Genome Biology

Meeting report

Unraveling the genomic diversity of small eukaryotes

Gilles Fischer*, Dawn Thompson†, Jennifer Russo Wortman‡ and Cécile Fairhead§

Addresses: *Microorganism Genomics, UPMC/CNRS FRE3214, Paris, France. †Broad Institute, Cambridge, MA 02139, USA. ‡Institute for Genome Sciences, University of Maryland, School of Medicine, Baltimore, MD 21201, USA. §Institut de Génétique et Microbiologie, Université Paris Sud 11, CNRS UMR8621, Orsay, France.

Correspondence: Cécile Fairhead. Email: cecile.fairhead@u-psud.fr

Abstract

A report of the meeting Comparative Genomics of Eukaryotic Microorganisms, 17-22 October 2009, San Feliu de Guixols, Spain.

The first meeting in a new series of EMBO meetings aimed at bringing together those working on genome-enabled research encompassing the great diversity of eukaryotic microorganisms was held recently in Spain. New technology such as high-throughput sequencing now allows less well studied eukaryotic microbes to come into the limelight, providing some fascinating glimpses into the eukaryotic world that lies outside multicellular plants and animals. Some of the highlights of the meeting are presented here.

In celebration of the 150th anniversary of the publication of Charles Darwin's *On the Origin of Species*, the meeting opened with 'The Darwin Lecture' delivered by Bernard Dujon (Institut Pasteur, Paris, France), who highlighted the fact that our current knowledge of eukaryotic genomes is highly biased towards the two supergroups of Unikonts (the animals and fungi) and Plantae (red and green algae and plants) and to a lesser extent towards the Chromalveolates (ciliates, brown algae, diatoms, and dinoflagellates, for example). Dujon pointed out the need for more genomics data representing the other two eukaryotic supergroups - the Excavata (which contains some important parasites of humans) and the Rhizaria (the pseudopodial amoeboids) - to foster a clearer understanding of eukaryotic diversity. A spectacular illustration of this diversity was presented by Gertraud Burger (University of Montreal, Canada), whose research on the mitochondrial genome of diplonemids (members of the Excavates) has shown that each core gene is split into several small modules scattered across multiple circular chromosomes. The concatenation of these modules occurs at the RNA level via an unusual *trans*-splicing mechanism, along with RNA editing at the junctions, in order to reconstruct a full transcript of each gene.

The vast majority of eukaryotic taxa are composed of unicellular organisms that, with the exception of yeasts, have largely been overlooked until recent advances in genomic technology. These eukaryotic microbes represent human and plant pathogens, as well as species of industrial and agricultural importance and key model organisms.

Sex and hybridization in yeasts and other fungi

Sexual reproduction is almost universal in eukaryotes and the diverse mating systems of fungi provide insights into its evolution. Joseph Heitman (Duke University, Durham, USA) raised the issue of the adaptive importance of the sexual cycle - indeed, asexual or rarely sexual species seem to be evolutionary dead ends. As he noted, one obvious benefit of sex is to increase genetic diversity. However, many fungal species have developed sophisticated mechanisms for selfing that, paradoxically, provide no opportunity for genetic exchange. In *Aspergillus nidulans*, for example, it is known that diploidization during the parasexual cycle allows the accumulation of mutations that individually incur fitness cost in haploids but are potentially adaptive when combined in segregants. Similarly, Heitman has shown that unisexual reproduction in *Candida albicans* and *Cryptococcus neoformans* can generate genetic diversity *de novo*.

Another route to increase genetic diversity is through interspecific hybridization events. Investigating hybrids of *Saccharomyces cerevisiae*/*Saccharomyces bayanus*, Daniela Delneri (University of Manchester, UK) and colleagues used Tap-tag complex purification combined with mass spectrometry identification of partners to show that about 60% of the proteins participate in the formation of hybrid complexes. Hybrids between *Saccharomyces* species are more fit than either of the parental species, and there is evidence that this is due to the trans-species protein complexes. Amparo Querol (Institute of Agrochemistry and Food Technology, Valencia, Spain) described a study combining transcriptomic and enzyme activity analysis to decipher the molecular basis of favorable enological (wine-making) traits in *S. cerevisiae* and *Saccharomyces kudriavzevii* hybrids, such as the ability to ferment at lower temperatures, increased glycerol production and lower ethanol yield. In a complementary

example, Sylvie Dequin (INRA, Montpellier, France) reported that a large piece of DNA originating from a major wine-contaminating yeast had introgressed into an enological strain of *S. cerevisiae* during the course of its adaptation to the wine ecosystem.

New technologies such as genome-wide proteomic profiling and metabolomics by mass spectrometry, in conjunction with high-throughput parallel sequencing to delineate transcriptome structure, is enabling an integrated approach to investigating the physiology of the eukaryotic cell through functional and comparative genomics. Dawn Thompson (Broad Institute, Cambridge, USA) described the use of an integrated transcriptomic and metabolomic approach to make a large-scale parallel reconstruction of the evolution of genetic networks in 15 yeast species. This now enables long-standing questions to be addressed, such as how central carbon metabolism is regulated in fungal species with different lifestyles. For example, respiro-fermentation (the fermentation of glucose in the presence of oxygen) has evolved at least twice in this phylogeny, and in both instances there has been convergent regulatory rewiring to repress mitochondrial functions in the presence of glucose.

