



Table of contents

1. Reads summary	3
1.1 Summary statistics	3
1.2 Read length distribution	4
2. QC for sequencing reads	5
2.1 Coverage	5
2.2 Ambiguous base-content	7
2.3 Nucleotide contributions	8
2.4 GC-content	9
2.5 Quality distribution	12
2.6 Enriched 5-mer	13
2.7 Sequence duplication levels	14

1. Reads summary

1.1 Summary statistics

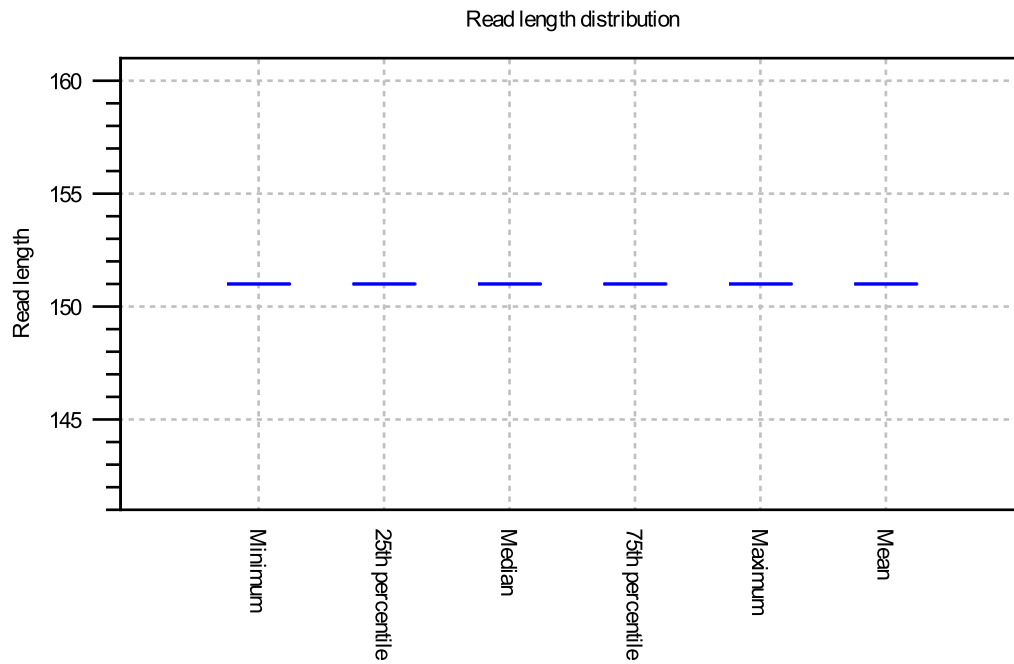
The table is based on 24 samples.

Sample name	Data sets (#)	Reads (#)	Paired reads (%)	Bases (#)
AM24_1_1 (paired) - graphical QC report	1	69,633,566	100	10,514,668,466
AM24_2_1 (paired) - graphical QC report	1	76,150,974	100	11,498,797,074
AM24_3_1 (paired) - graphical QC report	1	73,612,726	100	11,115,521,626
AM24_INF_1_1 (paired) - graphical QC report	1	62,650,674	100	9,460,251,774
AM24_INF_2_1 (paired) - graphical QC report	1	70,523,514	100	10,649,050,614
AM24_INF_3_1 (paired) - graphical QC report	1	79,934,394	100	12,070,093,494
AM48_1_1 (paired) - graphical QC report	1	62,248,176	100	9,399,474,576
AM48_2_1 (paired) - graphical QC report	1	73,062,270	100	11,032,402,770
AM48_3_1 (paired) - graphical QC report	1	65,555,178	100	9,898,831,878
AM48_INF_1_1 (paired) - graphical QC report	1	74,715,412	100	11,282,027,212
AM48_INF_2_1 (paired) - graphical QC report	1	73,339,746	100	11,074,301,646
AM48_INF_3_1 (paired) - graphical QC report	1	65,201,204	100	9,845,381,804
NM24_1_1 (paired) - graphical QC report	1	68,055,724	100	10,276,414,324
NM24_2_1 (paired) - graphical QC report	1	76,255,700	100	11,514,610,700
NM24_3_1 (paired) - graphical QC report	1	75,139,100	100	11,346,004,100

