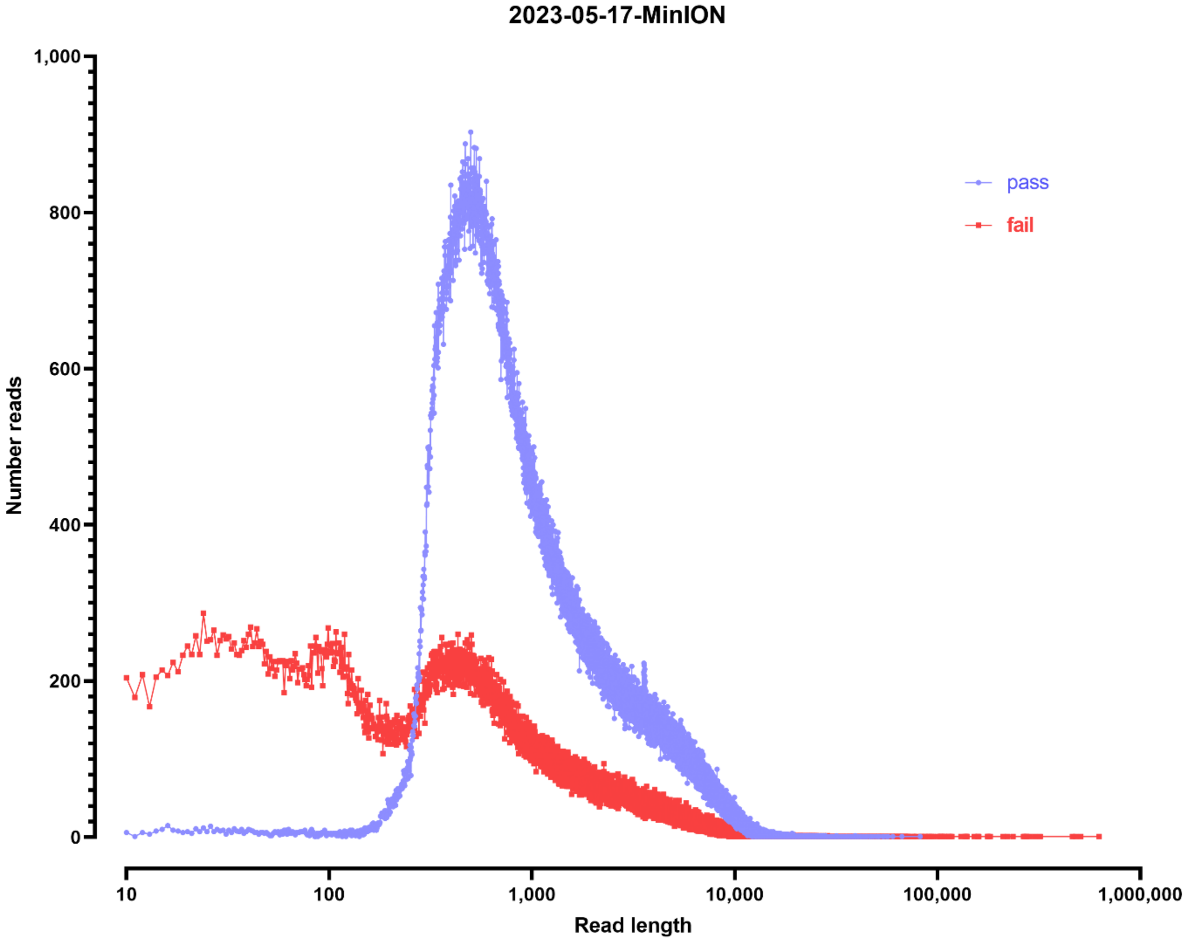
# PREVIOUS ASSEMBLIES

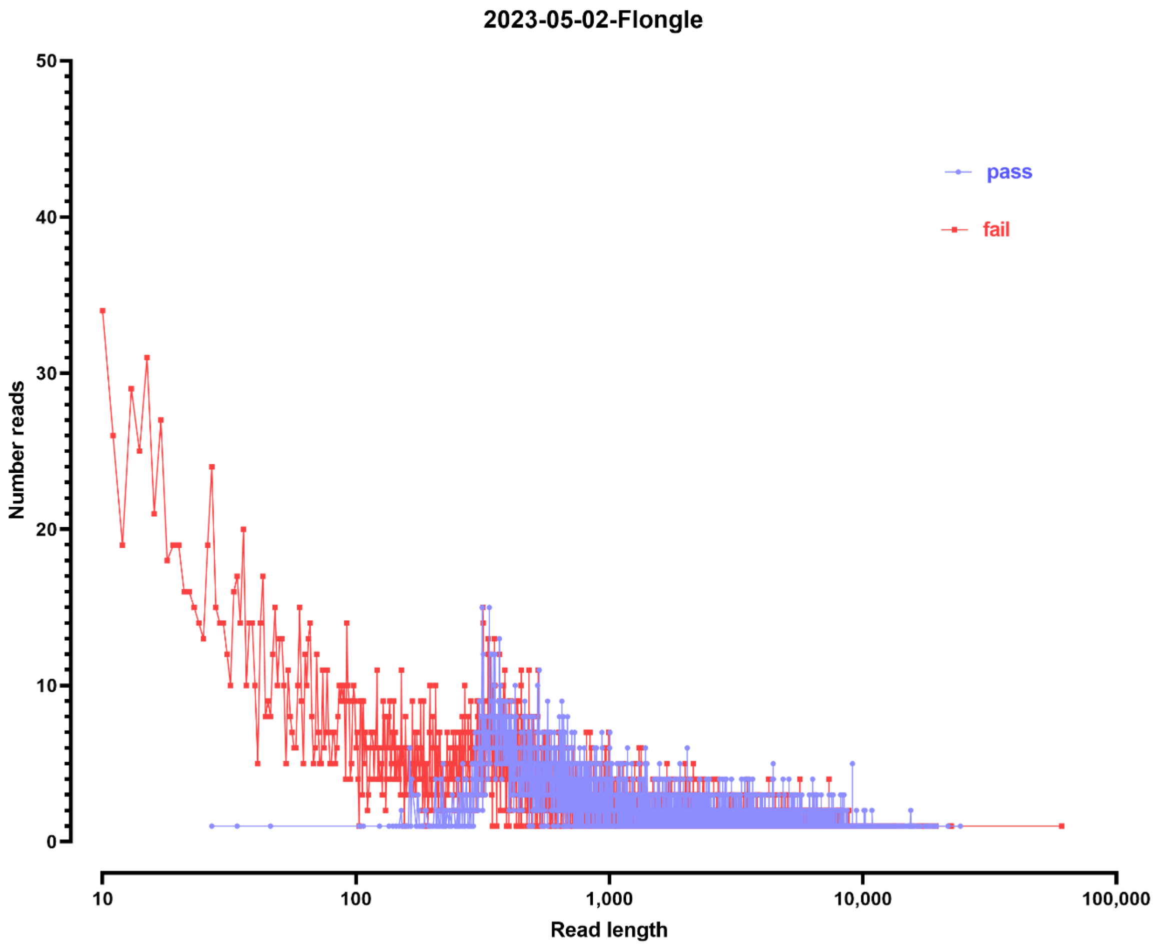
# MinION read length distribution

~7 Gbp



# Flongle read length distribution

~60 Mbp



# Short read data:

| | | | | |
|---|---|---|---|---|
| L2789 | yeast_YAC_Matt_1 | 230522_NB551514_0056_AHNC33BGXN | 2x38 | 3,736,315 |
| L2790 | yeast_YAC_Matt_1 | 230522_NB551514_0056_AHNC33BGXN | 2x38 | 3,194,013 |

# SPAdes assembly Illumina + ONT

```
N50         11,250 bp

NODE_1_length_58913_cov_3.402649    58913
NODE_2_length_55548_cov_3.600666    55548
NODE_3_length_43828_cov_3.742916    43828
NODE_4_length_42124_cov_3.727020    42124
NODE_5_length_41786_cov_3.544105    41786
NODE_6_length_39554_cov_2.970320    39554
NODE_7_length_39293_cov_2.902369    39293
NODE_8_length_38432_cov_3.002448    38432
NODE_9_length_38014_cov_3.463258    38014
NODE_10_length_36870_cov_2.472840   36870
```

This is weird – there is so much nanopore data…

# SPAdes assembly Illumina only:

```
N50         3,068 bp

NODE_1_length_21614_cov_6.438580    21614
NODE_2_length_16447_cov_5.123797    16447
NODE_3_length_14171_cov_4.914274    14171
```

