ThruPLEX HV outperforms NEBNext Ultra II for DNA-seq of cell-free and FFPE DNA

Introduction

A next-generation sequencing (NGS) library preparation system must have a simple, streamlined workflow accommodating a range of sample inputs without compromising accuracy. ThruPLEX DNA-Seq HV satisfies these requirements, with a complete, fast, and accurate NGS library preparation system that enables the generation of complex libraries. The kit's unique single-tube workflow makes this one of the fastest, most consistent, and highest-throughput library preparation workflows on the market. Additionally, the use of a single tube helps prevent loss of precious samples by eliminating the need for time-consuming bead purification found in competitor kits (Figure 1). Furthermore, ThruPLEX DNA-Seq HV decreases hands-on time by eliminating the need for adapter dilution and protocol optimization (Figure 1).

By increasing the starting input volume (30 µl) and range of input (up to 200 ng), we have eliminated the need for concentrating inputs while producing high-complexity libraries. The libraries generated by ThruPLEX DNA-Seq HV can be used directly for whole-genome sequencing applications or enriched using a custom panel for the leading target enrichment platforms.

A

	ThruPLEX DNA-Seq HV	Kapa Hyper Prep	NEBNext Ultra II
Hands-on time	15 min	20 min	20 min
Total time	2.4–2.6 hr	2.5–2.7 hr	3.1–3.2 hr
Single-tube workflow	Yes	No	No
Adapter dilution	No	Yes	Yes
Intermediate cleanup	No	Yes	Yes
Post-ligation size selection	No	No	Yes (>100 ng)