Sample name	Data sets (#)	Reads (#)	Paired reads (%)	Bases (#)
NM24_INF_1_1 (paired) - graphical QC report	1	78,618,562	100	11,871,402,862
NM24_INF_2_1 (paired) - graphical QC report	1	76,115,730	100	11,493,475,230
NM24_INF_3_1 (paired) - graphical QC report	1	74,225,174	100	11,208,001,274
NM48_1_1 (paired) - graphical QC report	1	73,239,196	100	11,059,118,596
NM48_2_1 (paired) - graphical QC report	1	67,070,276	100	10,127,611,676
NM48_3_1 (paired) - graphical QC report	1	72,191,834	100	10,900,966,934
NM48_INF_1_1 (paired) - graphical QC report	1	69,123,512	100	10,437,650,312
NM48_INF_2_1 (paired) - graphical QC report	1	72,931,746	100	11,012,693,646
NM48_INF_3_1 (paired) - graphical QC report	1	62,374,538	100	9,418,555,238
Minimum	1.00	62,248,176.00	100.00	9,399,474,576.00
Median	1.00	72,997,008.00	100.00	11,022,548,208.00
Maximum	1.00	79,934,394.00	100.00	12,070,093,494.00
Mean	1.00	71,332,038.58	100.00	10,771,137,826.08
Standard deviation	0.00	5,096,063.12	0.00	769,505,530.61

1.2 Read length distribution

The box plot is based on 24 samples.



2. QC for sequencing reads

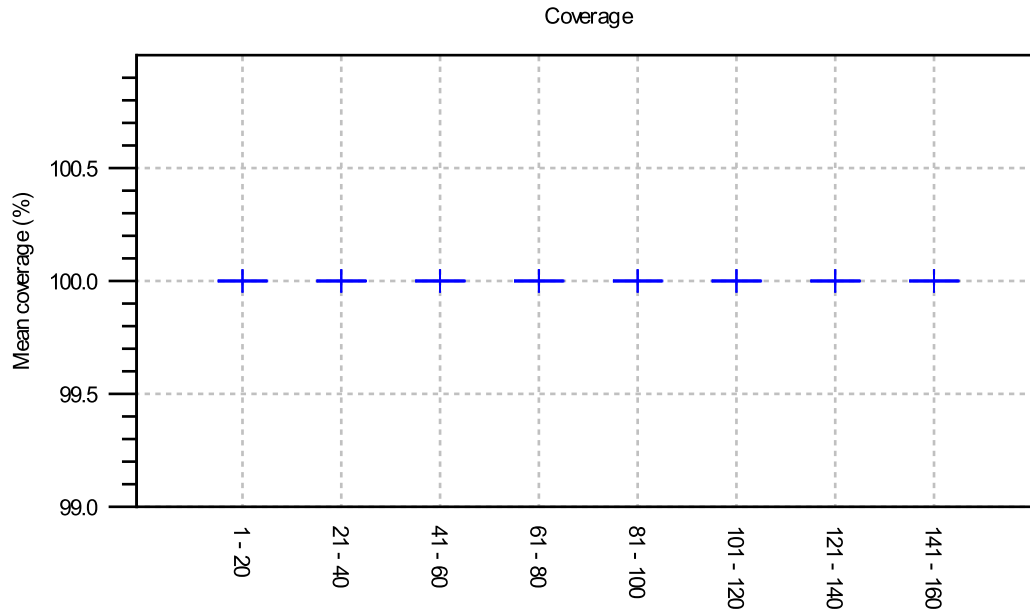
2.1 Coverage

Summarizes the number of sequences that support (cover) the individual base positions.

p - r: The base positions in the read

y: the mean number of sequences covering the respective base positions normalized to the total number of sequences

The box plot is based on 24 samples.



	Outliers
1 - 20	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
21 - 40	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
41 - 60	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
61 - 80	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
81 - 100	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
101 - 120	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
121 - 140	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report
141 - 160	NM24_2_1 (paired) - graphical QC report, NM24_INF_1_1 (paired) - graphical QC report, NM24_INF_2_1 (paired) - graphical QC report

