



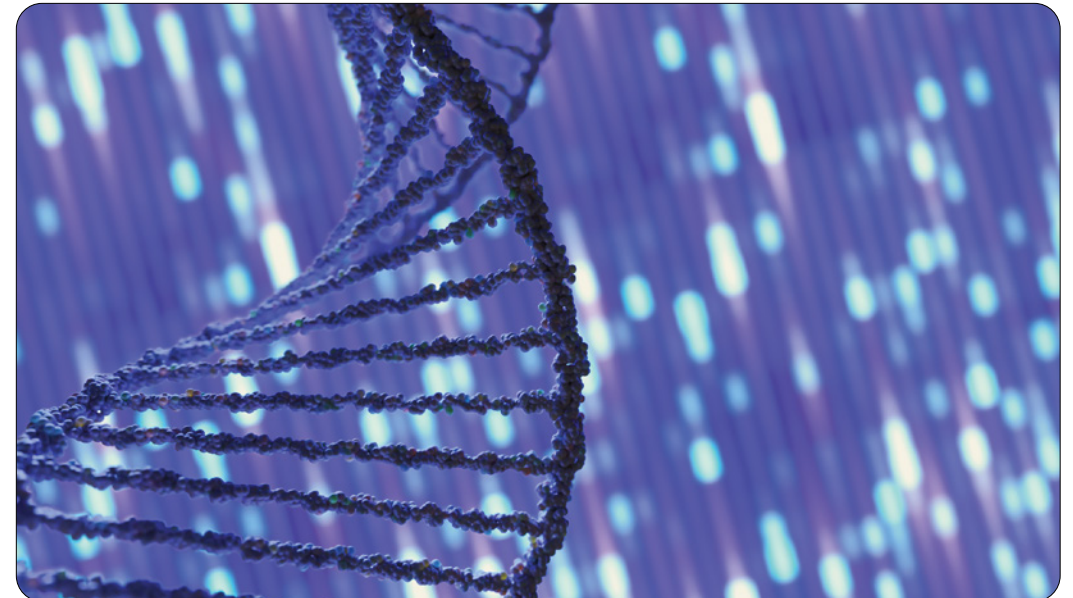
Decoding library complexity

for quantitative sequencing and rare variant analysis.

NEXTFLEX™ UDI-UMI Adapters

- Set of 96 full length UDI (10 bp) - UMI (9 bp) barcodes for multiplexing NGS libraries for Illumina® and Element® sequencers
- Compatible with Revvity library prep and 'T-A' overhang (Illumina® TruSeq® style barcodes) ligation-based workflows
- Compatible with PCR-free workflows
- Every lot is functionally verified and tested for index purity by sequencing
- Available in 8, 24, 48 and 96 sample formats

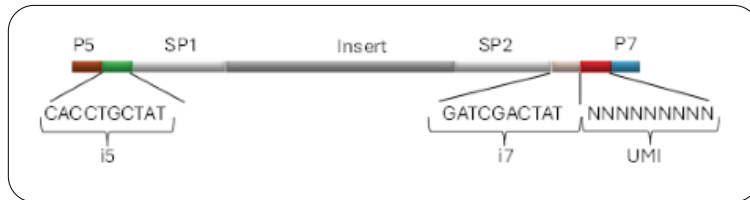
The NEXTFLEX™ UDI-UMI Adapters consist of 96 unique dual-index (UDI) barcodes, each with a 10-base pair UDI and a 9-base pair unique molecular identifier (UMI). These adapters are designed for multiplexing NGS libraries on Illumina® and Element® sequencers. Unique Dual Indexes (UDIs) and Unique Molecular Identifiers (UMIs) are nucleotide sequences or "barcodes" that are incorporated during library preparation for Next-Generation Sequencing (NGS) and provide several advantages.



Unique Dual Indexes (UDIs)

Revvity's 10 bp UDI (represented as i5 and i7 below) are incorporated on both ends of an NGS library molecule, enabling researchers to sequence multiple samples in parallel, which is often referred as multiplexing.

