

The only software to provide comprehensive and end-to-end support for multiple **Unique Molecular Identifier** protocols

UMI Support in DNA-Seq

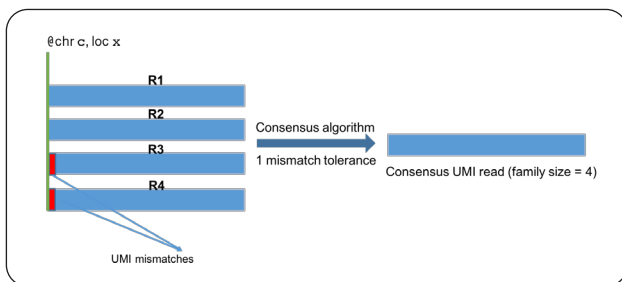
Strand NGS supports the alignment and analysis of reads containing Unique Molecular Identifiers (UMIs). You can start from Aligned BAM with UMIs (BC tag) or fastq with UMIs. The DNA-Seq workflow for UMIs in Strand NGS includes the following steps.

1. Import and align UMI reads.

Reads containing UMIs are imported into Strand NGS after identifying the UMI protocol. The UMI segment in each read is isolated before aligning it to the reference.

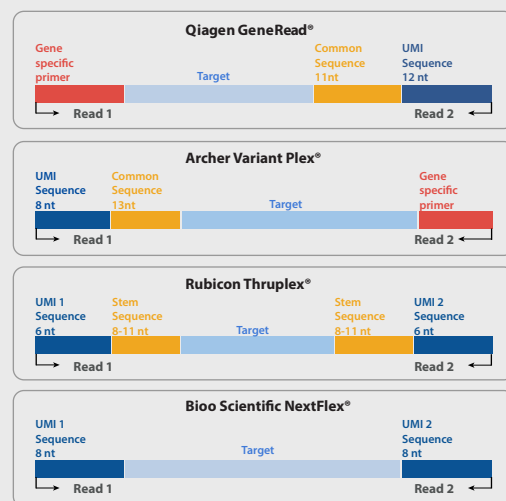
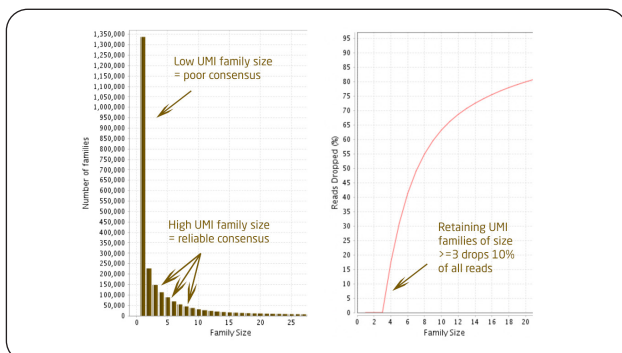
2. Create UMI consensus.

After sorting, UMI families are identified. Each family is replaced by a consensus read, generating the consensus aligned read list. Each read in the consensus list contains the UMI family size as well as information about its UMI and that of its mate.



3. UMI Quality Control.

Following alignment and consensus read generation, a set of quality control (QC) plots visualize family size and UMI distributions.



+Fully supported custom UMI protocols

4. UMI filters.

The QC plots can be used to assist filtering reads less than a specified family size, as well as reads with infrequently occurring UMIs.

This filter removes family with the specified family size threshold.

Choose Read List:

Remove reads with family size below:

This filter allows only reads whose Unique molecular barcode is in top K UMIs ordered by their read counts in descending order.

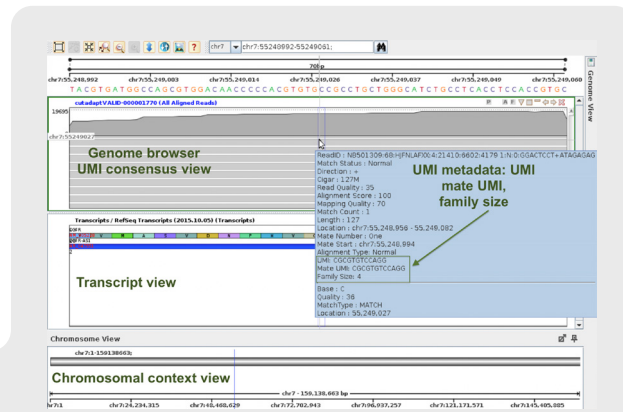
Choose Read List:

Remove reads with UMI not in top:

5. Low-frequency variant calling.

A low-frequency variant caller calls variants on the filtered read list.

	G	T	C	C	A	G	C	T	G	T	A	Size
+	+	+	+	+	+	+	+	+	+	+	+	7251
-	+	+	+	+	+	+	+	+	+	+	+	70
-	+	+	+	+	+	+	+	+	+	+	+	29
-	+	+	+	+	+	+	+	+	+	+	+	22
-	+	+	+	+	+	+	+	+	+	+	+	16
-	+	+	+	+	+	+	+	+	+	+	+	15
-	+	+	+	+	+	+	+	+	+	+	+	14



6. Visualization, interpretation, and discovery.

Every UMI read is associated with UMI-related metadata. This metadata is visible in the genome browser, as are the consensus reads themselves. Additionally, for variant calls, non-supporting fragments can be removed from view, allowing easy visual access to those reads that support the mutant allele.

Discover low allele frequency variants at high sequencing depths from somatic/ tumor samples

Features

- Protocol diversity.** Strand NGS supports data analysis from UMI protocols
 - Qiagen GeneRead®
 - Archer VariantPlex®
 - Rubicon Thruplex®
 - Bioo Scientific NextFlex®
 - A robust interface to specify custom UMIs
- End-to-end or point-to-point.** Users can go from reads to variants, can start at aligned BAMs containing the BC tag, or start/end at any reasonable point in the alignment/analysis workflow.
- Workflow diversity.** Strand NGS supports UMI protocols in DNA-, RNA- and small RNA-Seq workflows
- Somatic- and UMI-ready visualizations.** The genome browser visualizes consensus read lists. Each read contains UMI-related metadata, such as family size, UMI and mate UMI. A filter allows the easy exclusion of wild-type reads. This is useful at high sequencing depths and low allele frequencies, typical of data from somatic/tumor samples.

Retain supporting fragments (for low allele frequency /high coverage)

10000 reads: high sequencing depth for a somatic sample... but where are the variant alleles?

Remove all wild-type fragments

Only reads supporting variant allele visible