Prokaryotic-like horizontal gene transfer and eukaryotic-like gene duplication shape Asgard archaeal genomes

Saioa Manzano-Morales[1] [2], Toni Gabaldón[1] [2] [3] [4]
[1]Barcelona Supercomputing Centre (BSC-CNS), Plaça Eusebi Güell, 1-3, 08034, Barcelona, Spain
[2]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028, Barcelona, Spain
[3]Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain.
[4]Centro de Investigación Biomédica En Red de Enfermedades Infecciosas (CIBERINFEC), Barcelona, Spain

INTRODUCTION

Asgard archaea are the **closest prokaryotic relatives of eukaryotes** from within which the latter emerged, and therefore key to our understanding of eukaryogenesis. Asgard archaea have a markedly larger genome size than other archaea (Figure 1); however, we know little about how **genome content** has been shaped across asgard evolution, in particular the role of **inter-domain Horizontal Gene Transfer.**

METHODS

Here we assessed duplication, loss and transfer dynamics over **31 sequenced asgard species,** spanning all **11 major lineages**, the 2 cultured isolates (*Ca.* L. sp. B-35 and *Ca.* P. syntrophicum MK-D1) combining a metaphylome (Fig.1) approach and AleRax.We tthen performed an **in-depth analysis of inter-domain HGT** (Fig.2) by combining screening with HGTector and a phylogenetic approach.