B

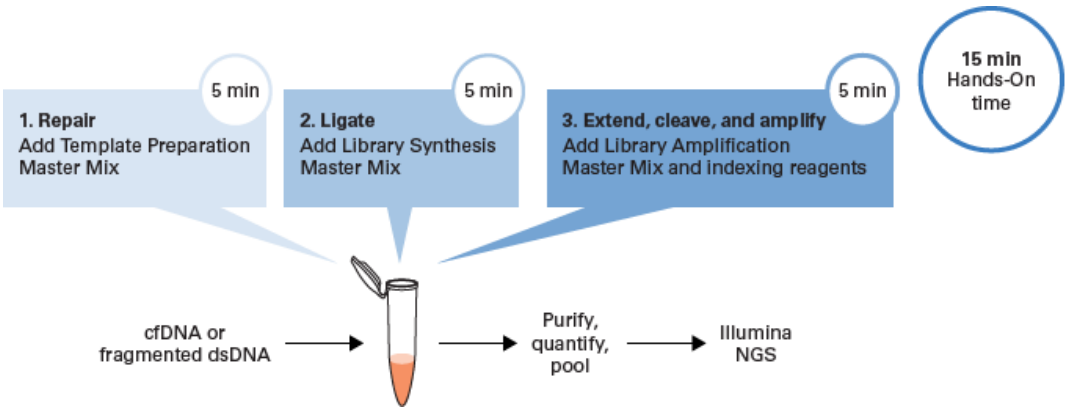


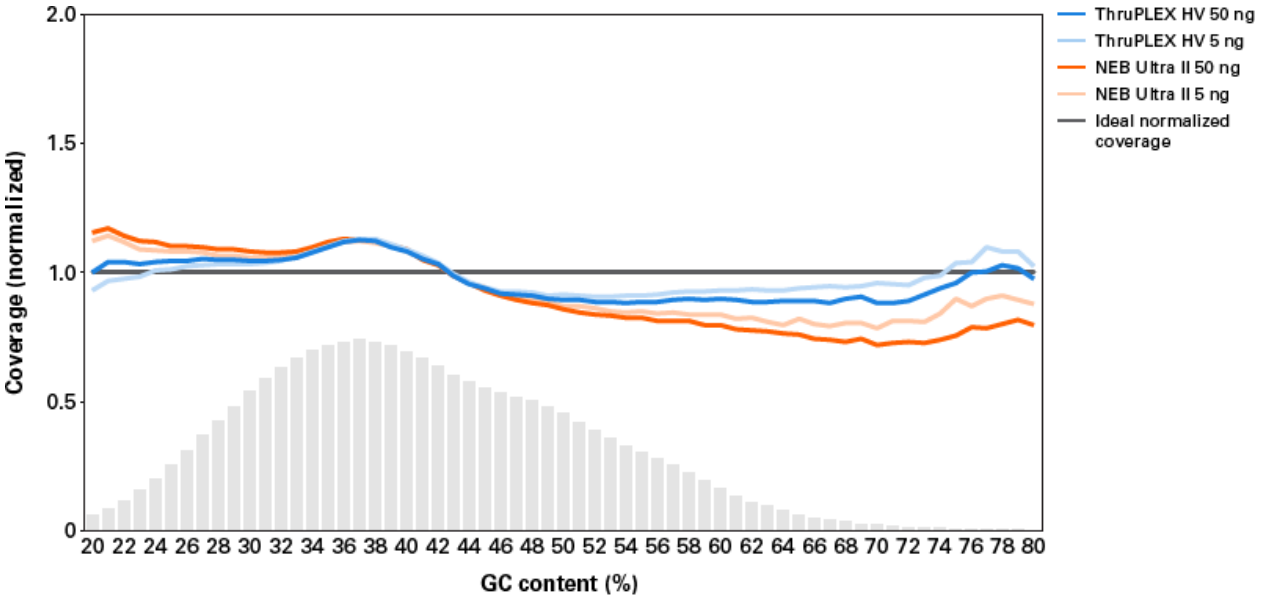
Figure 1. Comparison of three leading NGS library preparation chemistries. Panel A. Total time is representative of the time required to amplify inputs of 200 and 5 ng with each chemistry to yield enough Illumina-compatible dual-indexed libraries for target enrichment. ThruPLEX DNA-Seq HV is the only single-tube workflow and the only chemistry which does not require adapter dilution, intermediate cleanup, or post-ligation size selection. This allows ThruPLEX DNA-Seq HV to provide the quickest path to sequenceable libraries with the least amount of hands-on time. **Panel B.** The ThruPLEX HV workflow consists of three simple steps that take place in the same well or PCR tube, eliminating the need to purify or transfer the sample material.

Results

Competitive library coverage uniformity

The ThruPLEX DNA-Seq HV workflow consists of three simple addition steps that take place in a single well or PCR tube—with just 15 minutes of hands-on time—to yield indexed libraries from fragmented DNA within two hours. Increasing the input volume (30 µl) and range of input (up to 200 ng) at the start of the protocol enables the generation of higher-complexity libraries for sequencing or target enrichment. Generation of high-complexity libraries is critical to achieving even coverage throughout the genome for whole-genome sequencing (Figure 2). When compared to NEBNext Ultra II, libraries generated using ThruPLEX DNA-Seq HV show coverage much closer to ideal normalized coverage.

A



B

	Input	Total reads	Reads aligned	Duplicate
ThruPLEX HV	50 ng	7,868,884	97%	0.73%
	5 ng	7,796,764	97%	0.84%
NEBNext Ultra II	50 ng	7,922,699	97%	0.98%
	5 ng	7,978,001	97%	0.89%

Figure 2. Superior coverage uniformity. Panels A and B. Libraries were prepared in triplicate from 5-ng and 50-ng inputs of a quantitative multiplex reference standard consisting of gDNA pooled from HCT116, RKO, and SW48 cell lines (Horizon Discovery). Libraries were generated following ThruPLEX DNA-Seq HV (ThruPLEX HV) or NEBNext Ultra II (NEBNext) protocols. Paired-end sequencing was performed on a NextSeq® 500/550 Mid Output Kit v2.5 (150 Cycles), and total reads were downsampled to 8 million total reads. The vertical gray bars represent the expected GC content distribution using 100-bp windows.

