

# Are essential genes conserved?

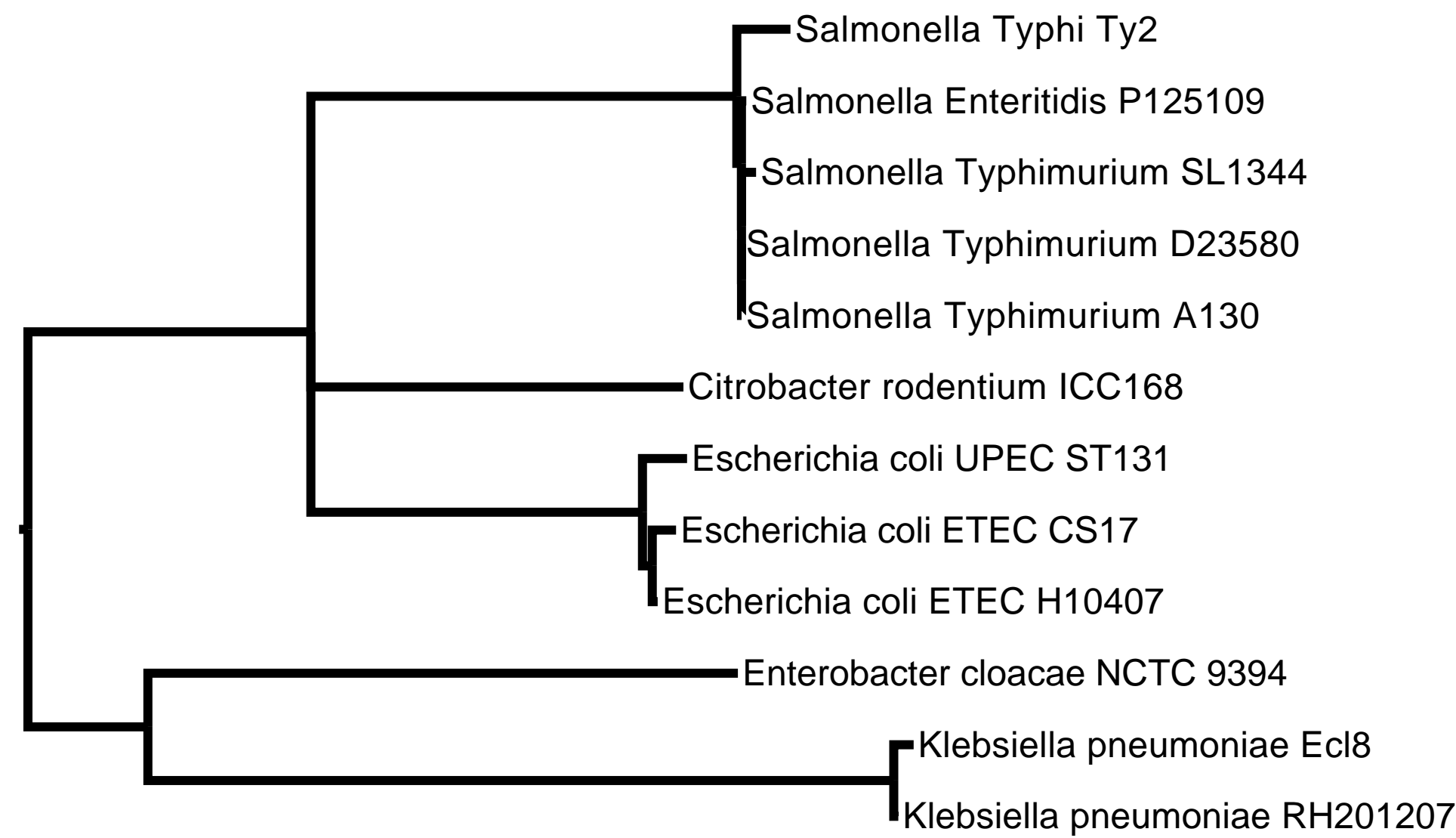
Fatemeh Ashari Ghomi<sup>1</sup>, Paul Gardner<sup>1</sup>, Lars Barquist<sup>2</sup>

<sup>1</sup> School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

<sup>2</sup> Institute for Molecular Infection Biology, University of Würzburg, Würzburg, Germany

## Introduction

Transposon-directed insertion-site sequencing is an approach for studying the essentiality of genes in prokaryotes. In this method, pools of single insertion mutants are constructed using transposon mutagenesis and the effect of each mutation on the mutants' survival is evaluated by sequencing the survivors. This can lead to the identification of essential genes. We have used transposon-directed insertion-site sequencing to study the essentiality of genes in 12 strains from Enterobacteriaceae which are depicted in the figure. For this, we have studied different biases that can affect our transposon insertion experiment. After correcting for the biases, we have studied the relation between the essentiality of genes and their conservation.

