Adaptation through horizontal gene transfer in the cryptoendolithic red alga *Galdieria phlegrea*

Huan Qiu1, Dana C. Price1, Andreas P. M. Weber2, Valérie Reeb3, Eun Chan Yang4, Jun Mo Lee5, Su Yeon Kim6, Hwan Su Yoon5,* Andreas P. M. Weber2, Valérie Reeb3, Huan Qiu1, Dana C. Price1

**Correspondences**

Thriving in the hot, acidic, and metal-rich environments associated with geothermal areas is possible for only a few eukaryotes, with the Cyanidiophyta red algae (*Cyanidium, Galdieria, and Cyanidoschyzon*) being a famous example. These unicellular taxa can live in pH 0–4 and temperatures reaching up to 56°C [1,2]. Because *Cyanidiophyta* is sister to a vast array of mesophilic red algae (the Rhodophyta), such as the unicellular *Porphyridium* and the seaweed *Chondrus* [3], the genetic basis of their adaptation to extreme environments is of great interest from both the perspective of biotechnology and of evolution. The recently completed 13.7 Mbp genome sequence from the hot-spring dwelling *Galdieria sulphuraria* demonstrated that horizontal gene transfer (HGT) from prokaryotic sources provided this taxon with remarkable metabolic versatility (e.g., glycerol metabolism) and the ability to survive in its hostile environment (e.g., genes to detoxify mercury and arsenic) [4]. To explore the role of HGT in other members of this genus, we generated an 11.4 Mbp draft genome assembly from the sister taxon *G. phlegrea* DBV 009 [5]. In contrast to *G. sulphuraria*, this species is adapted to dry habitats near fumaroles such as fissures between rocks or cryptoendolithic environments [5,6]. Here, we provide evidence for extensive gene loss in the common ancestor of *Cyanidiophyta* that includes the eukaryote-derived loci required for urea utilization. Surprisingly, we find that *G. phlegrea* has regained the complete set of genes required for urea hydrolysis through HGT from eubacteria. The unlinked nature of these genes is likely explained by multiple gene transfers that resulted in assembly of the pathway in *G. phlegrea*. Our study demonstrates that genome reduction, a common outcome in eukaryotes for adaptation to a specialized niche, can be ameliorated by the gain of once lost, or novel functions through HGT.

Protein divergence between the two *Galdieria* taxa is similar to that between human and teleost fishes (Figure 1A) and about twice that between the unicellular green alga *Chlamydomonas reinhardtii* and its multicellular sister Volvox carteri (Figure S1A, see Supplemental Information). This suggests that genome-wide impacts of adaptation to different environments should be discernible in the *Galdieria* genomes. Comparison of genome data from Cyanidiophyta with complete (*Porphyridium purpurescens* [7]) or partial (*Calliarthus tuberculatus* and *Porphyra umbilicalis*) genome data from mesophilic Rhodophyta shows that of the 6801 orthologous gene families present in the most recent common ancestor of Rhodophyta (Figure 1B), 1448 were lost in the Cyanidiophyta common ancestor, compared to 456 in Rhodophyta. This pattern is most evident in *Cyanidoschyzon merolae* that underwent 1312 additional gene losses resulting in a gene-poor lineage (4775 nuclear protein coding genes) [8] restricted to aqueous environments [9].

Given severe genome reduction, we postulated that HGT provides a mechanism to gain adaptive functions in *G. phlegrea*. Consistent with this idea, BLASTp and phylogenetic analysis of the *G. phlegrea* predicted proteins turned up 11 instances of prokaryote-derived HGT unique to this taxon within Cyanidiophyta (Figures S1B–S1G). These transferred genes include seven urease genes (i.e., *UreA–G*; Figures S1C–S1G) that encode all components of urease, a multi-subunit enzyme that catalyzes the hydrolysis of urea into carbon dioxide and ammonia. This reaction provides an alternative source of nitrogen in N-limited environments. In eubacteria, the seven genes are typically encoded in a single operon (Figure 1C). However, in *G. phlegrea*, urease genes are located on independent DNA contigs except for the linked *UreB* and *UreC* (Figure S1C). Phylogenetic analysis shows that all urease genes have an eubacterial origin and are flanked by eukaryotic or red algal-specific genes, with the exception of *UreD*, which is flanked by a eukaryotic gene only in the upstream region (Figures S1C–S1G). An example of this gene structure is *UreA* located on contig 977 along with several genes that show micro-synteny with *G. sulphuraria* and are of eukaryotic provenance (Figure 1D; Figure S1H). The position of *G. phlegrea* *UreA* in the *G. sulphuraria* genome is occupied by a gene encoding formamidase, an enzyme with a function related to nitrogen production. The *G. phlegrea* formamidase homolog has been moved to another genomic region (contig 1594, Figure S1I). These results suggest that the seven genes required for the complete urease metabolic pathway have originated in *G. phlegrea* through HGT.