Target enrichment with FFPE and formalin-compromised inputs

Processing of clinical research materials for long-term storage may include fixation with formalin. Exposing samples to formalin can lead to significant damage of the nucleic-acid content, which is often present in limited quantities. Due to this damage, the construction of libraries can prove challenging and, therefore, generally requires a kit with a robust repair mechanism to produce enough post-PCR product for target enrichment. ThruPLEX DNA-Seq HV was designed to accommodate a large input volume and a higher amount of starting material, which improves coverage and mutation detection by increasing the complexity of the input. Another important consideration for the confident calling of low-frequency mutations is achieving even coverage throughout the genome in order to ensure optimal read depth at all relevant loci. To facilitate the necessary even coverage, our system has been optimized for improved coverage uniformity across a broad range of inputs with varying

levels of damage and GC contents.

We compared the performance of ThruPLEX DNA-seq HV and NEBNext Ultra II (Figure 3) on libraries generated in triplicate with 50-ng and 5-ng inputs of Horizon DNA references, including formalin-compromised material with severe damage, as well as with 30-ng and 5-ng of formalin-fixed paraffin-embedded material. ThruPLEX HV libraries outperformed NEBNext libraries in mean target coverage for the formalin-compromised DNA and exhibited comparable mean target coverage for the FFPE samples (Figures 4). Both kits performed similarly in the detection of positive variants, on the two types of formalin-treated samples (Figure 5).

Input type	Chemistry	Input amount	Reads	Reads aligned	Bases aligned	On-target bases
Formalin compromised (severe)	ThruPLEX DNA-Seq HV	50 ng	4.93×10^6	98.0%	3.6×10^8	1.8×10^8
		5 ng	4.86×10^6	97.6%	3.4×10^8	1.3×10^8
	NEBNext Ultra II	50 ng	5.00×10^6	98.7%	3.1×10^8	1.6×10^8
		5 ng	5.00×10^6	98.6%	3.1×10^8	1.2×10^8
Formalin-fixed paraffin-embedded	ThruPLEX DNA-Seq HV	30 ng	4.85×10^6	97.6%	3.0×10^8	8.7×10^8
		5 ng	4.66×10^6	97.0%	2.9×10^8	8.2×10^8
	NEBNext Ultra II	30 ng	5.00×10^6	98.3%	3.1×10^8	8.9×10^8
		5 ng	5.00×10^6	98.3%	3.1×10^8	9.0×10^8

Figure 3. Excellent target-capture efficiency across a range of DNA quality. Libraries were generated in triplicate with 50-ng and 5-ng inputs of Horizon DNA references using formalin-compromised material with moderate and severe damage (Cat. # HD803), as well as with 30-ng and 5-ng inputs of formalin-fixed paraffin-embedded material (Cat. # HD200). Libraries were amplified with ThruPLEX DNA-Seq HV chemistry or NEBNext Ultra II. Furthermore, on-target efficiency was roughly equivalent between the two kits (data not shown).