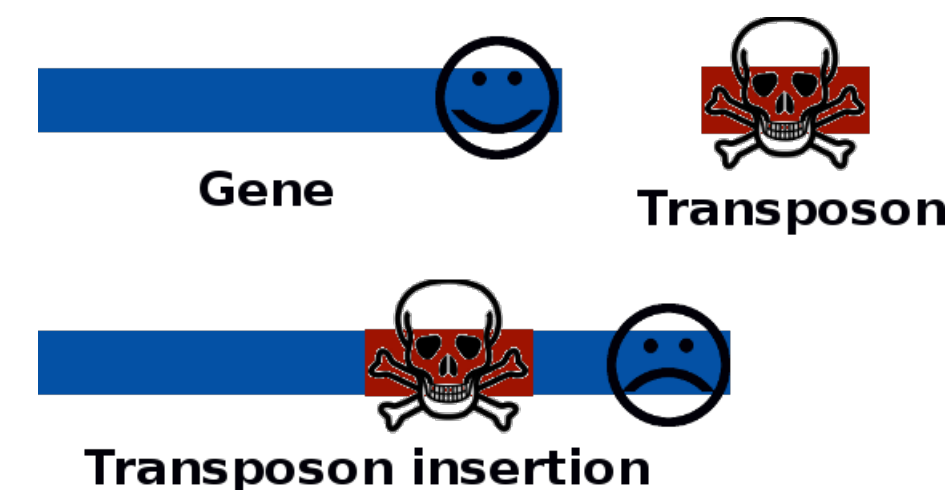
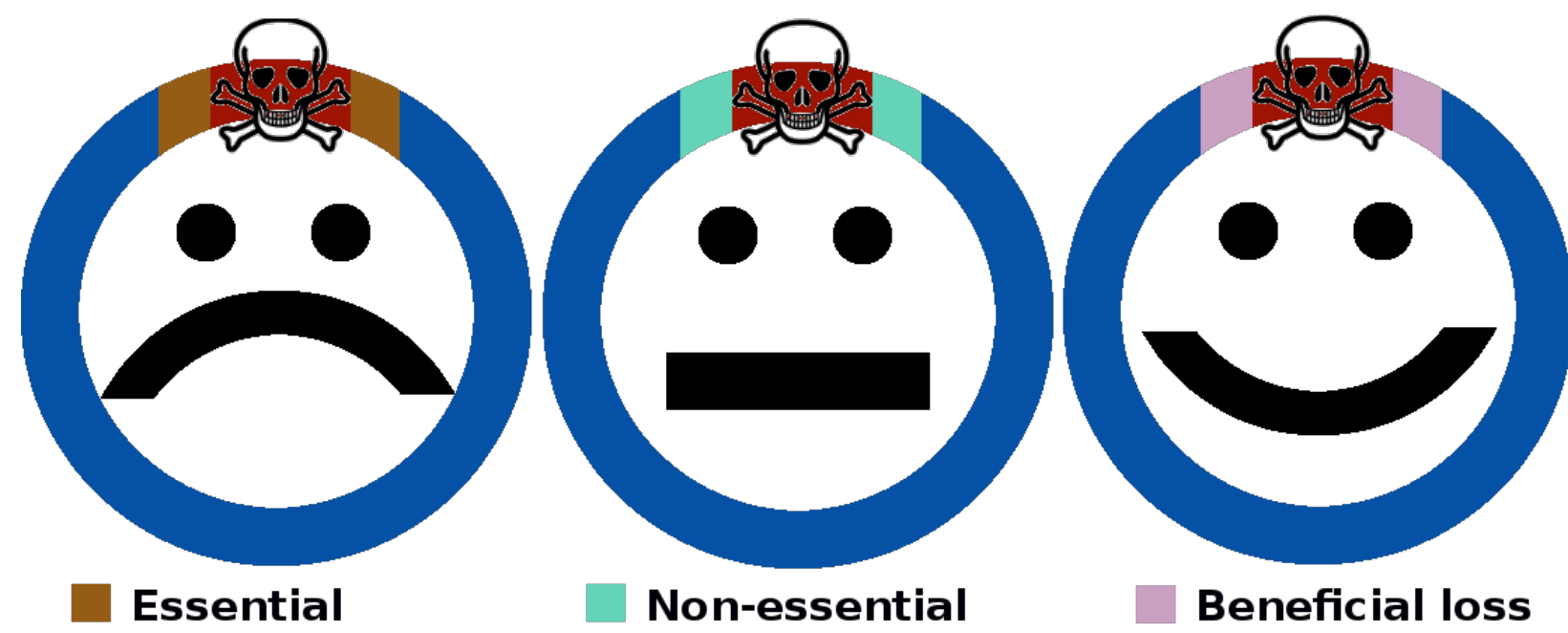


## Questions

1. Are there any biases that affect the results of transposon insertion experiments?
2. Is the conservation of essentiality consistent with the species tree?
3. Are essentiality of genes and their conservation related?

