

Allegro Data Delivery Specifications

Output data info

- FASTQ.GZ files containing raw sequences.

No quality clipping nor adapter masking is provided with raw reads delivery. Depending on library type, from one to three FASTQ files per sample will be delivered, see table below for details:

Library type	First read sequence	Second read sequence	UMI sequence
Single end	R1	No	No
Single end plus UMI	R1	No	R2
Paired end	R1	R2	No
Paired end plus UMI	R1	R3	R2

Allegro standard bioinformatics analysis

- Alignments in BAM format (For UMI-bearing libraries, alignments are deduplicated).
- Variants in VCF format.
- Sequencing and coverage metrics in tabular format (total reads per sample, sample mean coverage, probe mean coverage).
- A REPORT file describing the library preparation and analysis flow.

FAQ

Are reads quality trimmed?

Delivered raw data are not quality trimmed. However, our internal analysis pipelines always rely on a quality trimming and adapter filtering steps which are described in the delivery REPORT.

Are raw reads depleted from duplicates?

With standard delivery, deduplicated raw reads are NOT provided. For UMI-bearing libraries deduplicated FASTQ files can be generated and delivered as an additional service.