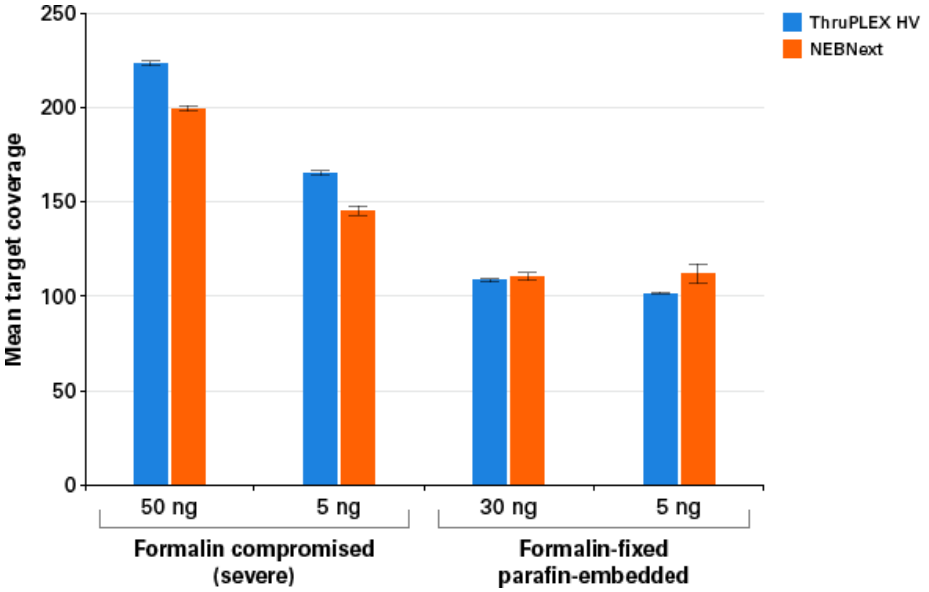


Figure 4. ThruPLEX HV has better mean target coverage for formalin-compromised and FFPE-treated references than NEBNext Ultra II. Libraries were constructed in triplicate with ThruPLEX DNA-Seq HV and NEBNext Ultra II using formalin-compromised material with moderate and severe damage, as well as with formalin-fixed paraffin-embedded material. Amplified libraries were purified with AMPure beads and pooled for target capture with IDT's xGEN Pan Cancer Panel covering 800 kb of the human genome. Paired-end sequencing was performed on a NextSeq 500/550 Mid Output Kit v2.5 (150 Cycles). Each library was downsampled to 5 million total reads and aligned to hg19 using bowtie2. Variants were called using VarDict Variant Caller and sorted to include a minimum of 30X coverage and allele frequencies above 0.5%.

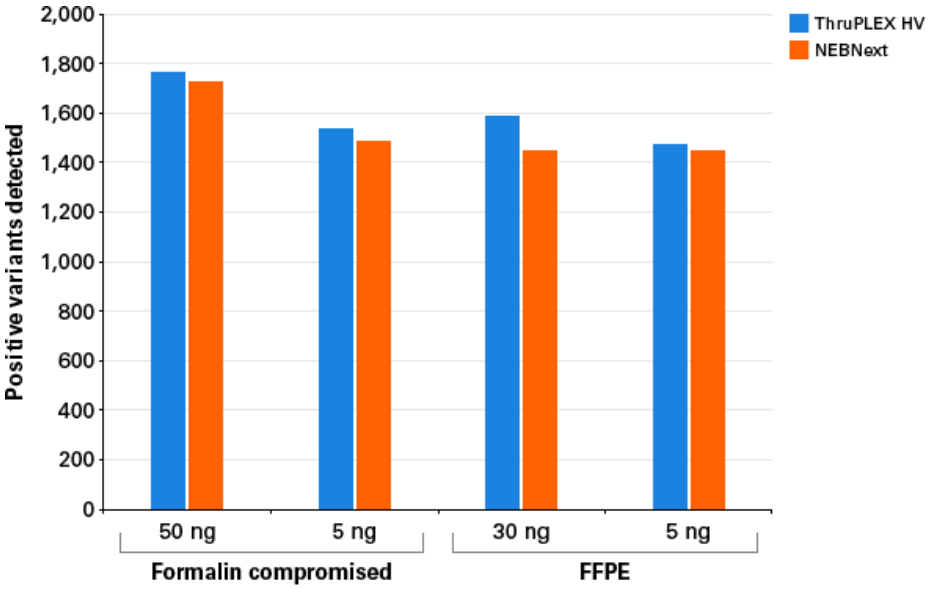


Figure 5. ThruPLEX HV detects more mutations in FFPE samples than NEBNext Ultra II. Libraries were constructed in triplicate with ThruPLEX DNA-Seq HV and NEB Next Ultra II using formalin-compromised material with moderate and severe damage, as well as with formalin-fixed paraffin-embedded material. Targets were enriched using the xGEN Pan Cancer Panel (IDT). A positive detection occurs when at least two of three replicates positively detect a variant.

Target enrichment with cell-free DNA

Sequencing of cell-free DNA (cfDNA) faces similar challenges to that of formalin-compromised and FFPE samples. We, therefore, set out to test the performance of ThruPLEX DNA-Seq HV with wild-type cfDNA and cfDNA containing eight confirmed single-nucleotide variants (SNVs) occurring at a 5% allelic frequency. While mean target coverage was comparable between ThruPLEX HV and NEBNext Ultra II libraries (Figure 6), ThruPLEX HV detected more positive variants in both the wild-type cfDNA and cfDNA reference containing confirmed SNVs (Figure 7).

