**Table S4.**

|  |  |  |  |
| --- | --- | --- | --- |
| Assembly name | Set | Notes | Reads |
| 927mc\_III-1b07\_assembly | 3 | 3' incomplete; A752del in the NTD-encoding domain of the assembly was corrected according to expressed clones | 927mc\_III-1b07.p1k |
|   |   |   | 927mc\_III-1b07.q1k-rev |
|   |   |   | 927mc\_III-01b01.q1k-rev |
| tryp\_X-54b12.q1c-rev\_assembly | 4 | 3' incomplete; cloned and sequenced from gDNA (KC434956) | tryp\_X-54b12.q1c-rev |
|   |   |   | tryp\_XI-910b10.q1k |
| 927mc\_VI-6b02\_assembly | 5 | 5' and 3' incomplete | 927mc\_VI-6b02.p1k |
|   |   |   | 927mc\_VI-6b02.q1k-rev |
|   |   |   | 927MC\_V-01a09.p1k |
|   |   |   | 927MC\_VI-01c10.q1k-rev |
|   |   |   | 927MC\_V-01a09.q1k-rev |
| tryp\_XI-1034g11\_assembly | 14 | cloned and sequenced from gDNA (KC434958) | tryp\_IXa-24h09.q1c-rev |
|   |   |   | tryp\_XI-1034g11.p1k |
|   |   |   | tryp\_XI-1034g11.q1k-rev |
| tryp\_XI-934d12\_assembly | 16 |   | tryp\_XI-934d12.q1k-rev |
|   |   |   | tryp\_X-334d06.p1c |
| tryp\_XI-1058d01\_assembly | 17 | 3' incomplete; cloned and sequenced from gDNA (KC434957) | tryp\_XI-1058d01.q1k |
|   |   |   | tryp\_XI-1121e11.p1k |
|   |   |   | 927mc\_VI-8a07.p1k-rev |
| trypA24a5.p1p | 24 | 5' and 3' incomplete | trypA24a5.p1p |
| tryp\_IXa-29b02.q1c\_assembly | 25 | 5' incomplete, not contiguous | tryp\_IXa-29b02.q1c |
|   |   |   | tryp\_IXb-301g08.p1c |
| 927mc\_IV-5e11\_assembly | 29 | 5' and 3' incomplete | 927mc\_IV-5e11.p1k |
|   |   |   | 927mc\_IV-5e11.q1k |
| tryp\_IXa-10f02\_assembly | 32 | 3' incomplete; cloned and sequenced from gDNA (KC434955) | 927mc\_IX-10d02.q1k |
|   |   |   | tryp\_IXa-10f02.p1c |
| 927mc\_VIII-11d07.p1k | 38 |   | 927mc\_VIII-11d07.p1k 927MC\_VIII.0.118 |
| tryp\_XI-1068b03.p1k-rev | 38 | 5' and 3' incomplete | tryp\_XI-1068b03.p1k-rev |
| tryp\_XI-909g03.q1k-rev | 49 | 5' and 3' incomplete | tryp\_XI-909g03.q1k-rev |
| tryp\_IXa-20a05\_assembly | 51 | 5' and 3' incomplete; cloned and sequenced from gDNA (KC434959) | tryp\_IXa-20a05.p1c |
|   |   |   | tryp\_IXa-6d11.q1c |
|   |   |   | tryp\_IXb-153c03.p1c-rev |
| tryp\_IXa-29e12\_assembly | 54 | 3' incomplete | tryp\_IXa-29e12.p1c |
|   |   |   | tryp\_IXa-4a12.p1c |
|   |   |   | tryp\_IXb-361b07.p1c |
| tryp\_X-99f10\_assembly | 54 |   | 927mc\_IV-3d11.p1k-rev |
|   |   |   | 927mc\_VI-10d07.p1k |
|   |   |   | tryp\_X-190c10.q1c-rev |
|   |   |   | tryp\_X-214a05.q1c-rev |
|   |   |   | tryp\_X-268e08.q1c |
|   |   |   | tryp\_X-277b10.q1c |
|   |   |   | tryp\_X-338f01.q1c-rev |
|   |   |   | tryp\_X-36c03.q1c |
