Introduction

Ammonia-oxidising archaea of the phylum Thaumarchaeota are important organisms in the nitrogen cycle, but the mechanisms driving their radiation into diverse ecosystems remain underexplored. Here, existing thaumarchaeotal genomes were complemented with 12 novel genomes reconstructed from a large co-assembly of river sediment metagenomic sequences and belonging to the previously under-sampled Nitrososphaerales order. Two lateral gene transfer (LGT) events were identified into early ancestors of the Nitrososphaerales and the fate of these gene families was highly lineage-specific, being lost in some descendant lineages, but undergoing extensive duplication in others, suggesting niche-specific roles.

Methods

12 new Thaumarchaeota genomes were co-assembled from river sediment metagenomes. A robust species tree was estimated for 152 Thaumarchaeota genomes using newly established tree comparison techniques and was probabilistically reconciled against 5683 individual Thaumarchaeota gene family trees, enabling the prediction of gene family histories and the probabilistic reconstruction ancestral genomes.

Results

The new genomes presented in this work are the first genomes for three of the six known Nitrososphaerales families. The new robust phylogeny was used to establish a uniform taxonomic stratification for the phylum.

While many of the genes acquired in early stages of Nitrososphaerales evolution are uncharacterised, several genes involved in molybdoenzymes synthesis, carbohydrate degradation and stress response mechanisms, such as osmoprotectant synthesis and cation osmoregulation, could be identified.

Results cont.

Extensive gene family origination in early ancestors of Nitrososphaerales (NS LCA and NS-2 LCA in Figure 3) was followed by a high level of gene duplication and loss in their descendants. While gene duplication was numerically the greatest contributor to the larger proteome size observed in some Nitrososphaerales, gene family origination introduced extensive proteome novelty.

Gene families that were acquired in two early ancestors of Nitrososphaerales had different fates, undergoing extensive duplications in some lineages and extensive losses in others. This duplication of laterally acquired genes, as well as ancestral genes, led to genome expansion in the NS-Gamma, -Beta and -Epsilon families of Nitrososphaerales.

Conclusions

- This work improves genome representation of Nitrososphaerales at the family level.
- Nitrososphaerales have acquired metabolic and stress response genes, likely aiding the colonization of terrestrial and sediment environments.
- Nitrososphaerales evolution was driven by early LGT events followed by lineage-specific gene duplication-mediated genome expansion.


Acknowledgments. This work and P.O.S. were financially supported by UKRI through the NERC grant NE/R001529/1. In addition, C.G.-R. and T.A.W. were both supported by a Royal Society University Research Fellowship (URF150571 and UF140626). C.Q. was funded through an MRC fellowship (MR/M50161X/1) as part of the Cloud Infrastructure for Microbial Genomics (CLIMB) consortium (MR/M015080/1). S.R. was funded through the BBRC grant BB/R005171/1. The Thames Metagenome Database was funded through the NERC grants NE/M011674/1, NE/M011259/1 and NE/M01133X/1.

Email: p.sheridan@abdn.ac.uk
Twitter: @Paul_Sheridan