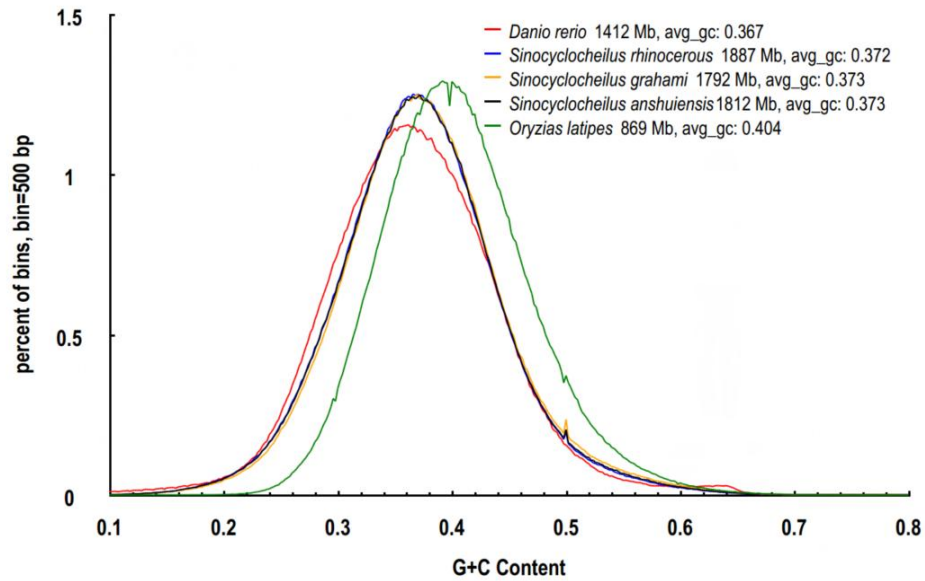


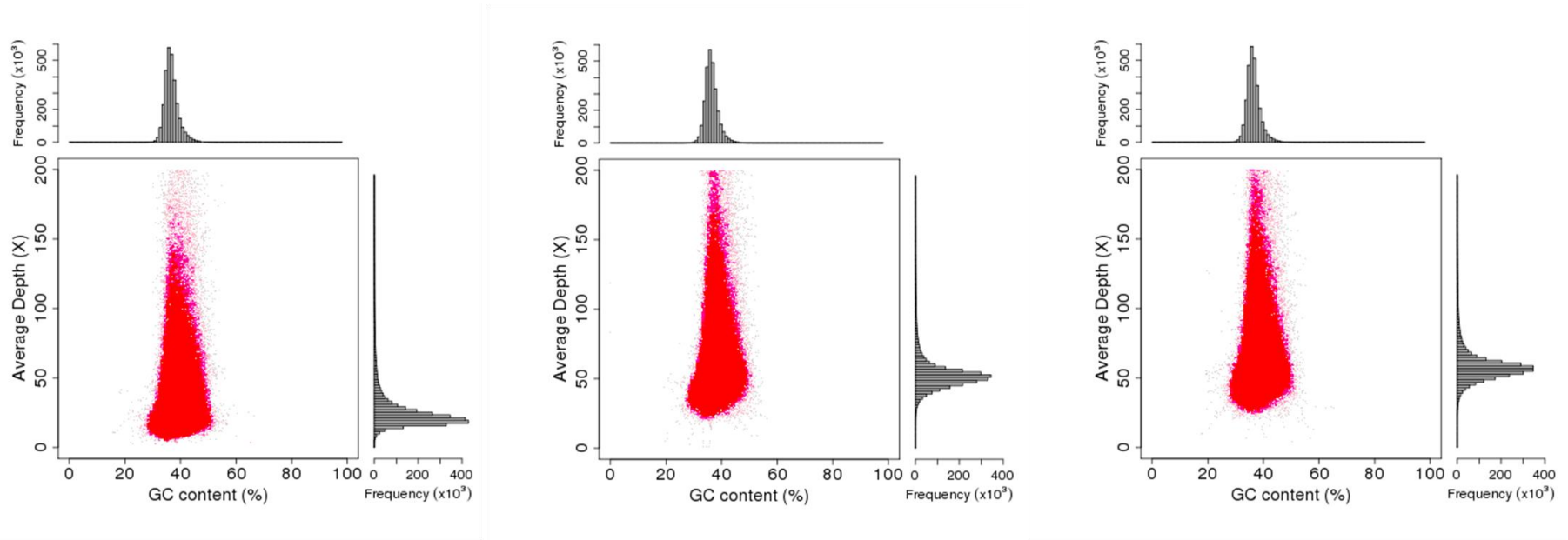
1
 2 **Figure S2. 17-mer frequency distribution in the three *Sinocyclocheilus* genomes.** The lines show the
 3 peak depth in the sequenced fishes: Sg, 43; Sr, 37; Sa, 42. Existence of the second peaks indicates that all
 4 of the three genomes have a fairly high proportion of repetitive sequences.
 5



