

1
2

TABLE S1: ASSEMBLY METRICS (GENERATED USING WEB QUASt), ALL STATISTICS ARE BASED ON CONTIGS OF SIZE \geq 500 BP

# contigs (\geq 0 bp)	53
# contigs (\geq 1000 bp)	48
# contigs (\geq 5000 bp)	36
# contigs (\geq 10000 bp)	33
# contigs (\geq 25000 bp)	25
# contigs (\geq 50000 bp)	19
Total length (\geq 0 bp)	2440756
Total length (\geq 1000 bp)	2436775
Total length (\geq 5000 bp)	2399806
Total length (\geq 10000 bp)	2380854
Total length (\geq 25000 bp)	2244693
Total length (\geq 50000 bp)	2017446
# contigs	53
Largest contig	210871
Total length	2440756
GC (%)	32.01

N50	108515
N75	60954
L50	9
L75	16
# N's per 100 kbp	2.91
Average coverage	155

3
4

5 **TABLE S2: SIMILARITY MATRIX FOR EACH STRAIN IN THIS STUDY, G6 2 AND REFERENCE STRAINS**

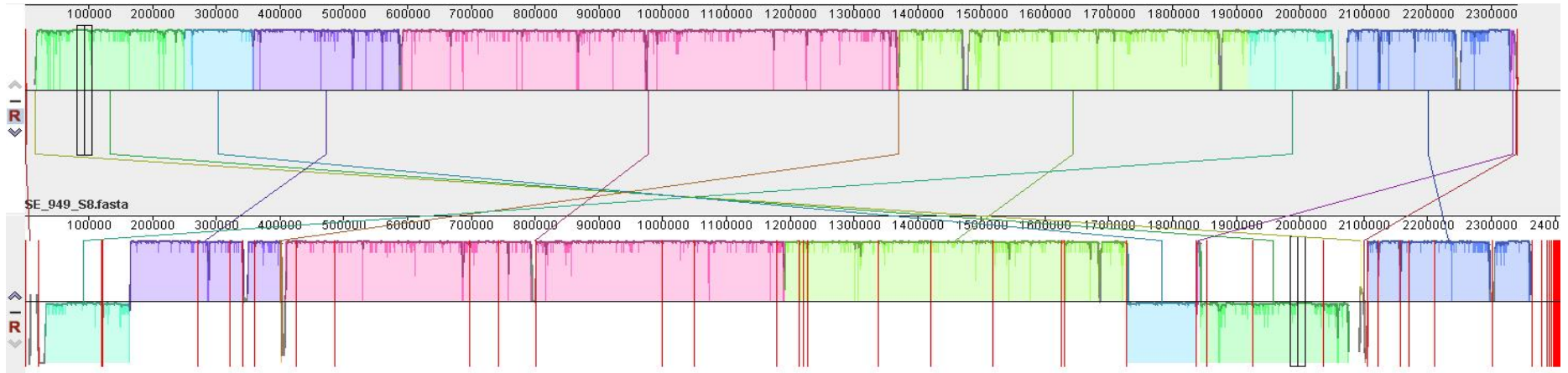
	<u>ATCC12228</u>	<u>BPH0662</u>	<u>PM221</u>	<u>RP62A</u>	<u>SEI</u>	<u>949 S8</u>	<u>G6 2</u>
<u>ATCC12228</u>	0	0.0984105	0.0672375	0.082899	0.0691486	0.0546022	0.0564311
<u>BPH0662</u>	0.0984105	0	0.103142	0.0781847	0.110307	0.10355	0.0963388
<u>PM221</u>	0.0672375	0.103142	0	0.0928811	0.0725552	0.0646446	0.070051
<u>RP62A</u>	0.082899	0.0781847	0.0928811	0	0.0966468	0.0755111	0.0834425
<u>SEI</u>	0.0691486	0.110307	0.0725552	0.0966468	0	0.065689	0.0689457
<u>949 S8</u>	0.0546022	0.10355	0.0646446	0.0755111	0.065689	0	0.0482764
<u>G6 2</u>	0.0564311	0.0963388	0.070051	0.0834425	0.0689457	0.0482764	0

