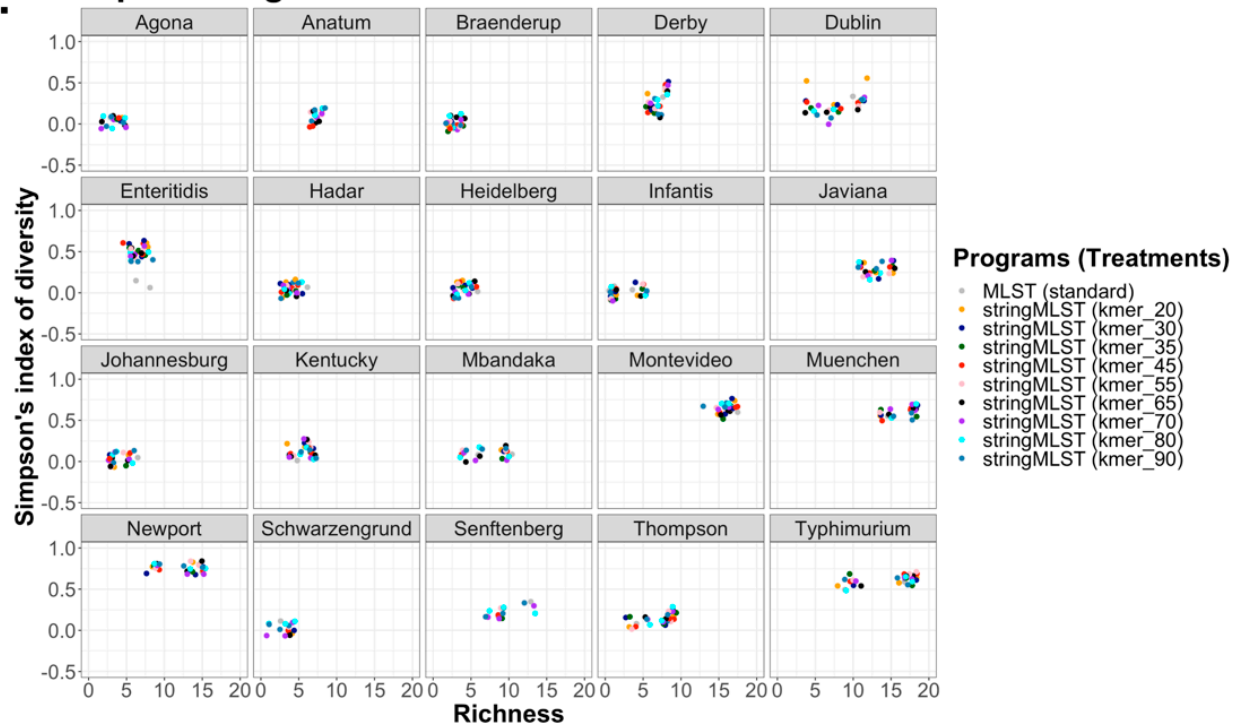
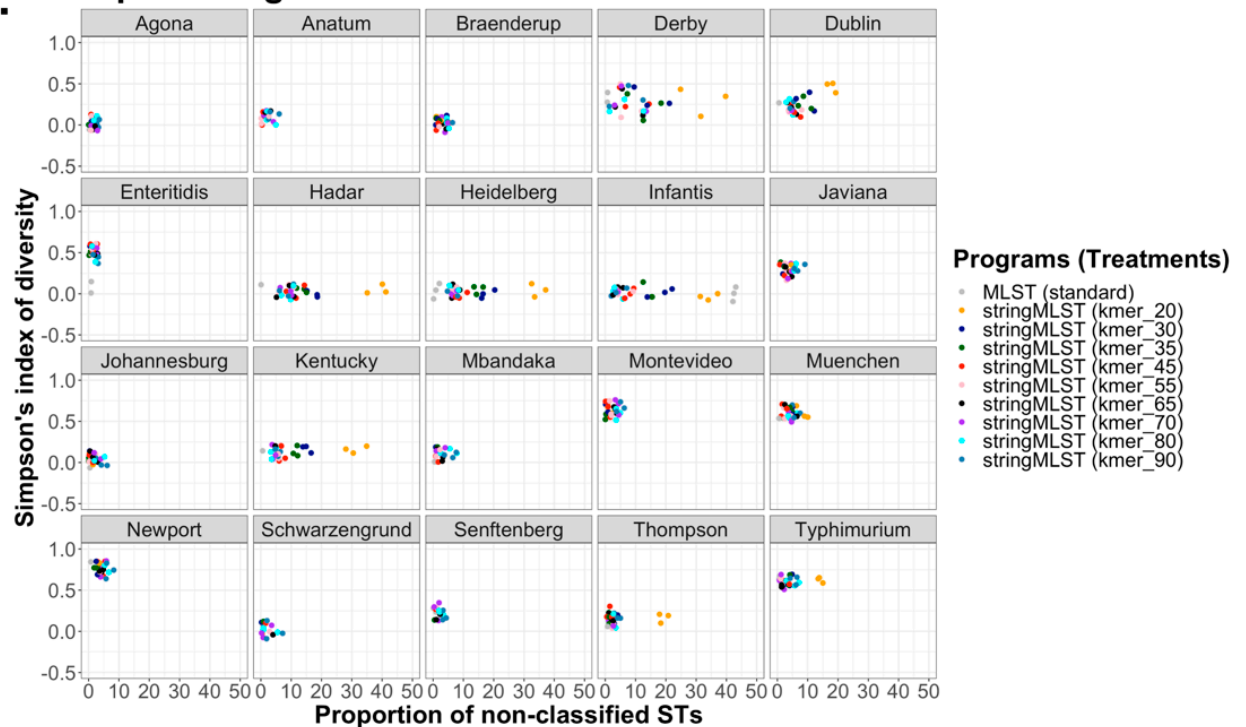