2.2 Ambiguous base-content

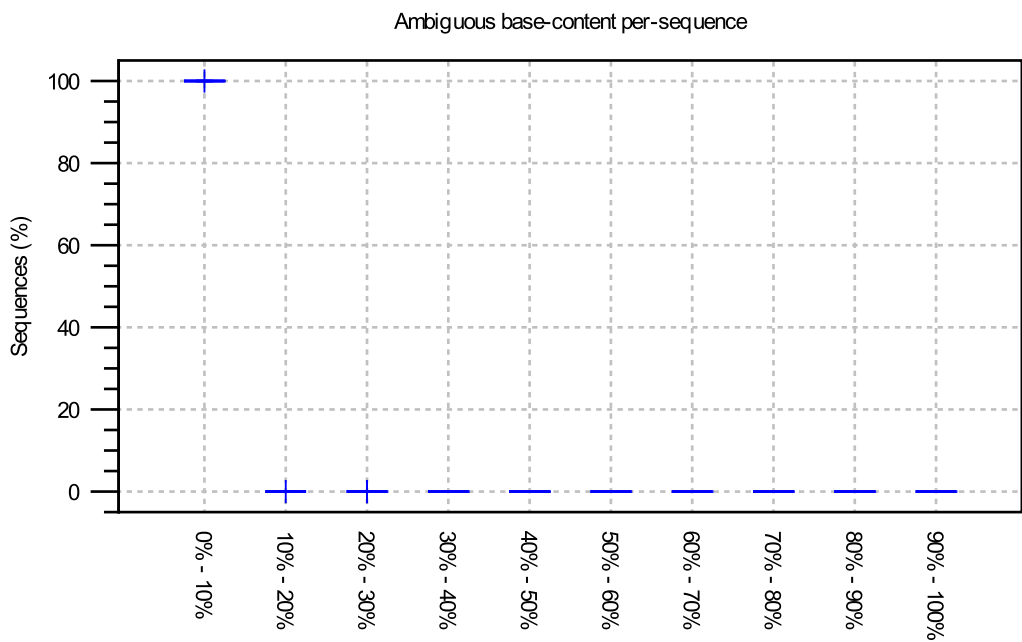
Ambiguous base-content per-sequence

Summarizes the distribution of N-contents. The N-content of a sequence is calculated as the number of ambiguous bases compared to all bases.

x: p% - r%: percentage range

y: the number of sequences featuring particular N-percentages normalized to the total number of sequences

The box plot is based on 24 samples.



	Outliers
0% - 10%	AM48_INF_1_1 (paired) - graphical QC report
10% - 20%	AM48_INF_1_1 (paired) - graphical QC report
20% - 30%	NM48_3_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report

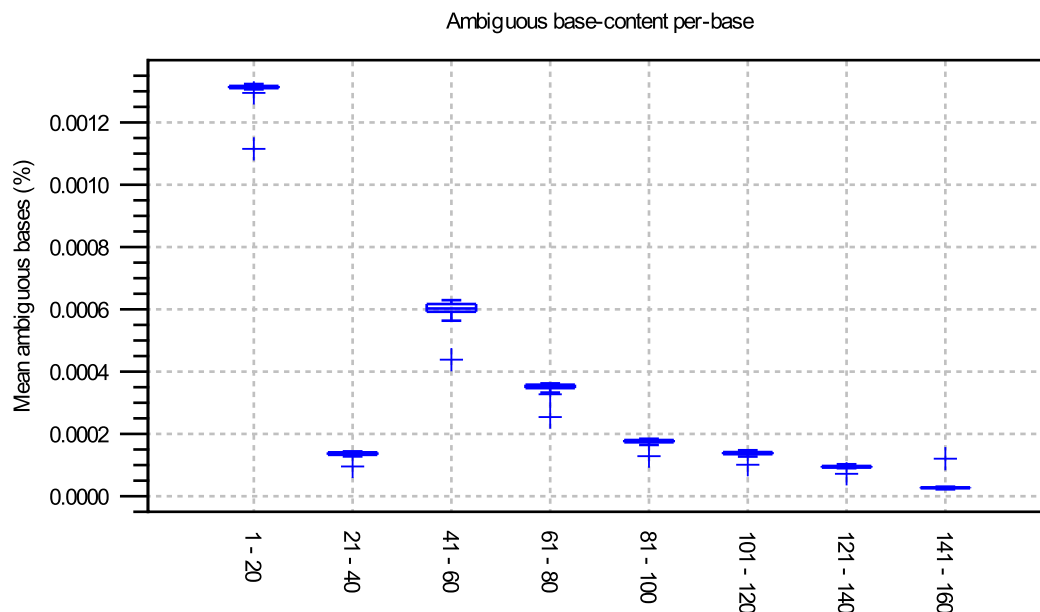
Ambiguous base-content per-base

Summarizes the cumulative coverage of ambiguous bases.

p - r: The base positions in the read

y: the mean number of ambiguous bases observed at the base positions normalized to the total number of bases observed at those positions

The box plot is based on 24 samples.



