

Genomic and mechanistic insights of convergent transcription in bacterial genomes

Kashish Chetal¹, Sarath Chandra Janga^{1*}

¹Indiana University School of Informatics and Computing, Indianapolis



Abstract

Convergent gene pairs with overlapping head-to-head configuration are widely spread across both eukaryotic and prokaryotic genomes. They are believed to contribute to the regulation of genes at both transcriptional and post-transcriptional levels, although the factors contributing to their abundance across genomes and mechanistic basis for their prevalence are poorly understood. In this study, we explore the role of various factors contributing to convergent overlapping transcription in bacterial genomes. Our analysis shows that the proportion of convergent overlapping gene pairs (COGPs) in a genome is affected by endospore formation, bacterial habitat and the temperature range. In particular, we show that bacterial genomes thriving in specialized habitats such as thermophiles exhibit a high proportion of COGPs. Our results also show that the density distribution of COGPs across the genomes is high for shorter overlaps with increased conservation of distances for decreasing overlaps. Our study also reveals that COGPs frequently contain stop codon overlaps with the middle base position exhibiting mismatches between complementary strands. Functional analysis using COGs (Cluster of Orthologous groups) annotations suggested that cell motility, cell metabolism, storage and cell signaling are enriched among COGPs suggesting their role in processes beyond regulation. Our analysis provides genomic insights into this unappreciated regulatory phenomenon, allowing a refined understanding of their contribution to bacterial phenotypes.

Introduction

In the process of gene expression transcription is the first step in which DNA is converted into RNA. Convergent transcription is the simultaneous induction of the sense and antisense transcripts through two different opposing promoters. Many studies have been done to document the role of convergent transcription to provide the evidence about the functionality and gene regulation occurring in different organisms¹⁻⁵. Convergent gene transcription also allows understanding the process of transcriptional gene silencing facilitated by RNA interference mechanism³.

In a recent study by Chatterjee *et al.*^{1,5}, convergent transcription has been shown to act as a bistable switch explaining its conservation across species. This bistable switch behavior has been understood in the *prgX/prgQ* operons through various mathematical models for *Enterococcus faecialis*. Studies also show that head on collision of the RNA polymerases II can hinder transcription⁴. In yeast, it has been shown that convergent transcription can lead to RNAP collision which can stop the transcription when there is head to head collision, strengthening its role in the process of gene regulation⁴.

Given these recent observations indicating the importance of convergent regulation, we performed a global survey of thousands of bacterial genomes, to study the contribution of various bacterial life style parameters and genome size to the abundance of this form of regulation and to uncover the functional pathways associated with convergent gene pairs.