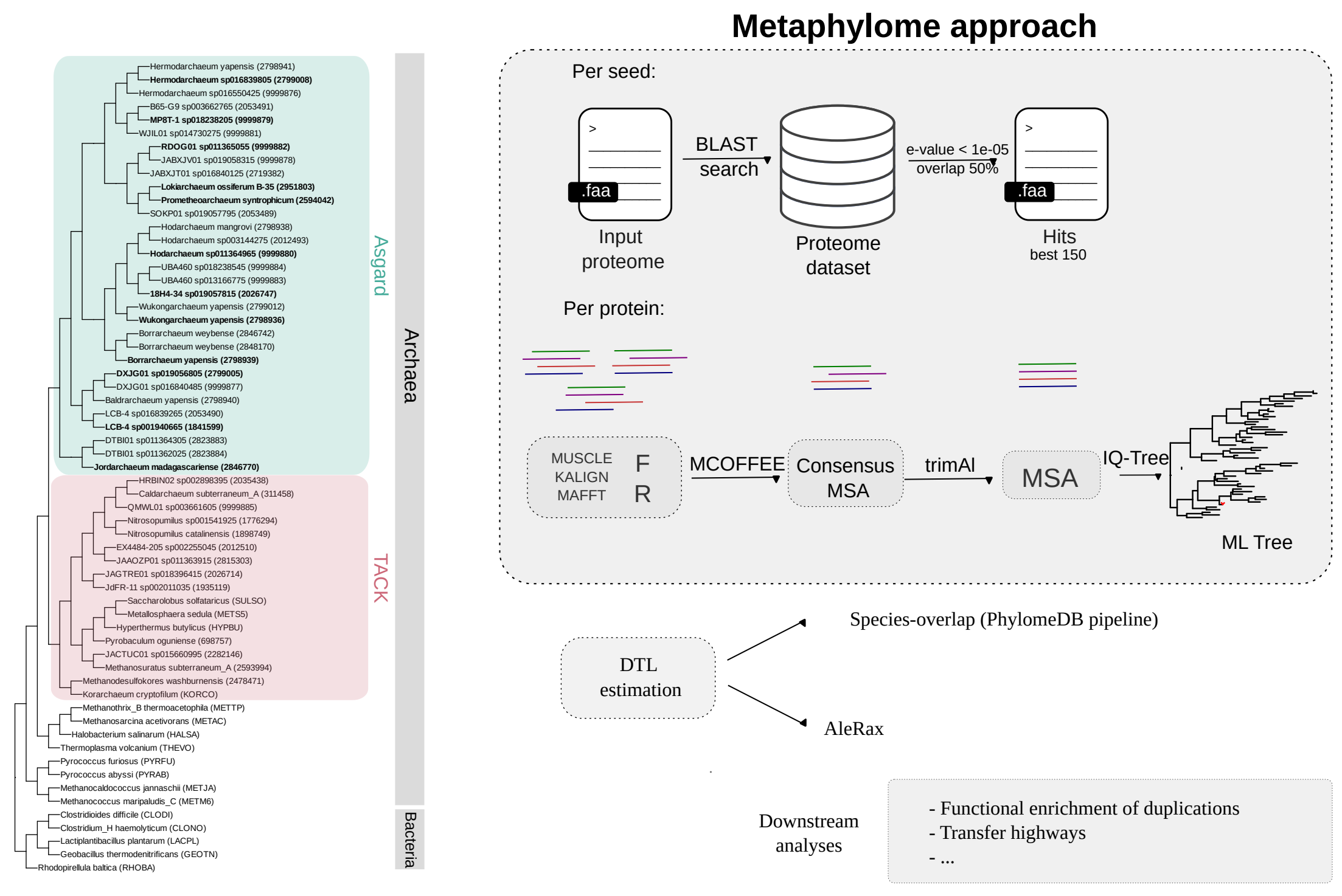

Figure 1: Workflow for DTL assessment.