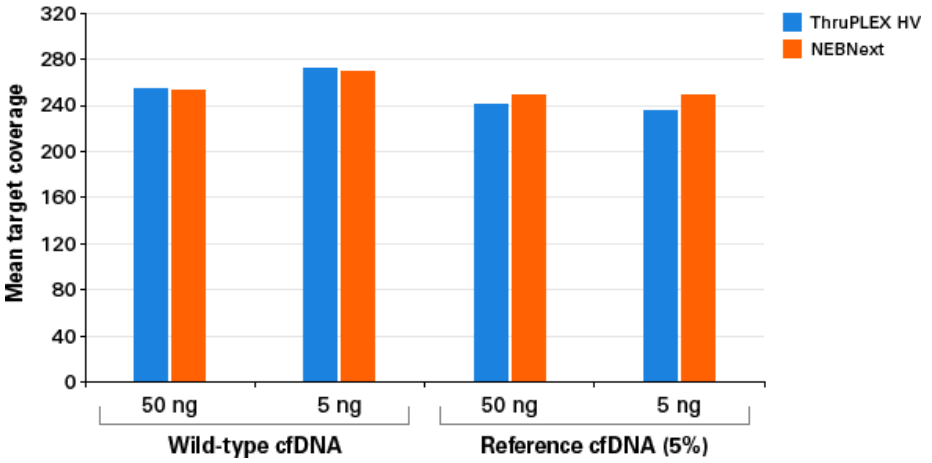


Figure 6. Mean target coverage is comparable for ThruPLEX HV and NEBNext Ultra II libraries. Libraries were constructed in triplicate with ThruPLEX DNA-Seq HV and NEBNext Ultra II using 50 and 5 ng of a wild-type multiplex cfDNA reference standard (Horizon Discovery; Cat. # HD776) and a reference standard with eight additional mutations occurring at a 5% allelic frequency. Targets were enriched using the xGEN Pan Cancer Panel (IDT).

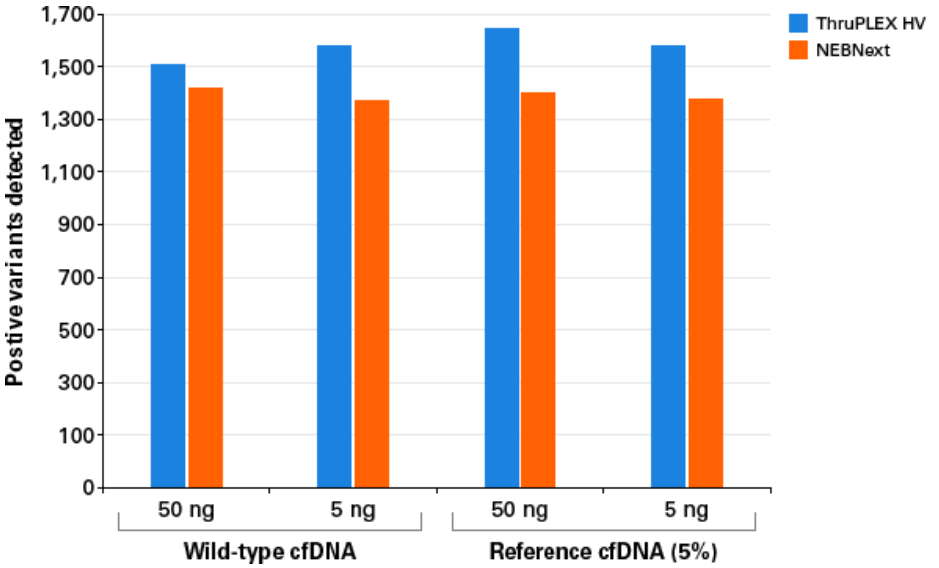


Figure 7. ThruPLEX HV detects more positive variants in wild-type cfDNA and reference cfDNA with SNVs confirmed 5% allele frequencies. Libraries were constructed in triplicate with ThruPLEX DNA-Seq HV and NEBNext Ultra II using 50-ng and 5-ng inputs of a wild-type multiplex cfDNA reference standard and a reference standard with eight additional mutations at a 5% allelic frequency. Targets were enriched using the xGEN Pan Cancer Panel (IDT). A positive detection occurs when at least two of three replicates positively detect a variant.