|   |   |   | tryp\_X-99f10.p1c |
|   |   |   | tryp\_XI-143d03.q1ca-rev |
| tryp\_XI-1157a04\_assembly | 59 | not contiguous | tryp\_XI-1157a04.p1k |
|   |   |   | tryp\_XI-1157a04.q1k-rev |
| 927mc\_VIII-3d10\_assembly | 61 | 3' incomplete, not contiguous | 927mc\_VIII-3d10.p1k-rev |
|   |   |   | 927mc\_VIII-3d10.q1k-rev |
|   |   |   | 927mc\_VIII-13b05.q1k |
|   |   |   | 927mc\_VIII-14d10.q1k-rev |
| tryp\_XI-1007d10.q1k | 63 | 3' incomplete | tryp\_XI-1007d10.q1k |
| tryp\_XI-1084f09.p1k | 67 | 5' and 3' incomplete | tryp\_XI-1084f09.p1k |
| tryp\_IXa-10d01\_assembly | 71 | 5' and 3' incomplete | tryp\_IXa-25g04.q1c-rev |
|   |   |   | tryp\_IXa-10d01.q1c-rev |
| tryp\_XI-1009a05\_assembly | 72 | 5' and 3' incomplete | tryp\_IXa-7a11.p1c-rev |
|   |   |   | tryp\_XI-1009a05.p1k-rev |
|   |   |   | tryp\_XI-1055f01.p1k-rev |
| tryp\_XI-393f08\_assembly | 74 | 3' incomplete | tryp\_XI-393f08.q1k |
|   |   |   | tryp\_XI-949b02.q1k-rev |
|   |   |   | tryp\_XI-1098g01.q1k-rev |
|   |   |   | trypA31g9.q1t-rev |
|   |   |   | trypA29a10.p1p-rev |
|   |   |   | tryp\_XI-204f04.p1c-rev |
|   |   |   | tryp\_X-48f04.p1c-rev |
|   |   |   | tryp\_XI-393f08.p1k-rev |
| tryp\_IXb-375c10\_assembly | 76 |   | tryp\_IXb-375c10.p1c |
|   |   |   | tryp\_IXb-62e09.q1c-rev |
|   |   |   | tryp\_IXb-334g12.p1c-rev |
|   |   |   | tryp\_XI-1117h04.q1k |
|   |   |   | tryp\_IXb-274b10.q1c-rev |
|   |   |   | tryp\_IXb-1e12.q1c-rev |
|   |   |   | tryp\_IXb-282d11.q1c |
|   |   |   | tryp\_IXb-342g05.p1c-rev |
|   |   |   | tryp\_IXb-375c10.q1c-rev |
|   |   |   | tryp\_XI-1117h04.p1k-rev |
|   |   |   | tryp\_IXb-125h06.q1c-rev |
|   |   |   | tryp\_IXb-282d11.p1c-rev |
| tryp\_XI-325h02\_assembly | 77 | 3' incomplete | tryp\_XI-1024b10.p1k-rev |
|   |   |   | tryp\_XI-1128e06.p1kw-rev |
|   |   |   | tryp\_XI-1139e06.q1k |
|   |   |   | tryp\_XI-238f01.p1c-rev |
|   |   |   | tryp\_XI-325h02.q1k-rev |
|   |   |   | tryp\_XI-325h02.q1ka-rev |
|   |   |   | tryp\_XI-330f02.q1k |
|   |   |   | tryp\_XI-330f02.q1ka |
|   |   |   | tryp\_XI-912a01.p1k-rev |
| 927mc\_III-7e05\_assembly | 79 | 3' incomplete, not contiguous | 927mc\_III-7e05.p1k |
|   |   |   | 927mc\_III-7e05.q1k |
| 927mc\_VI-4a12\_assembly | 80 | 3' incomplete | 927mc\_VI-4a12.p1k |
|   |   |   | 927mc\_VI-5d03.p1k |
|   |   |   | 927mc\_V-8d11.p1k |
|   |   |   | 927mc\_V-1d03.q1k-rev |
|   |   |   | 927mc\_VI-12c01.q1k-rev |
|   |   |   | 927mc\_VI-9c10.q1k |
|   |   |   | 927mc\_VI-11g04.p1k-rev |
| 927mc\_VI-4g03\_assembly | 81 |   | 927mc\_VI-4g03.q1k |
|   |   |   | 927mc\_V-6d03.q1k |