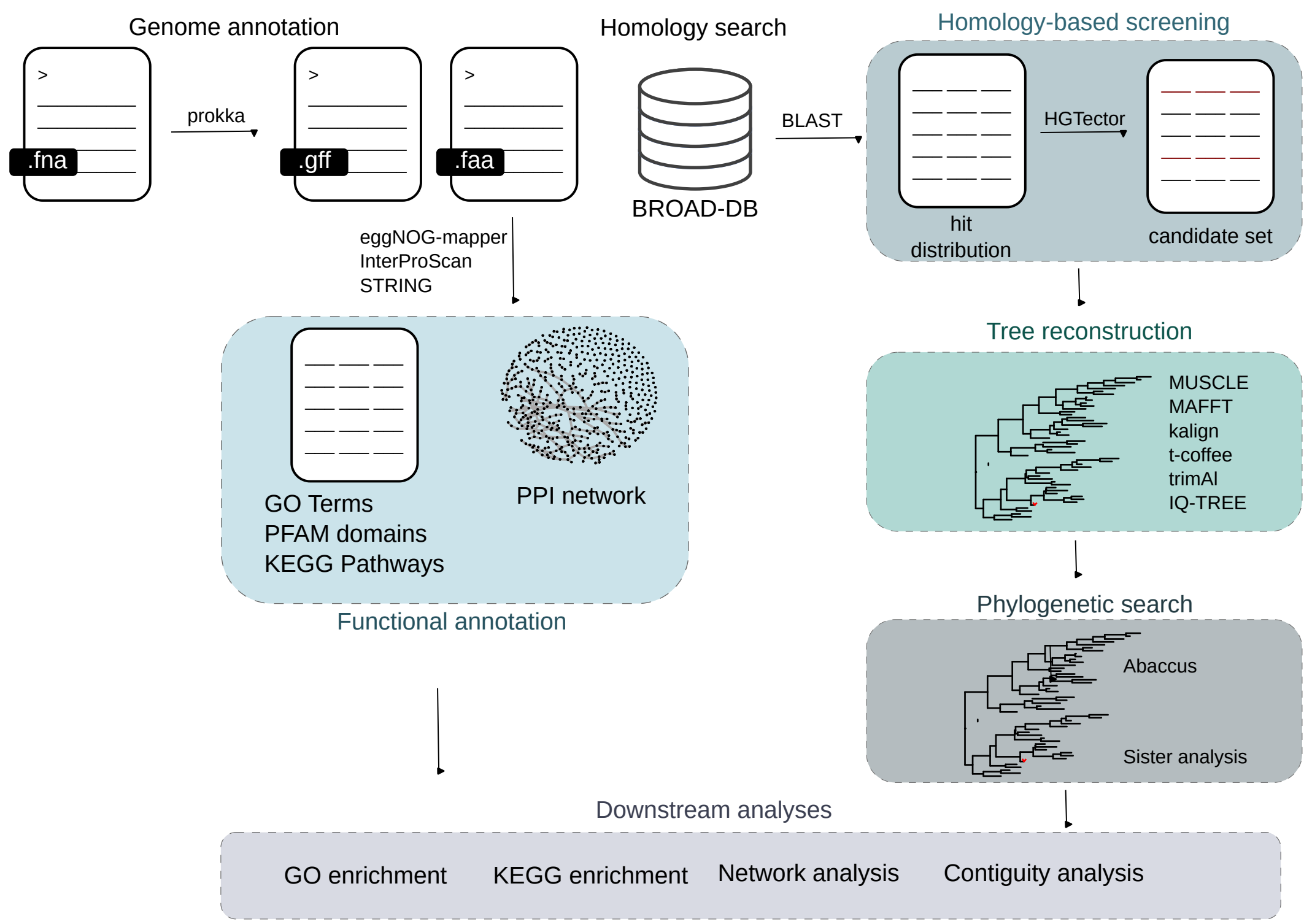

Figure 2: Workflow for inter-domain HGT analysis.