Conclusion

ThruPLEX DNA-Seq HV is a simple, fast, accurate system with only three addition steps that can be completed in a single tube in two hours. The ThruPLEX DNA-Seq HV library preparation kit for Illumina sequencing elevates the ThruPLEX family by increasing starting input volume and expanding the amount of starting material when compared to previous ThruPLEX DNA-seq kits. Along with these improvements, ThruPLEX DNA-Seq HV retains the coveted ThruPLEX single-tube workflow with no intermediate cleanup steps. Through workflow optimization and reformulation, ThruPLEX DNA-Seq HV outperforms NEBNext Ultra II in coverage of regions with increasing GC content as well as in the detection of variants in both FFPE and cell-free DNA.

Methods

DNA preparation

Human genomic DNA from Horizon Discovery (Cat. #s HD701, HD803, and HD200) was sheared on a Covaris M220 following the 250-bp shearing protocol. Sheared input material and cfDNA material not requiring shearing (Horizon Discovery, Cat. #s HD776 and HD777) were evaluated for correct size on an Agilent 2100 BioAnalyzer using Agilent High Sensitivity DNA Reagents. The concentration of these samples was measured using a Qubit 2.0 Fluorometer with the Quant-iT dsDNA Assay kit, high sensitivity (Thermo Fisher Scientific).

Library preparation

Libraries were prepared following the manufacturer's instructions using the ThruPLEX DNA-Seq HV kit or NEB Next Ultra II kit. All libraries were generated using dual indexes. Amplified libraries were purified using AMPure XP beads (Beckman Coulter) and eluted in low-TE buffer for whole-genome sequencing (WGS). Purified library size was assessed on the Agilent 2100 BioAnalyzer using Agilent High Sensitivity DNA Reagents. Libraries were quantified by qPCR using the Library Quantification Kit (Takara Bio, Cat. # 638324) or Qubit 2.0 Fluorometer with the Quant-iT dsDNA Assay kit, high sensitivity (Thermo Fisher Scientific).

Target capture

Amplified libraries were purified with AMPure beads and pooled for target capture with the IDT's xGEN Pan Cancer Panel covering 800 kb of the human genome.

Illumina sequencing

Quantified post-PCR libraries were pooled and loaded onto an Illumina NextSeq 500/550 v2.5 flow cell for sequencing. Libraries were loaded

following Illumina recommended loading concentrations.

Data analysis

Raw sequencing reads were downsampled to equal numbers across all samples using seqtk (v1.3-r106) and quality processed to remove adapters and low-quality bases using trimmomatic (v0.36). Quality processed reads were aligned to the UCSC hg19 reference genome with bowtie2 (v2.3.4.1) with default parameters. Resulting SAM files were coordinate sorted using Picard SortSam (v2.18.3) and converted to BAM files with samtools view (v1.8). Duplicate reads were identified and marked from sorted BAM files with picard MarkDuplicates (v2.18.3) and used as input to collect alignment, insert size, GC bias, and various WGS metrics with Picard AlignmentSummaryMetrics (v2.18.3), Picard CollectInsertSizeMetrics (v2.18.3), Picard CollectGcBiasMetrics (v2.18.3), and Picard CollectWgsMetrics (v2.18.3), respectively. Variants were called using VarDict with a minimum of 30X coverage and a minimum of 0.5% allele frequency.

Related Products

Cat. #	Product	Size	License	Quantity	Details
R400740	ThruPLEX® DNA-Seq HV	96 Rxns		*	
	ThruPLEX DNA-Seq HV uses a simple, three-step workflow to generate high-complexity DNA libraries from standard samples and challenging sample sources such as FFPE and cell-free plasma DNA. This product contains reagents for 96 reactions and includes unique dual indexing (UDI) primers.				
	Documents | Components				
R400741	ThruPLEX® DNA-Seq HV	24 Rxns		*	

Add to Cart

Takara Bio USA, Inc.
United States/Canada: +1.800.662.2566 • Asia Pacific: +1.650.919.7300 • Europe: +33.(0)1.3904.6880 • Japan: +81.(0)77.565.6999