6

7

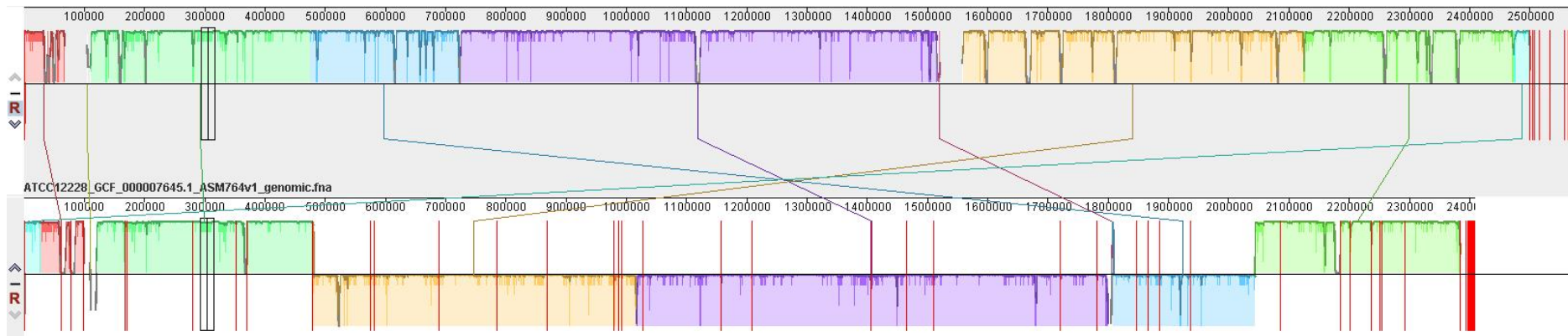
8

a)



G6 2

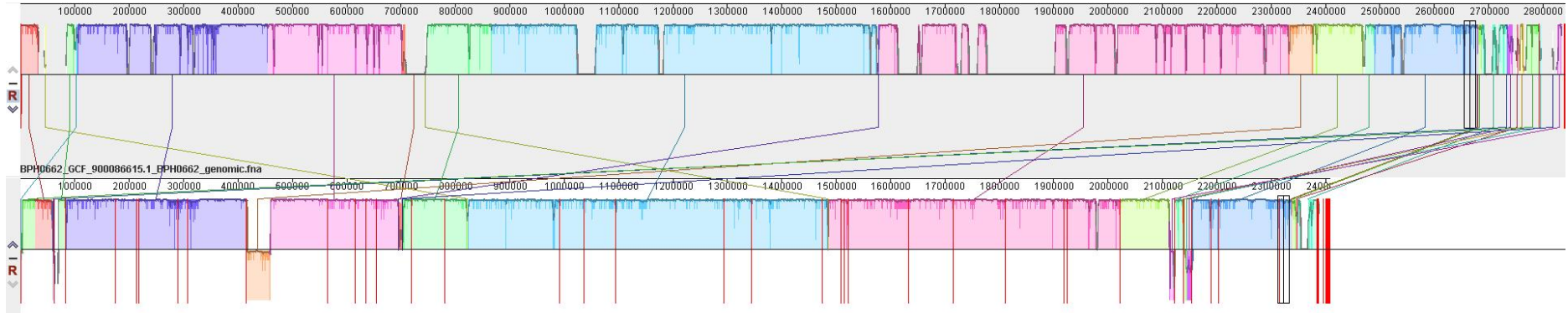
b)



G6 2

9

c)