	Outliers
1 - 20	AM48_INF_1_1 (paired) - graphical QC report, NM24_INF_3_1 (paired) - graphical QC report
21 - 40	AM48_INF_1_1 (paired) - graphical QC report
41 - 60	AM48_INF_1_1 (paired) - graphical QC report
61 - 80	AM48_INF_1_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report
81 - 100	AM48_INF_1_1 (paired) - graphical QC report
101 - 120	AM48_INF_1_1 (paired) - graphical QC report
121 - 140	AM48_INF_1_1 (paired) - graphical QC report
141 - 160	AM48_INF_1_1 (paired) - graphical QC report

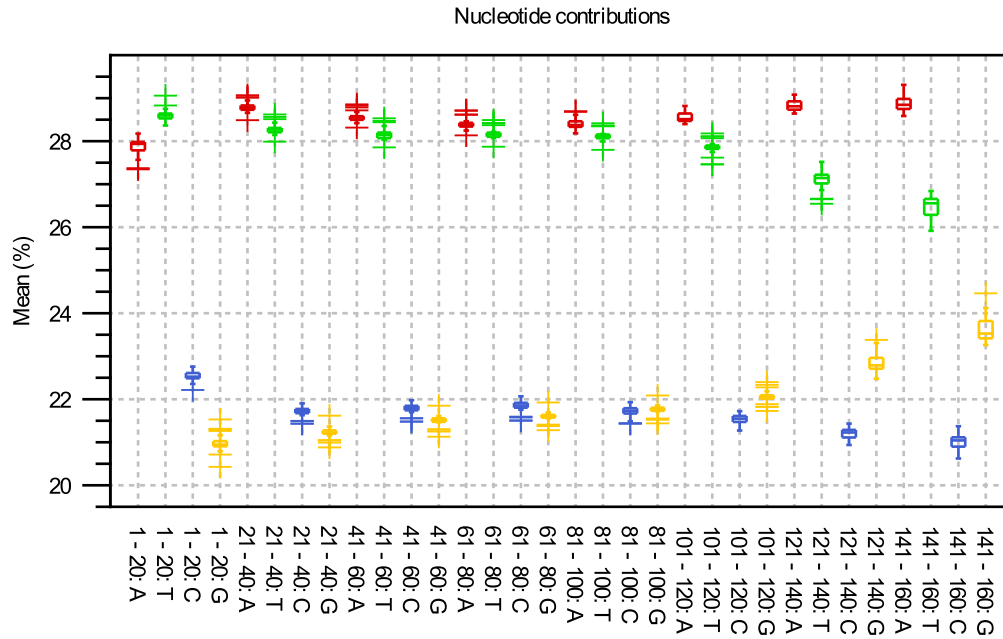
2.3 Nucleotide contributions

Summarizes the coverages for the four DNA nucleotides.

p - r: The base positions in the read

y: the mean number of nucleotides observed per type normalized to the total number of nucleotides observed at those positions

The box plot is based on 24 samples.