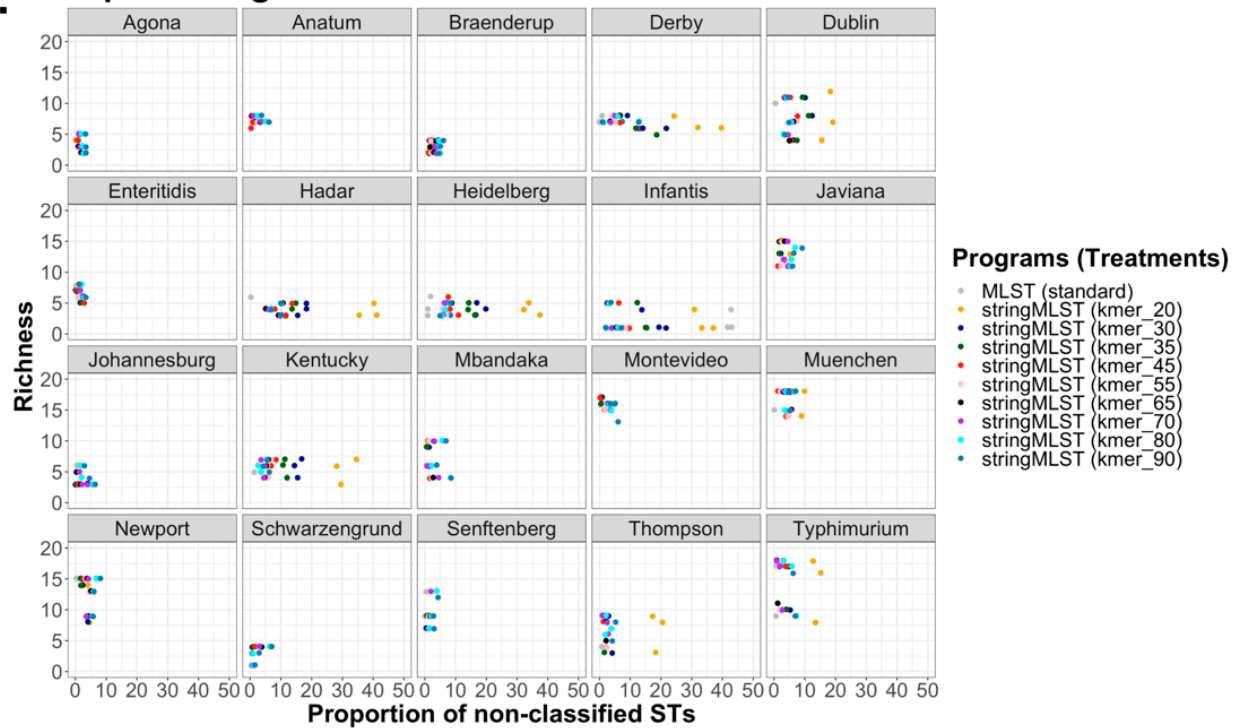
## A. Biplot using mean of indexes across serovars



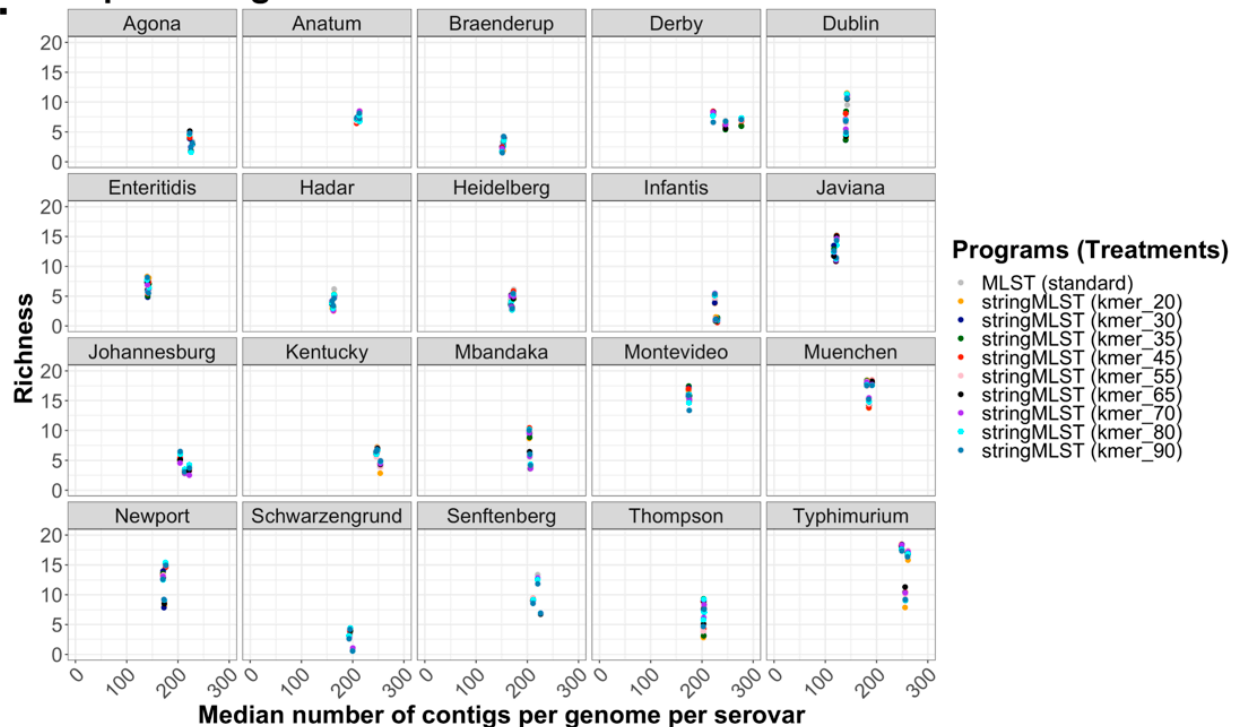
## B. Biplot using mean of indexes across serovars



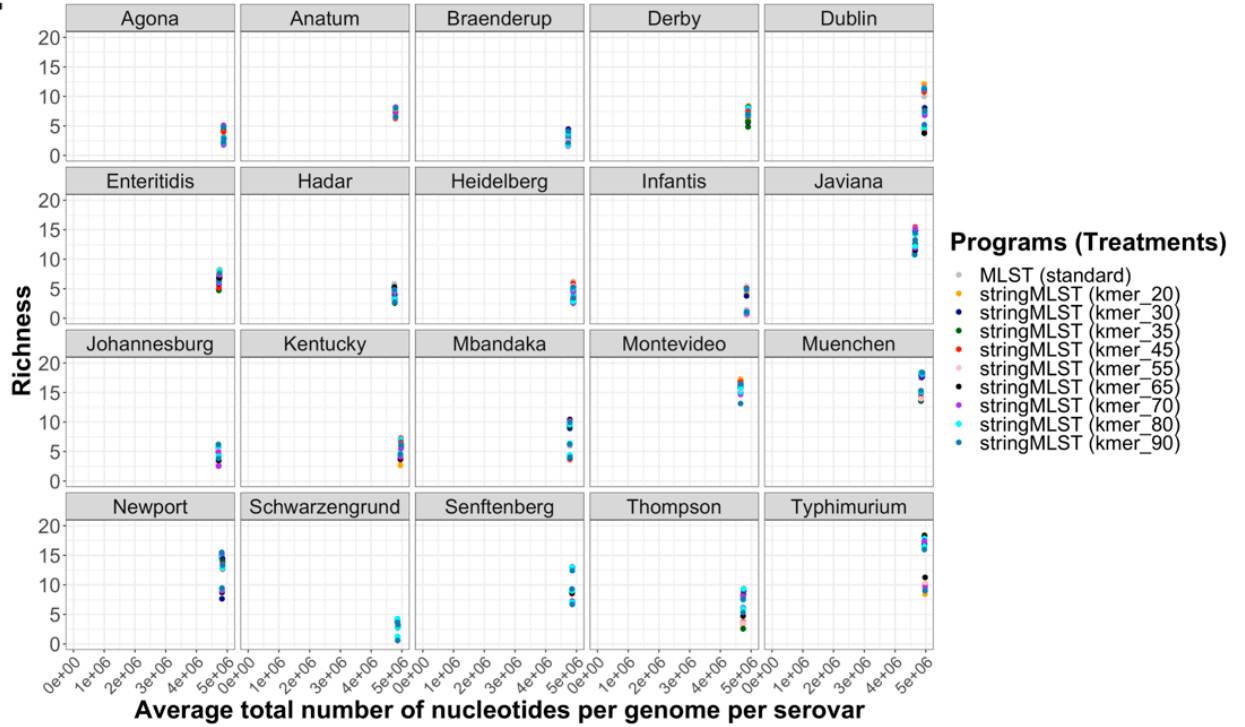
### C. Biplot using mean of indexes across serovars