|   |   |   | 927mc\_VI-2g12.q1k |
|   |   |   | 927mc\_V-6d05.p1k |
|   |   |   | 927mc\_VI-10e07.q1k |
| 927mc\_VI-6c05\_assembly | 82 |   | 927mc\_VI-6c05.p1k-rev |
|   |   |   | 927mc\_VI-6c05.q1k-rev |
|   |   |   | 927mc\_VI-4a05.p1k-rev |
|   |   |   | 927mc\_VI-5e11.p1k |
|   |   |   | 927mc\_VI-5e11.q1k-rev |
|   |   |   | 927mc\_VI-5h02.p1k |
|   |   |   | 927mc\_VI-7e06.q1k-rev |
|   |   |   | 927mc\_VI-13c03.q1k |
| 927mc\_VII-13h07\_assembly | 83 | 3' incomplete | 927mc\_VII-13h07.p1k-rev |
|   |   |   | 927mc\_VII-13h07.q1k-rev |
|   |   |   | 927mc\_VII-5h01.p1k-rev |
|   |   |   | 927mc\_VII-7g03.q1k-rev |
|   |   |   | 927mc\_VIII-1f08.q1k-rev |
|   |   |   | 927mc\_VII-11c05.p1k |
|   |   |   | 927mc\_VI-11f11.p1k-rev |
| 927mc\_VIII-14d09\_assembly | 84 |   | 927mc\_VIII-14d09.p1k |
|   |   |   | 927mc\_VIII-14d09.q1k |
|   |   |   | 927mc\_IV-5h11.p1k |
|   |   |   | 927mc\_VII-5a05.q1k-rev |
| tryp\_XI-224e03\_assembly | 85 | not contiguous | tryp\_XI-224e03.q1c |
|   |   |   | tryp\_XI-224e03.p1c-rev |
|   |   |   | tryp\_XI-943h03.q1k-rev |
| tryp\_IXa-28d11\_assembly | 86 |   | tryp\_IXa-28d11.q1c-rev, |
|   |   |   | tryp\_IXa-17g02.p1c-rev |
|   |   |   | tryp\_X-316d12.q1c-rev |
| tryp\_IXb-338f09\_assembly | 87 |   | tryp\_IXb-68a11.q1c-rev |
|   |   |   | tryp\_IXb-18c04.q1c-rev |
|   |   |   | tryp\_IXb-257c04.q1c |
|   |   |   | tryp\_IXb-257g03.q1c-rev |
|   |   |   | tryp\_IXb-296f08.p1c |
|   |   |   | tryp\_IXb-338f09.p1c |
|   |   |   | tryp\_IXb-302g08.p1c |
| tryp\_X-187a04\_assembly | 87 | 5' incomplete | tryp\_X-187a04.p1c |
|   |   |   | tryp\_X-187a04.q1c-rev |
|   |   |   | tryp\_X-402h07.p1c |
| tryp\_X-274d02\_assembly | 88 | 3' incomplete, not contiguous | tryp\_X-350d12.q1c-rev |
|   |   |   | tryp\_X-246a07.p1c-rev |
|   |   |   | tryp\_X-440d01.q1c |
|   |   |   | tryp\_X-274d02.p1c-rev |
|   |   |   | tryp\_X-135c02.q1c-rev |
| tryp\_XI-1087c04.q1k-rev | 89 | 5' and 3' incomplete | tryp\_XI-1087c04.q1k-rev |
| tryp\_IXa-3a03\_assembly | 91 | 5' incomplete, not contiguous | trypB62c6.q1t-rev |
|   |   |   | tryp\_IXa-3a03.p1c |
|   |   |   | tryp\_IXa-3a03.q1c-rev |
|   |   |   | tryp\_IXb-89d01.p1c-rev |
|   |   |   | tryp\_IXb-89d01.q1c |
|   |   |   | tryp\_XI-982g12.q1k |
| tryp\_XI-1077d06\_assembly | 93 | 3' incomplete | tryp\_XI-1077d06.q1k |
|   |   |   | tryp\_XI-170f03.p1k |
| tryp\_IXb-321a12\_assembly | 94 | 3' incomplete | tryp\_IXb-321e12.q1c |
|   |   |   | tryp\_IXb-321e12.p1c-rev |

aThe ‘notes’ column provides GenBank accession numbers for those donors cloned and sequenced from gDNA.

