Additional File 10:
The two diagrams of the proposed reasons for why TSPs are enriched in GI regions.

(A), Evolution of regulatory region. To eliminate the fitness cost by non-optimal codon usage or absonant transcripts of GI, accelerated evolution in regulatory regions will result in TSPs accumulation in GI region.

(B), Re-shape of GI region. Following the event of HGT, the GI region might undergo gene loss/acquisition and genetic rearrangements. During this process, former operon was broken and more transcription units probably formed. Newly formed transcription units carry or call for the new TSPs during evolution.
The original TSP used by the donor, T0, was transcriptionally incompatible with the host’s transcription regulation system and no mRNA could be synthesized. A genetic mutation occurred and a new TSP (T1) appeared. However, the host system still could not use T1, either.

Accelerated evolution in regulatory regions kept occurring, and new TSP kept appearing. Until one day, a functional TSP (T4) arose and an mRNA was synthesized (mRNA-4).

Random genetic mutations in regulatory regions continued occurring. Several functional TSPs might appear, but only the optimal ones were kept and used for initiating transcript synthesis (T4 and T7).
After HGT, a mobile genetic element might develop into a GI following gene loss (or acquisition, or genetic rearrangement). The original T0 carried by the GI became inactive and new TSP(s) might evolve (in this case, T1).

Consecutive insertions of DNA elements, such as insertion sequence (IS) and transposons (Tns), might come along. This urged the GI to evolve further if its genes are about to express, and more TSPs (T2, T3, T4) might be introduced.

Not all newly-introduced TSPs were functional. Only the ones (T4, T5) functionally compatible with the host transcription regulatory system would signal the start point of a transcript.

Before they can be coordinately reorganized into a new operon, multiple transcription units will be maintained.