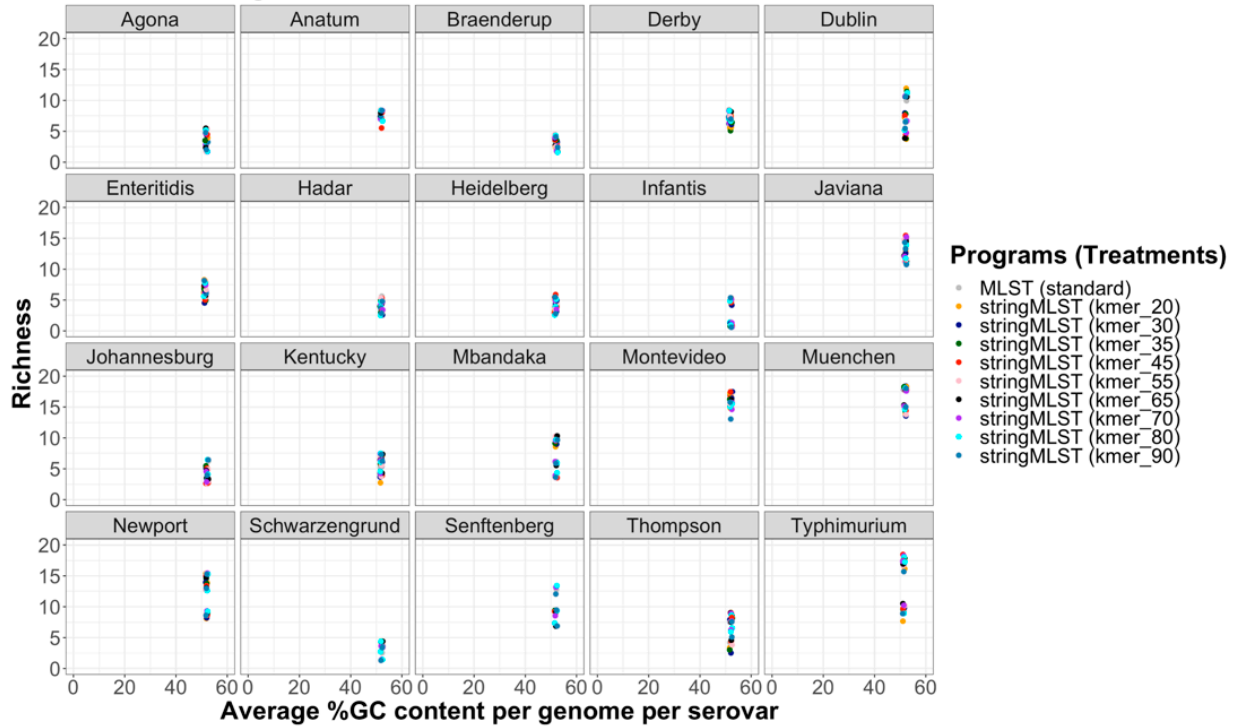
### D. Biplot using mean of indexes across serovars



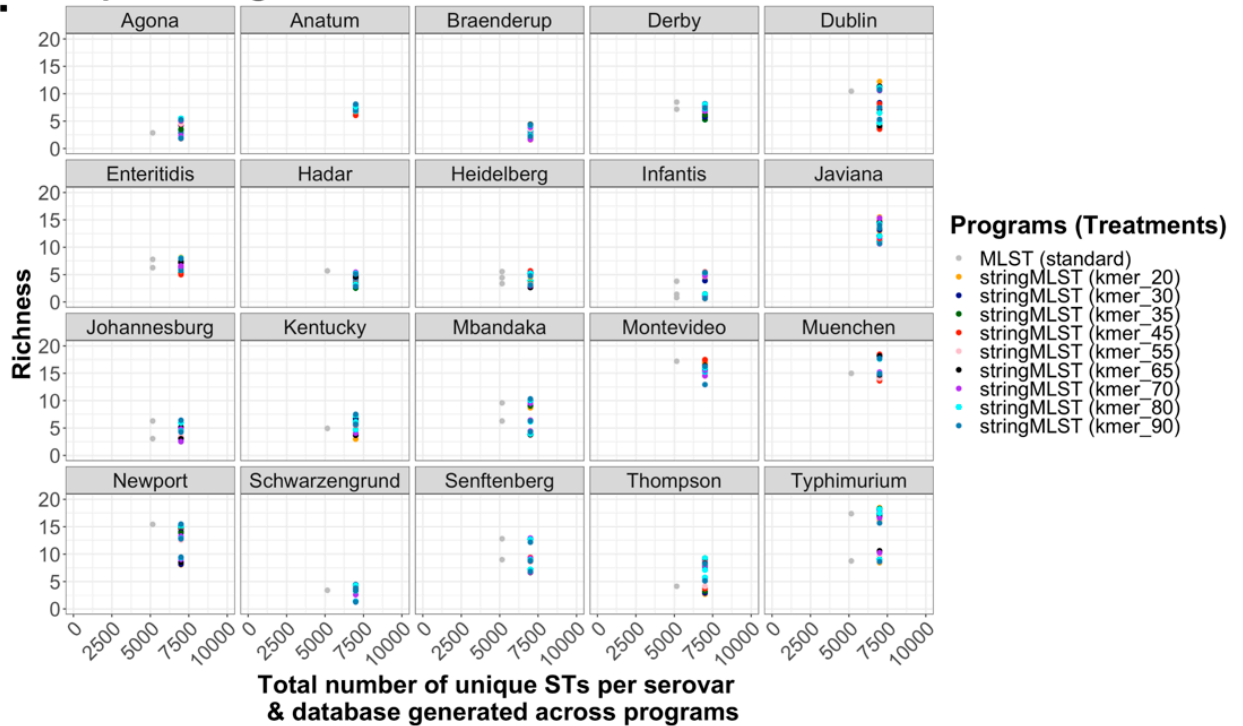
## E. Biplot using mean of indexes across serovars



