



Table of contents

| | |
|--------------------------------------|---|
| 1. Selected input sequences | 3 |
| 1.1 Sequence reads | 3 |
| 1.2 Reference Sequences | 3 |
| 2. References | 3 |
| 3. Read quality control | 3 |
| 3.1 Strand specificity | 3 |
| 3.2 Adapter read-through | 4 |
| 3.3 Paired distance | 4 |
| 4. Mapping statistics | 4 |
| 4.1 Paired reads | 5 |
| 4.2 Match specificity (paired) | 5 |
| 5. Fragment statistics | 5 |
| 5.1 Fragment counting (total) | 5 |
| 6. Transcript length coverage | 6 |

1. Selected input sequences

1.1 Sequence reads

| Name | Number of reads | Longest read | Paired |
|---------------------------------------|-----------------|--------------|--------|
| SRR10571737_1 (paired, trimmed pairs) | 112,931,722 | 51 | yes |

For paired data, there are two reads in a pair.

1.2 Reference Sequences

| References | Length | Genes |
|------------|-------------|---------|
| 164,291 | 597,293,918 | 164,291 |

2. References

3. Read quality control

3.1 Strand specificity

| | |
|--------------------------------|------------|
| Strand specific setting | Both |
| Forward % of reads mapped | 0.94 |
| Reverse % of reads mapped | 99.06 |
| Reads with known strand | 82,782,881 |
| Ignored reads (wrong strand) | 0 |
| Ignored reads % (wrong strand) | 0.00 |

Strand specificity refers to whether a read is generated by sequencing an RNA fragment in the Forward (5' to 3') or Reverse (3' to 5') orientation. In a strand-specific protocol almost all reads are generated from a specific orientation. In other protocols a mix of both orientations is expected.

Strand specific setting: >90% of reads were mapped in the same orientation. Consider re-running the tool with a strand specific setting ("Forward"/"Reverse").

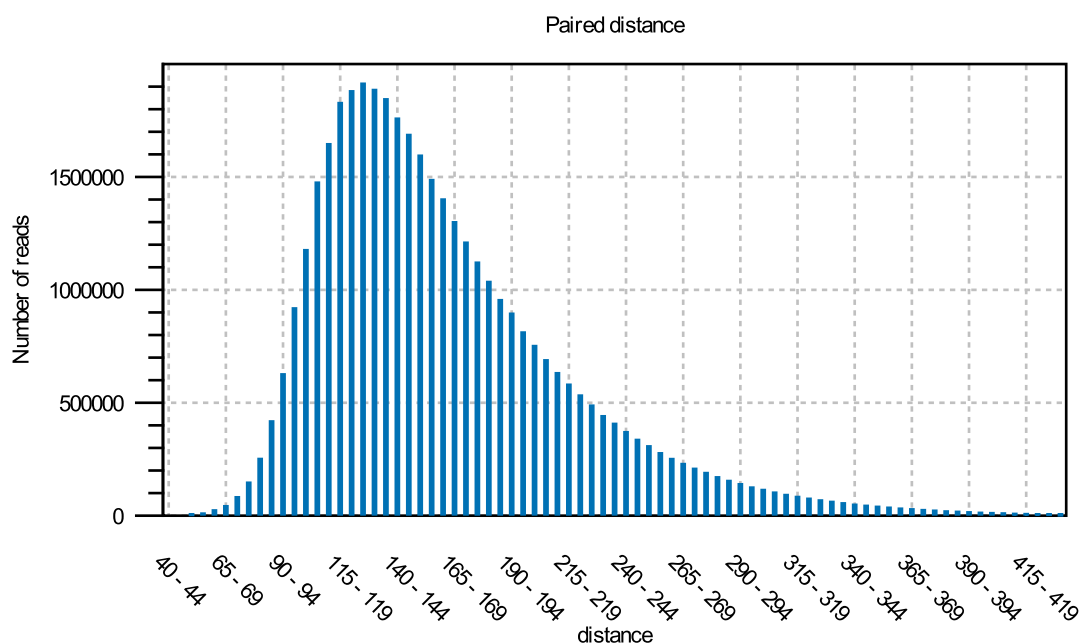
3.2 Adapter read-through

| | |
|--|------|
| Estimated % containing read-through adapters | 0.00 |
|--|------|

This value is estimated from the number of mapped paired-end reads where one read is longer than the sequenced RNA fragment.