## Transposon-directed insertion-site sequencing

Transposon insertion is the process of inserting a nucleotide sequence into a gene so that it disrupts the gene and causes the gene lose its functionality.



- If the gene is *essential* the organism will not be able to survive.
- If it is *non-essential* the organism will be able to survive.
- If it is a *beneficial loss* the organism will benefit from losing it.

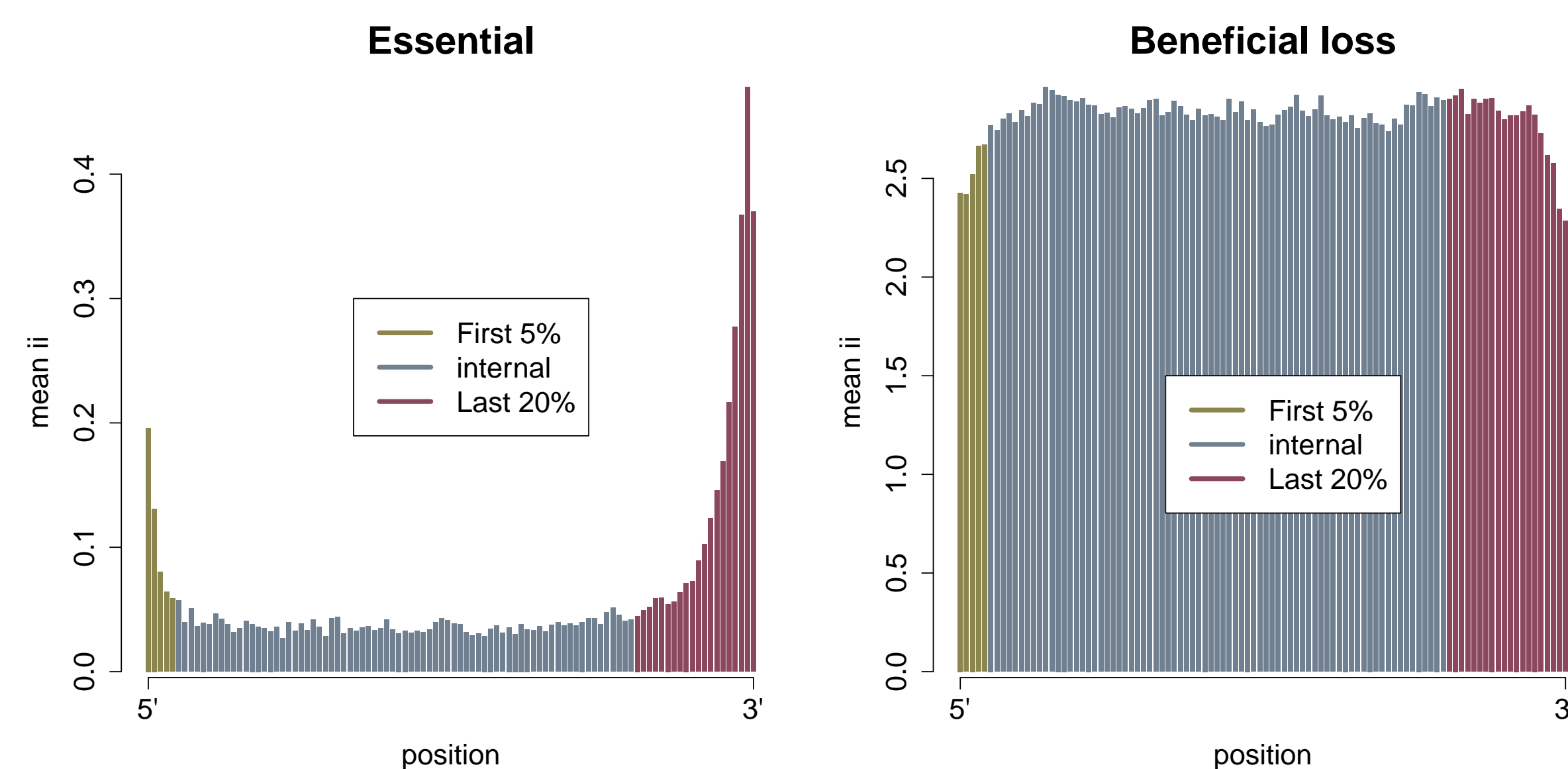


After genome sequencing:

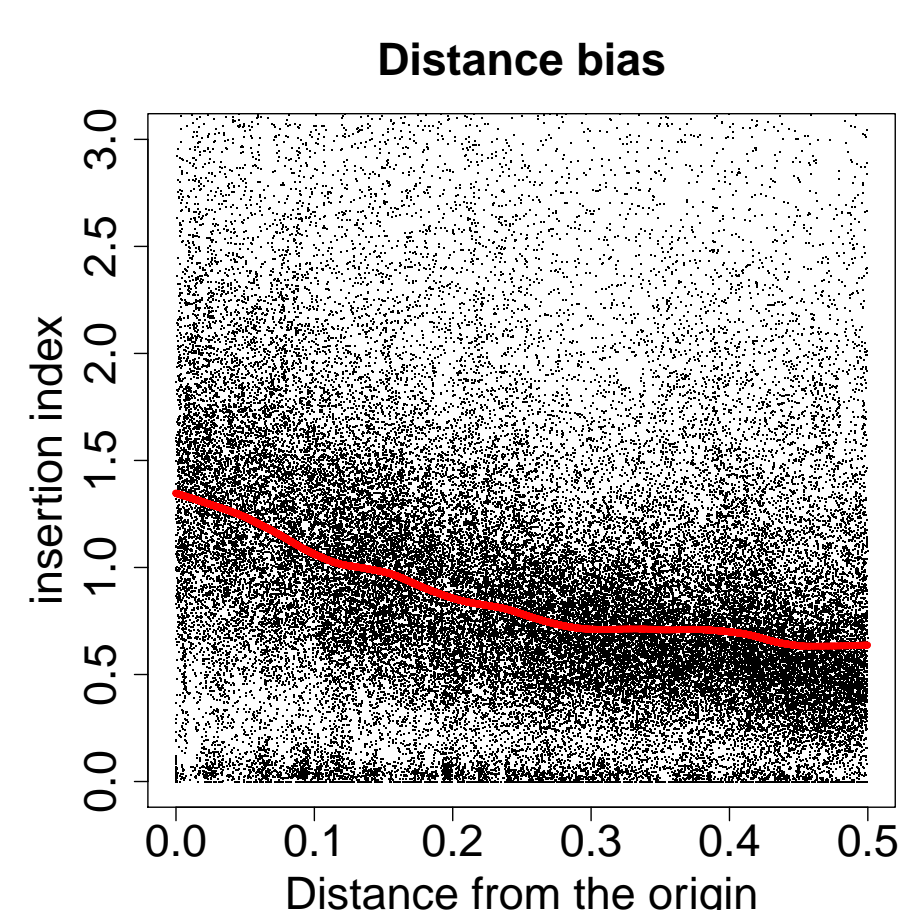
- No or few transposon insertions are spotted in essential genes.
- An intermediate number of transposon insertions are detected in non-essential genes.
- Many transposon insertions are observed in beneficial losses.