GAIN-LOSS DYNAMICS

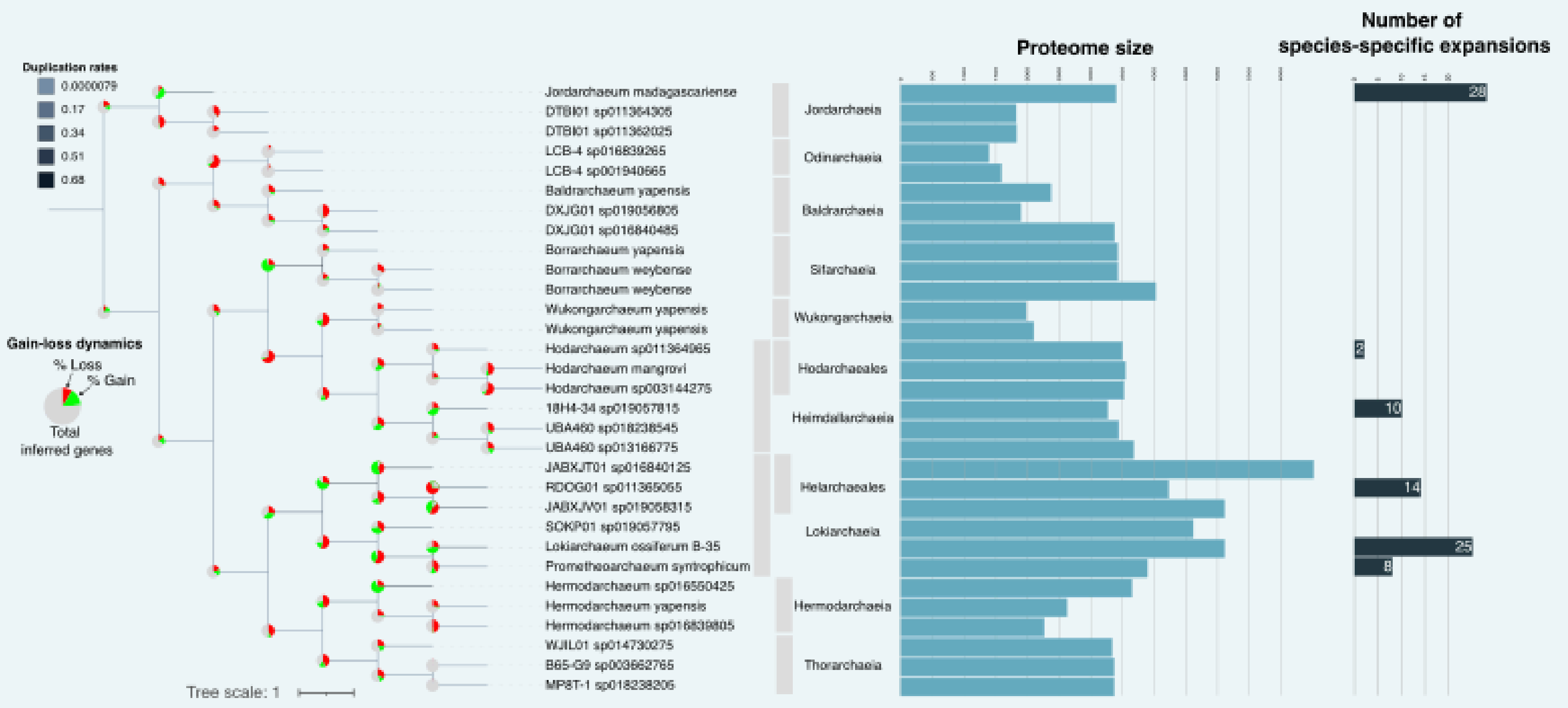

Figure 3: Gene gain-loss dynamics across Asgard archaea. Mean duplication rate at each internal node, counted per seed phylome as the ratio between the number of duplications by the total number of gene trees that contain that node. Total protein count per proteome and species-specific duplications displayed by color scale in terminal nodes. Gain percentage (green) is the sum of duplications and originations, relative to the total number of gene trees in that node, and gain loss (red) is the percentage of losses relative to total inferred gene trees

An analysis of gene gain-loss dynamics across Asgard-archaea (Fig.3) reveals **lineage-dependent dynamics** that reflect in their proteome size, and their number of species-specific expansions. These expansions are enriched in GO terms related to signal transduction systems; whereas gene duplications are enriched in transposases and transposon activities, hinting at a mechanism for the gene duplication, as well as lineage-dependent processes.

Asgard archaea have, on average, significantly higher duplication rates than their sister clade Thermoproteota (Fig.4A) despite comparable rates of HGT (Fig. 4B). This points to a significant role of gene duplication in the evolution of genome size in the Asgard clade.

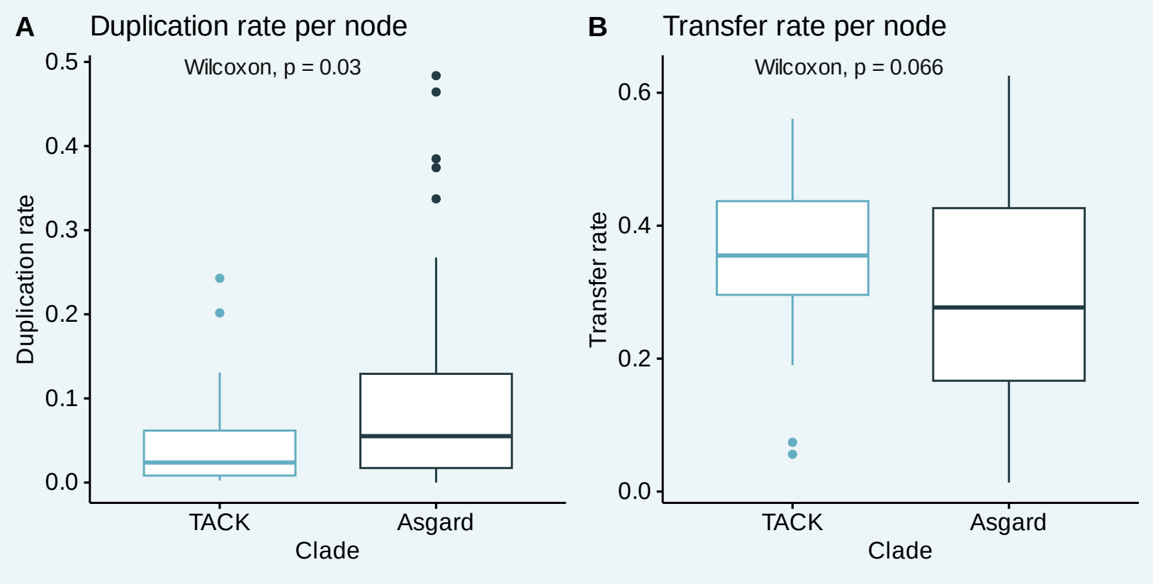

Figure 4: Duplication rate (A) and transfer rate (B) per node for Asgard and TACK archaea. Duplication/Transfer rate (number of duplications/transfers normalized by the number of genes inferred to be present in the node) for nodes stemming from the MRCA of Asgard and TACK archaea

INTER-DOMAIN HGT

HGT has been a **continuous (and on-going) process** across Asgard evolution as seen by the array of acceptor clades (Fig. 5), involving a **large number of donor phyla**, which tend to correlate with co-occurrence patterns from metagenomic samples.