2.4 GC-content

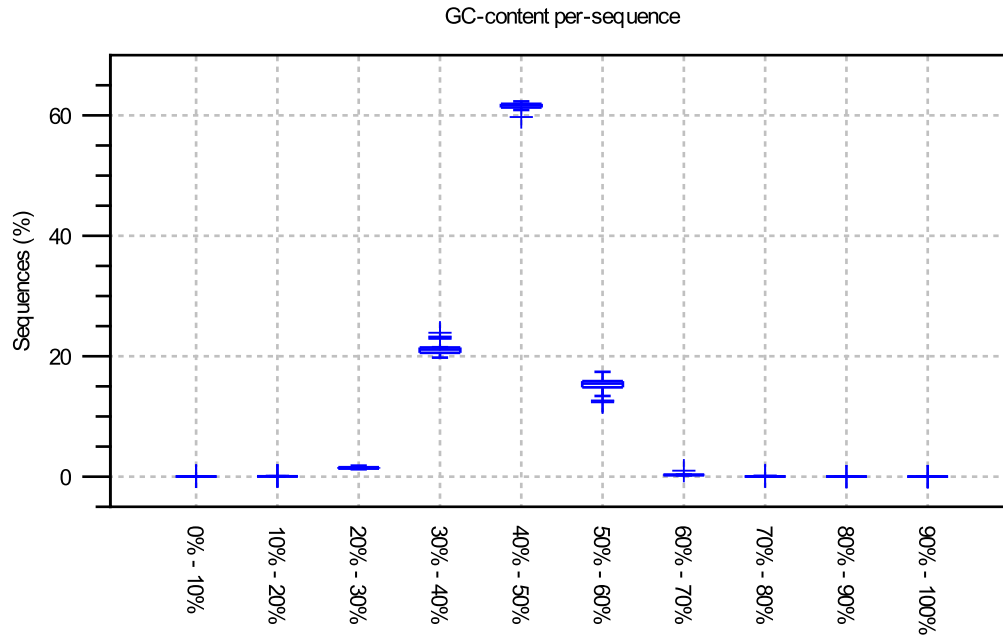
GC-content per-sequence

Summarizes the distribution of GC-contents. The GC-content of a sequence is calculated as the number of GC-bases compared to all bases (including ambiguous bases).

x: p% - r%: percentage range

y: the number of sequences featuring the particular GC-percentage range normalized to the total number of sequences

The box plot is based on 24 samples.



	Outliers
0% - 10%	AM24_1_1 (paired) - graphical QC report, AM24_INF_1_1 (paired) - graphical QC report, AM48_2_1 (paired) - graphical QC report, AM48_INF_2_1 (paired) - graphical QC report, NM48_1_1 (paired) - graphical QC report
10% - 20%	AM24_1_1 (paired) - graphical QC report, AM24_INF_1_1 (paired) - graphical QC report, AM48_2_1 (paired) - graphical QC report, AM48_INF_2_1 (paired) - graphical QC report, NM48_1_1 (paired) - graphical QC report
30% - 40%	AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
40% - 50%	AM24_INF_1_1 (paired) - graphical QC report
50% - 60%	AM48_INF_1_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
60% - 70%	AM24_INF_1_1 (paired) - graphical QC report
70% - 80%	AM24_1_1 (paired) - graphical QC report, AM24_INF_1_1 (paired) - graphical QC report
80% - 90%	AM24_1_1 (paired) - graphical QC report, AM24_INF_1_1 (paired) - graphical QC report, NM48_1_1 (paired) - graphical QC report
90% - 100%	AM24_1_1 (paired) - graphical QC report, AM24_INF_1_1 (paired) - graphical QC report, AM48_2_1 (paired) - graphical QC report, AM48_INF_2_1 (paired) - graphical QC report, NM48_1_1 (paired) - graphical QC report

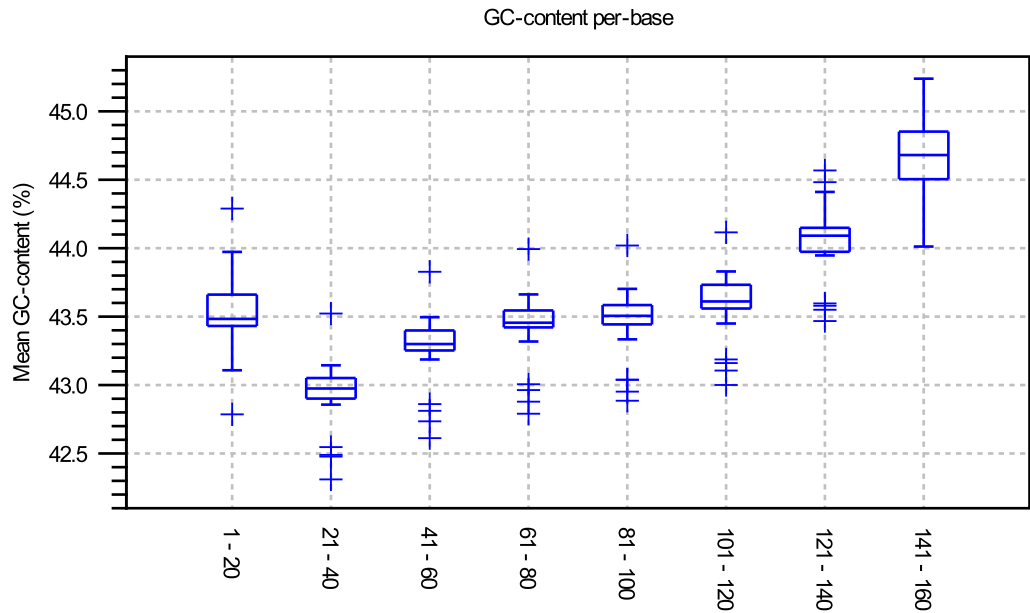
GC-content per-base

Summarizes the cumulative coverage of G- and C-bases.

p - r: The base positions in the read

y: the mean number of G- and C-bases observed at the base positions normalized to the total number of bases observed at those positions

The box plot is based on 24 samples.