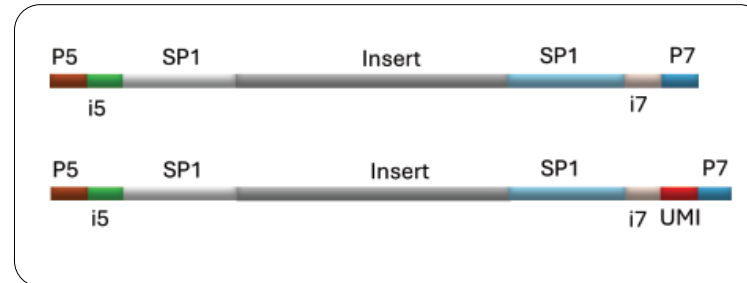
The presence of two indices minimizes the effects of index hopping, a phenomenon where a read is assigned to the wrong index as consequence of errors that appear during sequencing process.



Example of a library containing NEXTFLEX® UDI-UMI Adapters, where Insert = DNA or RNA fragment from a sample, P5 and P7 = Flow cell binding sites for Illumina® platforms, SP1 and SP2 = Binding sites for sequencing primers, i5 and i7 = Short pair of 10 bp sequences used to identify a particular sample (UDI), and UMI = 9 bp sequence used to uniquely tag each molecule within a library.

Unique molecular identifiers (UMIs)

A single UMI is incorporated on each library. Revvity's UMI is composed of 9 random bases, providing hundreds of thousands of combinations to uniquely tag each molecule in a sample library.



Average base percentage found across the UMI found in 24 libraries constructed with different inputs of human gDNA (0.1-1000 ng) using the NEXTFLEX Rapid XP v2 DNA-seq Kit and sequenced on an Illumina® MiSeq® system.

UMIs are also referred as “Molecular Barcodes” or “Random Barcodes”. The presence of UMI allows differentiation of PCR duplicates (which contain same UMI sequence) from true copy number (each copy contains a different UMI sequence). They are an excellent tool to assess library complexity. UMI sequence information in conjunction with alignment coordinates enables grouping of sequencing data into read families representing individual sample DNA or RNA fragments. UMIs are used mainly in the context of quantitative sequencing and rare variant analysis.

Quantitative sequencing

When starting from low input samples, stochastic effects in the first rounds of the PCR and the biases of the PCR polymerases against GC content can potentially cause erroneous quantitation data. Removal of PCR duplicates using alignment coordinate information is known to be inefficient for low input situations. It is also known to be inefficient for deep sequencing of RNA-seq or ChIP-Seq libraries (>80 million reads per sample). In the latter case alignment coordinate-based de-duplication will remove large numbers of biological duplicate reads, especially for the most abundant transcripts. As indicated above, UMIs alleviate the PCR duplicate problem by adding unique molecular tags to the sequencing library molecules before amplification.

Rare variant analysis

NGS sequencing provides data with low error rates (~0.1 to 0.5%) for most applications. These low error rates interfere with the confident identification of low abundance variants. UMI-less data can't distinguish between these and sequencing errors. By incorporating individual barcodes on each original DNA fragment, variant alleles present in the original sample (true variants) can be distinguished confidently from errors introduced during library preparation, target enrichment, or sequencing. Any identified errors can be removed by bioinformatics methods before final data analysis. Applications include sequencing of heterogeneous tumor samples, cfDNA sequencing including ctDNA sequencing and deep exome sequencing.

The incorporation of full-length UDI and UMI sequences by ligation instead of PCR reduces the possibility of introducing errors and allows taking advantage of these features in PCR-free workflows.

For enhanced variant calling, gene expression analysis, sequencing of heterogeneous tumor samples, ctDNA sequencing, deep exome sequencing, single-cell RNA-seq, and haplotyping via linked reads, we suggest utilizing the NEXTFLEX UMI-UDI barcodes. In summary, UMIs provide error correction, enhance variant calling, and improve overall data in genomics research.

revvity