

VIB Nucleomics Core*- SEQUEL HGAP4 report

8plex-EColi_HGAP4

sample: Ecoli_8plex_demo

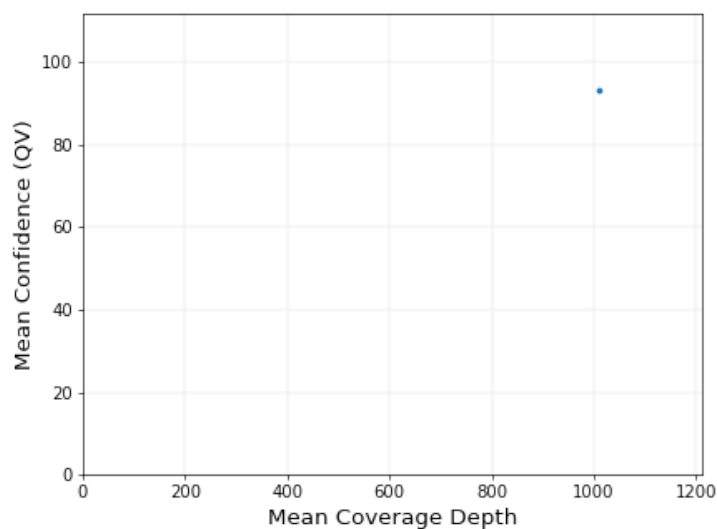
dataset: /data/NC_projects/PB_Coli_8-plex/Ecoli_8plex_demo.barcoded.subreads.bam

target haploid asm size: 5 Mb

Polished Assembly

Value	Analysis Metric
1	Polished Contigs
4'635'027	Maximum Contig Length
4'635'027	N50 Contig Length
4'635'027	Sum of Contig Lengths
4'635'027	E-size (sum of squares / sum)

