Supernova assembly reports

P assembly

--------------------------------------------------------------------------------

SUMMARY

--------------------------------------------------------------------------------

- Tue Oct 30 07:26:40 2018

- [PF\_15\_M1\_S3]

- software release = 2.1.1(6bb16452a)

- likely sequencers = HiSeq2000/2500

- assembly checksum = 4,448,637,804,504,588,006

--------------------------------------------------------------------------------

INPUT

- 156.27 M = READS = number of reads; ideal 800M-1200M for human

- 139.50 b = MEAN READ LEN = mean read length after trimming; ideal 140

- 52.65 x = RAW COV = raw coverage; ideal ~56

- 42.31 x = EFFECTIVE COV = effective read coverage; ideal ~42 for raw 56x

- 92.20 % = READ TWO Q30 = fraction of Q30 bases in read 2; ideal 75-85

- 361.00 b = MEDIAN INSERT = median insert size; ideal 350-400

- 89.13 % = PROPER PAIRS = fraction of proper read pairs; ideal >= 75

- 1.00 = BARCODE FRACTION = fraction of barcodes used; between 0 and 1

- 448.19 Mb = EST GENOME SIZE = estimated genome size

- 22.50 % = REPETITIVE FRAC = genome repetitivity index

- 0.43 % = HIGH AT FRACTION = high AT index

- 38.70 % = ASSEMBLY GC CONTENT = GC content of assembly

- 0.13 % = DINUCLEOTIDE FRACTION = dinucleotide content

- 46.41 Kb = MOLECULE LEN = weighted mean molecule size; ideal 50-100

- 327.59 = P10 = molecule count extending 10 kb on both sides

- 215.00 b = HETDIST = mean distance between heterozygous SNPs

- 9.20 % = UNBAR = fraction of reads that are not barcoded

- 118.00 = BARCODE N50 = N50 reads per barcode

- 10.86 % = DUPS = fraction of reads that are duplicates

- 57.77 % = PHASED = nonduplicate and phased reads; ideal 45-50

--------------------------------------------------------------------------------

OUTPUT

- 2.40 K = LONG SCAFFOLDS = number of scaffolds >= 10 kb

- 20.46 Kb = EDGE N50 = N50 edge size

- 76.33 Kb = CONTIG N50 = N50 contig size

- 3.11 Mb = PHASEBLOCK N50 = N50 phase block size

- 4.93 Mb = SCAFFOLD N50 = N50 scaffold size

- 14.86 % = MISSING 10KB = % of base assembly missing from scaffolds >= 10 kb

- 259.47 Mb = ASSEMBLY SIZE = assembly size (only scaffolds >= 10 kb)

--------------------------------------------------------------------------------

ALARMS

--------------------------------------------------------------------------------

FM assembly

--------------------------------------------------------------------------------

SUMMARY

--------------------------------------------------------------------------------

- Thu Nov 01 01:36:51 2018

- [OT18\_12\_MP4\_re-do\_S2]

- software release = 2.1.1(6bb16452a)

- likely sequencers = HiSeq2000/2500

- assembly checksum = -2,535,080,572,152,179,051

--------------------------------------------------------------------------------

INPUT

- 151.74 M = READS = number of reads; ideal 800M-1200M for human

- 139.50 b = MEAN READ LEN = mean read length after trimming; ideal 140

- 55.69 x = RAW COV = raw coverage; ideal ~56

- 44.06 x = EFFECTIVE COV = effective read coverage; ideal ~42 for raw 56x

- 92.43 % = READ TWO Q30 = fraction of Q30 bases in read 2; ideal 75-85

- 338.00 b = MEDIAN INSERT = median insert size; ideal 350-400

- 90.33 % = PROPER PAIRS = fraction of proper read pairs; ideal >= 75

- 1.00 = BARCODE FRACTION = fraction of barcodes used; between 0 and 1

- 411.44 Mb = EST GENOME SIZE = estimated genome size

- 18.80 % = REPETITIVE FRAC = genome repetitivity index

- 0.44 % = HIGH AT FRACTION = high AT index

- 38.71 % = ASSEMBLY GC CONTENT = GC content of assembly

- 0.15 % = DINUCLEOTIDE FRACTION = dinucleotide content

- 52.29 Kb = MOLECULE LEN = weighted mean molecule size; ideal 50-100

- 396.88 = P10 = molecule count extending 10 kb on both sides

- 11.85 Kb = HETDIST = mean distance between heterozygous SNPs

- 9.62 % = UNBAR = fraction of reads that are not barcoded

- 114.00 = BARCODE N50 = N50 reads per barcode

- 12.37 % = DUPS = fraction of reads that are duplicates

- 66.87 % = PHASED = nonduplicate and phased reads; ideal 45-50

--------------------------------------------------------------------------------

OUTPUT

- 2.38 K = LONG SCAFFOLDS = number of scaffolds >= 10 kb

- 11.37 Kb = EDGE N50 = N50 edge size

- 68.49 Kb = CONTIG N50 = N50 contig size

- 10.04 Kb = PHASEBLOCK N50 = N50 phase block size

- 3.57 Mb = SCAFFOLD N50 = N50 scaffold size

- 10.43 % = MISSING 10KB = % of base assembly missing from scaffolds >= 10 kb

- 266.26 Mb = ASSEMBLY SIZE = assembly size (only scaffolds >= 10 kb)

--------------------------------------------------------------------------------

ALARMS

--------------------------------------------------------------------------------