## Are transposon insertions evenly distributed within genes?

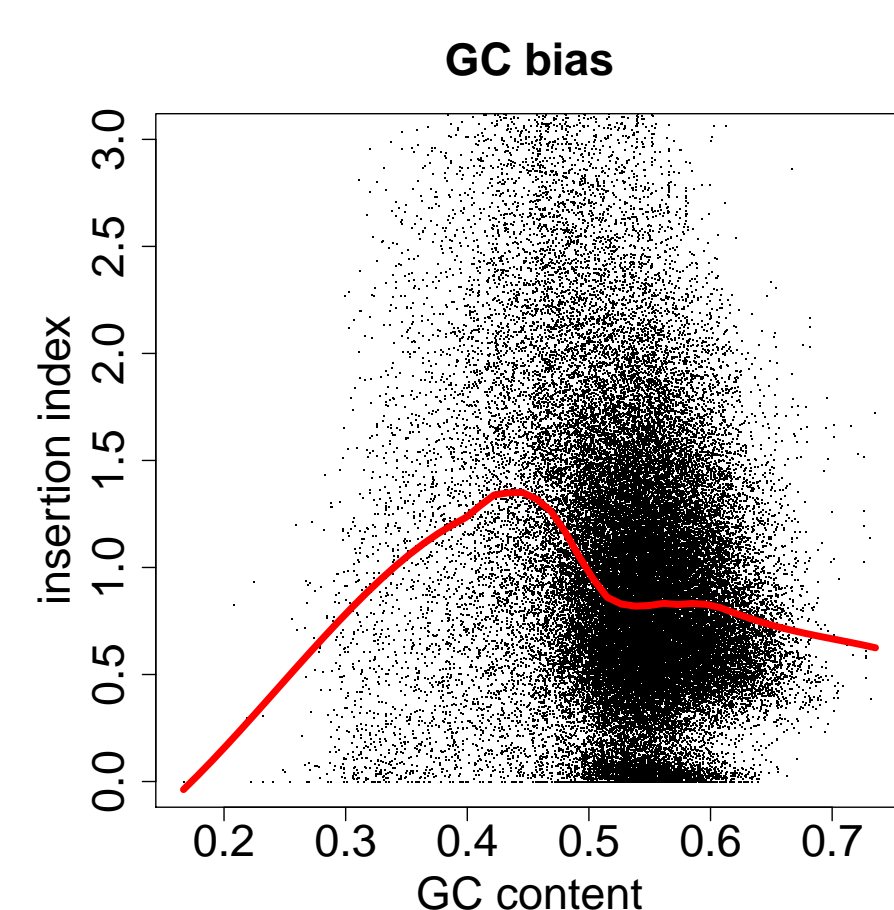
We have divided our genes into 3 segments: 5% of the genes on the 5' end, 20% of the genes on the 3' end, and the rest in the middle. The figure shows that the number of insertions on the 3' and 5' ends is more than the internal region in essential genes and less than the internal region in beneficial losses.



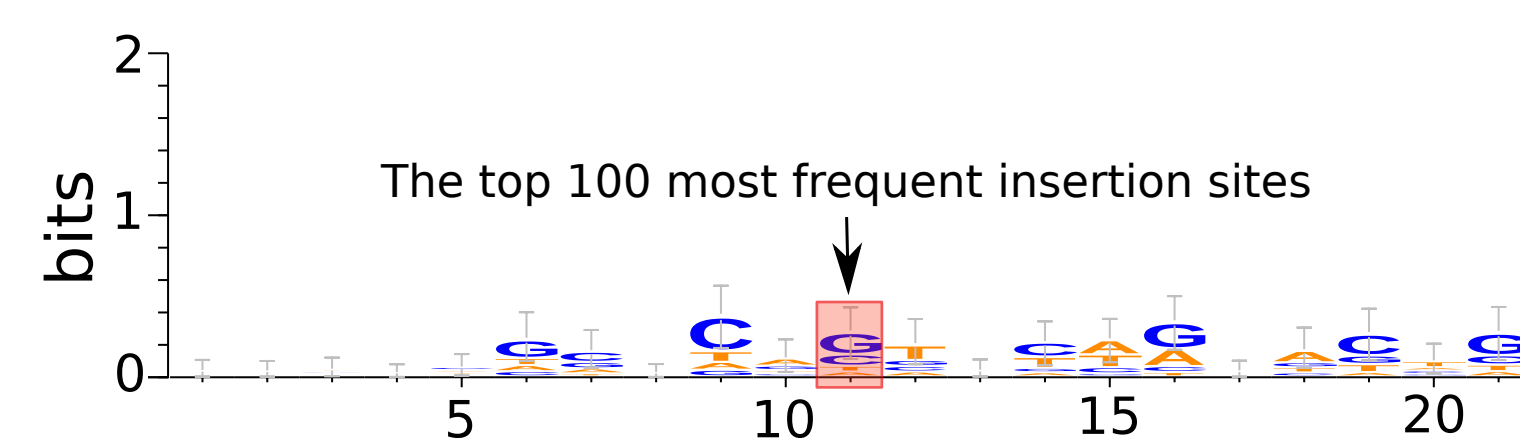
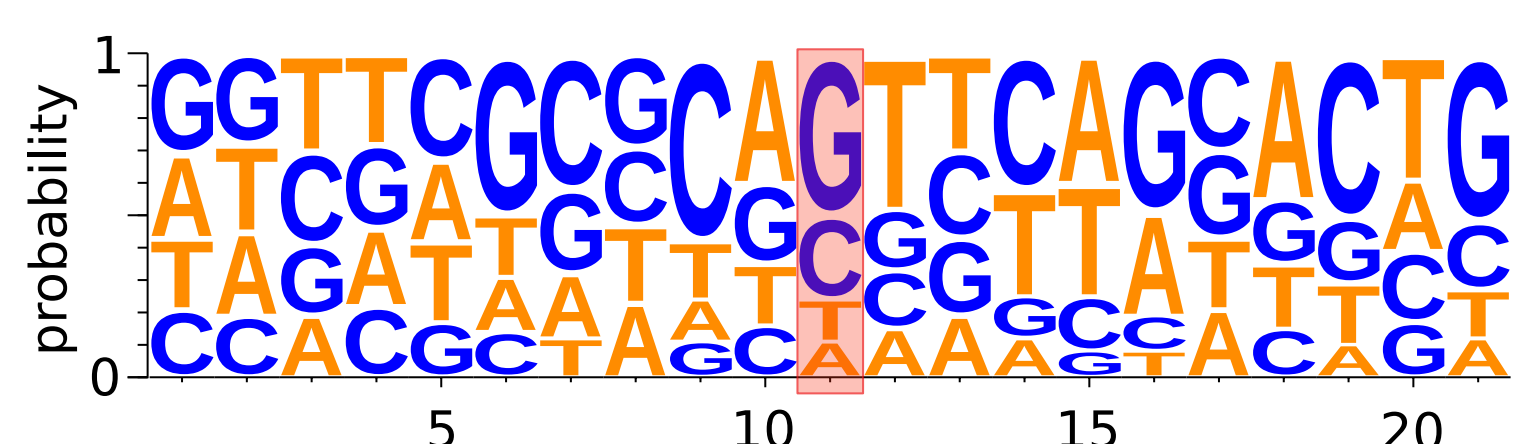
## Are transposons biased towards certain positions in the genome?