# Shasta 0.10.1 assembly, ONT only

N50 4,167 bp

# wtdbg-2.5 assembly, ONT only

N50 31,354 bp

# NECAT assembly, ONT only

N50 954,962 bp

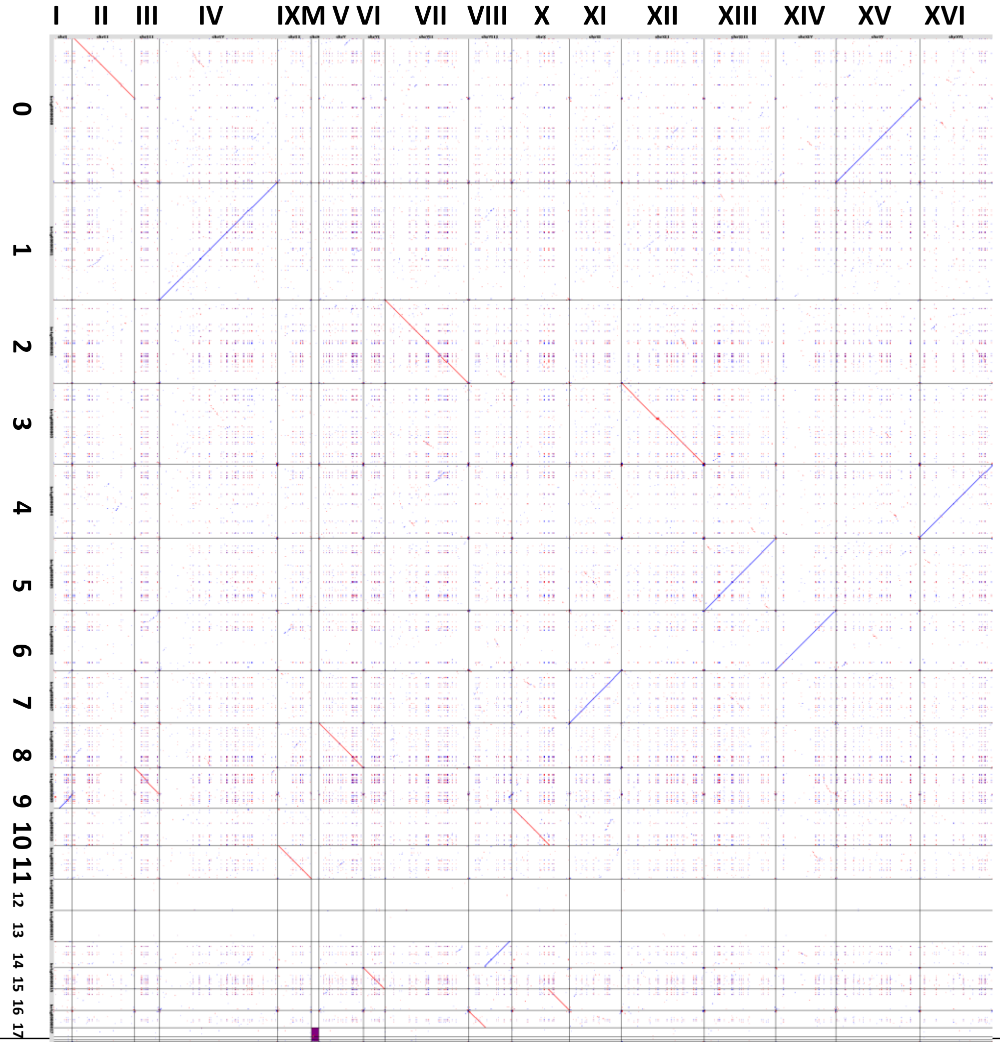
```
bctg00000019    30247
bctg00000020    25288
bctg00000021    13832
bctg00000022    968
bctg00000018    111478
bctg00000016    271877
bctg00000017    217462
bctg00000015    277415
bctg00000012    401330
bctg00000014    331861
bctg00000013    399974
bctg00000011    429026
bctg00000009    524798
bctg00000010    477224
bctg00000008    579349
bctg00000007    679205
bctg00000006    779327
bctg00000004    954962
bctg00000005    939887
bctg00000002    1085228
bctg00000003    1053385
bctg00000001    1529316
bctg00000000    1873469
```

# NECAT assembly, ONT only; BLAST results:

```
bctg00000019      30247    mitochondrion
bctg00000020      25288    Expression vector pSB229
bctg00000021      13832    mitochondrion
bctg00000022      968      chrIV
bctg00000018      111478   mitochondrion
```