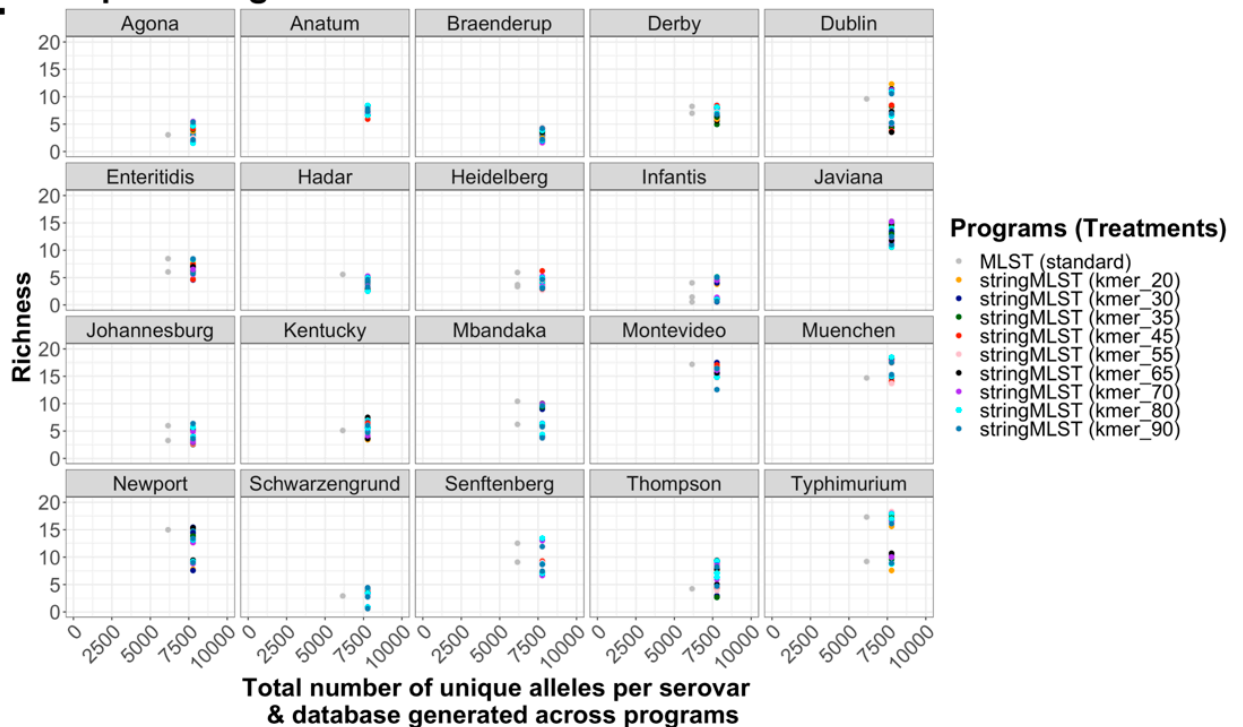
## F. Biplot using mean of indexes across serovars



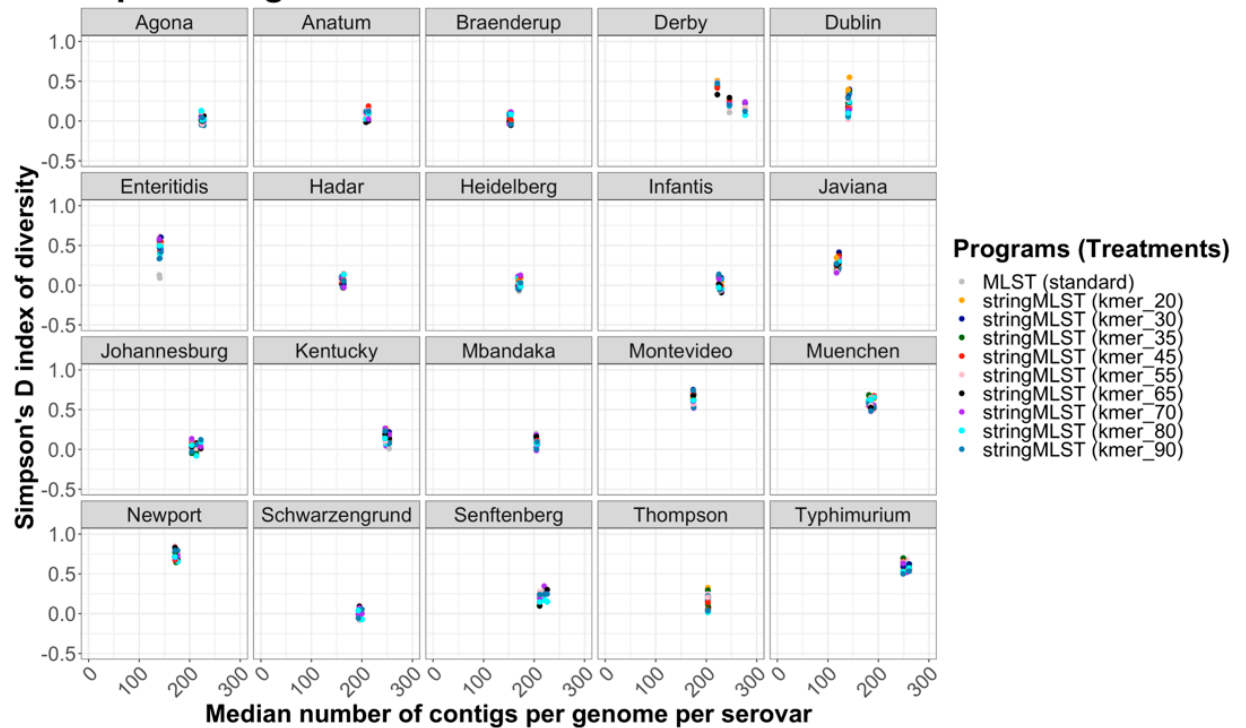
## G. Biplot using mean of indexes across serovars