Contig Coverage vs. Confidence



©SP/NC 2018-06-09, v1.1.0

printed: Wed Jun 27 03:38:33 PM 2018



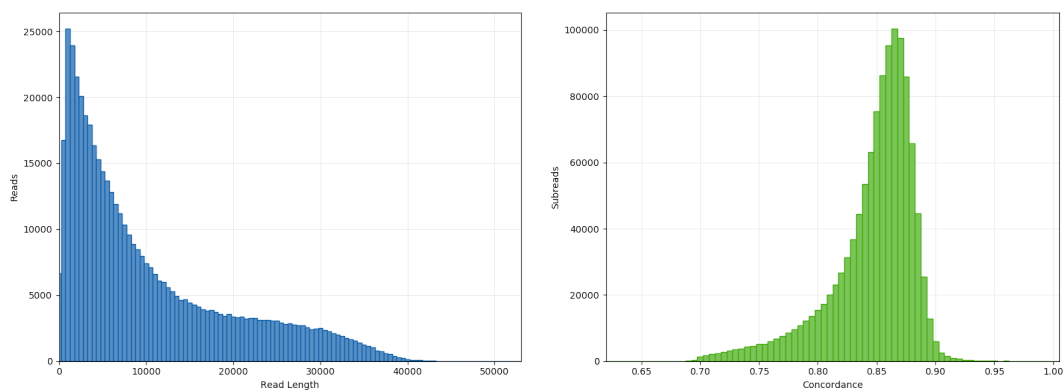
more at <http://www.nucleomics.be>

*VIB - Nucleomics Core, nucleomics@vib.be

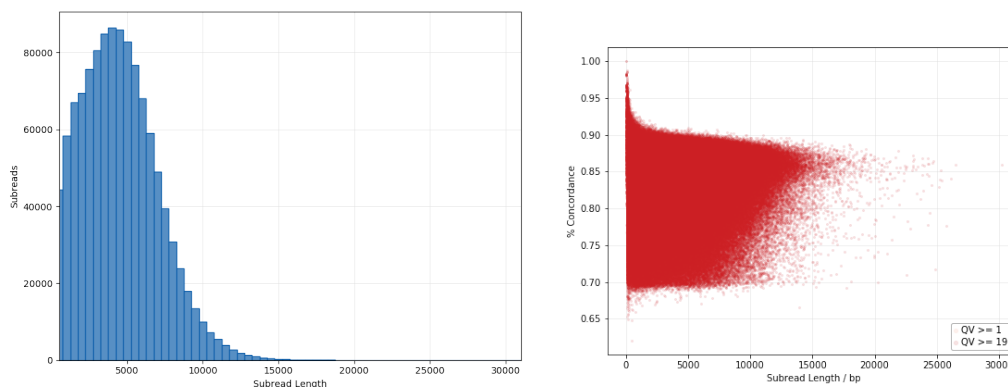
Realignment to Draft Assembly

Value	Analysis Metric
0.93	Percent Realigned Bases
0.84	Mean Concordance (realigned)
1'150'464.00	Number of Subreads (realigned)
4'906'748'082.00	Number of Subread Bases (realigned)
4'265.00	Subread Length Mean (realigned)
5'547.00	Subread Length N50 (realigned)
8'760.00	Subread Length 95% (realigned)
30'233.00	Subread Length Max (realigned)
475'472.00	Number of Polymerase Reads (realigned)
10'560.00	Polymerase Read Length Mean (realigned)
19'037.00	Polymerase Read N50 (realigned)
30'650.00	Polymerase Read Length 95% (realigned)
52'771.00	Polymerase Read Length Max (realigned)

Mapped Polymerase Read Length - Realigned Subread Concordance



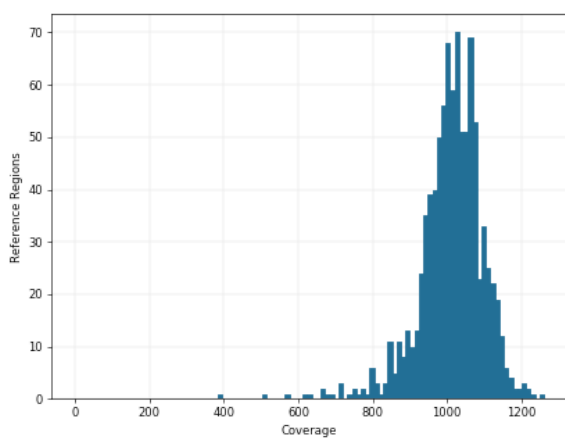
Realigned Subread Length - Realigned Concordance vs. Read Length



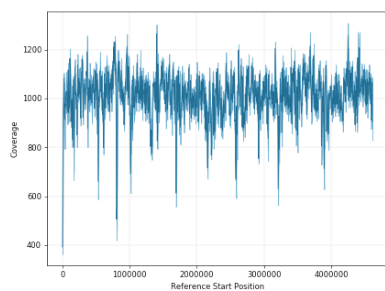
Coverage

Value	Analysis Metric
1'010	Mean Coverage
0	Missing Bases

Coverage plots



coverage_plot_000000F



Preassembly

Value	Analysis Metric
5'000'000.00	Genome Length (user input)
1'152'407.00	Number of Filtered Subreads
4'558.00	Filtered Subread Length Mean
5'699.00	Filtered Subread Length (N50)
9'188.00	Filtered Subread Length 95%
6'105.66	Filtered Subread E-size
5'253'062'991.00	Number of Filtered Subread Bases
1'050.61	Filtered Subread Coverage
12'800.00	Length Cutoff (user input or auto-calc)
9'371.00	Number of Seed Reads
16'008.00	Seed Read Length Mean
15'217.00	Seed Read Length (N50)
23'944.00	Seed Read Length 95%
16'940.74	Seed Read E-size
150'006'344.00	Number of Seed Bases (total)
30.00	Seed Coverage (bases/genome_size)
9'936.00	Number of Pre-Assembled Reads
6'942.00	Pre-Assembled Read Length Mean
8'340.00	Pre-Assembled Read Length (N50)
13'571.00	Pre-Assembled Read Length 95%
8'860.24	Pre-Assembled E-size (sum of squares / sum)
68'974'223.00	Number of Pre-Assembled Bases (total)
14.00	Pre-Assembled Coverage (bases/genome_size)
0.46	Pre-Assembled Yield (bases/seed_bases)
1.43	Average Number of Reads That Each Seed Is Broken Into
5'310.72	Average Number of Bases Lost from Each Seed

Analysis Parameters

	Value
HGAP_AggressiveAsm_bool	FALSE
HGAP_FalconAdvanced_str	
HGAP_GenomeLength_str	5000000
HGAP_SeedCoverage_str	30
HGAP_SeedLengthCutoff_str	-1
save_las_for_unzip	FALSE
algorithm	best
gff2bed_purpose	variants
masking	TRUE
min_confidence	40
min_coverage	5
track_description	PacBio: snps, insertions, and deletions derived from consensus calls against reference
track_name	variants
use_score	0
algorithm_options	-minMatch 12 -bestn 10 -minPctSimilarity 70.0
concordant	FALSE
hit_policy	randombest
min_accuracy	70
min_length	50
no_split_subreads	FALSE
other_filters	rq >= 0.7
read_length	0
force_num_regions	FALSE
max_contigs	25
max_region_size	100000
num_regions	1000
region_size	0