The rest is too large to run with the online BLAST tool

# NECAT assembly, ONT only; whole genome alignment against sacCer3



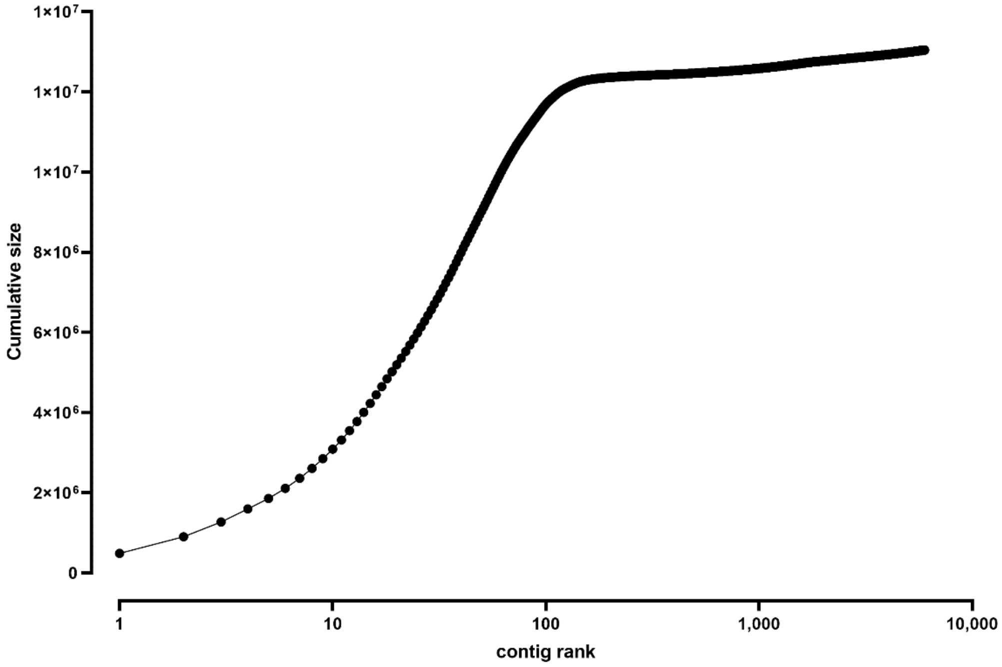
# WITH UPDATED NOVASEQ DATA

## Short read data:

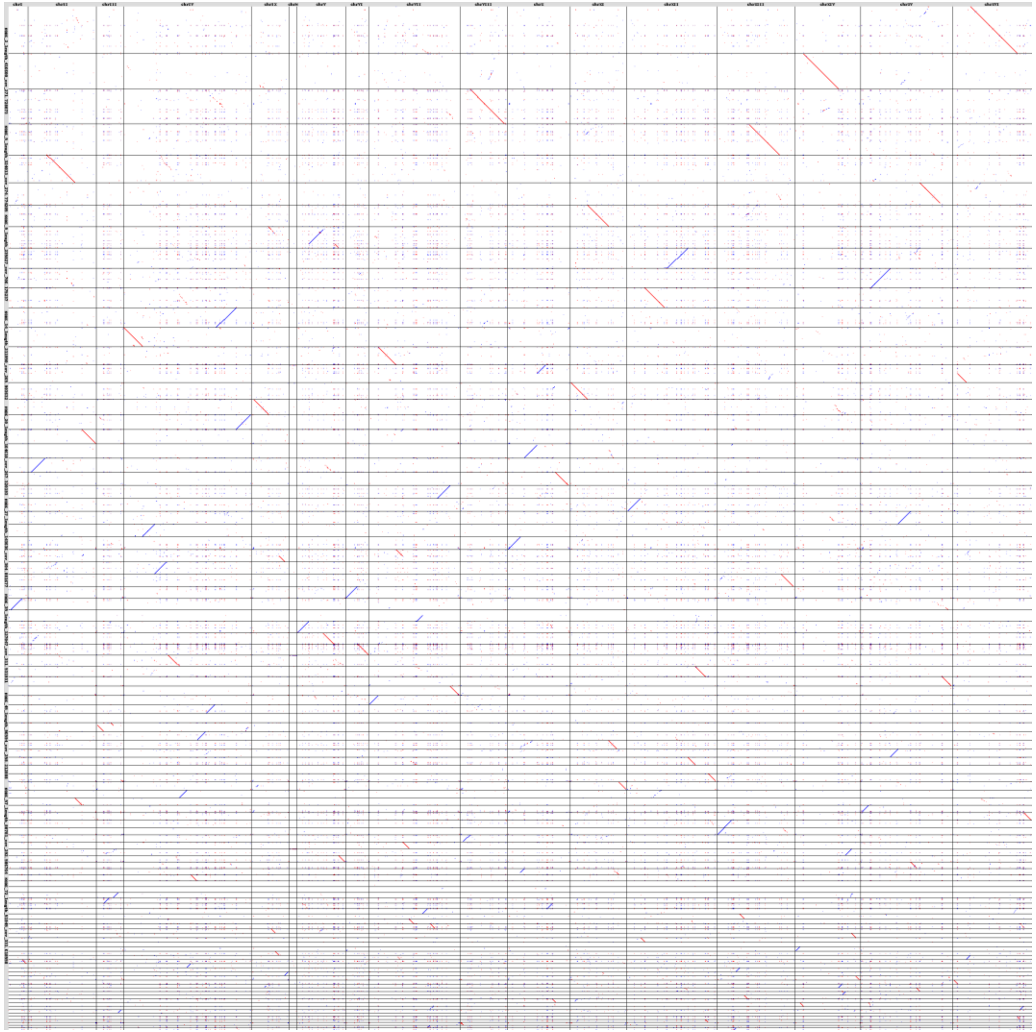
| | | | | |
|---|---|---|---|---|
| L2789 | yeast_YAC_Matt_1 | 230522_NB551514_0056_AHNC33BGXN | 2x38 | 3,736,315 |
| L2790 | yeast_YAC_Matt_1 | 230522_NB551514_0056_AHNC33BGXN | 2x38 | 3,194,013 |
| | | CKDL230018741-1A_225KKJLT3_L2 | 2x150 | 80,916,164 |

