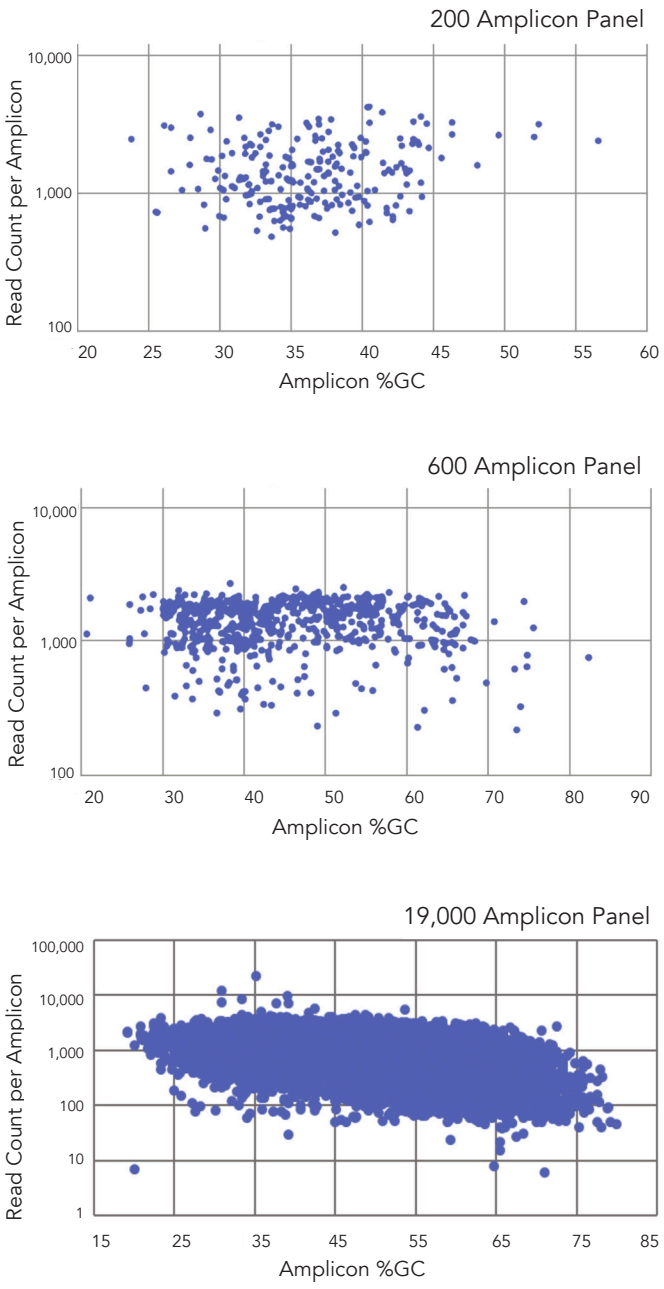


Figure E. Minimal GC bias across panels of different size. Libraries were prepared with CleanPlex NGS Panels containing 200, 600, and 19,000 amplicons and sequenced on an Illumina platform. For all three panels, amplicon read counts were evenly distributed across GC content, indicating minimal preferential amplification or dropout due to differences in GC content.

Cost-effective sequencing of target-enriched libraries

The combination of low PCR background, low GC bias and high performance (high mapping rate, high on-target rate and high coverage uniformity), translates directly to fewer sequencing reads being wasted on sequencing off-target and unmappable sequences. It also means that fewer sequencing reads are required to ensure all targets are covered at a minimally required depth to make confident base calls. When comparing a 207-amplicon panel amplified using CleanPlex versus a competing multiplex PCR chemistry (Figure F), we found that 60% fewer sequencing reads would be required to achieve the same quality of data using CleanPlex. This means 2.5X more samples can be sequenced on the same sequencing. Overall, CleanPlex NGS Panels allow efficient use of sequencing reads so that sequencing can be performed cost-effectively by allowing more samples to be sequenced at a time.