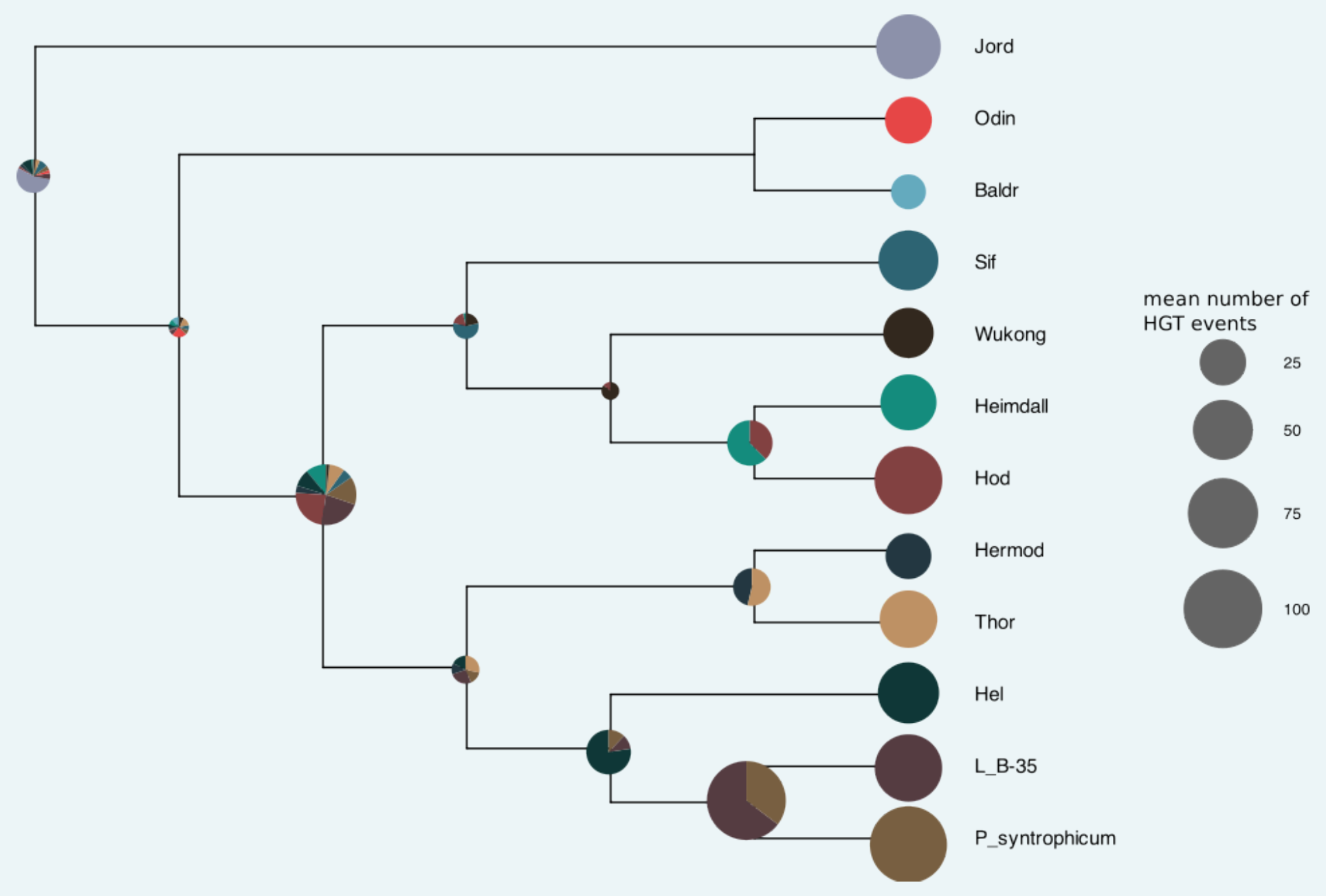

Figure 5: Acceptor analysis. Dot sizes represent number of HGT gene trees where the acceptor is placed at that given node. Colors represent the lineage used as seed for the tree. The acceptor was defined as the MRCA of the biggest Asgard monophyletic group stemming from the seed protein. Eukaryotic sequences were considered *de facto* Heimdall. Sub-lineage transfers are collapsed to the tip node.