	Outliers
1 - 20	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report
21 - 40	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
41 - 60	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
61 - 80	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
81 - 100	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report

	Outliers
101 - 120	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
121 - 140	AM24_INF_1_1 (paired) - graphical QC report, AM48_2_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, AM48_INF_3_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report

2.5 Quality distribution

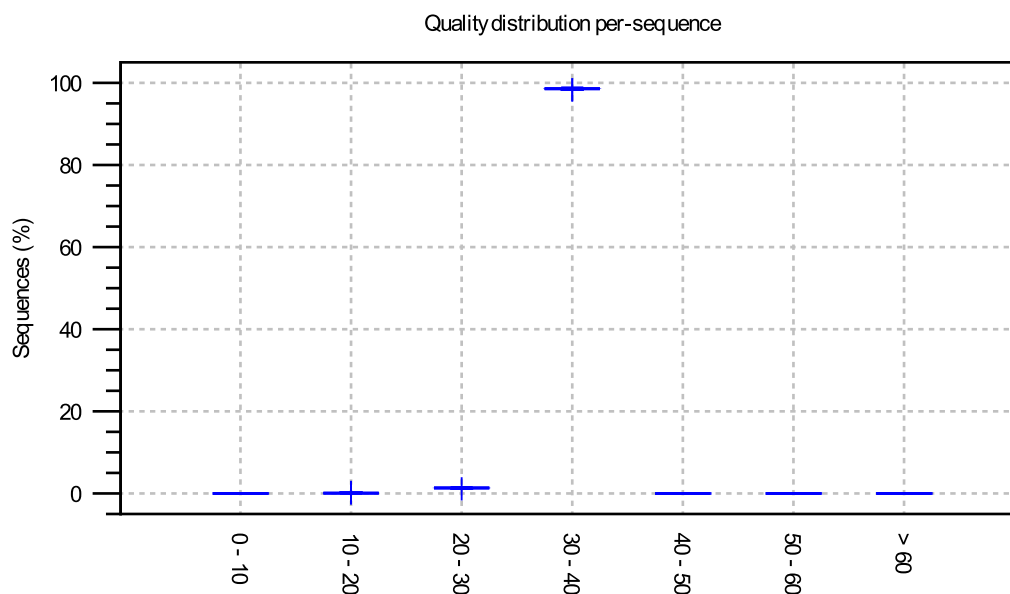
Quality distribution per-sequence

Summarizes the distribution of average sequence quality scores. The quality of a sequence is calculated as the arithmetic mean of its base qualities.

x - y: PHRED range

y: the number of sequences observed at that qual. score normalized to the total number of sequences

The box plot is based on 24 samples.



	Outliers
10 - 20	AM24_1_1 (paired) - graphical QC report, AM24_INF_1_1 (paired) - graphical QC report, AM48_2_1 (paired) - graphical QC report, AM48_INF_2_1 (paired) - graphical QC report, NM48_1_1 (paired) - graphical QC report
20 - 30	AM24_INF_2_1 (paired) - graphical QC report

	Outliers
30 - 40	AM24_INF_1_1 (paired) - graphical QC report, NM48_1_1 (paired) - graphical QC report

