Defining the sequence requirements for the positioning of base J in DNA using SMRT sequencing

Paul-Andre Genest, Loren Baugh, Alex Taipale, Wanqi Zhao, Sabrina Jan, Henri G.A.M. van Luenen, Jonas Korlach, Tyson Clark, Khai Luong, Matthew Boitano, Steve Turner, Peter J. Myler and Piet Borst

SUPPLEMENTARY DATA

Supplementary Figure S1. Selection of IPD ratio cut-off for calling Js. Based on the synthetic oligonucleotides containing J, any set of peaks with an IPD ratio >1.6 at a T position or at 2 or 6 nucleotides downstream from a T, is scored as J. The 1.6 cut-off was chosen to minimize false-positive and false-negative rates (FPR, FNR). **(A)** The distribution of single molecule IPD ratios (in 0.1 bins) at all positions in the J-less 25.2S plasmid (grown in the JBP2-null strain, blue bars) was subtracted from the total distribution from all samples at all positions (red bars), to estimate the distribution of IPD ratios due to J (green bars). FPR was defined as the percentage of IPD ratios greater than the cut-off value in the JBP2-null sample; FNR was defined as the percentage of IPD ratios smaller than the cut-off value in the "Difference" population due to J. **(B)** A cut-off of 1.6 was chosen to minimize the estimated false positive and false negative rates (FPR and FNR).

Supplementary Figure S2. Partial DNA digest with restriction enzyme DdeI. Supplementary Figure S2 shows blots of plasmid digests with restriction endonuclease DdeI, hybridized with probes corresponding to the genomic *Leishmania* DNA, inserted into the plasmid. Here we present our interpretation of the main partial digest products observed. This interpretation is based on the assumption that all partial digest products are due to a partial J-modification of the DdeI sites investigated, as we do not see these products in plasmid DNA from *E. coli*. **(A)** Overview of the DdeI restriction sites in the 25.2L plasmid and the potential hybridization products when the DdeI sites are fully or partially cut because of the presence of J. **(B)** Plasmid DNA was purified from *L. tarentolae* and digested with DdeI and increasing amounts were loaded on gel. Also genomic DNA from wild type and JBP2-null *L. tarentolae* was isolated and digested with DdeI. We also isolated DNA from JBP2-null cultures grown in the presence of BrdU, which lowers J-levels further (1). The DNA was size-fractionated in an agarose gel, blotted and probed.

The main partial digest band visible with both probes is a 1358-bp band spanning the two DdeI sites at positions 4238 and 4308 in the plasmid containing the 25.2L insert. SMRT sequencing shows that these sites are indeed partially modified as shown by the sequence maps in Supplementary Figure S3 which result in an estimated modification of 30% (site 4238) and 34% (site 4308) based on the analysis of individual molecules. Note that these two sites are not part of detectable "doublet", in contrast to the surrounding J- residues that are >75% modified and part of a doublet.

There are several other partial digest bands in Supplementary Figure S2. The weak 1054 band can be explained by partial cutting of site 4238, and the prominent 374 band, only seen by the 25.2S probe, by partial cleavage of site 4308. The larger bands, seen by both probes, cannot be unambiguously allocated, as we do not see any additional modified DdeI sites in the SMRT sequence. Notably, the 2602 band might also be located between DdeI sites 1991 and 4612. The larger bands only seen by the 25.2S probe, also require modified DdeI sites, not visibly modified in the SMRT sequence.

Genomic DNA was also digested with DdeI. The partially cut 1.4 kb fragment is in wild-type DNA and disappears in the JBP2-null mutant.

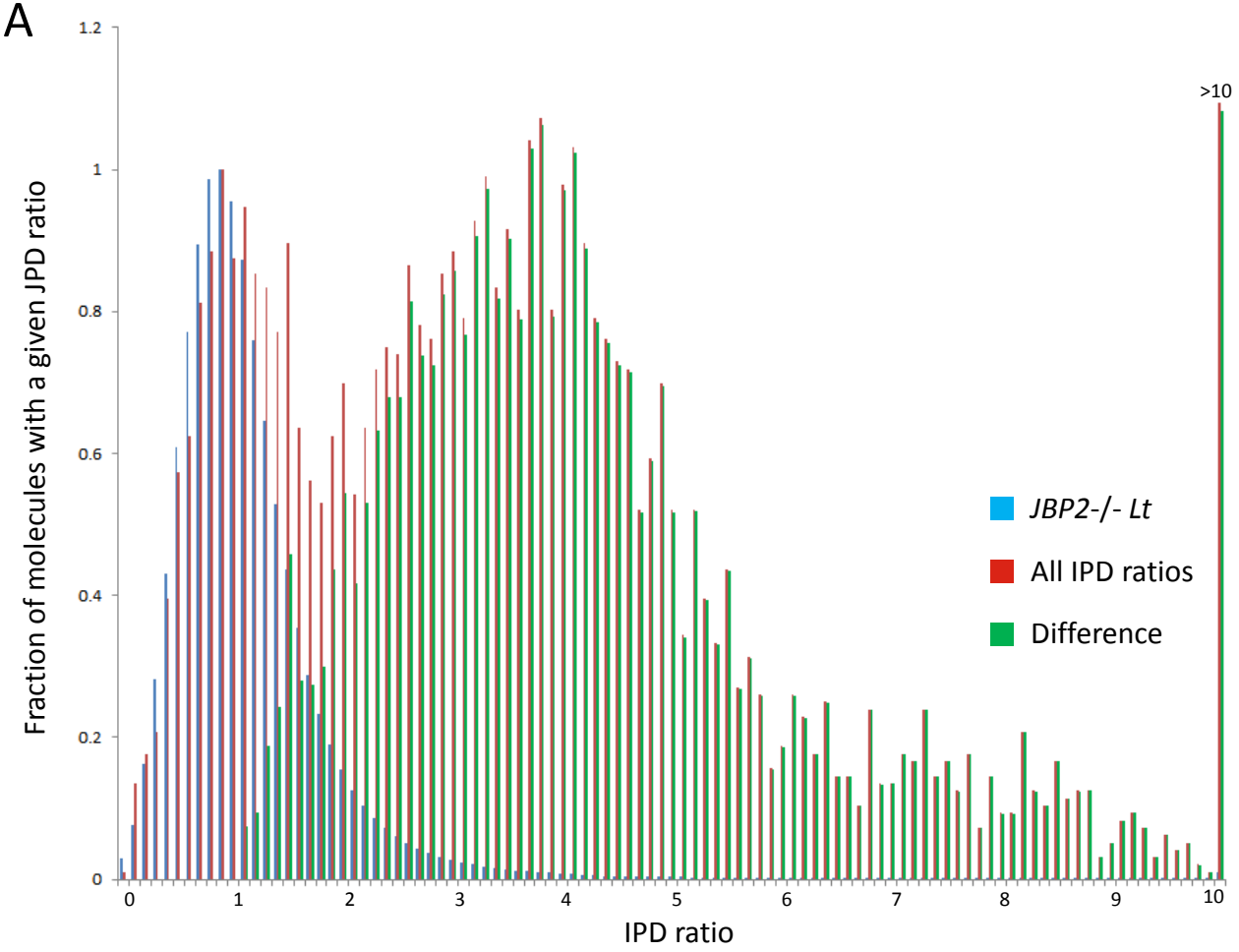
Supplementary Figure S3. Nucleotide-resolution kinetograms of plasmids isolated from *L. tarentolae*. Plasmids containing a cSSR (25.2S, 25.2L 28.2 or 12.1) or the telomeric repeat sequence $(GGGTTA)_{10}$ were grown in *L. tarentolae*, purified and SMRT sequenced. The IPD ratio plots for the complete plasmids are shown. The sequences have been aligned in the vector sequences and the relevant restriction sites, plasmids fragments, and G-quadruplexes have been indicated

Supplementary Figure S4. Average kinetic signature for base J. All cases in which there was only one T within 7 nucleotides with an IPD ratio >1.6 at relative positions 0 (T), +1, or +2 were included in the calculation; thirteen cases were found (sequences listed above). The direction of the sequencing polymerase is from left to right.

References

1. Vainio, S., Genest, P.A., Ter Riet, B., Van Luenen, H.G.A.M. and Borst, P. (2009) Evidence that J-binding protein 2 is a thymidine hydroxylase catalyzing the first step in the biosynthesis of DNA base J. *Ml. Biochem. Parasitol.*, **164**, 157-161.