Enriched functions (Fig. 6) deal with **metabolism of basic building blocks** (amino acids and lipids), but the specific pathways and proteins are organism-dependent.

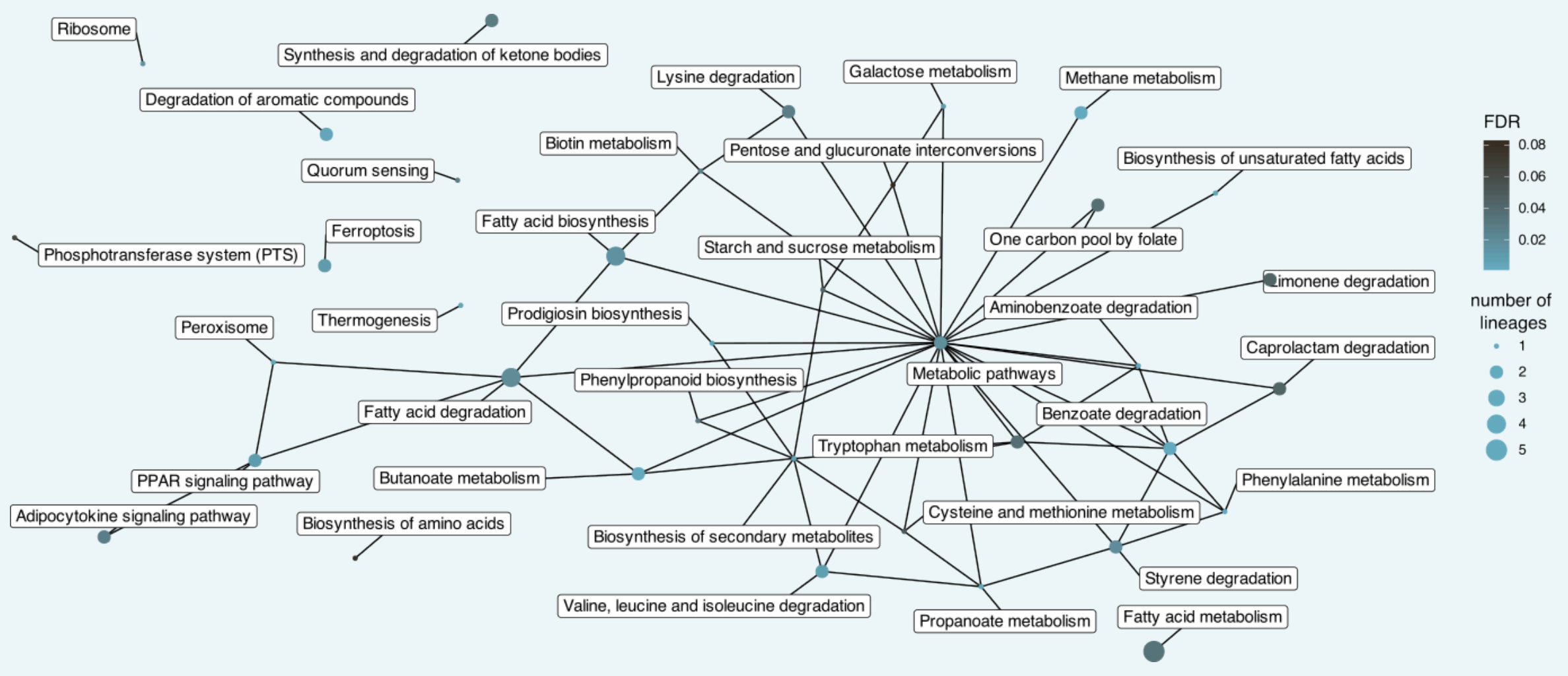

Figure 6: Network of functional relationships between KEGG enriched terms. Nodes are KEGG Terms enriched (FDR <=0.1) in at least one lineage, edges represent functional relationships between KEGG terms.

HGT proteins are well-integrated, if **peripheral** (Fig. 7) and seeminlgly **not modular** but stepwise additions to native pathways.