A small percentage of read-through adapters is unlikely to affect the quantification of expression, but their presence may lead to false positive variant calls. To remove a read-through adapter, run the "Trim Reads" tool on the input reads, using either Automatic read-through adapter trimming or a Trim Adapter List configured to remove the adapter sequence from the 3' end.

3.3 Paired distance



The graph shows the distribution of paired-end distances, which is equivalent to the distribution of sequenced RNA fragment sizes. There should be a single broad peak at the target fragment size. An asymmetric peak may indicate problems in size selection.

| Reads | Paired distance estimates | Messages |
|---------------------------------------|---------------------------|----------|
| SRR10571737_1 (paired, trimmed pairs) | 43 to 433 bp | OK |

4. Mapping statistics

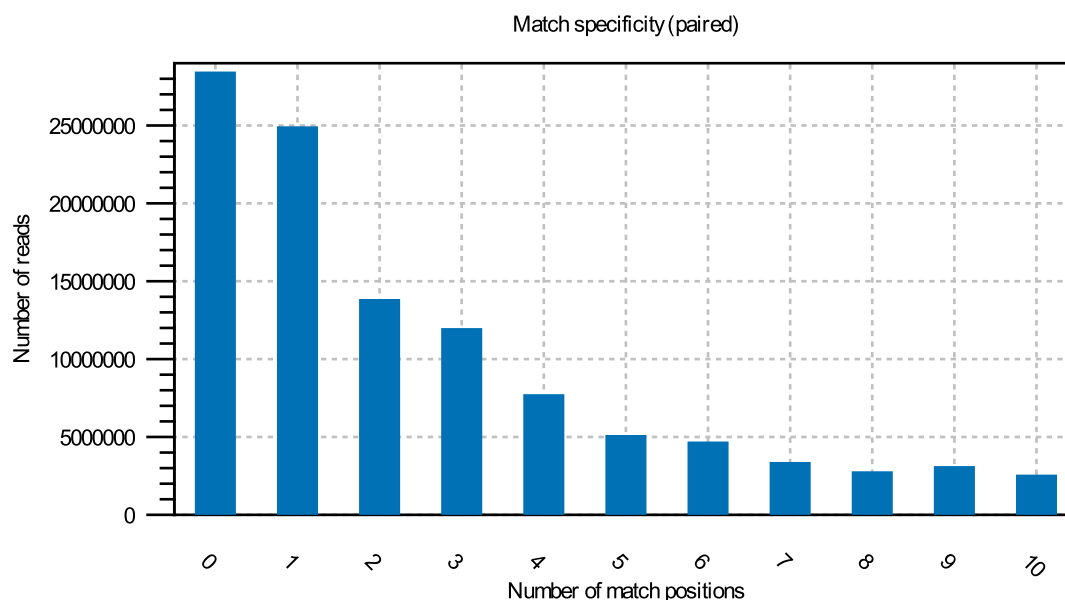
4.1 Paired reads

| | Number of sequences | % |
|------------------------------|---------------------|--------|
| Reads mapped in pairs | 80,282,978 | 71.09 |
| Reads mapped in broken pairs | 2,499,903 | 2.21 |
| Reads not mapped | 30,148,841 | 26.70 |
| Total | 112,931,722 | 100.00 |

For paired data, there are two reads in a pair.

If a strand specific setting was used, 'Reads not mapped' will include those reads that were ignored because they mapped to the wrong strand.