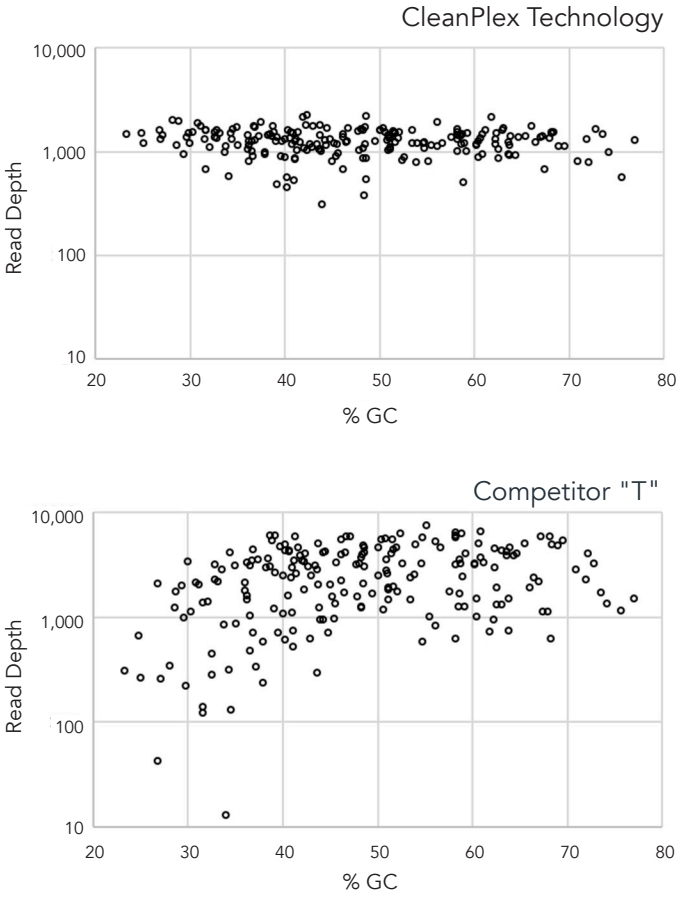


Figure F. High performance translates to cost-effective sequencing. For this comparison a 207-amplicon panel was used to generate target-enriched NSG libraries using either the CleanPlex or Competitor T's library preparation chemistry. Libraries were sequenced and analyzed for coverage uniformity across GC content. The results indicate that for a 207-amplicon panel, 60% less sequencing would be required using CleanPlex, which means 2.5X more samples can be sequenced on a flow cell. To achieve similar data quality, CleanPlex's mean read depth could be reduced to 600X coverage while Competitor T's would need to be increased to >1,500X coverage.

High performance NGS panels for every application

At Paragon Genomics, we understand that every researcher or assay developer has different challenges and requirements to meet. We offer the highest level of flexibility to our customers to support their specific applications. Most importantly, we deliver high quality custom NGS panels to help customers accomplish what they set out to do, whether it's answering a research question, replacing an existing laborious hybrid-capture workflow, or deploying a new NGS-based assay.

CleanPlex Custom NGS Panels have been successfully adopted by researchers and assay developers for a wide spectrum of applications, ranging from oncology and genetic health to phylogenomics[5] and CRISPR-based gene editing studies. Here we highlight some of the key applications covered by our CleanPlex Custom NGS Panels:

- Tumor Mutation Burden (TMB) profiling
- Early cancer detection via liquid biopsies
- Mitochondrial disease testing
- Non-Invasive prenatal screening (NIPS)
- Infectious disease research
- Immune repertoire profiling
- Agrigenomics high-throughput genotyping[5]

To demonstrate the performance of CleanPlex Custom NGS Panels, we randomly selected 15 panels from the hundreds of panels that we have delivered to customers and plotted their mapping rates and on-target rates (Figure G). Although these 15 panels were designed for different applications, they all produced high performance NGS libraries with mapping rates close to 100% and on-target rates above 90%.

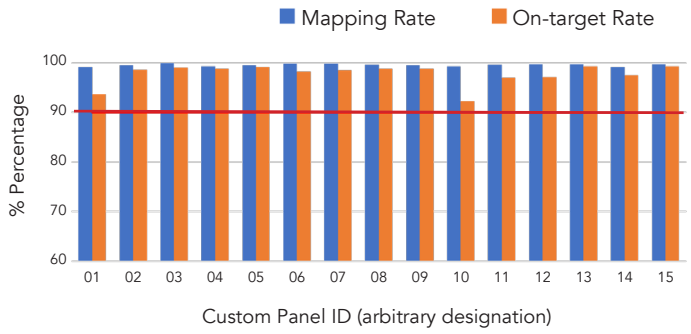


Figure G. High quality custom NGS panels. Fifteen CleanPlex Custom NGS Panels were randomly selected from Paragon Genomics' internal database, and their sequencing metrics were plotted to show the consistent performance delivered by our design team. All panels produced nearly perfect mapping rates and on-target rates well above 90%.

Conclusion

CleanPlex technology is an ultra-high multiplex PCR-based target enrichment technology for both Illumina and Ion Torrent NGS. It features a highly advanced proprietary primer design algorithm and an innovative, patented background cleaning chemistry. These give rise to important advantages that are critical to successful targeted sequencing, allowing researchers to break the limits of existing target enrichment technologies.

Features	Benefits
Low sample input (as low as 1 ng)	• Compatible with challenging samples (FFPE and/or cfDNA)
Tunable amplicon size	• Flexible assay options • Compatible with intact high quality gDNA and degraded DNA (FFPE and/or cfDNA)
High multiplexing capacity (> 20,000 amplicons in one reaction)	• Interrogate more genes and loci in a single reaction • Scalable assay development as new targets can be added without affecting protocol and performance
Streamlined, fast, single-tube (3 hrs total time with 30 mins hands-on time)	• Single-tube workflow reduces errors and preserves sample complexity • Save time and get sequencing-ready libraries quickly
High uniformity of coverage	• Fewer sequencing reads needed to cover all targets at required depth • Cost effective sequencing • More sample pooling on each of the sequencing runs
High specificity	• Utilize sequencing reads effectively without sequencing off-target regions
High sensitivity	• Successful target amplification with maximum coverage and minimum dropouts

Table H. Features and benefits of CleanPlex NGS Panels

References

1) Library construction for next-generation sequencing: Overviews and challenges. Head et al. Biotechniques. (2014) 56(2): 61-77
2) Whole-genome sequencing for an enhanced understanding of genetic variation among South Africans. Choudhury et al. Nat. Commun. (2017) 8:Article 2062
3) Target-enrichment strategies for next-generation sequencing. Mamanova et al. Nat. Methods. (2010) 7(2):111-8
4) Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing. Samorodnitsky E et al. Hum Mutat. (2015) 36(9):903-14
5) HiMAP: Robust phylogenomics from highly multiplexed amplicon sequencing. Dupuis JR et al. (2018) Mol Ecol Resour. 18(5):1000-1019

Discover more with less™

Paragon Genomics, Inc. | 3521 Investment Blvd Suite 1, Hayward CA 94545, USA | + 1.650.822.7545
www.paragongenomics.com | techsupport@paragongenomics.com

PARAGON GENOMICS