

Figure S1

```
1   CATATGGCAT CAGTTGATCC GTTGGGTGTT GTGGGATCTG GAGCTGATGT GTATCTCTAC TTCCTGTCTG CTGGGAACGA GAATTTGATT TCAAGGATCA
101  TCGAAAATCA CGAATCAAAA GCGGACATCA AGAAGATAGT CGATCGAACA ACTGCGGTCT ACGGTGCCTT CTTTGCAAGA TCCAAAGAAT TCAGACTCTT
201  TGGATCTGGA TCTTACCCCT ATGCTTTTAC AAATCTTATT TTCTCGCGTT CAGATGGGGT GGCAGGTACA AAAACCGAAC ACGGTATCAC ATATTACGAG
301  AGCGAGCATA CTGATGTCTC TATTCCTGCT CCTCATTCTT CTTGTGTGAT TTTTGGATCG TCTAAAAGGG AACGAATGAG CAAGATGTTG AGCAGATTAG
401  TTAACCCCTGA CAGACCGCAA TTGCCTCCTA GATTGAAAA GGAATGCACG TCAGAAGGAA CATCACAAC TGTCGCTTTA TACATAAAAA ATGGAGGTCA
501  TTTTCATCACT AAAGTGTAA ATTTCCACA GTTGAACCTC CCTTAGGTG CTATGAACT CTATCTCACA GCTCGTCGAA ATGAGTACCT CTACACACTA
601  AGCCTCCAAC TAGGAAATGC CAAGATTAAC TTCCCGATAC AGTTCCTGAT TTCACGAGTT TTGAATGCTC ATATTCATGT GGAGGGAGAC AGACTGATCA
701  TTGAGGATGG AACTATCTCG GCGGAACGTT TGGCCTCGTT AATCTCCTCC CTGTACTCTA AAAAGGGCTC TTCTGGTGGG GGTGGCAGTG GTGGAGGAGG
801  AAGCGGCGGG GGTGGGTAC AAGCAAATGA TAATCGGTAC GAAGGTAAAC CTATATCAGC AATATCATTC GAGGGGCTCG AATACATCGC ACGTGGGCG
901  TTAGATACGA TCTTCTCCCA GTATAAAGGG CAGAAGTGA CCTACGAACT TTACTGGAG ATTCTTCAA AGGTTTACGA TTTGGAATAT TTTTCTGAAG
1001 TGCTCTCCAA AGTGTGTCCC ACTGATCCTG AATACCAATA CGTTATGTTA CAATTCAGTG TGAAGAAGG GCCCTCAGTT AAAGGAATTA AGATGTTGG
1101 GAATTCCTCAG ATTAGGAGTG GAGATTTGCT CAGTAAAATT CTCTTAAAA AGGGTGATAT CTATAATGAG GTTAAGATGA AAGTGGATCA AGAAAGTCTT
1201 AGAAGACATT ACTTGACCA GGGATACGCT GCAGTGAAAA TTTCTTGTGA AGCTAAGACA GAGGCTGGCG GCGTTGTTGT CCAGTTTCCAT ATCCAAGAAG
1301 GGAAGCAAAC GGTGTGCAGT AGAATCCAGT TCAAAGGAAA CAAGGCTTTC ACCGAAAGTG TCCTTAAGAA GGTACTAAGC ACTCAAGAAG CTAGATTCTT
1401 AACCTCTGTT GTGTTTAAGG AGAATGCTCT AGAGGCCGAC AAAGCTGCAG TGCATTCTTA CTATGCTGAA AGAGGATATA TCGATGCTAG GGTGAGGGT
1501 GTGGCTAAGA CGGTTGATAA GAAGACAGAT GCATCTAGAA ATCTCGTAAC CTGACCTAC ACTGTGGTGC AGGGAGAGCA ATACCGATAC GGTGGAGTCA
1601 CTATAGTGGG AAACCAAATA TTCAGCACAG AAGAACTACA AGTAAAATT AGACTTAAAC GAGGAGCTAT TATGAATATG GTTGCTTTTG AGCAGGGCTT
1701 TCAAGCCCTC GCTGACGCAT ACTTTGAAAA CGGTTACACA AGTAACTACT TGAACAAGGA AGAACACAGG GACACAGCTG AAAAAACGTT ATCTTTTAAA
1801 ATTACAGTAG TGGAGAGAGA AAGATCTCAT GTTGAACATA TCATTATTAA AGGAACTAAG AATACTAAG ACGAGGTTAT CCTAAGAGAG ATGCTTTTGA
1901 AACCCGGGGA TGTGTTTAGT AAATCTAAAT TTAATGATAG TTTACGTAAC CTTTTC AATT TGAGATACTT CAGTAGCCTG GTTCCAGATG TAAGGCCAGG
2001 CAGCGAGCAA GATTTAGTCG ATATTATACT AAACGTTGAA GAACAATCTA CAGCAAACCTG ACTCGAG
```

Figure S1: Nucleotide sequence of the *tp0453-tp0326* chimera. Bold font represents the restriction sites NdeI and XhoI, respectively. Underlined region represents the glycine-serine [(GGGS)₃] repeat.