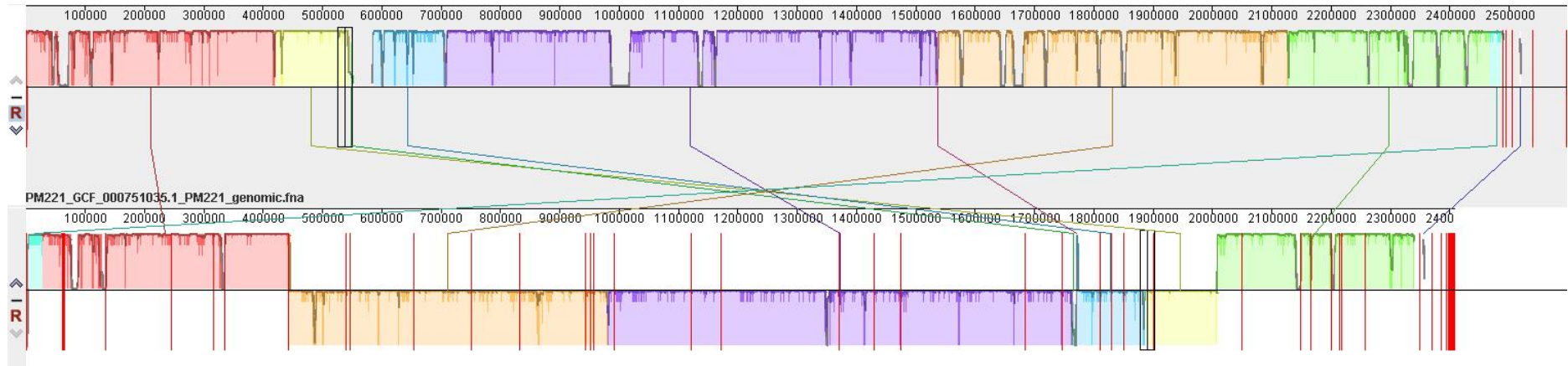


G6 2

10

11

d)



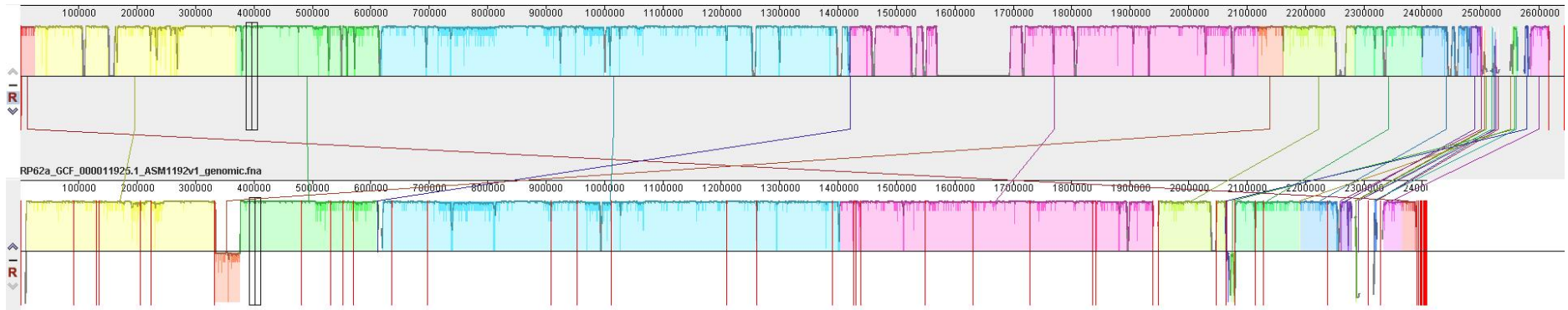
12

G6 2

13

14

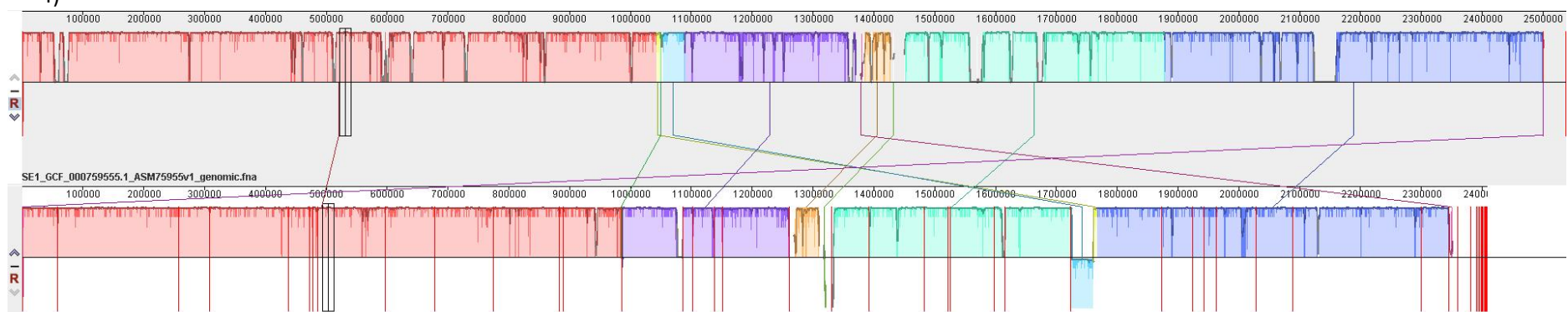
e)



G6 2

15

f)



G6 2

16

17 **Figure S1. Pairwise allignments of the genome presented in this study; G6_2 compared against a) 949_S8; b) ATCC12228; c) BPH 0662;**
18 **d) PM221, e) RP62A and e) SEI**