We have also investigated if the number of insertions in genes is related to the position of the gene within the genome or the G-C content of the gene. The results propose the further we get from the origin of replication, the fewer number of insertions we have (the left figure). Moreover, the right figure shows that when the G-C content is greater than 0.5, there is no bias.

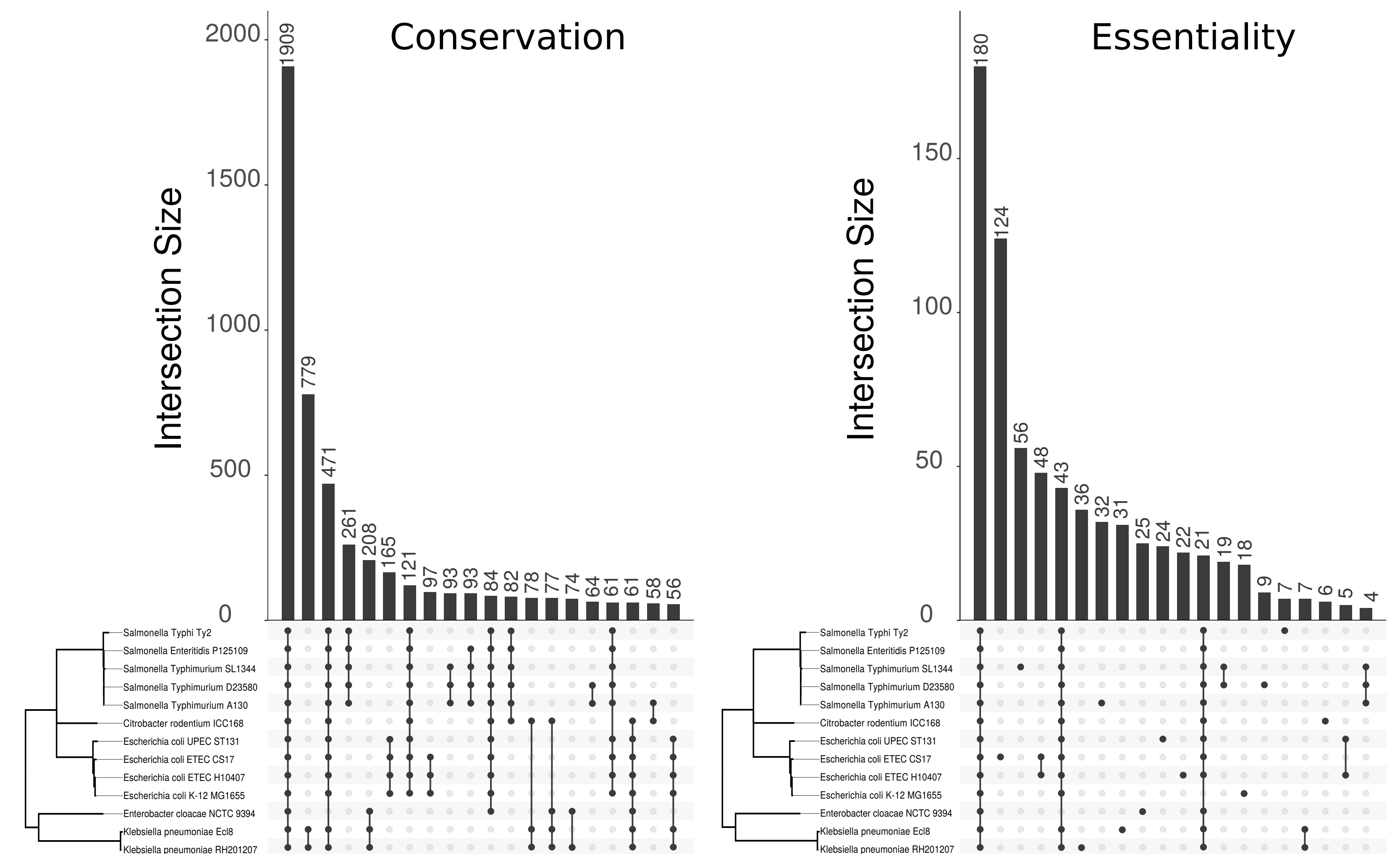


To test whether the results are biased towards certain motifs, we have generated logos from 10 nucleotides flanking the 100 top most frequent insertion sites. The analysis shows no significant bias.



