

**B.**

beta-tubulin `GTGTATGCCTTCCTCCCTTATTCGCTTCTTCTGTCGTGCACACCTCTTTCTCTCTCCCTTCCGCCTTTCTTTCAATCTTGTTTTCTGACCAG`

BES 1 `AGACAAACGGCGTGTTCGCCGCTGATGCTACAGAACCCAGCTTAAATTTCCAGAAGACGAAAAATTTGCATGTTTTCCCAATAATTTAACTACTTGAAG`

BES 2 `CGAAGCTTTAAATCTGTCTAAGGCGGCAGCGGGCCACTGCCACACAGCAAAAACCAAAAGCCTCGGATACTTTTACCTTTGTGTTCTACAAAAGTAAAAAG`

BES 3 `ATTAGCTGAAGGAAACCGGTGATGAAAAATAGATAATAAACATTTCCCAAAAAATATTACCACTAACTTCTTTGTTTTCTCTTTTCATGTTGCTAACTAG`

BES 5 `TAAATTCGGCTACCGTAGACCTTTGATAAATTTTCTATGCCGATAAGTTTCTAATAGAAATACCTTAAACGCATTTTAAATGTTTTTGGTCCGAAAAAAG`

BES 7 `GGTTGAGAACCCAGCAAAAAACAAAAAATGCCCAATAATAGGTGGTGGACACCCGAGATGAAAAAATCCACTATTCATGAGTCGTTGACCTAACTCTAG`

BES 8 `AGTGATAAAGCTAAGAAGCAGTCTACGTCAACAAAAAATTTCTTTATGGCAAGACCCGAGTAGATAGGGGTTTCTTTGAAATGCTGTCTTCTTTTACAG`

BES 10 `TGACCAATAACTGAACTTTGGGAATGCAAAAGCAACAGCAAAAGAACCTCGGCAGAGACCCCAATTTTTTGTTCCTCCCAAAATATCCATCATCAAG`

BES 11 `TATCTGTAAAACTACAGACTTCAGCTCCAAACAGCAAAATTCGGCGCTCTAATATGTTCTAACCAACAACACGATTCTGAGTGTATATAATAACAG`

BES 12 `AAAAGCAGAAAGTGTAAACACGGGCAGAGTACAGGCCACTCCTGGCAACTTACTTCAATTTCTTTTTTTCTTTTCCACCATCAAAACGAGCCAAACAG`

BES 13 `TCGAAAAACTAAAAATCCCAACGACACGAGGCCATAGTACGCTACATCTTACAACAGACAGGCCACATGCACCTCCTTTCTTTGAAAAAATCGACTATAG`

BES 15 `TAAAGCAGCCAAACTCGGTTTGTAGAACCCAGAACAGAGGTAAACCTCAGAGCAAAACGACAAAAAGGCAAAAGAAAAATTTCTACTTGTGGCATATGCAATAG`

BES 17 `AGCCGCTGTAAAAATGCAACCGTCAACAGTATTGTTACCCAACTAACAAACCCAGACAATGTCTAAAAAACAGTCTTTTCAACCCCTATCCTGAAG`

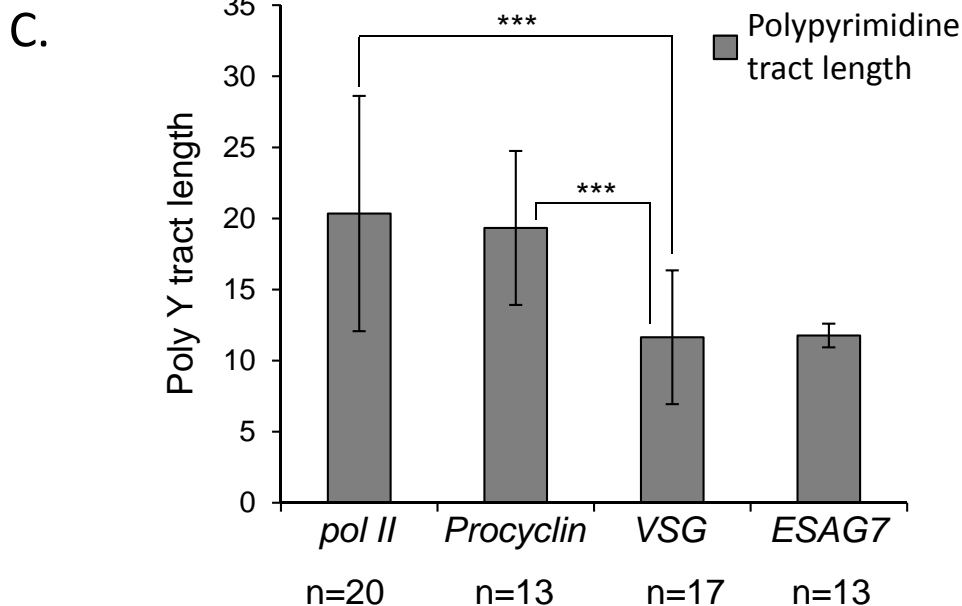
mVSG 397 `CGCACAGCAAAATATGCTTTGTAAACAAAAACAGGACACAGTCCAGCTCAGGCAATTACGCCTACCAACCCAGCCTTACTTACTTCACCTGCAAAAAATAG`

mVSG 531 `GCAGGTGAAACTGCAAAAAGTAAAGTTATTATGTTTGGCCCTGAAGCAACTGCCTACTTTTTGTCCCAAAATTTTTTTTTGAATGGGACGACGCAAAAG`

mVSG 639 `CTAGGCAAAAACCCAGCAACAGAGGCCCTGCCGCCAGCAAAAGCCAAAGTATGCCACAGAAAACTATTCGGAAATCTCCACCAGCGGACTTTTTTACCAG`

mVSG 653 `AATGTTCCGTTCAAAATGGAGCGGTGTTGCTTTTCTAACGAAAAAATTCGAAAAAATCATGTGATTGACGGCGTACCATCGAGCCCGCACTTCACTAAG`

mVSG 1954 `ACAGCCAGCAAAACTGGAGATTTCCGTTTCCAAACAACTAAAGAAACTGCAACAGAAAAAGTACAGAGCACAACTACTTTCTATTTTTATCAG`



**Additional file 6: VSG genes have significantly shorter polypyrimidine tracts upstream. A.** Logo plots of the 17 VSG *trans*-splice sites identified from our data set (top) and genes in the RNA Pol-I transcribed *procyclin* loci. Pyrimidines, red; purines, grey. **B.** Individual *trans*-splice sites from beta-tubulin (top), bloodstream VSG-ES linked VSGs and metacyclic VSGs. Pyrimidines, red; purines, grey. **C.** Bar charts show average polypyrimidine tract length comparisons over several gene groups. \*\*\*,  $P < 0.0005$ .