Methods

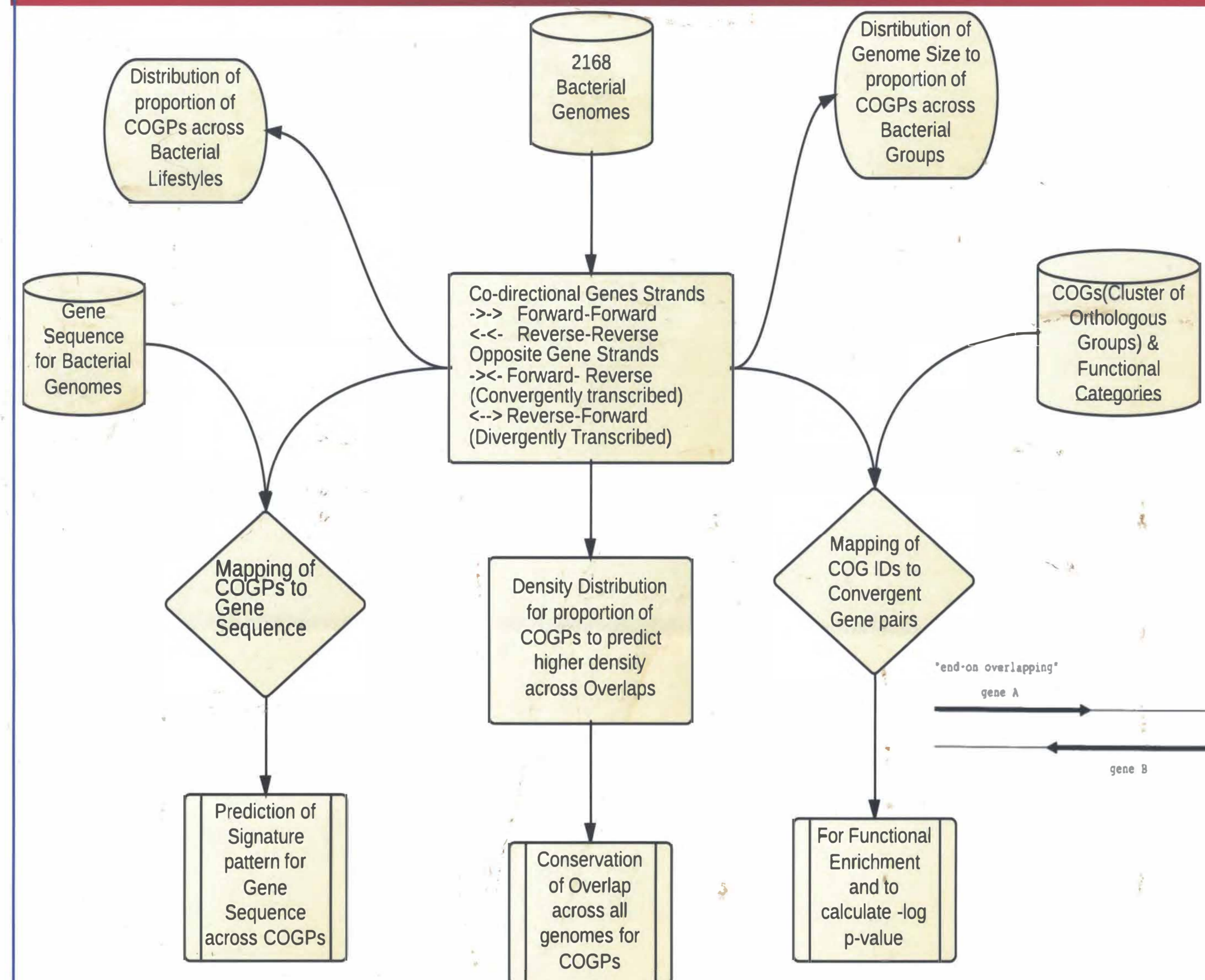


Figure: A representation of methodology to get the genomic and mechanistic insights of convergent transcription in bacterial genomes

