

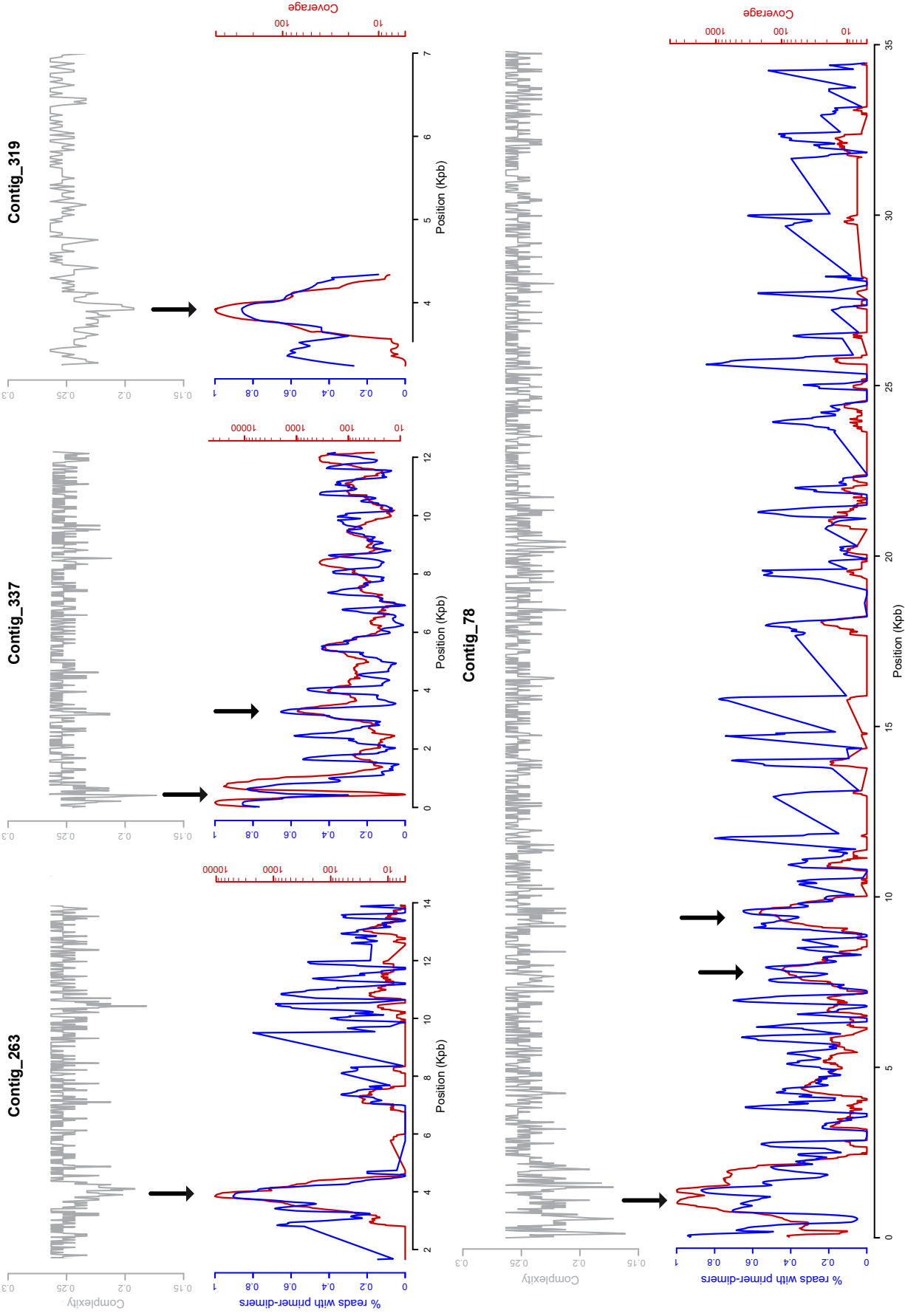


Figure S5. Profile of reads with primer-dimers in contigs with high coverage peaks at regions of low linguistic complexity. Sequences of primers used during SISPA were trimmed at the 5'-end of R1 reads before mapping to the contigs. Percentage of reads trimmed ≥ 35 bp calculated along 50 nt windows with steps of 20 nt are depicted with a blue colour line. Only windows with at least $> 5\times$ coverage are evaluated. Trifonov linguistic complexity (grey colour line) and coverage (red colour line) profiles are also shown. Black arrows point to coverage peaks that overlap with low sequence complexity and high % of reads with primer-dimers.