bWhen assembling reads, mismatches between the reads were treated as follows. If the mismatch occurred towards the ends of a read, it was considered to be the result of poor quality sequencing, and was corrected in favour of the other read(s). If the mismatch occurred in the middle of the read, but more than one other read did not contain the mismatch, it was corrected in favour of the other reads. If the mismatch occurred in the middle of the read, and there was only one other read with which to compare it, the mismatch was corrected in favour of the corresponding clone sequence. Read sequences were obtained from the following files obtained from ftp://ftp.sanger.ac.uk/pub/databases/T.brucei\_sequences/ : T\_brucei\_chrIXa\_reads.20Oct2001, T\_brucei\_chrIXa\_reads.29May03, T\_brucei\_chrIXb\_reads.20Oct2001, T\_brucei\_chrIXb\_reads.29May03, T\_brucei\_chrI\_reads.03Feb2000, T\_brucei\_chrXI\_reads.V1\_17Oct2002, T\_brucei\_chrXI\_reads.V2\_14Mar2003.fas, T\_brucei\_chrXI\_reads.V3\_11Jul03.fas, T\_brucei\_chrX\_reads.03Mar2003, T\_brucei\_chrX\_reads.20Oct2001, T\_brucei\_chrX\_reads.30Apr2001, T\_brucei\_reads.21Dec1999, TbchrIXa\_reads\_26May04.fas, TbchrIXb\_reads\_26May04.fas, Tb927\_IC\_II\_reads.fas, Tb927\_IC\_I\_reads.fas, 177bp\_repeat.dnaTb927\_MC\_cons.dna.crunch, 177bp\_repeat.dnaTb927\_MC\_test\_reads.dna.crunch, 927MC\_readsII.fas, 927MC\_readsIII.fas, 927MC\_readsIV.fas, 927MC\_readsV.fas, 927MC\_readsVI.fas, 927MC\_readsVII.fas, 927mcIII\_v1.cons, 927mcII\_v1.cons, 927mcIV\_v1.cons, 927mcIX\_v1.cons, 927mcVIII\_v1.cons, 927mcVII\_v1.cons, 927mcVI\_v1.cons, 927mcV\_v1.cons, Tb927MC\_telo1\_contigs.fas, Tb927MC\_telo2\_contigs.fas, Tb927MC\_telo2\_v2contigs.fas, Tb927MC\_telo9\_contigs.fas, Tb927MC\_telo9\_v2contigs.fas, Tb927MCtelo\_1\_test.fas, Tb927MCtelo\_2\_test.fas, Tb927MCtelo\_9\_test.fas, Tb927\_MC.cons, Tb927\_MCIX\_contigs\_6Mar08.fas, Tb927\_MC\_III\_reads.fas, Tb927\_MC\_II\_reads.fas, Tb927\_MC\_IV\_reads.fas, Tb927\_MC\_IX\_reads.fas, Tb927\_MC\_VIII\_reads.fas, Tb927\_MC\_VII\_reads.fas, Tb927\_MC\_VI\_reads.fas, Tb927\_MC\_V\_reads.fas, Tb927\_MC\_test\_reads.fas, Tb927\_MC\_testreads\_newlib.fas, uniprot-1-xTb927\_MC\_cons.dna.crunch, uniprot-1-xTb927\_MC\_test\_reads.dna.crunch, TPunknown\_May06.cons, Tbrucei\_extrachr\_testreads\_Dec05.fas, chr\_unknown\_test\_reads\_Jan06.fas