Results and Discussion

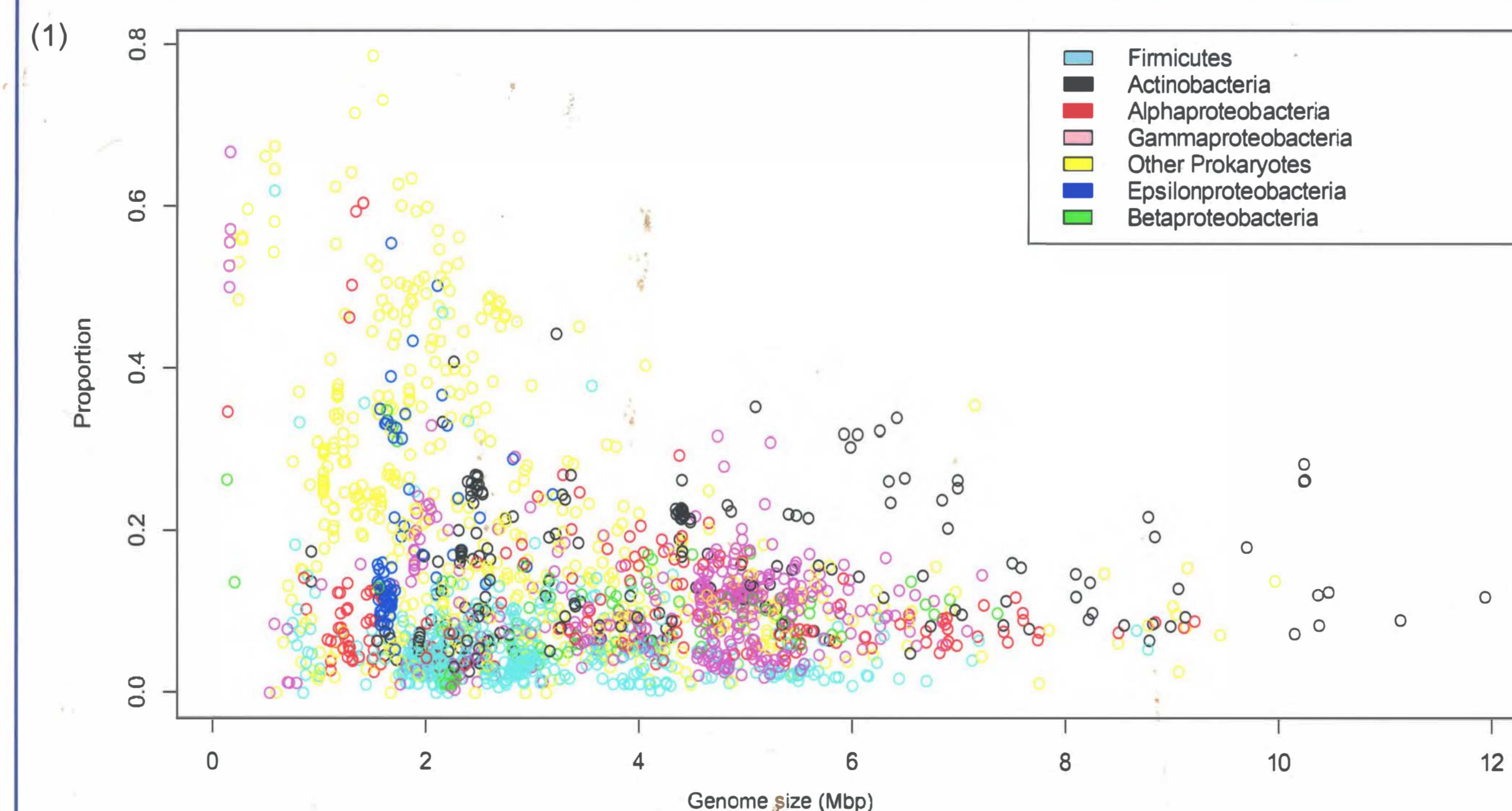


Figure 1: Distribution of proportion of COGPs (Convergent Overlapping gene Pairs) with increase in genome size for bacterial taxa