A

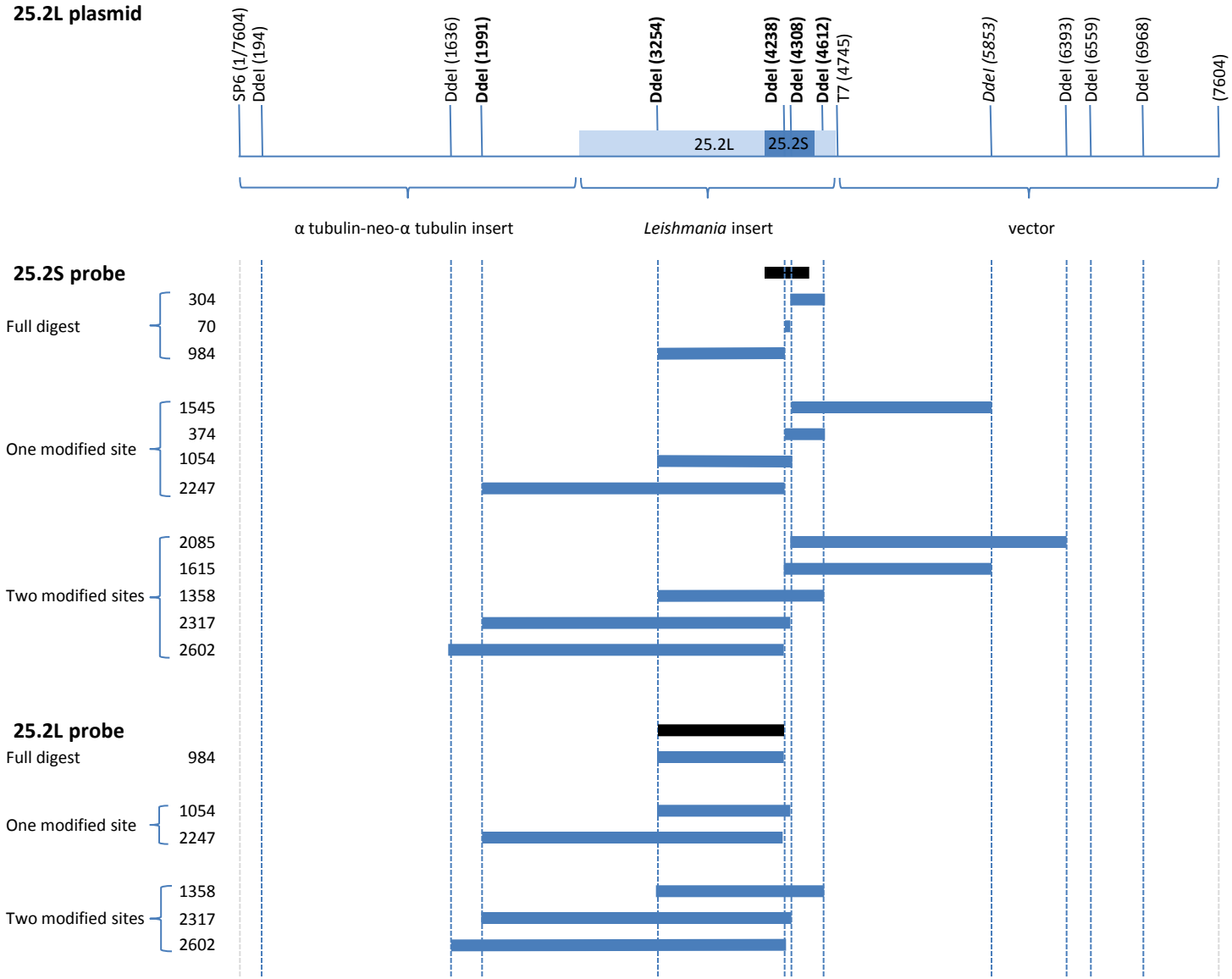


B

cut-off	FPR	FNR
1	17.2%	0.0%
1.1	15.2%	0.0%
1.2	13.2%	0.0%
1.3	11.3%	1.1%
1.4	9.6%	2.1%
1.5	8.1%	4.0%
1.6	6.9%	6.2%
1.7	5.8%	10.1%
1.8	4.9%	12.0%
1.9	4.2%	13.8%
2	3.6%	15.6%
2.1	3.1%	18.3%
2.2	2.7%	21.3%
2.3	2.4%	23.5%
2.4	2.1%	26.1%
2.5	1.9%	28.9%
2.6	1.7%	31.8%
2.7	1.5%	34.4%
2.8	1.4%	37.3%
2.9	1.3%	39.7%
3	1.2%	41.9%