6
7 **Figure S3. GC content distributions in the three *Sinocyclocheilus* genomes.** The x-axis is GC content
8 and the y-axis is the proportion of the bin numbers divided by the total windows. We used 500-bp bins
9 (with a 250-bp overlap) sliding along the genomes. Our data demonstrate that the three genomes have
10 similar GC content distributions.

11

12



13

14 **Figure S4. Correlation of GC content and sequencing depth.** The x-axis represents GC content and the y-axis represents the average depth. The scatter
15 diagram in the middle shows GC content and sequencing depth. The histogram on the top shows the GC content percentage and the histogram on the right
16 shows the sequencing depth percentage. We used 10-kb non-overlapping sliding windows and calculated the GC content and average depth among the
17 windows. (a): Sg, (b): Sr, (c): Sa.

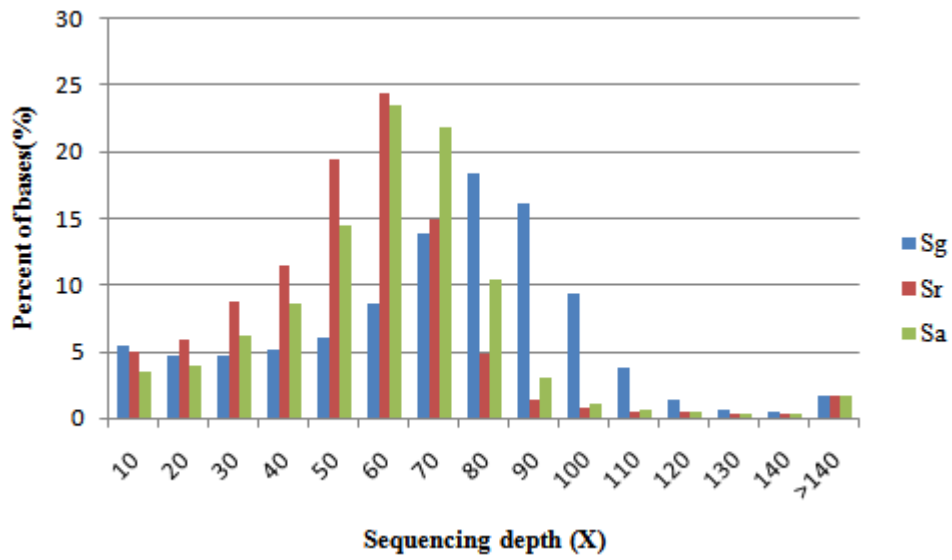


Figure S5. Sequencing depth of the three *Sinocyclocheilus* species. The filtered reads were aligned onto the assembled genome sequences using SOAPaligner for calculation of the percentage of bases with different depth frequency in the genomes.