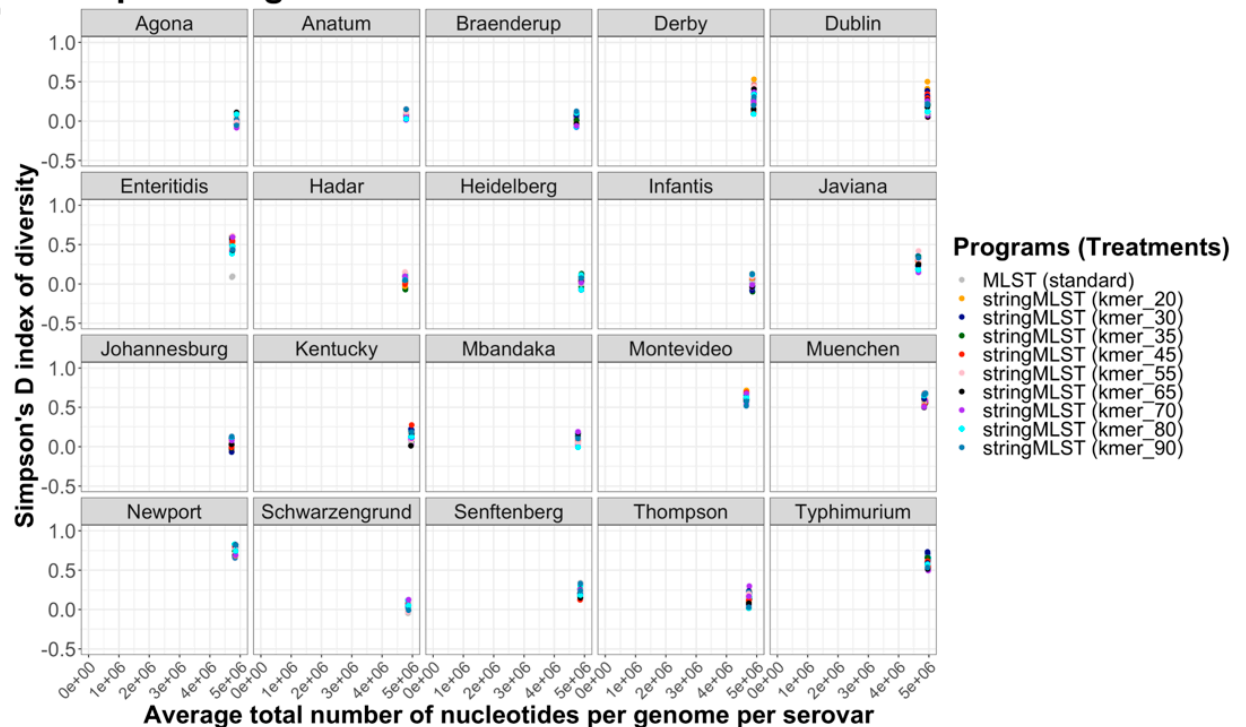
## H. Biplot using mean of indexes across serovars



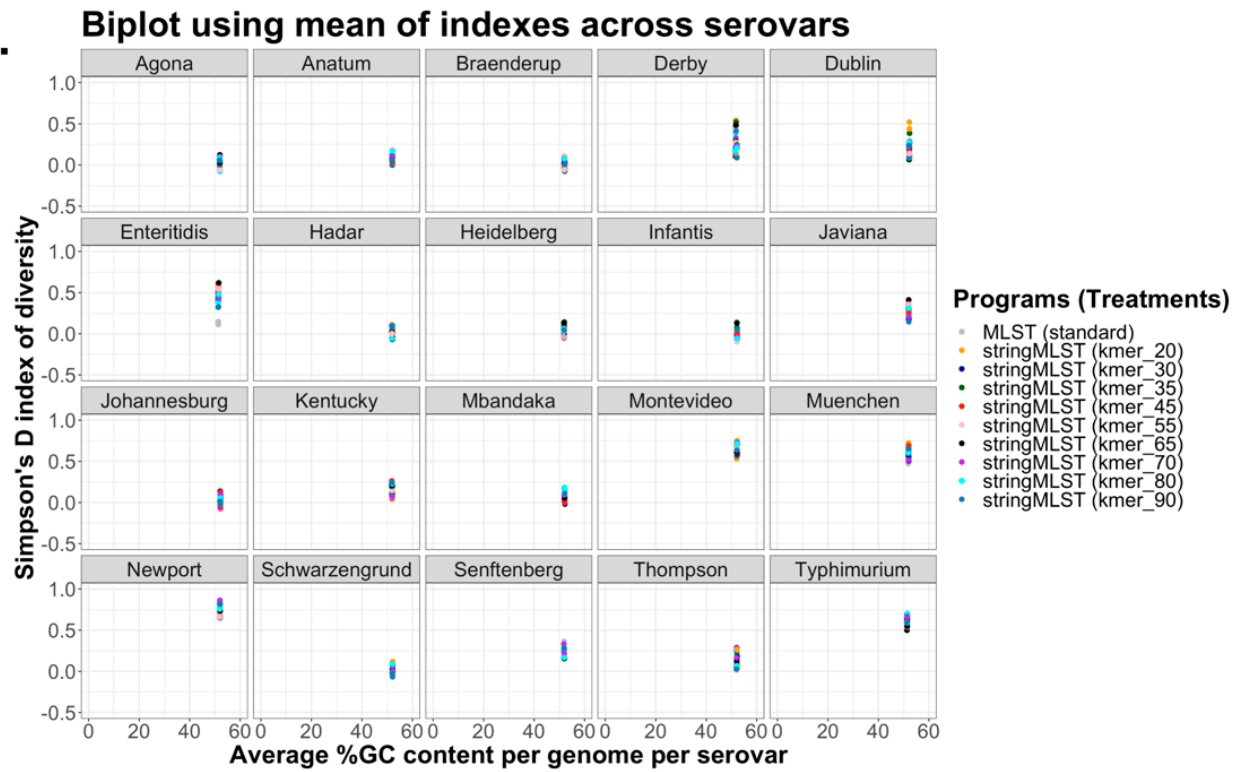
# I. Biplot using mean of indexes across serovars



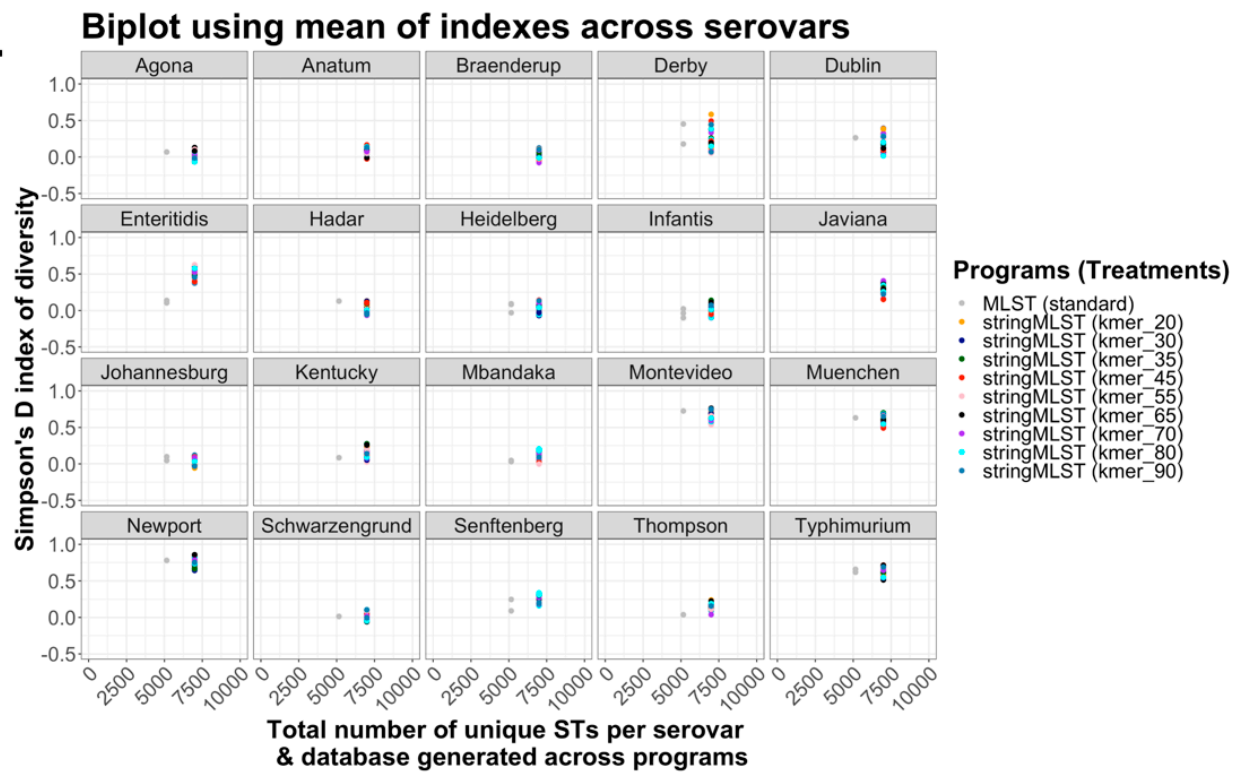
# J. Biplot using mean of indexes across serovars