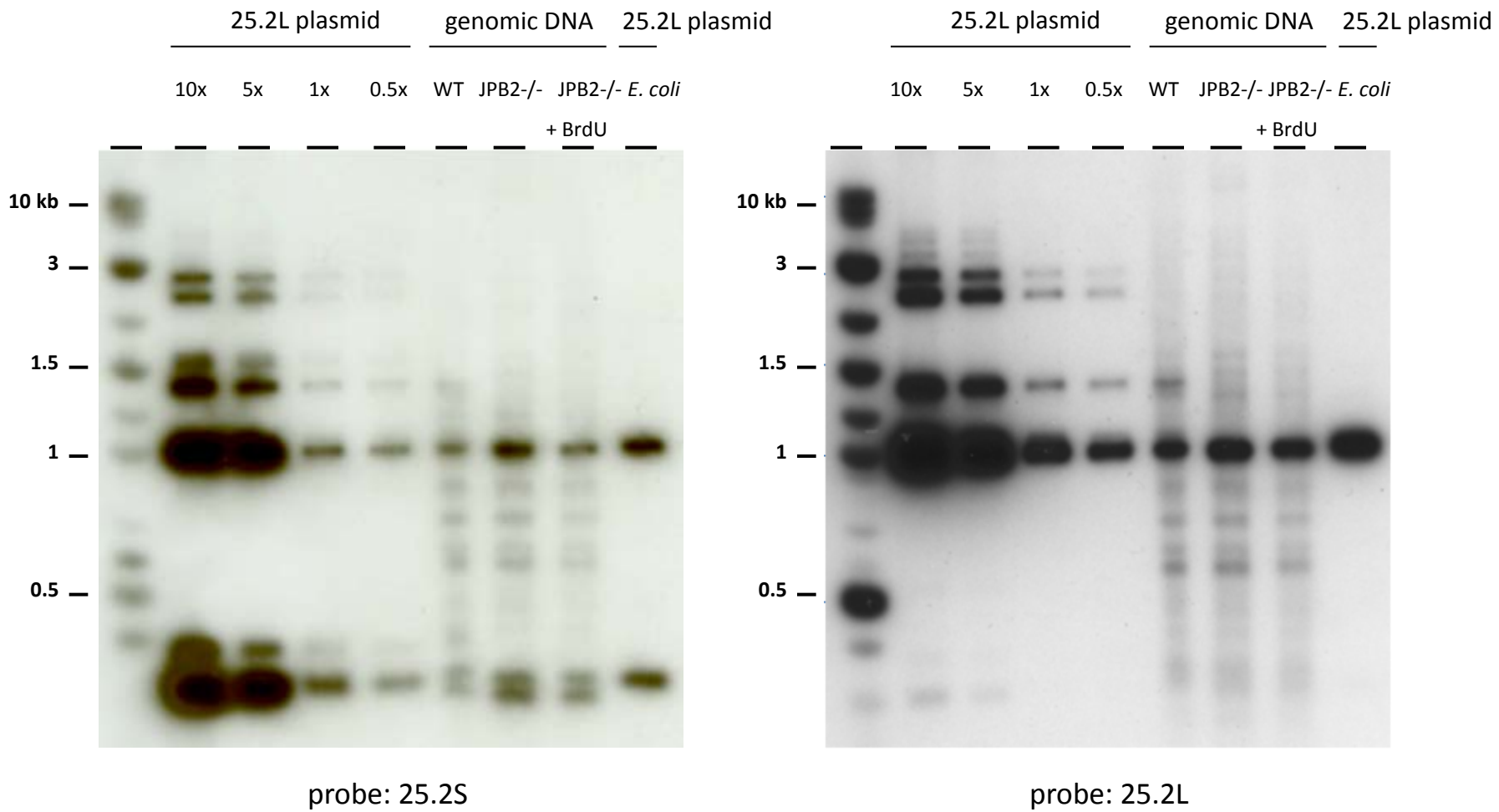
Supplementary Figure S1

A



Supplementary Figure S2

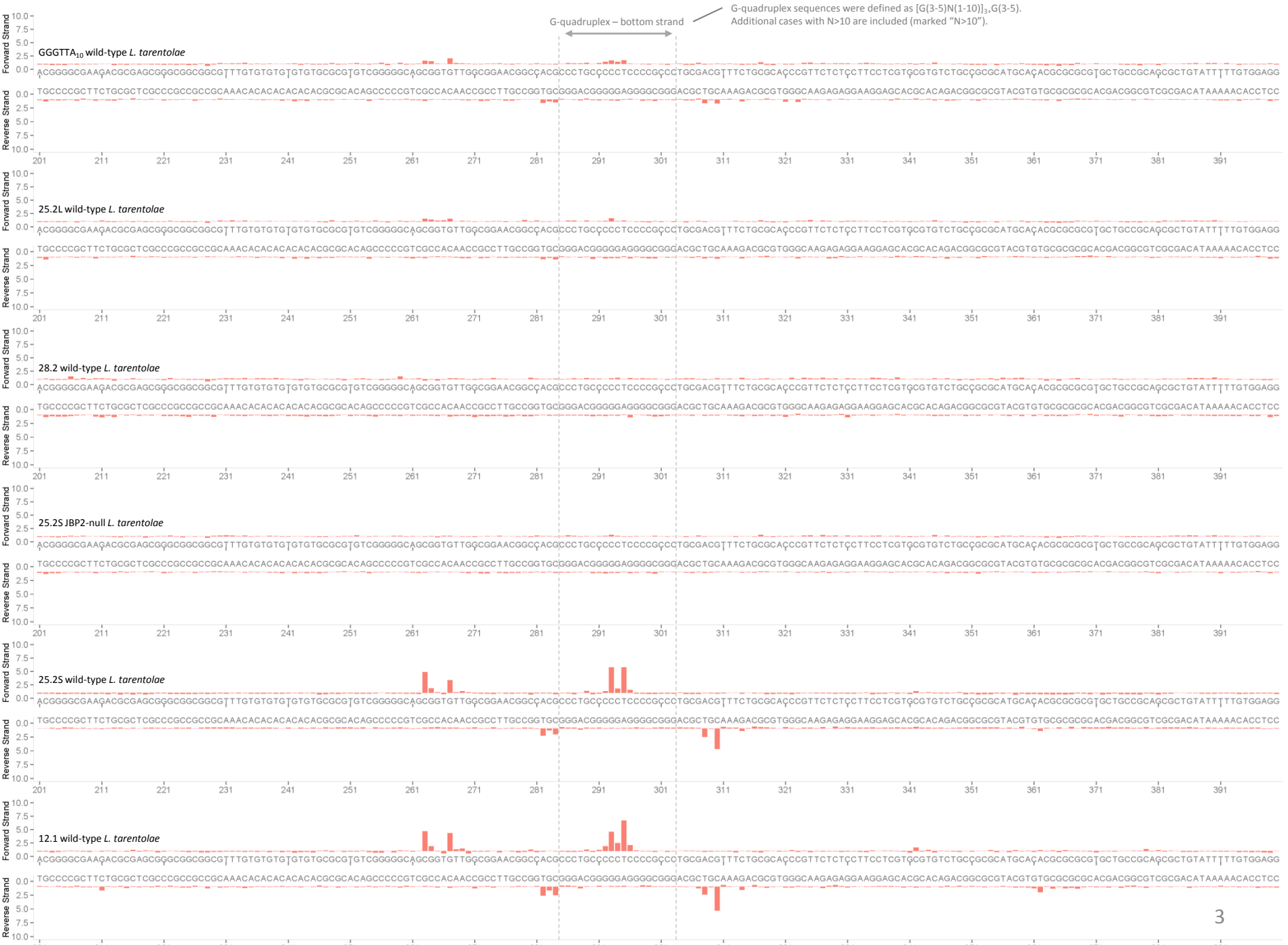
B

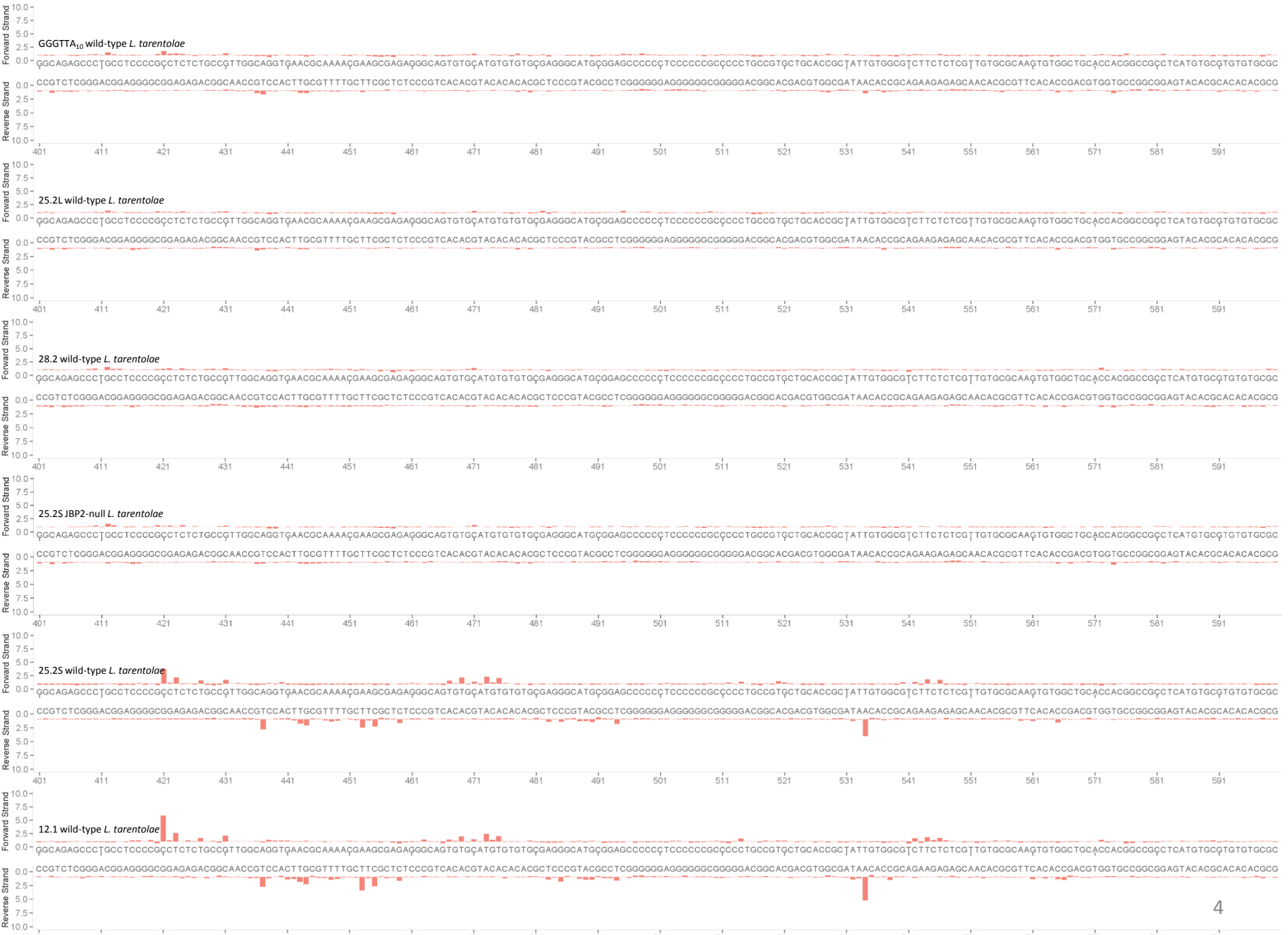


Supplementary Figure S2

Supplementary Figure S3. Nucleotide-resolution kinetograms.


BamHI
α-tubulin IR1 →
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


G-quadruplex – bottom strand
G-quadruplex sequences were defined as [G(3-5)N(1-10)]3,G(3-5).
Additional cases with N>10 are included (marked "N>10").
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
401
411
421
431
441
451
461
471
481
491
501
511
521
531
541
551
561
571
581
591


G-quadruplex – bottom strand
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
601 611 621 631 641 651 661 671 681 691 701 711 721 731 741 751 761 771 781 791


← α-tubulin IR1
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
801 811 821 831 841 851 861 871 881 891 901 911 921 931 941 951 961 971 981 991

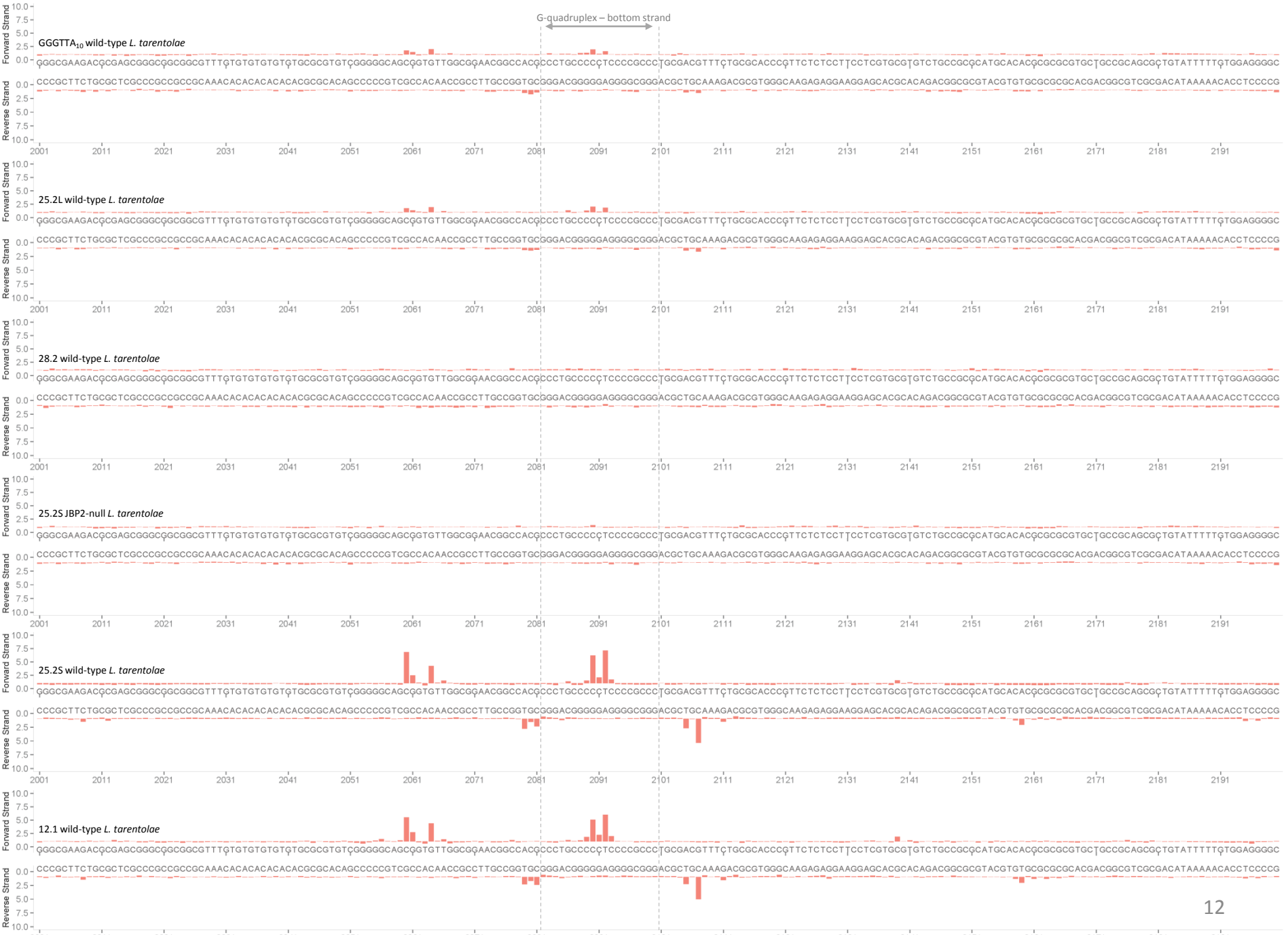

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
1001
1011
1021
1031
1041
1051
1061
1071
1081
1091
1101
1111
1121
1131
1141
1151
1161
1171
1181
1191

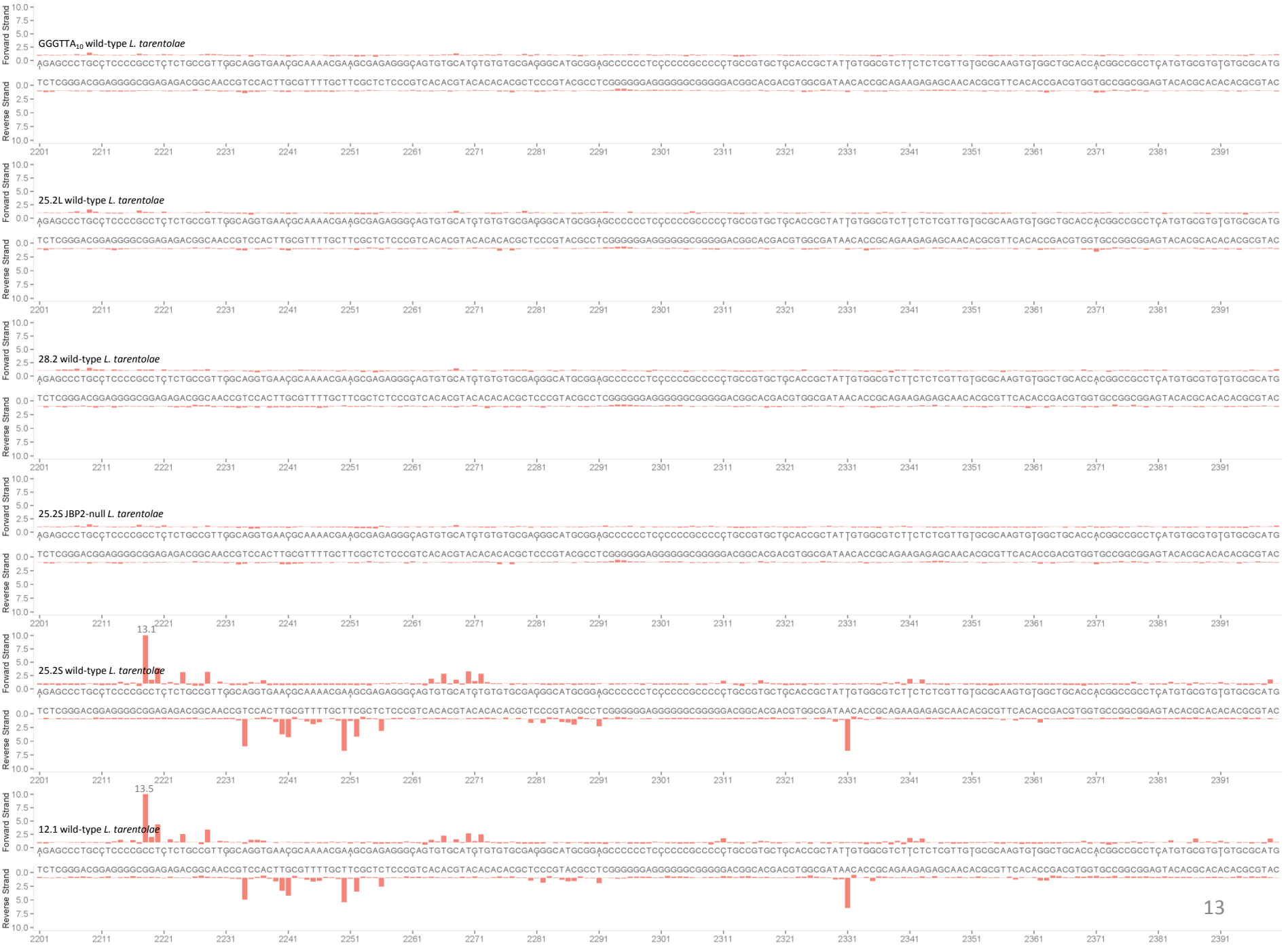

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
1201
1211
1221
1231
1241
1251
1261
1271
1281
1291
1301
1311
1321
1331
1341
1351
1361
1371
1381
1391