## SPAdes assembly Illumina + ONT

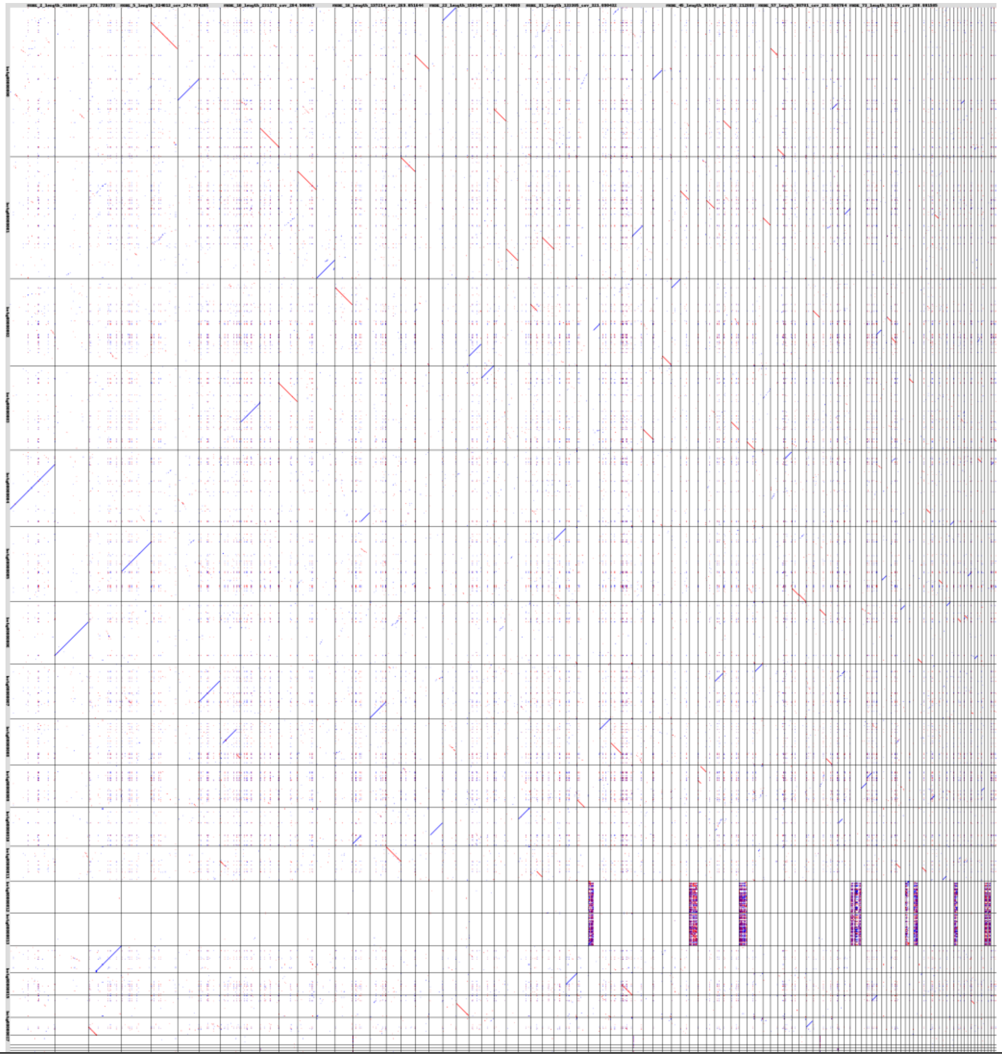
```
N50          141,094 bp

NODE_1_length_485139_cov_304.603745 485139
NODE_2_length_416600_cov_271.966767 416600
NODE_3_length_369332_cov_266.547291 369332
NODE_4_length_323787_cov_274.668778 323787
NODE_5_length_260118_cov_255.803113 260118
NODE_6_length_255486_cov_267.614961 255486
NODE_7_length_250130_cov_272.256015 250130
NODE_8_length_248254_cov_322.828513 248254
NODE_9_length_239762_cov_268.664903 239762
NODE_10_length_235827_cov_706.113321235827
```