K.



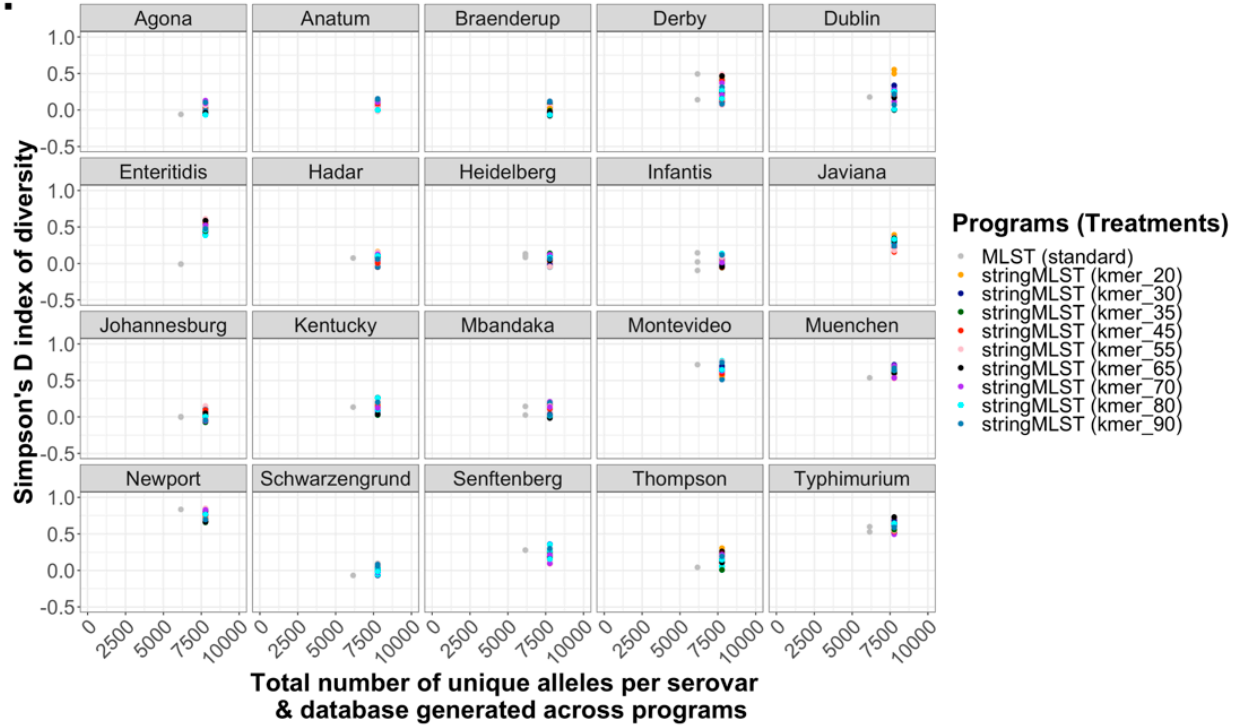
L.