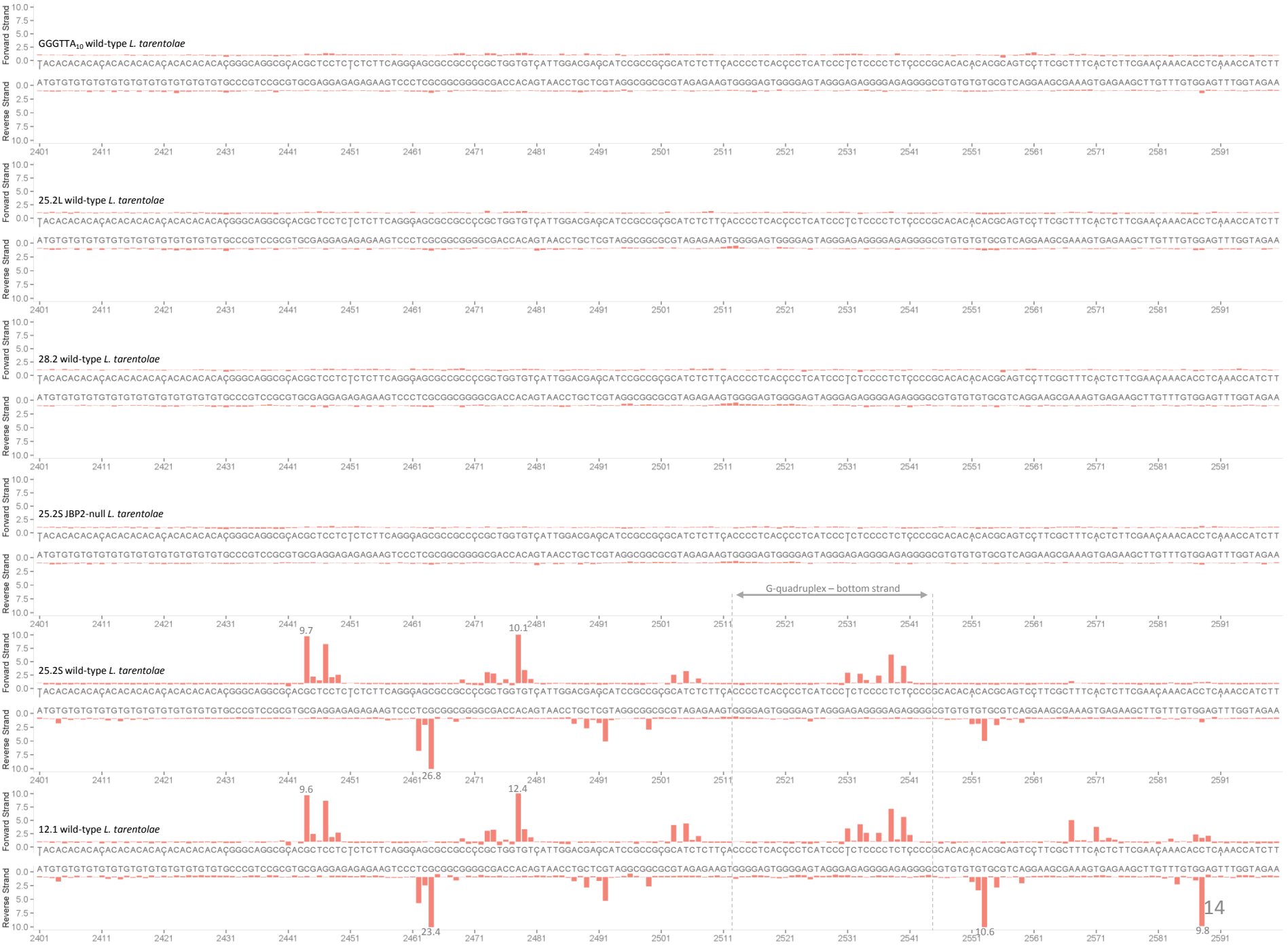

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
1401
1411
1421
1431
1441
1451
1461
1471
1481
1491
1501
1511
1521
1531
1541
1551
1561
1571
1581
1591

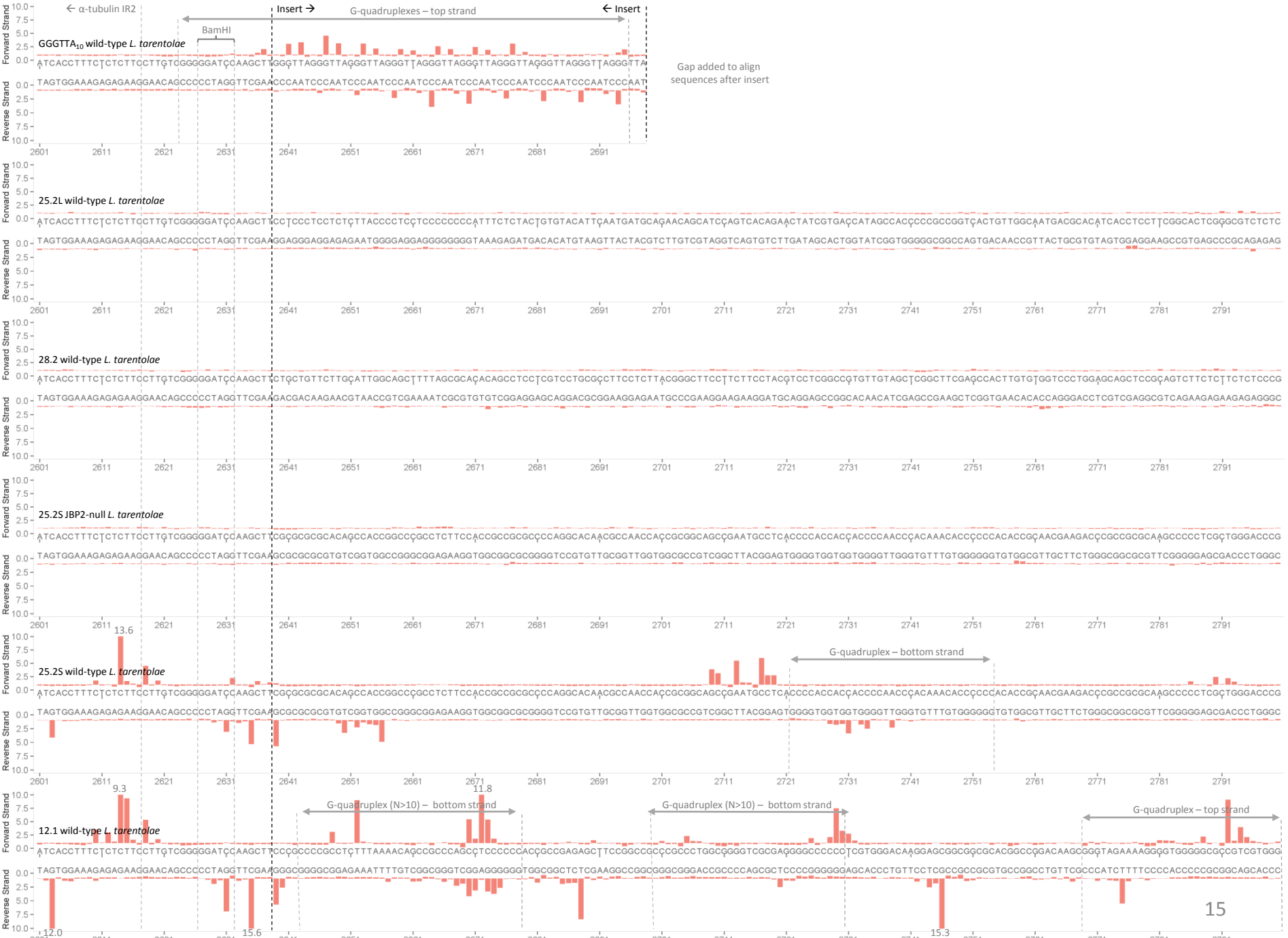

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
1601
1611
1621
1631
1641
1651
1661
1671
1681
1691
1701
1711
1721
1731
1741
1751
1761
1771
1781
1791