4.2 Match specificity (paired)



5. Fragment statistics

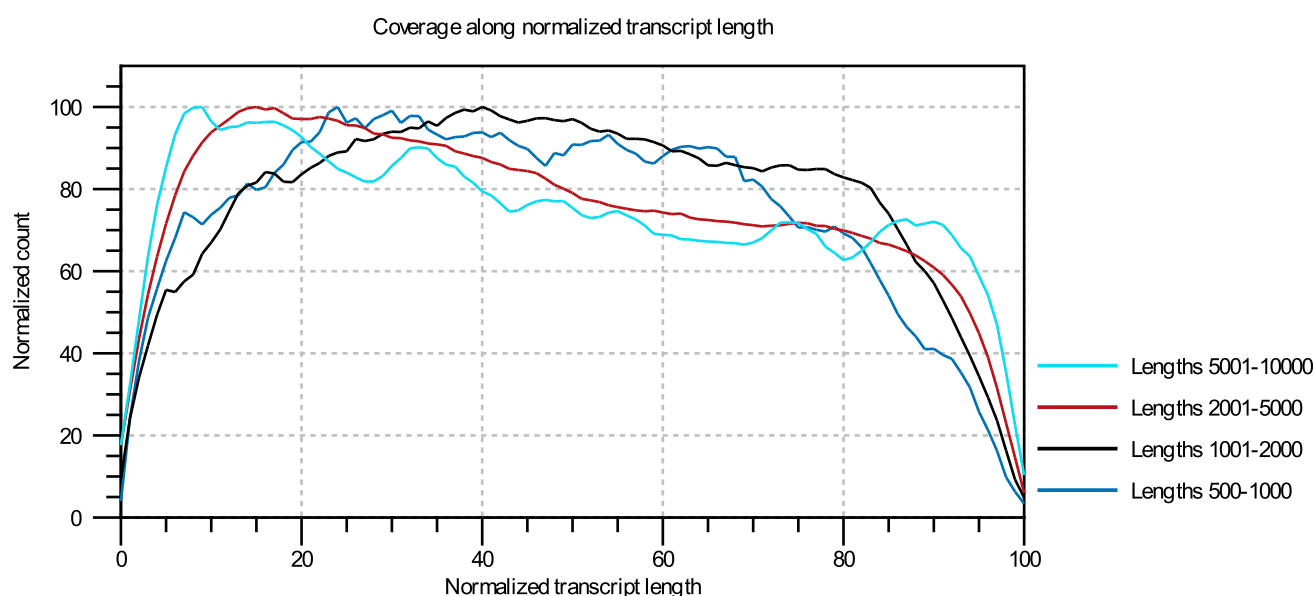
5.1 Fragment counting (total)

| | Fragments | % of total |
|---------------------|------------|------------|
| Counted fragments | 40,141,489 | 71.09 |
| - unique fragments | 12,474,317 | 22.09 |
| - non-specifically | 27,667,172 | 49.00 |
| Uncounted fragments | 16,324,372 | 28.91 |

| | Fragments | % of total |
|-----------------|------------|------------|
| Total fragments | 56,465,861 | 100.00 |

Default counting scheme ('Fragment counts'): An intact pair is counted as one, broken pairs are ignored.

6. Transcript length coverage



| | |
|--|----------|
| Expected coverage bias | Unbiased |
| Difference between average 3' and 5' normalized counts | -16.97 |
| % reads mapping to transcripts > 10000 bp | 3.28 |
| % reads mapping to transcripts 5001 - 10000 bp | 17.62 |
| % reads mapping to transcripts 2001 - 5000 bp | 42.85 |
| % reads mapping to transcripts 1001 - 2000 bp | 24.98 |
| % reads mapping to transcripts 500 - 1000 bp | 10.42 |
| % reads mapping to transcripts < 500 bp | 0.84 |

The plot shows the normalized coverage across a transcript body for four different groupings of transcript length. The lines should be flat in the center of the plot, and the plot should be approximately symmetric. An erratic line may indicate that there are few transcripts in the given length range.