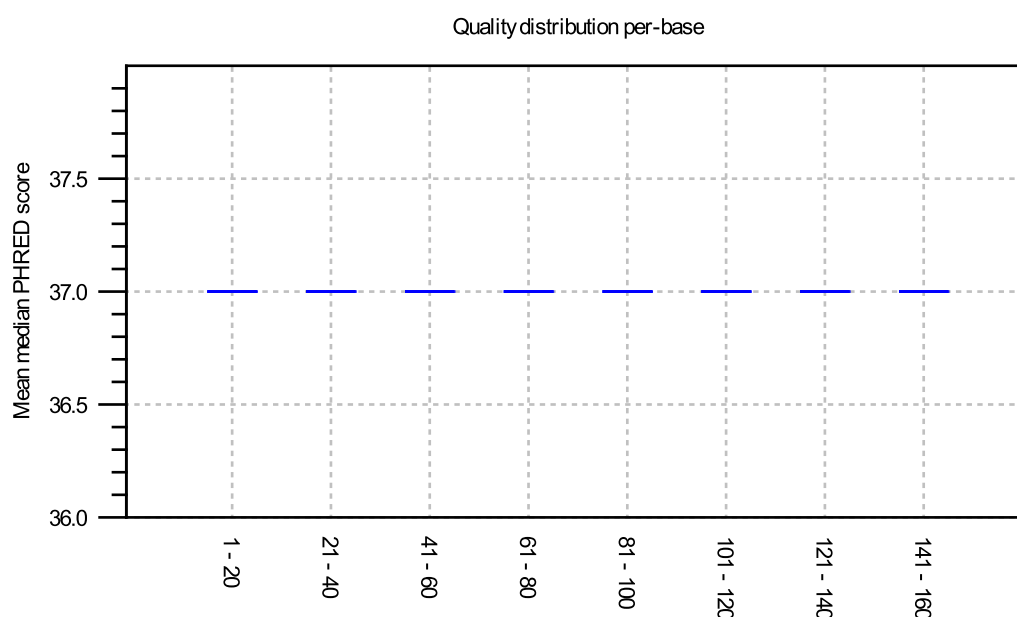
Quality distribution per-base

Summarizes the base-quality distribution along the base positions.

p - r: The base positions in the read

y: the mean of the median of quality scores observed at the base positions

The box plot is based on 24 samples.



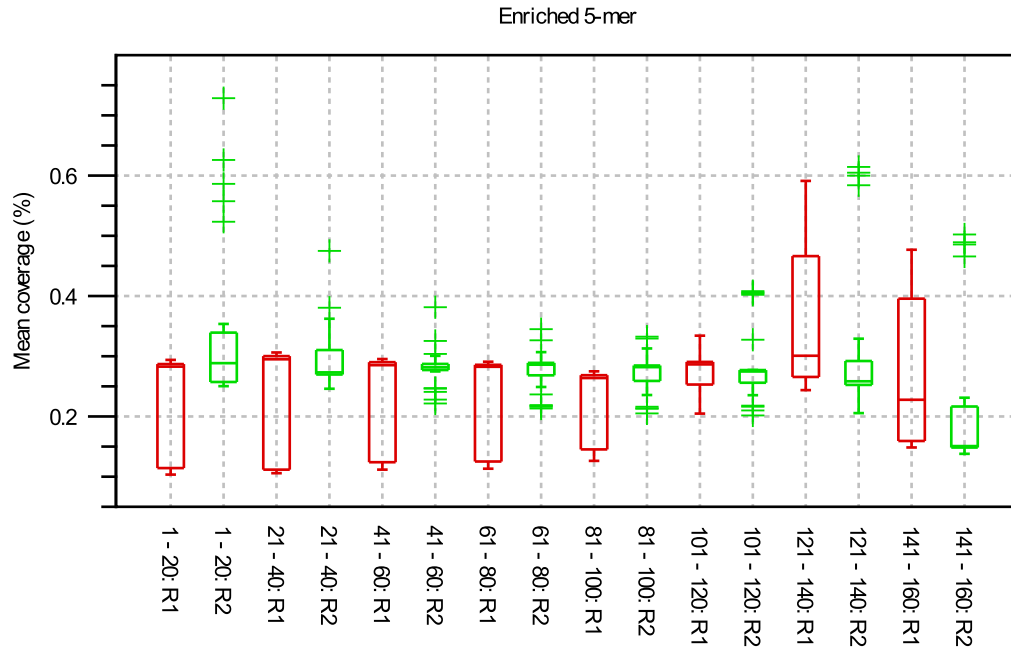
2.6 Enriched 5-mer

Summarizes the most-overrepresented 5-mer in the different types of reads. The over-representation of a 5-mer is calculated as the ratio of the expected and observed 5-mer frequency. The expected frequency is calculated as product of the empirical nucleotide probabilities that make up the k-mer. (5-mers that contain ambiguous bases are ignored)

p - r: The base positions in the read

y: the mean number of times the most enriched 5-mer has been observed at the base positions normalized to all 5-mers observed at those positions in the different types of reads

The box plot is based on 24 samples.