# whole genome alignment against sacCer3



# whole genome alignment against the NECAT assembly



# SPAdes assembly Illumina only:

```
N50          34,470

NODE_1_length_125259_cov_264.502819     125259
NODE_2_length_124980_cov_278.791355     124980
NODE_3_length_121603_cov_269.775266     121603
NODE_4_length_112055_cov_339.881679     112055
NODE_5_length_110676_cov_244.957739     110676
NODE_6_length_107579_cov_260.590715     107579
NODE_7_length_102685_cov_251.435847     102685
NODE_8_length_96069_cov_264.200669      96069
NODE_9_length_95322_cov_265.567815      95322
NODE_10_length_94610_cov_259.480429     94610
```

# SPAdes assembly Illumina + ONT + NECAT as trusted contigs

```
N50         361,639 bp

NODE_1_length_635429_cov_535.230310   635429
NODE_2_length_591528_cov_339.146962   591528
NODE_3_length_578907_cov_567.743736   578907
NODE_4_length_577267_cov_388.173784   577267
NODE_5_length_519707_cov_493.958617   519707
NODE_6_length_518279_cov_313.443229   518279
NODE_7_length_491309_cov_276.163712   491309
NODE_8_length_471186_cov_277.625418   471186
NODE_9_length_427514_cov_309.102679   427514
NODE_10_length_412351_cov_326.292513  412351
NODE_11_length_397030_cov_285.063425  397030
NODE_12_length_375451_cov_279.093658  375451
NODE_13_length_369321_cov_267.150369  369321
NODE_14_length_361639_cov_274.243166  361639
NODE_15_length_345757_cov_310.453605  345757
NODE_16_length_330076_cov_292.666961  330076
NODE_17_length_261094_cov_285.504495  261094
NODE_18_length_260115_cov_256.453688  260115
NODE_19_length_251113_cov_310.335341  251113
NODE_20_length_244809_cov_310.965181  244809
NODE_21_length_238063_cov_3534.724950 238063
NODE_22_length_236756_cov_308.451781  236756
NODE_23_length_219415_cov_550.835467  219415
```

```
NODE_24_length_196329_cov_437.542537  196329
NODE_25_length_184719_cov_286.156338  184719
NODE_26_length_176762_cov_495.501933  176762
NODE_27_length_167174_cov_685.479766  167174
NODE_28_length_164348_cov_444.778779  164348
NODE_29_length_135708_cov_1625.830199 135708
NODE_30_length_133388_cov_336.071850  133388
NODE_31_length_133295_cov_332.682625  133295
NODE_32_length_133288_cov_475.670307  133288
NODE_33_length_128538_cov_472.769152  128538
NODE_34_length_115653_cov_1009.856736 115653
NODE_35_length_114725_cov_323.455080  114725
NODE_36_length_103550_cov_274.350674  103550
NODE_37_length_102633_cov_373.677757  102633
```

This seems to actually break the contiguity…

# SPAdes assembly Illumina + ONT + NECAT as trusted contigs