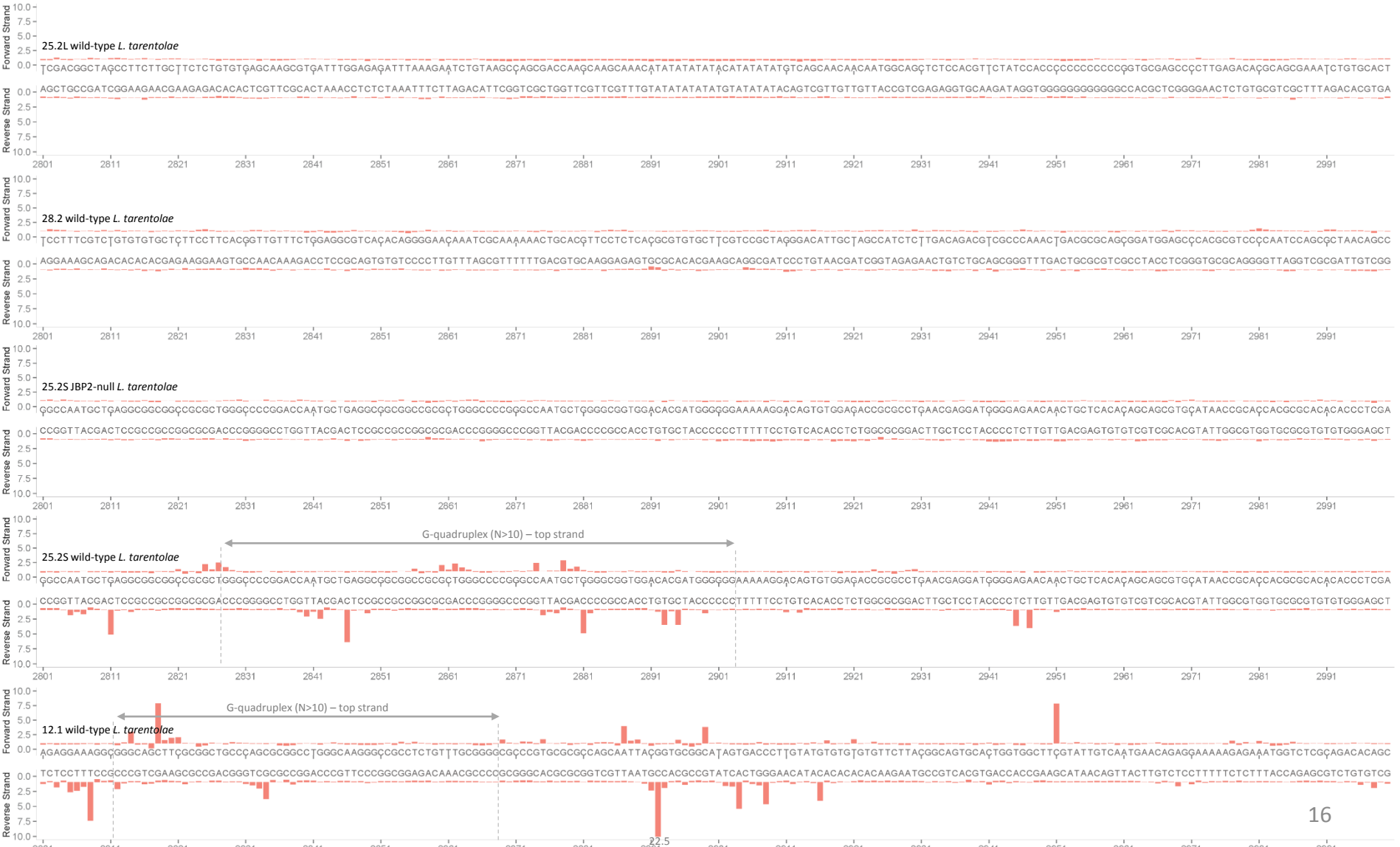

EcoRI
α-tubulin IR2 →
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
1801
1811
1821
1831
1841
1851
1861
1871
1881
1891
1901
1911
1921
1931
1941
1951
1961
1971
1981
1991

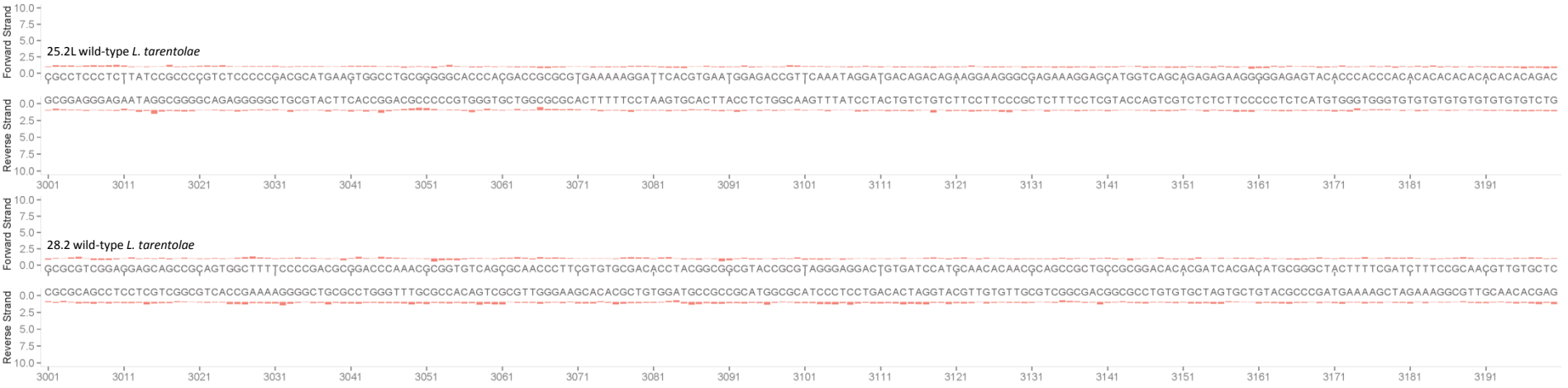

G-quadruplex – bottom strand
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
2001
2011
2021
2031
2041
2051
2061
2071
2081
2091
2101
2111
2121
2131
2141
2151
2161
2171
2181
2191

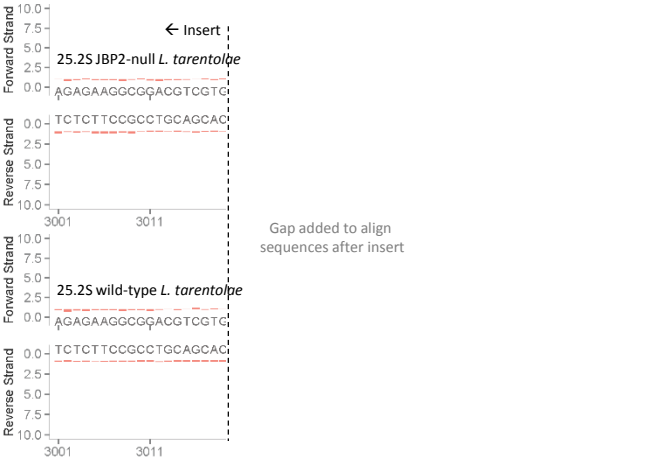

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
13.1
12.1 wild-type L. tarentolae
13.5
Forward Strand
Reverse Strand
2201 2211 2221 2231 2241 2251 2261 2271 2281 2291 2301 2311 2321 2331 2341 2351 2361 2371 2381 2391


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
G-quadruplex – bottom strand
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
9.7
10.1
26.8
9.6
12.4
23.4
10.6
9.8

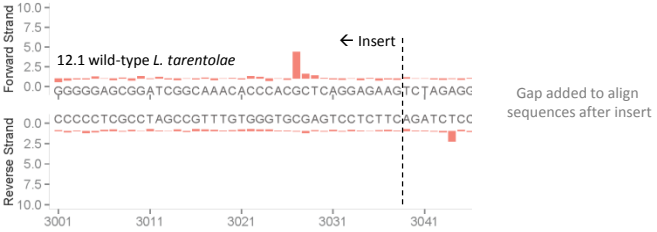

← α-tubulin IR2
Insert →
G-quadruplexes – top strand
← Insert
BamHI
GGGTTA10 wild-type L. tarentolae
Gap added to align sequences after insert
Forward Strand
Reverse Strand
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
G-quadruplex – bottom strand
13.6
12.1 wild-type L. tarentolae
G-quadruplex (N>10) – bottom strand
G-quadruplex (N>10) – bottom strand
G-quadruplex – top strand
9.3
11.8
12.0
15.6
15.3


25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
G-quadruplex (N>10) – top strand
25.2S wild-type L. tarentolae
G-quadruplex (N>10) – top strand
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand
2801
2811
2821
2831
2841
2851
2861
2871
2881
2891
2901
2911
2921
2931
2941
2951
2961
2971
2981
2991





Gap added to align sequences after insert



Gap added to align sequences after insert


25.2L wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3201
3211
3221
3231
3241
3251
3261
3271
3281
3291
3301
3311
3321
3331
3341
3351
3361
3371
3381
3391

28.2 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3201
3211
3221
3231
3241
3251
3261
3271
3281
3291
3301
3311
3321
3331
3341
3351
3361
3371
3381
3391


G-quadruplexes – bottom strand
25.2L wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3401
3411
3421
3431
3441
3451
3461
3471
3481
3491
3501
3511
3521
3531
3541
3551
3561
3571
3581
3591

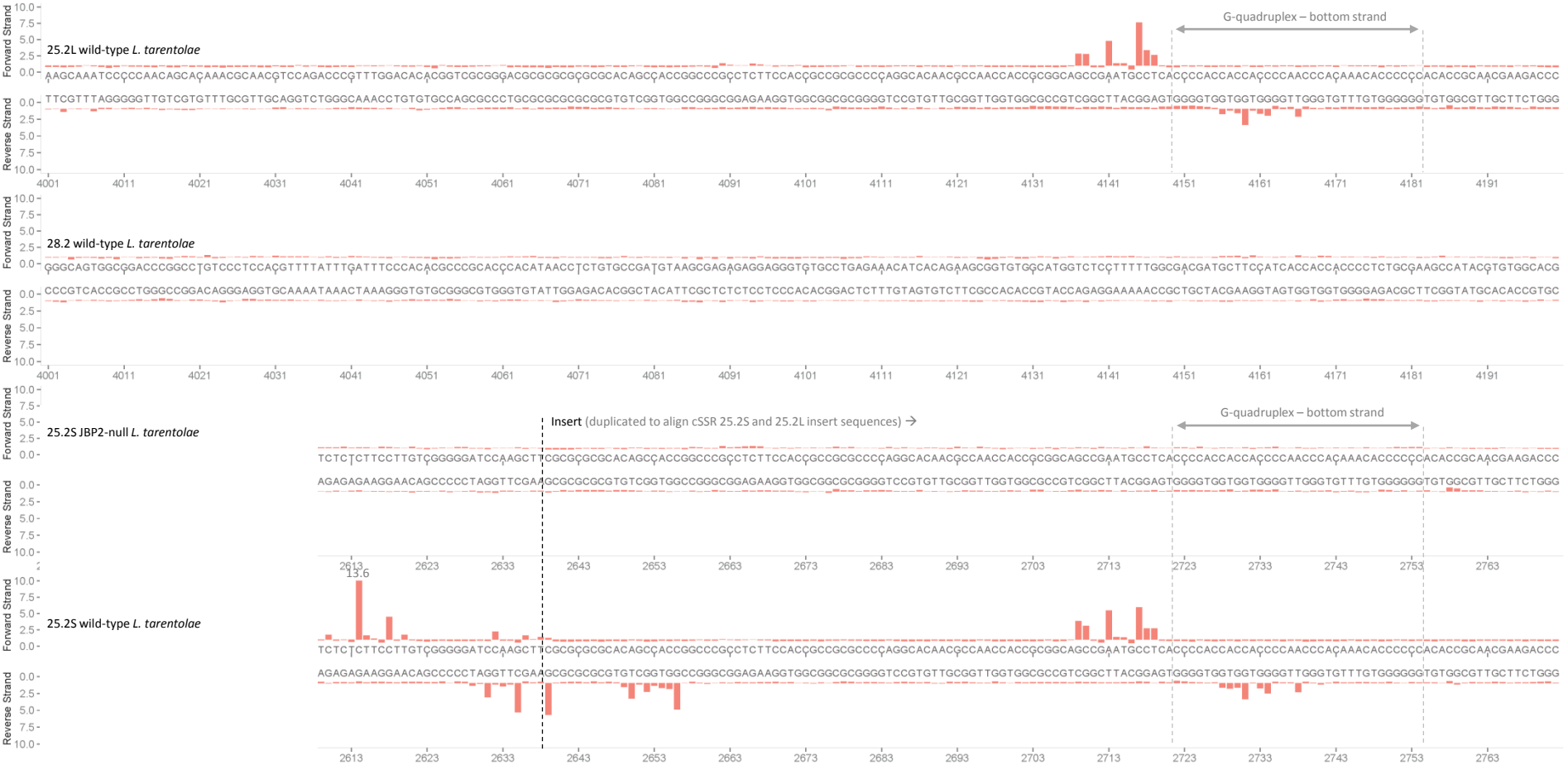
28.2 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3401
3411
3421
3431
3441
3451
3461
3471
3481
3491
3501
3511
3521
3531
3541
3551
3561
3571
3581
3591


