

ParaC Pan-genome 3. Gene gain/loss events and numbers of pseudogenes by sub-lineage.

Lineages	No. of Islands (categories)	Island gains	Island losses	Nos. of island genes	pseudogenes
Para C Lineage	227 (127)	100	211	3,123	1,251
Paratyphi C	78 (41)	30	84	1,052	439
Typhisuis	15 (14)	0	15	191	202
Choleraesuis	96 (57)	46	80	1,912	381
Root branch of Para C Lineage, including Lomita	55 (43)	24	32	596	333

Note: Islands *versus* categories: GI0008 is an island category that includes the islands GI0008A-GI0008D, which are variants of GI0008 that differ in their contents. Changes in the presence or absence of some islands or parts of islands and some pseudogenes represented repetitive, independent gain or loss (homoplasies) in different branches.

A list of all genomic islands is present in ParaC Pan-genome 7, and the branches on which these events occurred is illustrated in ParaC Pan-genome 1. The cumulative numbers of pseudogenes per branch can also be seen in Figure S3 C.

ParaC Pan-genome 4 provides a list of all genes and pseudogenes for each genome.