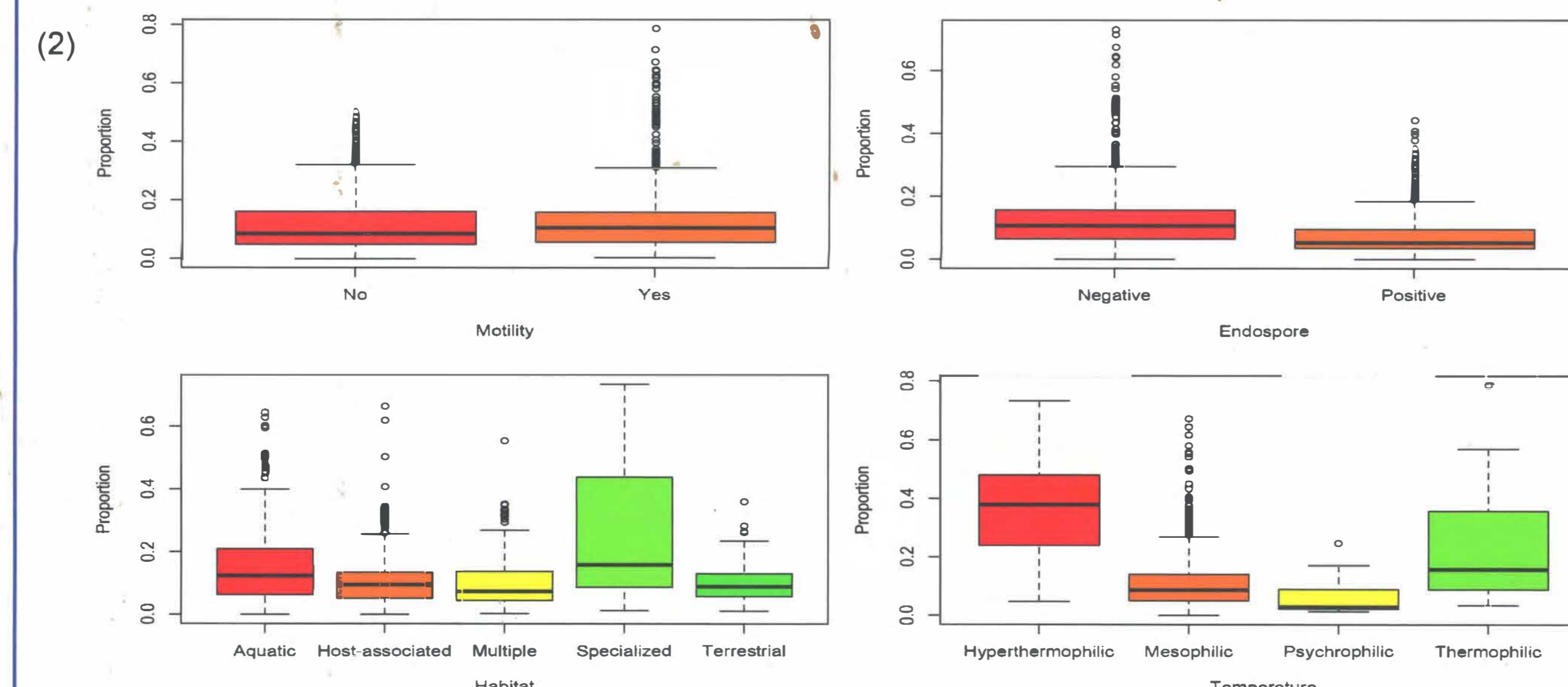


Figure 2: Comparison of proportion of COGPs for different lifestyle parameters across bacterial genomes