25.2L wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3601
3611
3621
3631
3641
3651
3661
3671
3681
3691
3701
3711
3721
3731
3741
3751
3761
3771
3781
3791

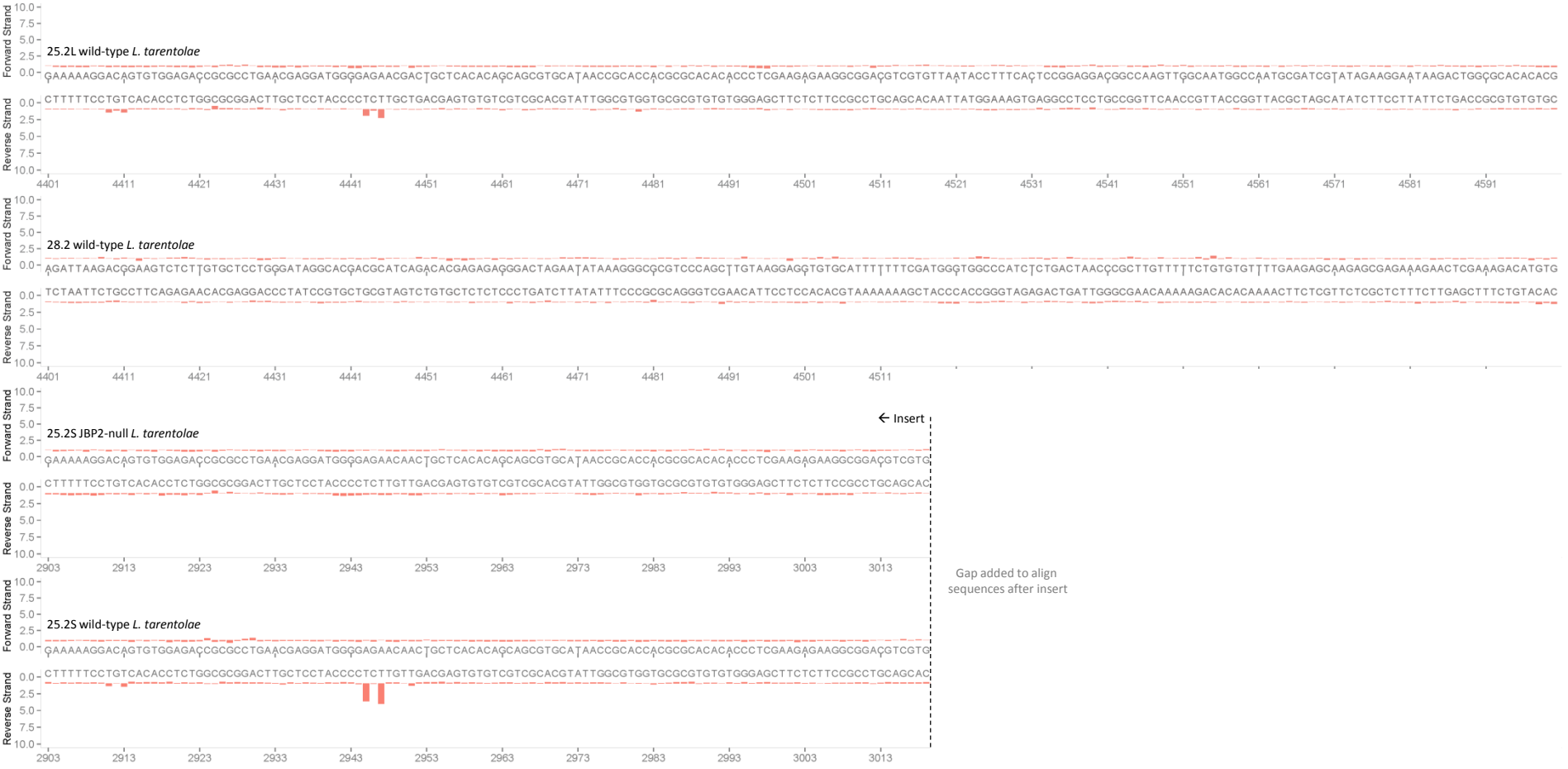
28.2 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3601
3611
3621
3631
3641
3651
3661
3671
3681
3691
3701
3711
3721
3731
3741
3751
3761
3771
3781
3791


25.2L wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3801
3811
3821
3831
3841
3851
3861
3871
3881
3891
3901
3911
3921
3931
3941
3951
3961
3971
3981
3991

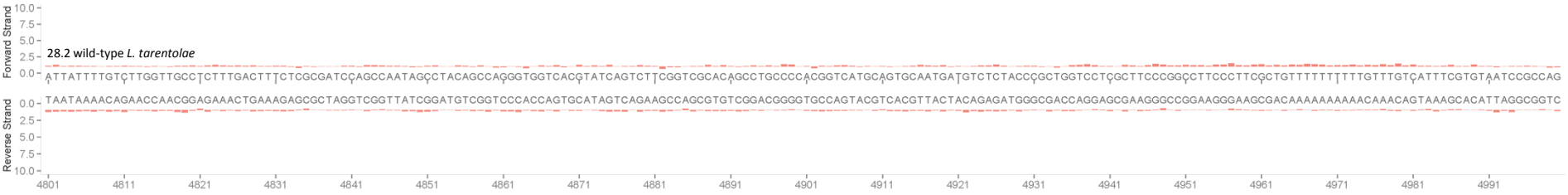
28.2 wild-type L. tarentolae
Forward Strand
Reverse Strand
10.0
7.5
5.0
2.5
0.0
3801
3811
3821
3831
3841
3851
3861
3871
3881
3891
3901
3911
3921
3931
3941
3951
3961
3971
3981
3991

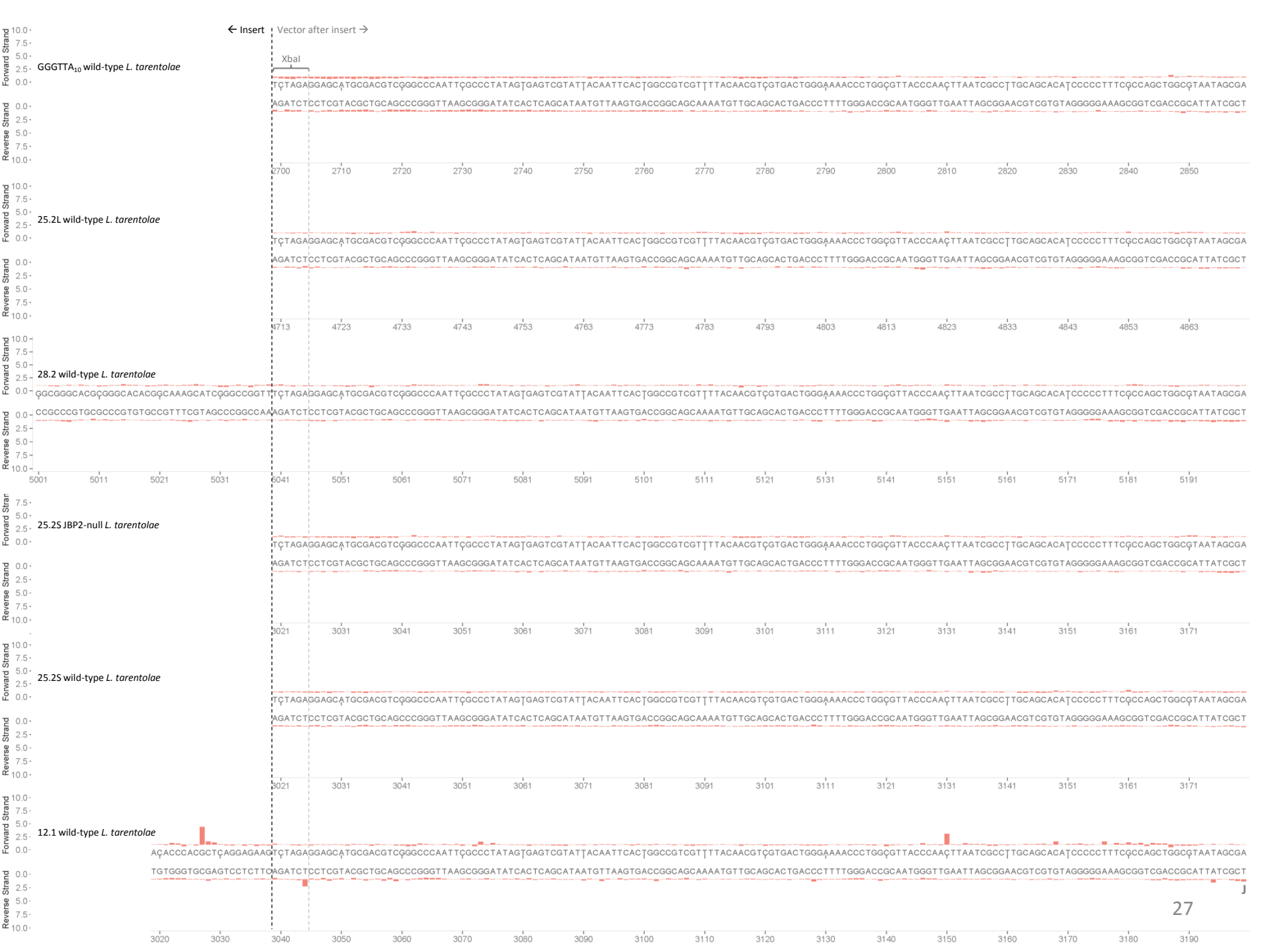

25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
G-quadruplex – bottom strand
Insert (duplicated to align cSSR 25.2S and 25.2L insert sequences) →
Forward Strand
Reverse Strand
13.6


