Shasta assembly summary

Shasta version

Shasta Release 0.13.0

Reads used in this assembly

Read representation	0 (Raw - no RLE)
Minimum read length	10000
Number of reads	2145809
Number of read bases	74632574696
Average read length	34780
Read N50	41843
Number of reads flagged as palindromic by self alignment	902
Number of reads flagged as chimeric	131

- Reads discarded on input are not included in the above table (see below).
- See ReadLengthHistogram.csv and Binned-ReadLengthHistogram.csv for details of the read length distribution of reads used in this assembly.

Reads discarded on input

	Reads	Bases
Reads discarded on input because they contained invalid bases	0	0
Reads discarded on input because they were too short	363962	2159212300
Reads discarded on input because they contained repeat counts greater than 255	0	0
Reads discarded on input, total	363962	2159212300
Fraction of reads discarded on input over total present in input files	0.145018	0.0281178

- Base counts in the above table are raw sequence bases.
- Here and elsewhere, "raw" refers to the original read sequence, as opposed to run-length encoded sequence.

Markers

Total number of markers on all reads, one strand	3733827269
Total number of markers on all reads, both strands	7467654538
Average number of markers per base	0.05003
Average base offset between markers	19.99
Average base gap between markers	-10.01

Alignments

Number of alignment candidates found by the LowHash algorithm	43649818
Number of good alignments	4393
Number of good alignments kept in the read graph	4393

Alignment criteria actually used for creation of the read graph

minAlignedMarkerCount	700
minAlignedFraction	0.99
maxSkip	20
maxDrift	10
maxTrim	20

Read graph

Number of vertices	4291618
Number of edges	8786

- The read graph contains both strands. Each read generates two vertices.
- Isolated reads in the read graph don't contribute to the assembly. See the table below for a summary of isolated reads in the read graph. Each isolated read corresponds to two isolated vertices in the read graph, one for each strand.

	Reads	Bases
Isolated reads	2141795	74461885646
Non-isolated reads	4014	170689050
Isolated reads fraction	0.9981	0.9977
Non-isolated reads fraction	0.001871	0.002287

Performance

Elapsed time (seconds)	1484
Elapsed time (minutes)	24.73
Elapsed time (hours)	0.4121
Average CPU utilization	0.139
Peak virtual memory utilization (bytes)	342481059840
Number of threads used	150
Total number of virtual CPUs available	255
Total physical memory available (bytes)	3921325645824