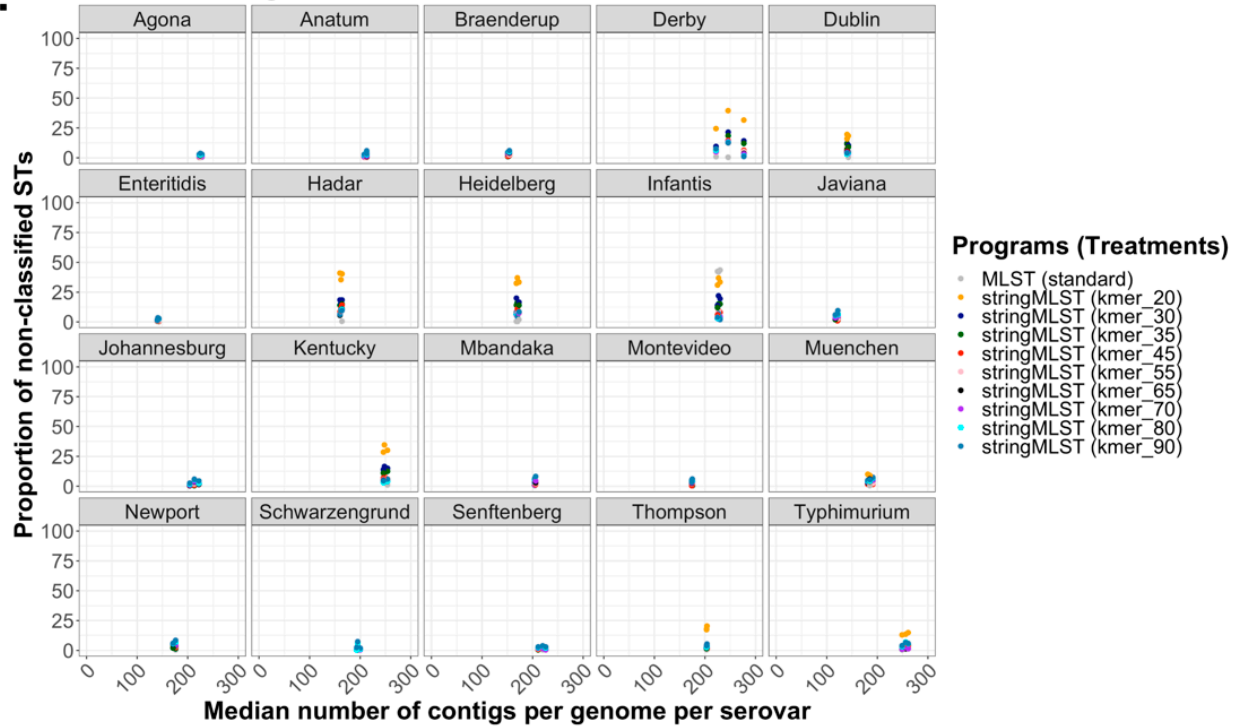
**M.**

### Biplot using mean of indexes across serovars

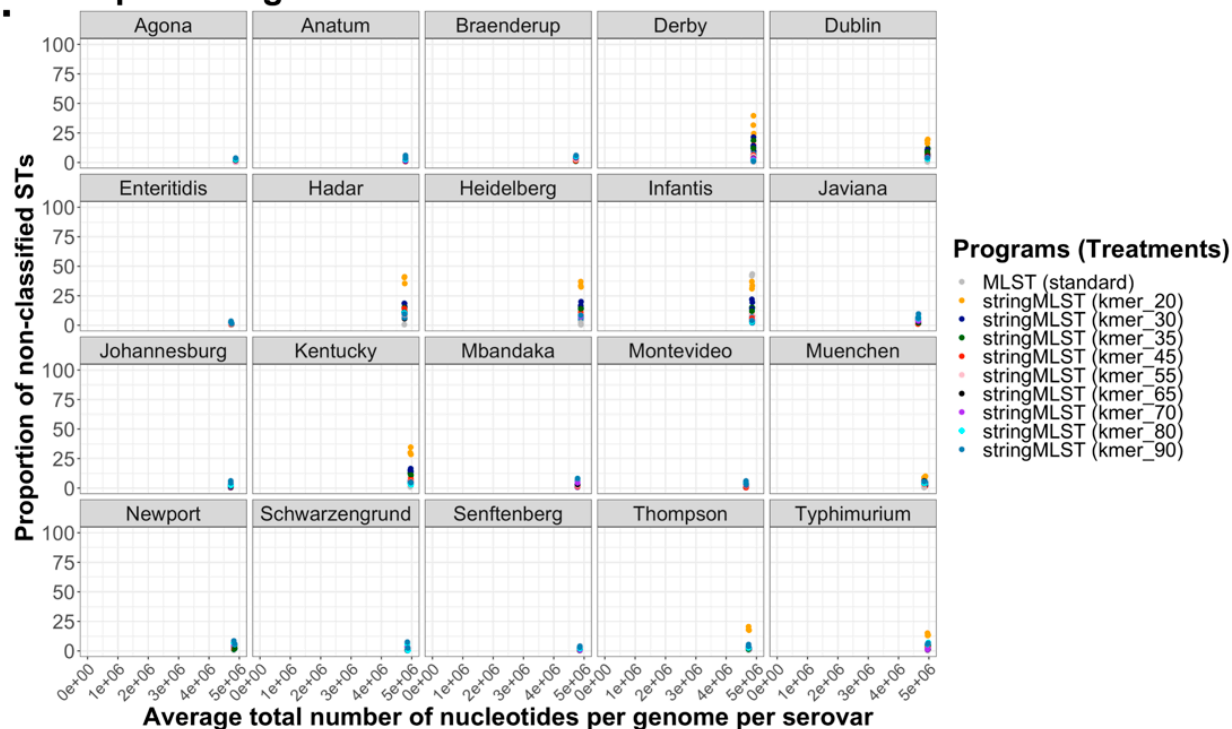


**N.**

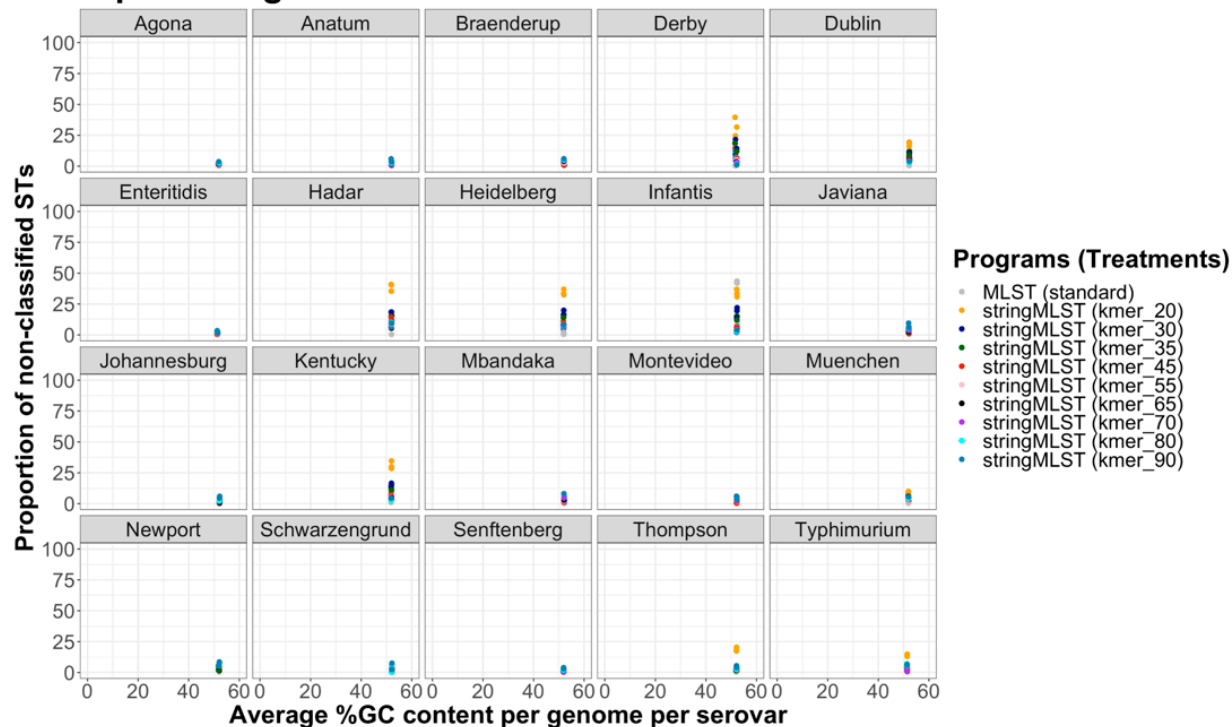
### Biplot using mean of indexes across serovars