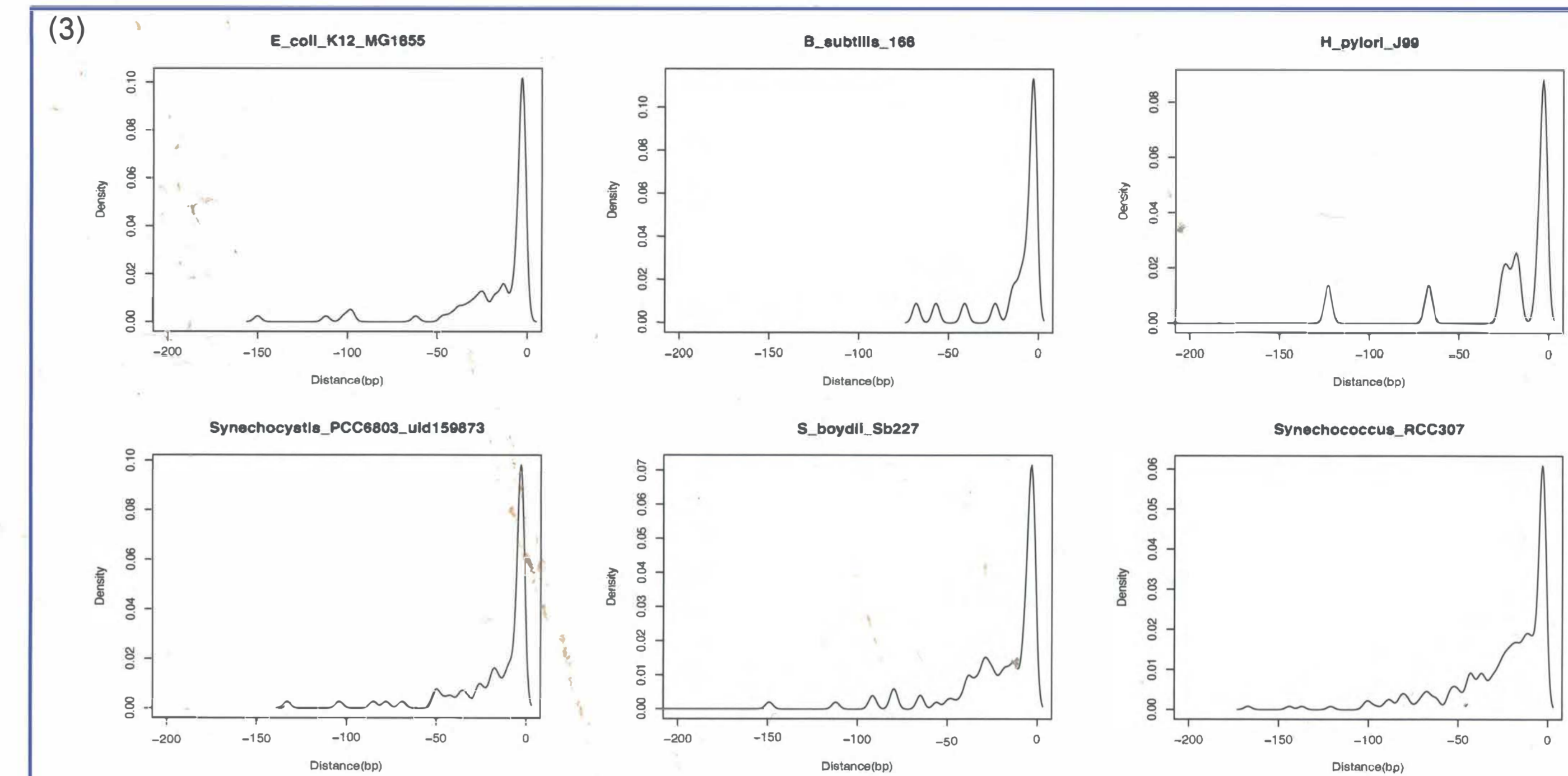


Figure 3: Density distribution of COGPs with inter-genic distance in selected bacterial genomes