More striking is that in *P. purpurescens* and green algae/plants, all urease genes (except *UreE*) have a eukaryotic origin (Figure 1C,D; Figures S1C–S1H). Furthermore, *UreA*, B, and C are a single gene (*UreABC*) (Figure 1C) that likely arose through a gene fusion in the Plantae ancestor. None of these genes are present in Cyanidiophyta except for a second *UreE* gene in *G. sulphuraria* (Figure S1E) shared with *G. phlegrea* that likely resulted from an independent bacterium-derived HGT event in *Galdieria* species. Because *G. sulphuraria* does not show evidence of massive gene loss, we interpret these results as indicating that *G. phlegrea* underwent lineage-specific HGTs, rather than complete loss of urease genes in *G. sulphuraria*. In contrast to *G. sulphuraria* and *C. merolae*, *G. phlegrea* grows in less extreme habitats (i.e., moderate pH and temperature inside of rock) [5]. Therefore, the presence of the urease pathway in *G. phlegrea* is likely to be an independent reacquisition of an ecologically important trait still present in Plantae. Extensive genome reduction is generally thought to represent a one-way, irreversible process; for example, in intracellular *Buchnera* [10]. Our results suggest that HGT provides a solution to the intractable problem of gene loss in extremophiles. In summary, although *G. sulphuraria* and *G. phlegrea* share a significant common gene pool, the eubacterium-derived HGTs in the latter provide not only the means to distinguish these two lineages, but more generally, to demonstrate that
Figure 1. Genome analysis of Cyanidiophytina.

(A) Protein sequence divergence within Cyanidiophytina (with G. phlegrea as the reference) in comparison to those within Metazoa (with human as the reference). (B) Orthologous gene family evolution in Cyanidiophytina with numbers at the branches indicating gene family gain (+) and loss (-). Two instances of massive gene loss are indicated with the thick arrows and the asterisk indicates the possible timing of urease gene losses. (C) Schematic representation of the urease operon in eubacteria and in different eukaryotes. Arrows indicate the presence of urease genes with black for bacterial (-derived) genes and olive-brown for genes that are descended from the Plantae ancestor. The ‘X’ indicates absence in different genomes. (D) Micro-synteny between the UreA gene-encoding contig 977 in G. phlegrea and the homologous region in G. sulphuraria (Stig28). Genes are represented using thick bars with the orientations indicated.

Like bacteria, microalgae may rely on foreign gene acquisition to adapt to changing and stressful environmental conditions.

Accession Numbers
The genome data, assembly, protein predictions, and annotations are available at http://cyanophora.rutgers.edu/gphlegrea/. Coordinates for the draft G. phlegrea genome have been deposited at the NCBI Sequence Read Archive (SRA) under the accession code SRR953992.

Supplemental Information
Supplemental Information includes experimental procedures, one figure, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.08.046.

Acknowledgements
This research was supported by a grant from the National Science Foundation awarded to D.B. and H.S.Y. (0936884 and 1317114) and the National Science Foundation awarded to D.B. and H.S.Y. (PJ009525) to H.S.Y. D.B. acknowledges generous support from the University of Iowa to produce the G. phlegrea draft genome data. We thank Claudia Ciniglia (Second University of Naples) for providing the G. phlegrea culture.

References

*Department of Ecology, Evolution, and Natural Resources, Rutgers University, New Brunswick, New Jersey 08854, USA. 1Institute for Plant Biotechnology, Heinrich-Heine-University, D-40225 Duesseldorf, Germany. 2Center for Genomics, University of Iowa, Iowa City, Iowa 52242, USA. 3Marine Biotechnology Research Center, Korea Institute of Ocean Science & Technology 787, Haean-ro, Sanggok-gu, Ansan, Gyeonggi-do, KOREA. 4Department of Biological Sciences, Sungkyunkwan University, Suwon 440-746, Korea. 5Department of Biology, Chungnam National University, Daejeon 305-764, Korea.

*E-mail: bhattacharya@aesop.rutgers.edu, hsyoon2011@skku.edu