G-quadruplexes (N>10) – top strand
35 nt repeat 1
35 nt repeat 2
35 nt repeat 3
35 nt repeat 4
Partial repeat
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
Forward Strand
Reverse Strand
+6
cSSR 25.2S has two fewer
35 nt repeats than 25.2L


25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
Forward Strand
Reverse Strand
← Insert
Gap added to align sequences after insert


← Insert
25.2L wild-type L. tarentolae
Gap added to align sequences after insert
Forward Strand
Reverse Strand
28.2 wild-type L. tarentolae


28.2 wild-type L. tarentolae
Forward Strand
10.0
7.5
5.0
2.5
0.0
ATTATTTTGTCTTGGTTGCCTCTTTGACTTTCTCGCGATCCAGCCAATAGCCTACAGCCAGGGTGGTCACGTATCAGTCTTCGGTCGCACAGCCTGCCCCACGGTCATGCAGTGCAATGATGTCTCTACCCGCTGGTCCTCGCTTCCCGGCCTTCCCTTCGCTGTTTTTTTTTTTGTTTGTCATTTCGTGTAATCCGCCAG
TAATAAAACAGAACCAACGGAGAAACTGAAAGAGCGCTAGGTCGGTTATCGGATGTCGGTCCCACCAGTGCATAGTCAGAAGCCAGCGTGTCGGACGGGGTGCCAGTACGTCACGTTACTACAGAGATGGGCGACCAGGAGCGAAGGGCCGGAAGGGAAGCGACAAAAAAAAAAACAAACAGTAAAGCACATTAGGCGGTC
0.0
2.5
5.0
7.5
10.0
Reverse Strand
4801
4811
4821
4831
4841
4851
4861
4871
4881
4891
4901
4911
4921
4931
4941
4951
4961
4971
4981
4991


← Insert
Vector after insert →
XbaI
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand

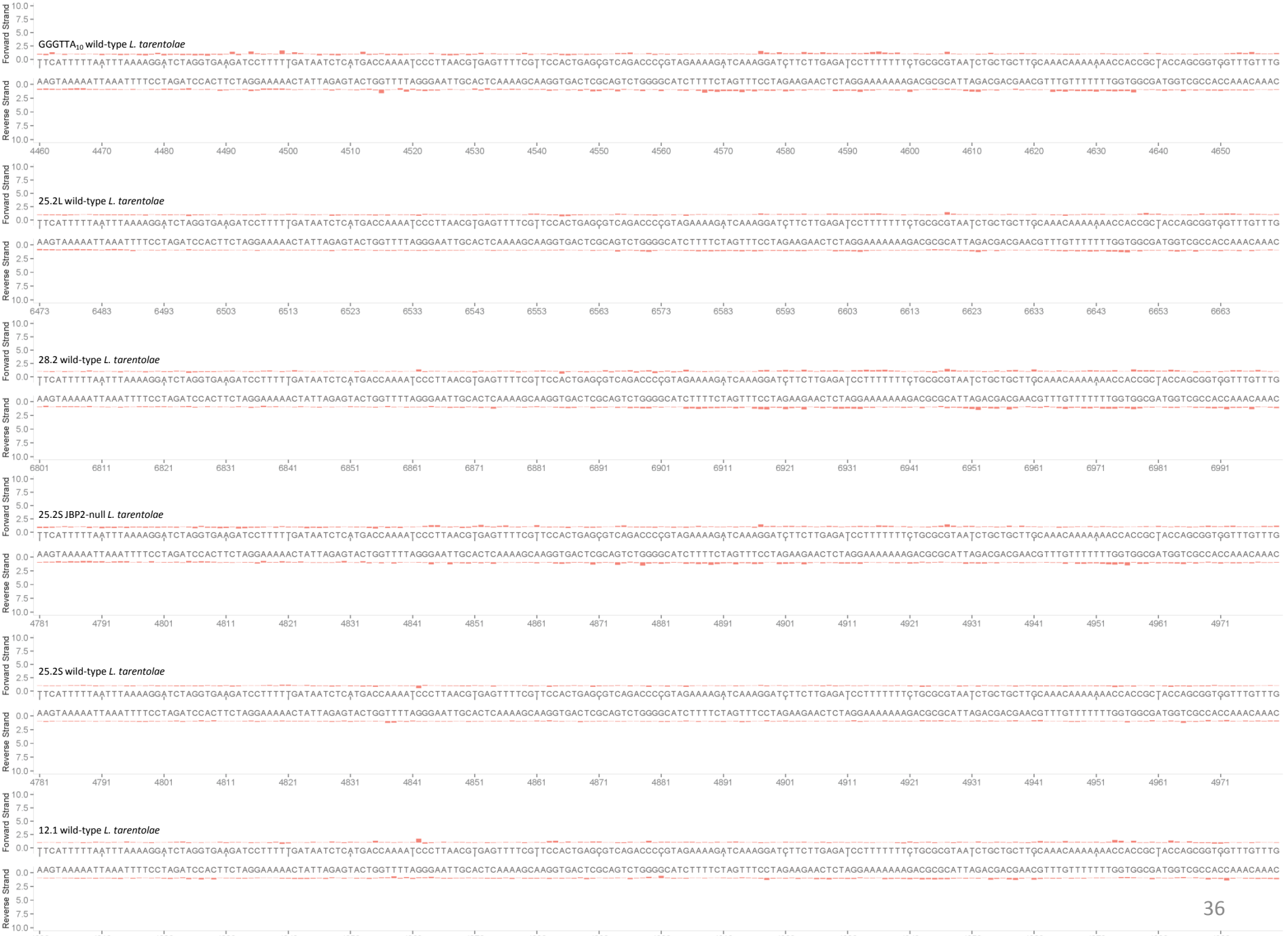

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


ScaI
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand

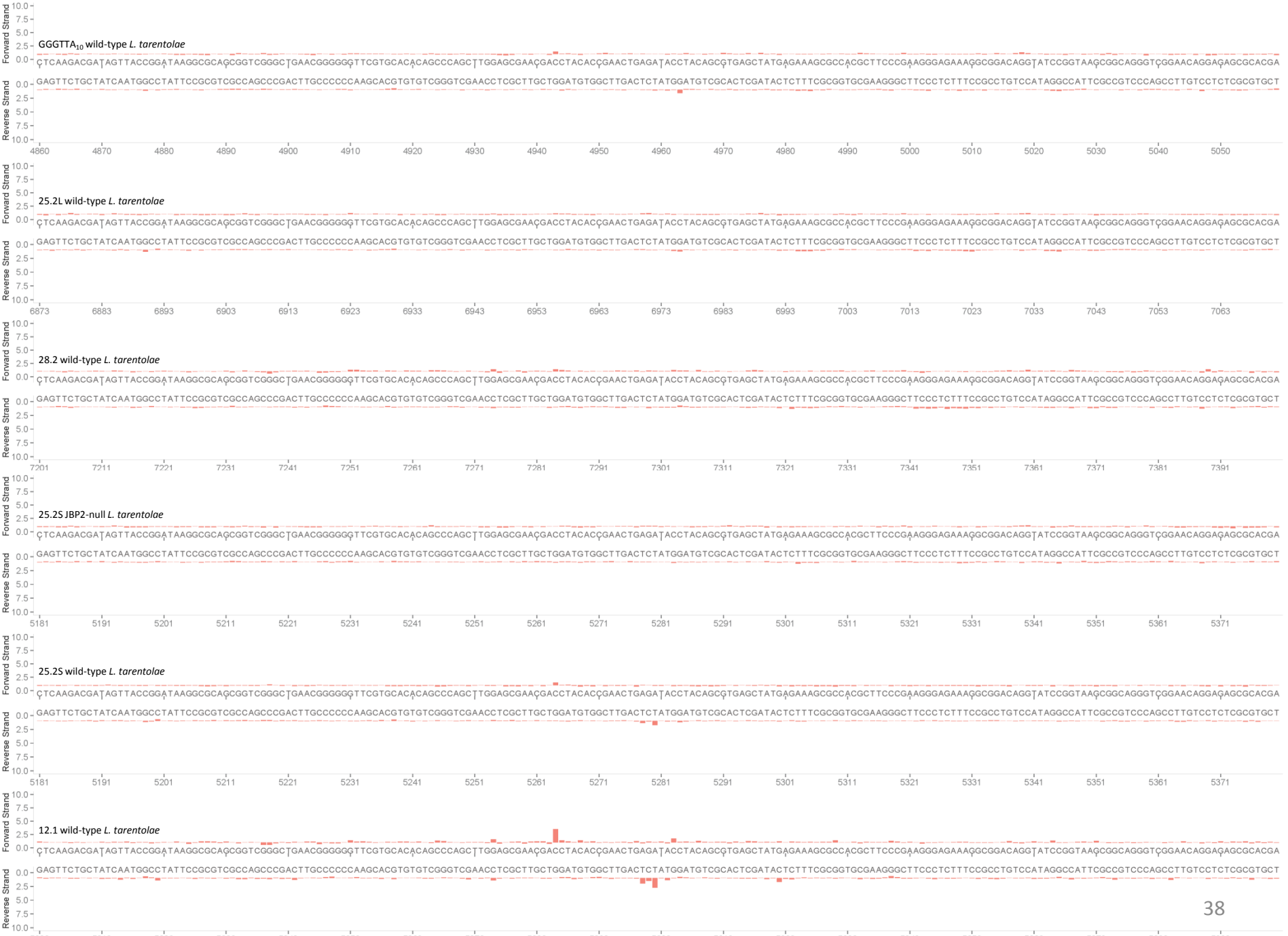

ScaI
GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
Forward Strand
Reverse Strand
TGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
4260 4270 4280 4290 4300 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440 4450
25.2L wild-type L. tarentolae
TGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
6273 6283 6293 6303 6313 6323 6333 6343 6353 6363 6373 6383 6393 6403 6413 6423 6433 6443 6453 6463
28.2 wild-type L. tarentolae
TGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
6601 6611 6621 6631 6641 6651 6661 6671 6681 6691 6701 6711 6721 6731 6741 6751 6761 6771 6781 6791
25.2S JBP2-null L. tarentolae
TGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
4581 4591 4601 4611 4621 4631 4641 4651 4661 4671 4681 4691 4701 4711 4721 4731 4741 4751 4761 4771
25.2S wild-type L. tarentolae
TGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
4581 4591 4601 4611 4621 4631 4641 4651 4661 4671 4681 4691 4701 4711 4721 4731 4741 4751 4761 4771
12.1 wild-type L. tarentolae
TGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
ACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTG
4600 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand