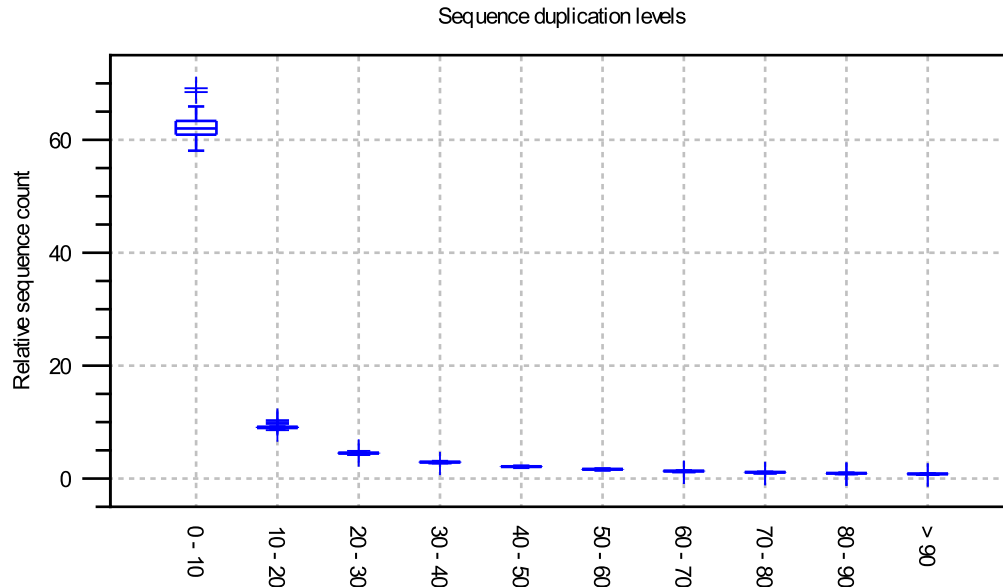
2.7 Sequence duplication levels

Summarizes the duplication level distribution. Duplication levels are simply the count of how often a particular sequence has been found.

x - y: sequence duplication range

y: the number of sequences that have been found that many times normalized to the number of unique sequences

The box plot is based on 24 samples.



	Outliers
0 - 10	NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
10 - 20	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
20 - 30	AM24_INF_1_1 (paired) - graphical QC report, AM48_INF_1_1 (paired) - graphical QC report
30 - 40	NM48_INF_3_1 (paired) - graphical QC report
60 - 70	NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
70 - 80	NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
80 - 90	AM48_INF_1_1 (paired) - graphical QC report, NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report
> 90	NM48_INF_1_1 (paired) - graphical QC report, NM48_INF_3_1 (paired) - graphical QC report

History for:



AM24_1_1 (paired) - graphical QC report (Combined report)

Combine Reports 1.1

29 Jun 2023 16:38:22

Version:	CLC Genomics Workbench 23.0.4
Modified by:	markop
Show summary information as plots	Yes
Include tables for outliers	Yes
Comments:	

Originates from:

	AM24_1_1 (paired) - graphical QC report
	AM24_1_1 (paired) - supplementary QC report
	AM24_2_1 (paired) - graphical QC report
	AM24_2_1 (paired) - supplementary QC report
	AM24_3_1 (paired) - graphical QC report
	AM24_3_1 (paired) - supplementary QC report
	AM24_INF_1_1 (paired) - graphical QC report
	AM24_INF_1_1 (paired) - supplementary QC report
	AM24_INF_2_1 (paired) - graphical QC report
	AM24_INF_2_1 (paired) - supplementary QC report
	AM24_INF_3_1 (paired) - graphical QC report
	AM24_INF_3_1 (paired) - supplementary QC report
	AM48_1_1 (paired) - graphical QC report
	AM48_1_1 (paired) - supplementary QC report
	AM48_2_1 (paired) - graphical QC report
	AM48_2_1 (paired) - supplementary QC report
	AM48_3_1 (paired) - graphical QC report
	AM48_3_1 (paired) - supplementary QC report
	AM48_INF_1_1 (paired) - graphical QC report
	AM48_INF_1_1 (paired) - supplementary QC report
	AM48_INF_2_1 (paired) - graphical QC report
	AM48_INF_2_1 (paired) - supplementary QC report
	AM48_INF_3_1 (paired) - graphical QC report
	AM48_INF_3_1 (paired) - supplementary QC report
	NM24_1_1 (paired) - graphical QC report
	NM24_1_1 (paired) - supplementary QC report
	NM24_2_1 (paired) - graphical QC report
	NM24_2_1 (paired) - supplementary QC report
	NM24_3_1 (paired) - graphical QC report
	NM24_3_1 (paired) - supplementary QC report
	NM24_INF_1_1 (paired) - graphical QC report
	NM24_INF_1_1 (paired) - supplementary QC report
	NM24_INF_2_1 (paired) - graphical QC report
	NM24_INF_2_1 (paired) - supplementary QC report