## Is the conservation of essentiality consistent with the species tree?

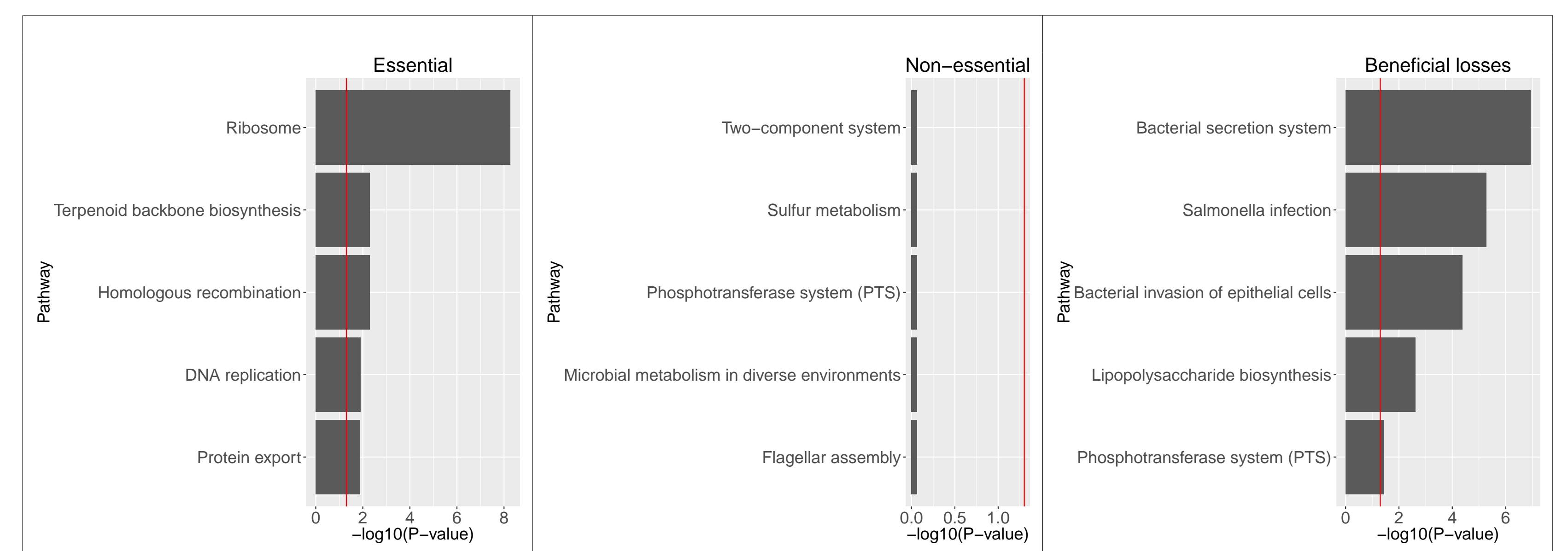
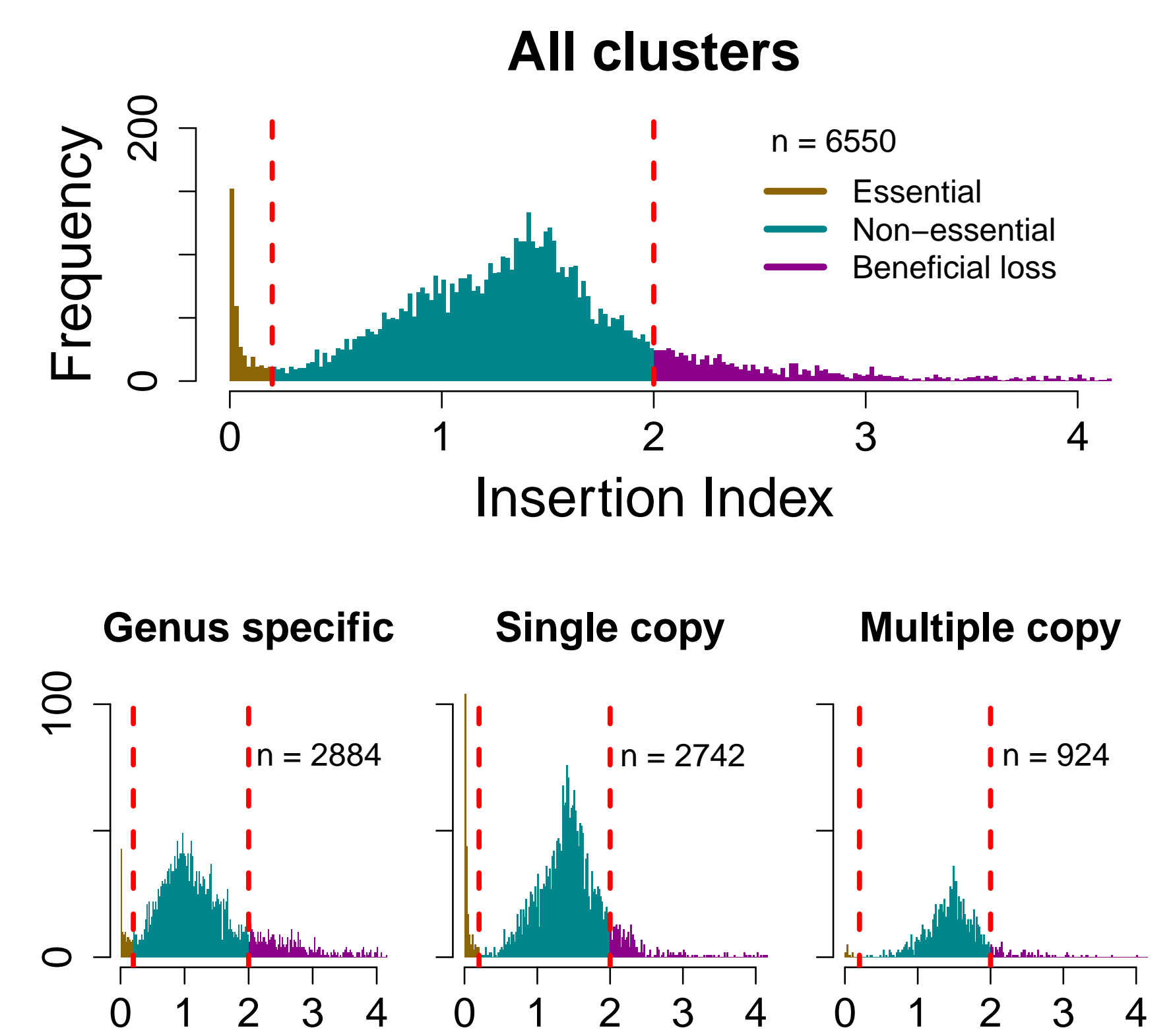
We have compared the number of genes that are conserved in different strains in our study and the number of genes that are essential in these strains. The results propose that although conservation of genes follows a tree-like trend, the essentiality does not show a tree-like signal.



## Are essential genes more likely to be conserved?

We have divided the genes in our 12 strains into 3 groups: genus specific genes, genes with one copy per genome, and genes with multiple copies per genome. The study of essentiality in these groups shows that most of the essential genes are copied once per genome and most of the beneficial losses are genus specific.

We have performed a pathway enrichment analysis on different groups of genes in *Salmonella* Typhi using KOBAS 2.0. The results indicate that essential genes are mostly involved in essential pathways such as replication and translation; the enrichment of the pathways related to non-essential genes is not statistically significant; and the beneficial losses are mostly involved in pathways that are not needed in nutrient-rich broth.



## Conclusions

- The 5' and 3' ends of genes have a different tolerance for insertions compared to the internal region in transposon-directed insertion-site sequencing.
- The number of transposons inserted to a gene is related to the distance of the gene from the origin of replication.
- The transposons are not biased towards certain motifs or G-C content of the gene.
- The conservation of essentiality is not consistent with the species tree.
- Essential genes are more likely to be conserved.

## Contact

Fatemeh Ashari Ghomi  
fatemeh.ashari ghomi@pg.canterbury.ac.nz