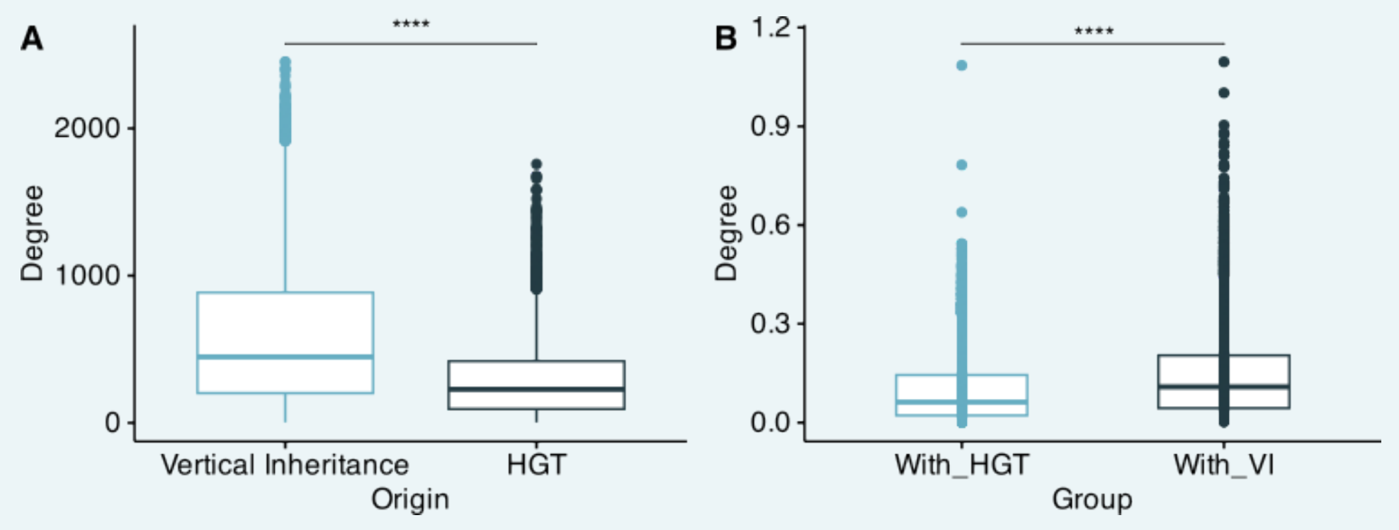

Figure 7: Network properties of HGT proteins within seed organism's interactome. (A) Number of connections (degree) of HGT proteins versus genes of vertical inheritance. (B) Number of connections of HGT proteins with other HGT proteins vs with proteins of vertical inheritance.

CONCLUSIONS

We found that **inter- (Fig.4B) and intra-domain (Fig.5) gene transfer is pervasive** throughout asgard evolution, involving mostly **metabolic genes** (Fig.6) peripheral in interaction networks (Fig.7), and bacterial **transfer partners that co-occur** in present-day metagenomic surveys. However, we found that **transfer levels are not higher than in other prokaryotes (Fig.4B)** , and that **gene duplications (Fig.4A)**, rather than transfers, likely **underlie increased genome sizes** in Asgard archaea. Hence, consistent with their phylogenetic position, asgard genomes showcase a **unique blend of prokaryotic-like transfer-driven and eukaryotic-like duplication-driven evolution.**

BIBLIOGRAPHY

1. Imachi H, Nobu MK, Nakahara N, et al. Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature*. 2020;577(7791):519-525.
2. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature*. 2017;541(7637):353-358.
3. Rodrigues-Oliveira T, Wollweber F, Ponce-Toledo RI, et al. Actin cytoskeleton and complex cell architecture in an Asgard archaeon. *Nature*. 2023;613(7943):332-339.
4. Eme L, Tamarit D, Caceres EF, et al. Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes. *Nature*. 2023;618(7967):992-999. doi:10.1038/s41586-023-06186-2
5. Zhu Q, Kosoy M, Dittmar K. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics*. 2014;15(1):717
6. Naranjo-Ortíz MA, Brock M, Brunke S, Hube B, Marcet-Houben M, Gabaldón T. Widespread Inter- and Intra-Domain Horizontal Gene Transfer of d-Amino Acid Metabolism Enzymes in Eukaryotes. *Front Microbiol*. 2016;7:2001

saioa.manzano@bsc.es
github.com/saimanzano

This research was supported by Gordon and Betty Moore Foundation (Grant GBMF9742)