

IN BRIEF

ENVIRONMENTAL MICROBIOLOGY**EnvDB, a database for describing the environmental distribution of prokaryotic taxa**

Pignatelli, M., Moya, A. & Tamames, J. *Environ. Microbiol. Rep.* **1**, 191–197 (2009)

Metagenomic sequencing allows researchers to generate vast amounts of data regarding the microbial genome sequences that are present in environments such as soils or ocean waters, yet the ability of researchers to effectively use these data can be hampered by a lack of information on the specific environmental conditions associated with different samples. Writing in the new journal *Environmental Microbiology Reports*, Miguel Pignatelli, Andrés Moya and Javier Tamames now describe a new database, EnvDB, which brings together all the 16S ribosomal DNA sequences available in GenBank and classifies them according to environment. EnvDB will be updated weekly and is available at <http://metagenomics.uv.es/envDB>. In the future, Pignatelli *et al.* hope to be able to include additional environmental, geographical and physicochemical details.

MICROBIAL ECOLOGY**Biogeography of the *Sulfolobus islandicus* pan-genome**

Reno, M. L. *et al. Proc. Natl Acad. Sci. USA* **106**, 8605–8610 (2009)

The pan-genome constitutes the global gene repertoire of a species, and is the sum of the core genome, which contains genes that are present in all strains, and the variable genome, which contains genes that can be present or absent in different species. In this study, Reno and colleagues undertook a population-scale analysis of the genomes of seven isolates of the thermoacidophilic archaeon *Sulfolobus islandicus* from three biogeographically isolated geothermal locations. Analysis of the variable genome found evidence that the differences in variable gene content were driven by geographical isolation, in contrast to previous studies which indicated that such differences are driven by environmental selection. The authors went on to look at the history of gene acquisition and loss, and found that most genomic variation observed can be accounted for by the movement of mobile genetic elements, such as viruses and plasmids. They were also able to calculate that the average rate of net gene gain by *S. islandicus* populations is accelerated compared with the rates that have been calculated for some bacterial species.

FUNGAL GENOMICS**Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes**

Butler, G. *et al. Nature* **459**, 657–662 (2009)

The first *Candida* genome sequence — that of *Candida albicans* SC5314 — was published in 2004. Five years on, an international consortium of researchers, writing in *Nature*, now report on the completion of six new *Candida* genome sequences that encompass a range of pathogenic and mating phenotypes: *Candida albicans* WO-1, *Candida tropicalis*, *Candida parapsilosis*, *Lodderomyces elongisporus*, *Candida guilliermondii* and *Candida lusitaniae*. They present a major comparative genome analysis of these new sequences along with the sequences of two other *Candida* species, *C. albicans* SC5314 and *Debaryomyces hansenii*. Highlights of the study include the identification of many new *Candida* genes and the discovery of unanticipated and intriguing differences in the mating and meiosis pathways.