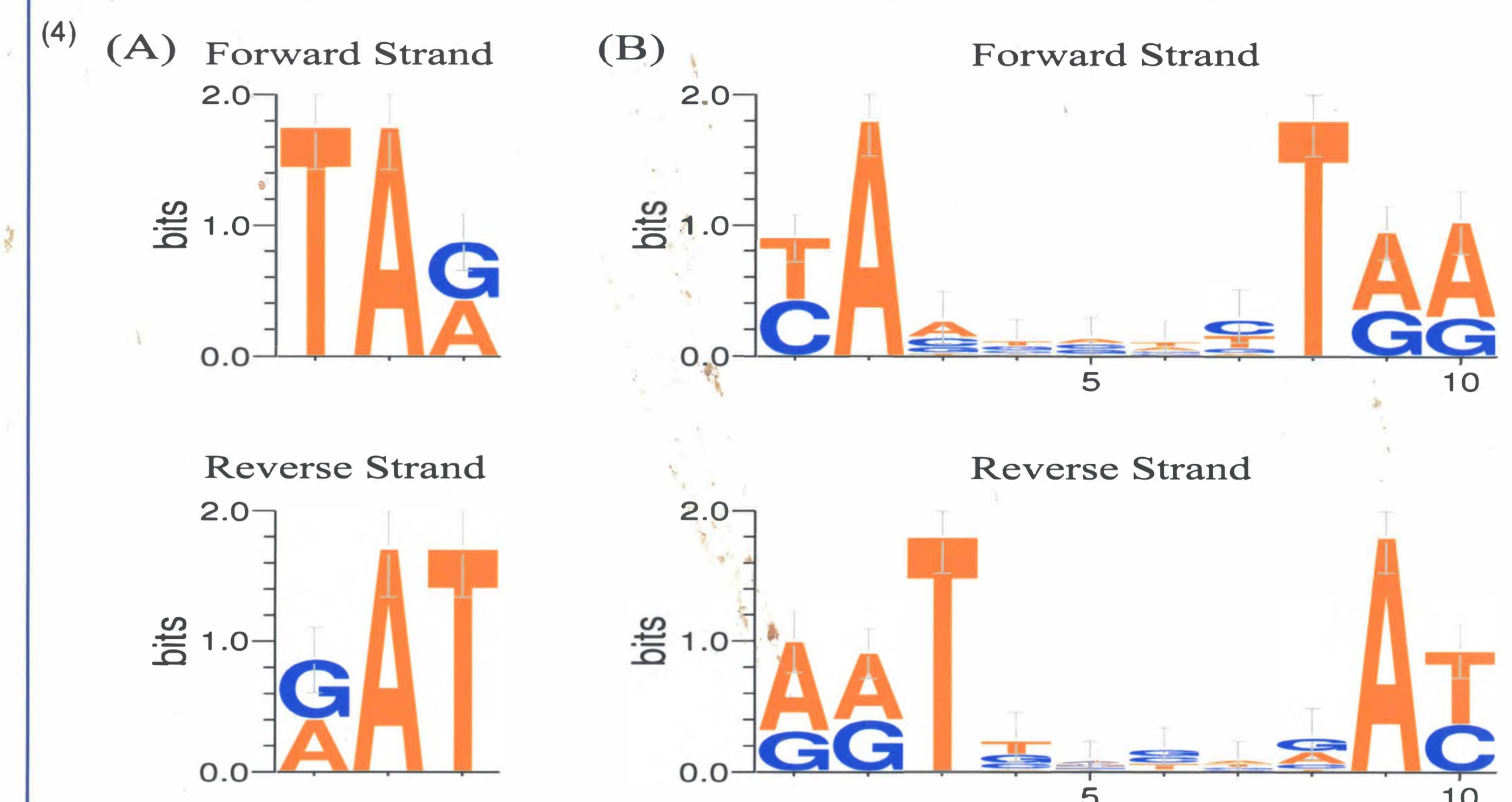


Figure 4: Sequence logo representing stop codons for the COGPs having inter-genic distance (A) -3 and (B) -10

Conclusion

In this study, we explore the role of various factors contributing to convergent overlapping transcription in bacterial genomes. Our analysis shows that the proportion of convergent overlapping gene pairs (COGPs) in a genome is affected by endospore formation, bacterial habitat and the temperature range. Our results also show that the density distribution of COGPs across the genomes is high for shorter overlaps with increased conservation of distances for decreasing overlaps. Our study also reveals that COGPs frequently contain stop codon overlaps with the middle base position exhibiting mismatches between complementary strands.

References

- Chatterjee, A., Drews, L., Mehra, S., Takano, E., Kaznessis, Y. N. & Hu, W. S. (2011). Convergent transcription in the butyrolactone regulon in *Streptomyces coelicolor* confers a bistable genetic switch for antibiotic biosynthesis. *PLoS One* 6, e21974.
- Crampton, N., Bonass, W. A., Kirkham, J., Rivetti, C. & Thomson, N. H. (2006). Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Res* 34, 5416-25.
- Gullerova, M. & Proudfoot, N. J. (2012). Convergent transcription induces transcriptional gene silencing in fission yeast and mammalian cells. *Nat Struct Mol Biol* 19, 1193-201.
- Hobson, D. J., Wei, W., Steinmetz, L. M. & Svestrup, J. Q. (2012). RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 48, 365-74.
- Chatterjee, A., Johnson, C. M., Shu, C. C., Kaznessis, Y. N., Ramkrishna, D., Dunny, G. M. & Hu, W. S. (2011). Convergent transcription confers a bistable switch in *Enterococcus faecalis* conjugation. *Proc Natl Acad Sci U S A* 108, 9721-6.

Contact

Kashish Chetal (kchetal@iupui.edu)

*Dr. Sarath Chandra Janga (scjanga@iupui.edu)