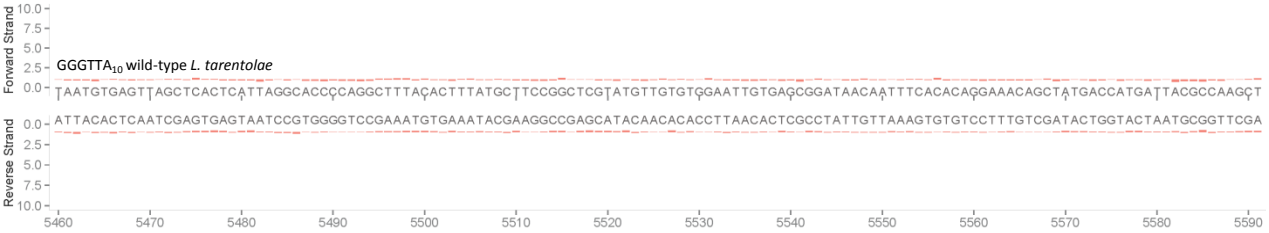

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand

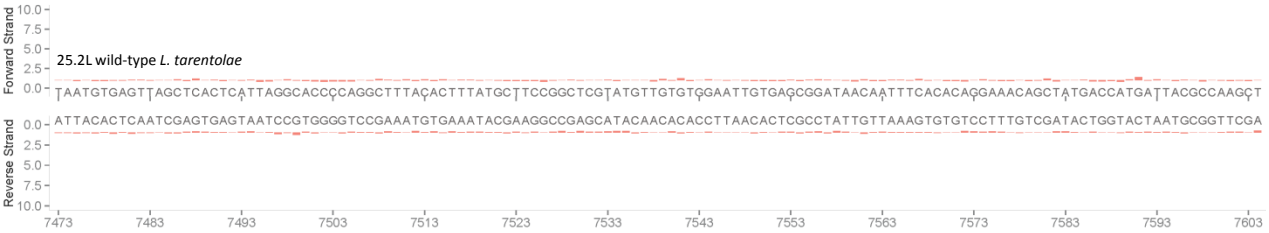

GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand


GGGTTA10 wild-type L. tarentolae
25.2L wild-type L. tarentolae
28.2 wild-type L. tarentolae
25.2S JBP2-null L. tarentolae
25.2S wild-type L. tarentolae
12.1 wild-type L. tarentolae
Forward Strand
Reverse Strand

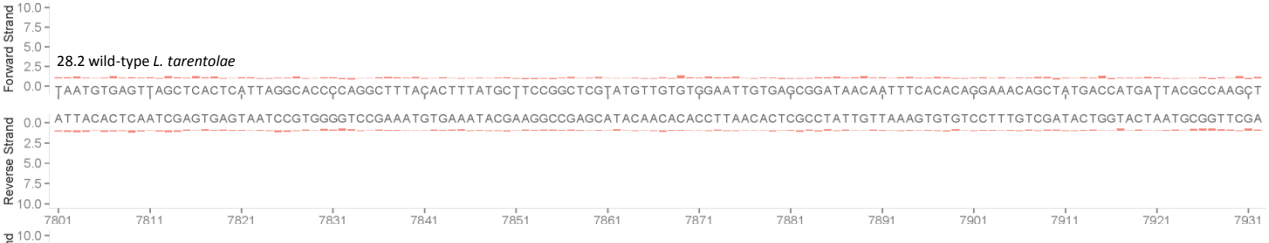
GGGTTA$_{10}$ wild-type *L. tarentolae*



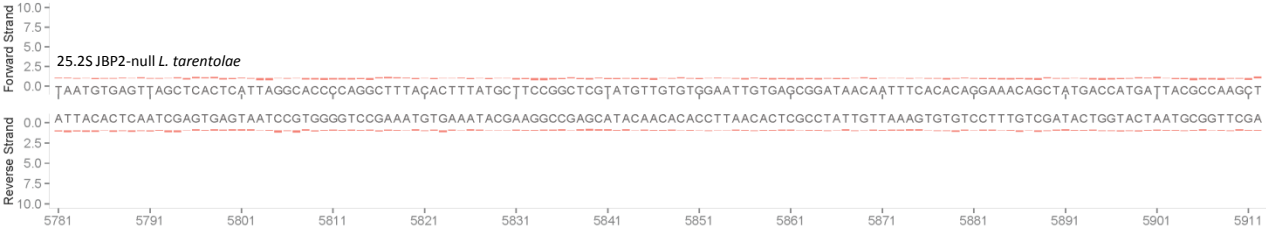
25.2L wild-type *L. tarentolae*



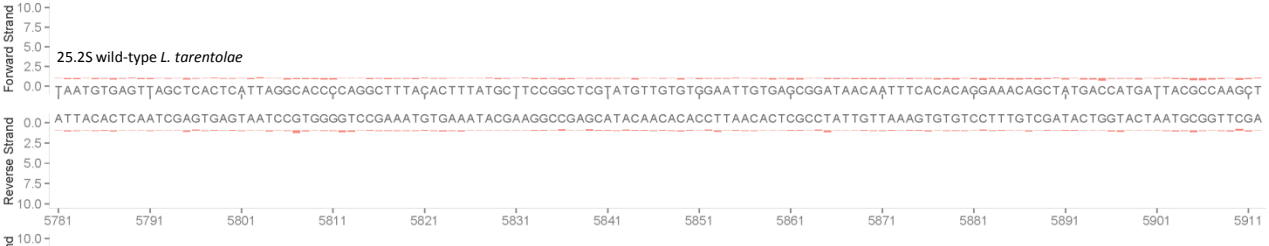
28.2 wild-type *L. tarentolae*



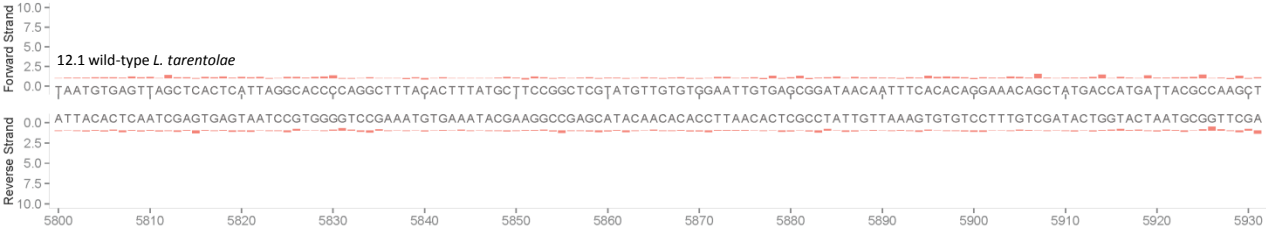
25.2S JBP2-null *L. tarentolae*

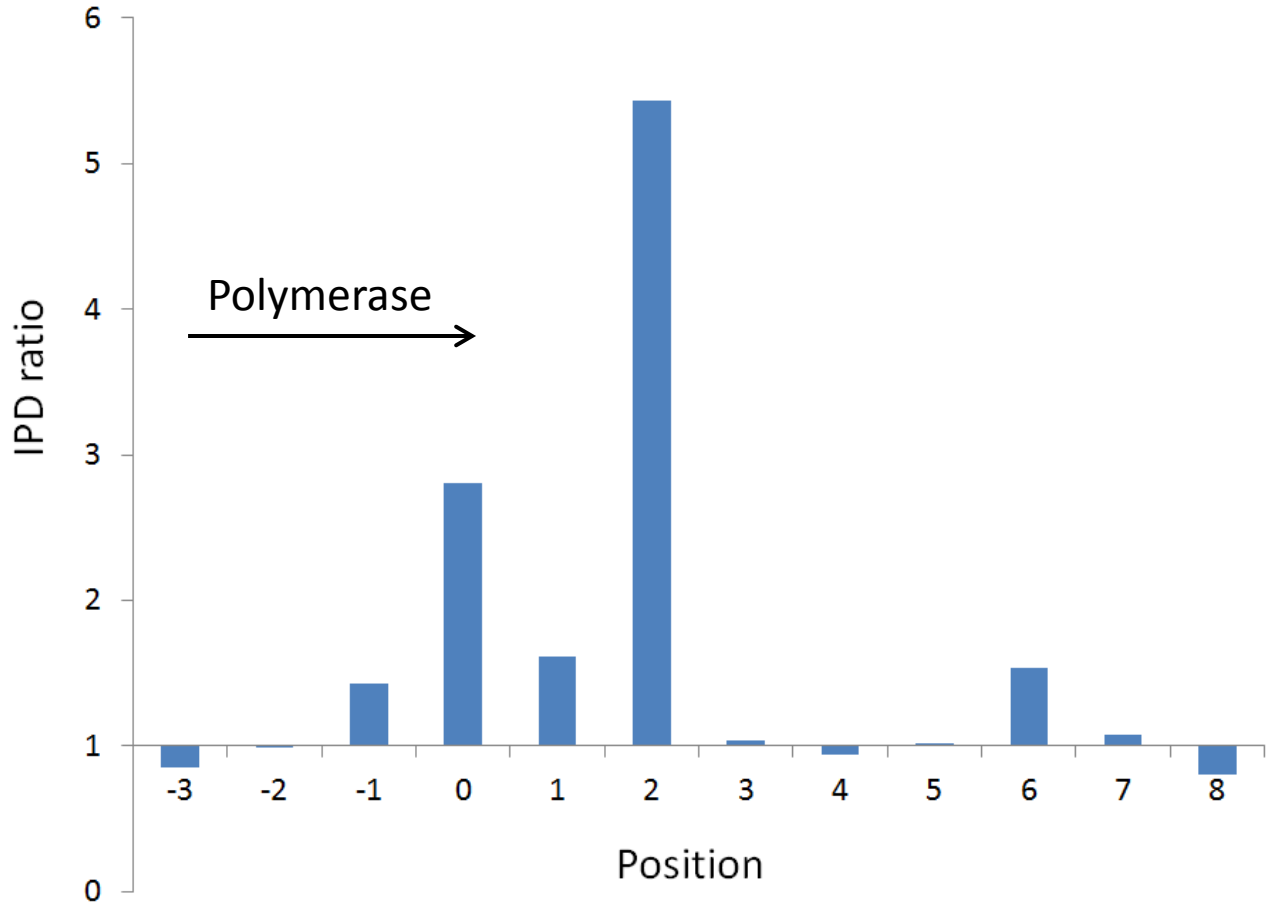


25.2S wild-type *L. tarentolae*



12.1 wild-type *L. tarentolae*





Sequences used

AACTCCACAAAC
CCCTCCCCCGTC
CCCTCCGACCCG
CCCTCGCGGCGG
CGATAACACCGC
CGCTGCAAAGAC
CGGTCCCGCCCG
CGGTGCGGGACG
GACTCGCACCCA
GGGTCGCGCCGG
GGGTCGGAGGGG
TAATGCCACGCC
TGATACGGCGTG

Supplementary Figure S4