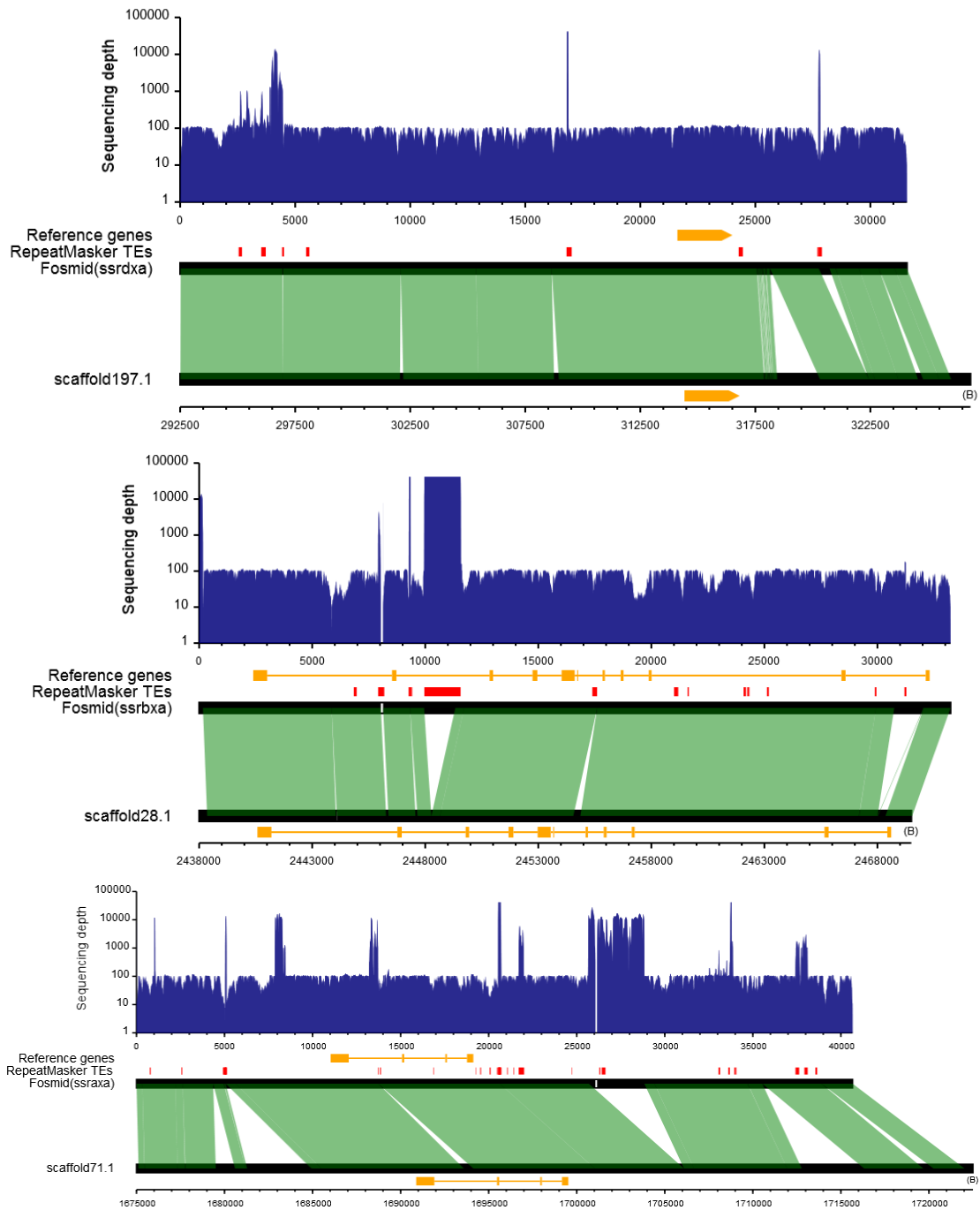


Figure S6. Comparison of the assembled Sg genome with 3 fosmid sequences. Sequencing depth of the fosmid sequences was calculated by mapping the Illumina short reads onto the fosmid sequences. Here, we performed single-end mapping and reported all the repetitive hits. The predicted genes and annotated TEs on the fosmid sequences are shown in orange and red respectively. The remaining unclosed gaps on the scaffolds and fosmid sequences are marked as white blocks. Note that most of the gaps are located in repeat regions, especially in recently evolved repeats with high sequencing depth.

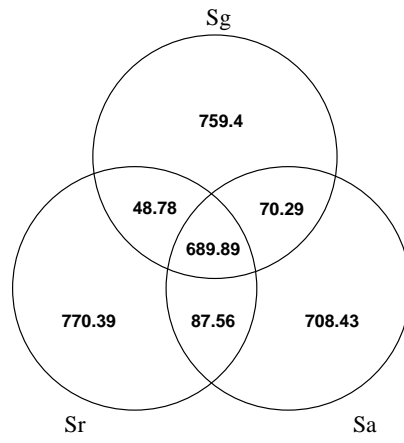


Figure S7. Conserved sequences (Mb) among the three *Sinocyclocheilus* genomes.