## O. Biplot using mean of indexes across serovars



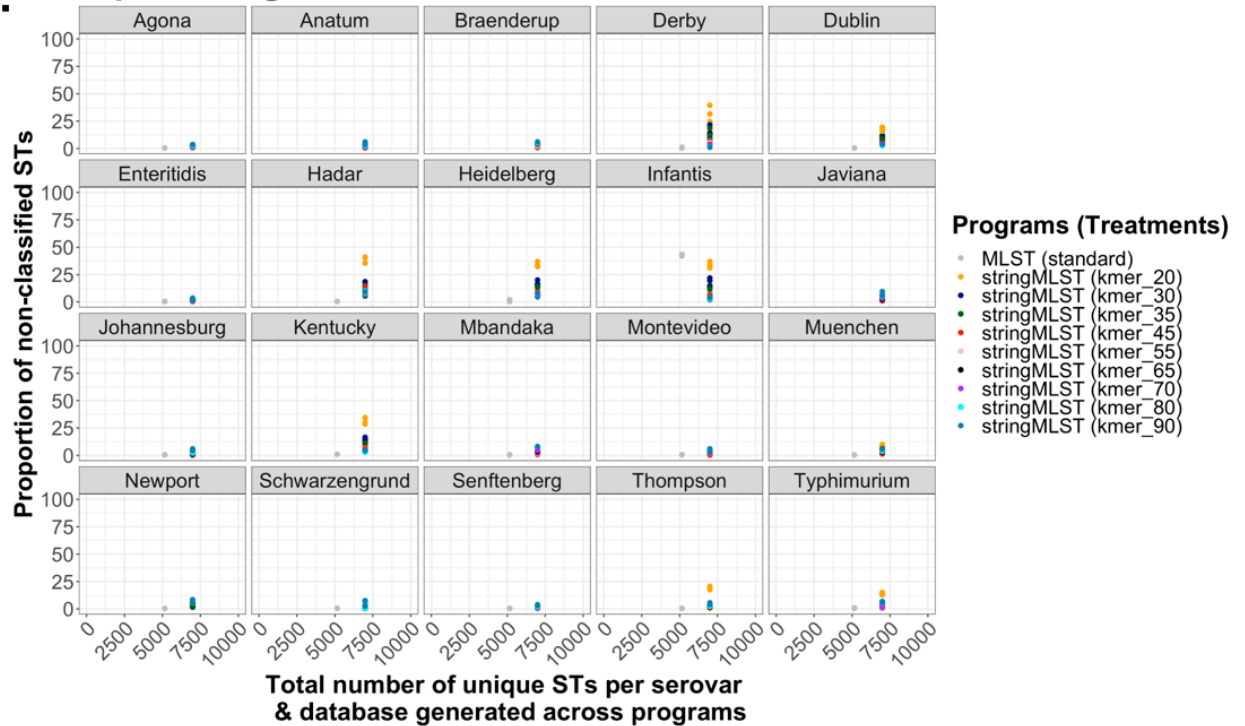
## P. Biplot using mean of indexes across serovars





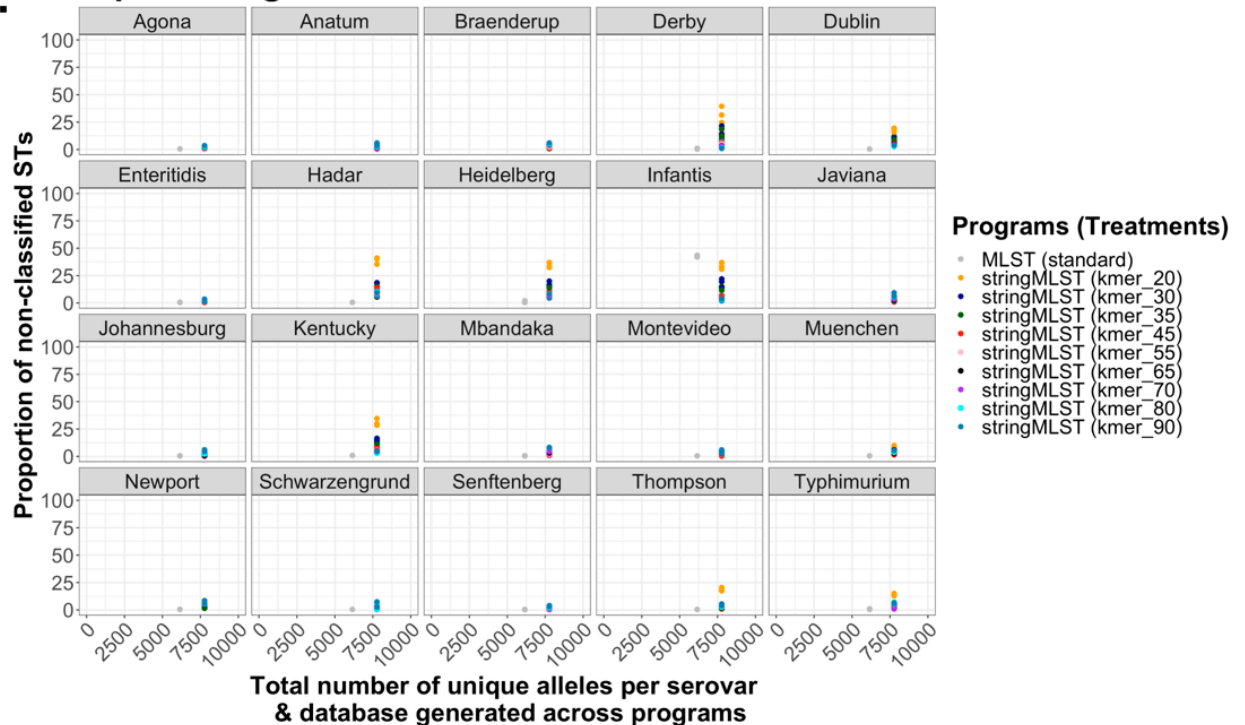
**Q.**

### Biplot using mean of indexes across serovars



**R.**

### Biplot using mean of indexes across serovars



**Figure S15.** Bivariate association between accuracy-based classification metrics and genome-intrinsic and –extrinsic variables across *S. enterica* serovars.

(A) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. ST richness (Richness). (B) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. proportion of non-classified STs. (C) Scatter plot demonstrating the relationship between the ST richness (Richness) vs. proportion of non-classified STs across *S. enterica* serovars and programs. (D) Scatter plot demonstrating the relationship between the ST richness (Richness) vs. median number of contigs per genome across *S. enterica* serovars and programs. (E) Scatter plot demonstrating the relationship between the ST richness (Richness) vs. the average total number of nucleotides per genome across *S. enterica* serovars and programs. (F) Scatter plot demonstrating the relationship between the ST richness (Richness) vs. the average GC% content per genome. (G) Scatter plot demonstrating the relationship between the ST richness (Richness) vs. the total number of unique STs in the database. (H) Scatter plot demonstrating the relationship between the ST richness (Richness) vs. the total number of unique alleles across all 7 loci in the database. (I) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. the median number of contigs per genome. (J) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. the average total number of nucleotides per genome. (K) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. the average GC% content per genome. (L) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. the total number of unique STs in the database. (M) Scatter plot demonstrating the relationship between the Simpson's D index of diversity ( $1 - D$ ) vs. the total number of unique alleles across all 7 loci in the database. (N) Scatter plot demonstrating the relationship between the proportion of non-classified STs vs. the median number of contigs per genome. (O) Scatter plot demonstrating the relationship between the proportion of non-classified STs vs. the average total number of nucleotides per genome. (P) Scatter plot demonstrating the relationship between the proportion of non-classified STs vs. the average GC% content per genome. (Q) Scatter plot demonstrating the relationship between the proportion of non-classified STs vs. the total number of unique STs in the database. (R) Scatter plot demonstrating the relationship between the proportion of non-classified STs vs. the total number of unique alleles across all 7 loci in the database. All plots were stratified by *S. enterica* serovars and data points were colored by the program (mlst vs. stringMLST including all k-mer lengths).