The emerging field of population genomics represents a powerful tool to dissect the genetic basis of regulatory variation underlying natural phenotypic diversity. Francisco Cubillos (University of Nottingham, UK) focused on the results of mapping quantitative trait loci (QTL) in crosses of four strains representing distinct lineages among the natural isolates of *S. cerevisiae* sequenced as part of the *Saccharomyces* Genome Resequencing Project. All segregants were extensively phenotyped under several conditions to identify the major QTLs, with those affecting high-temperature growth and resistance to sodium arsenite being detected in the greatest numbers.

Protist genomes and metagenomes

Several presentations sought to further our understanding of a major transition in evolutionary history - the origin of multicellularity. This aim is being supported by UNICORN (UNICellular Opisthokonts Research iNitiative), an international genome project targeting ten unicellular relatives of multicellular animals and fungi [http://www.broadinstitute.org/annotation/genome/multicellularity_project/MultiHome.html]. The Opisthokonts comprise the Metazoa and their unicellular relatives, the Fungi and their unicellular relatives and several independent lineages of unicellular microorganisms. Communications from Nicole King (University of California, Berkeley, USA), Franz Lang (University of Montreal, Canada) and Inaki Ruiz-Trillo (University of Barcelona, Spain) highlighted studies on the evolutionary significance of the Choanoflagellata, the Nuclearia and the Blastocladiomycota that support the importance of this focused effort. King reported work with colleagues that has detected a diverse array of protein domains (cadherins and lectins) encoded in the genome of the choanoflagellate *Monosiga brevicollis*. Such domains are involved in cell adhesion, cell contact or cell sorting functions that were previously thought to be restricted to metazoans. By analyzing the genome sequence of the nuclearian *Capsaspora owczarzaki*, which belongs to an independent opisthokont lineage, Ruiz-Trillo and colleagues have confirmed that some of the genes involved in cell signaling and adhesion in metazoans were already present in the ancestor of metazoans. Lang reported on work with colleagues on the zoosporic fungus-like *Allomyces macrogynus* (a member of the Blastocladiomycota, which produce uniflagellate spores resembling animal sperm) that aims to reveal genes present in our opisthokont ancestor more than a billion years ago. This genome contains more than 17,500 genes and is very rich in gene and segmental duplications. The origin of the Metazoa is a landmark event in the history of life and we are now taking critical steps in the understanding of this major transition.

The oomycetes are another group formerly considered as fungi, but now classified within the Chromalveolates. Sophien Kamoun (Sainsbury Laboratory, Norwich, UK) described the 240-Mb genome of the oomycete *Phytophthora infestans*, the agent of potato blight, which contains 74% repeated sequences. Oomycete plant pathogens secrete inhibitor proteins (called effectors) that carry conserved motifs downstream to the signal peptide, which mediate delivery inside host plant cells. These effectors interfere with the activity of extracellular plant hydrolases and in some cases suppress plant immunity. Kamoun reported studies with colleagues that effectors involved in pathogenicity in *P. infestans* are encoded by genes within repeat-containing regions. This genomic organization allows repeat-driven expansion and rapid evolution of these effectors, pointing to a mechanism for host jumps by oomycetes, as described by Marco Thines (University of Hohenheim, Germany). His talk highlighted insights gleaned from the study of several species of the clade *Hyaloperonospora*, including two closely related species that exist on unrelated host species of the plant family Brassicaceae, while a third has made a host jump across plant families and is a parasite of members of the family Resedaceae.

The sequences of metagenomes - the total DNA of microbial populations of environmental samples - allows access to the wealth of diversity represented by unculturable microbial species. Early metagenomics projects focused almost exclusively on the dominant bacterial populations of microbial communities, but there are now a number of initiatives sampling the eukaryotic microbes that are important in these environments. Roland Marmeisse (CNRS, University of Lyon, France)

described a metatranscriptomic approach to investigating organic matter degradation and adaptation of microbial communities to heavy-metal contamination in forest soils. To detect eukaryotic, rather than bacterial, transcripts, polyadenylated mRNA is purified from the environmental sample, and is coupled with functional screening in yeast. Patrick Wincker (Institut de Genomique du CEA, Evry, France) presented the metagenomics platform being developed by the Tara-Oceans expedition [http://oceans.taraexpeditions.org/?id_page=1] to sample protist diversity in the oceans, which will link reference genome sequencing, 18S diversity estimates and whole-genome shotgun sequencing, along with imaging of microorganisms.

High-throughput parallel sequencing is facilitating a new era in eukaryotic population genomics, and several talks described current and planned projects in this area. The amount of sequence data generated and the computational analyses required will necessitate increasingly sophisticated bioinformatics resources. Present bioinformatics capabilities in comparative genomics were highlighted in several talks, including those of Jason Stajich (University of California, Riverside, USA) and Toni Gabaldón (CRG, Barcelona, Spain). Stajich presented work that is complementary to the projects in UNICORN, showing how the systematic comparative analysis of core and accessory genes of chytrids can shed light on the origin of complex eukaryotic features such as the centriole/spindle pole body functional homologs [http://fungalgenomes.org]. Gabaldón presented new bioinformatics methods and analysis pipelines (integrated series of methods) to decipher 'phylomes' - a newly coined term for the 'forests' of phylogenies of genes derived from comparative genomics. Among these tools are a database to explore phylomes [http://phylomeDB.org] and the Environment for Tree Exploration [http://ete.cgenomics.org], a software tool that, among other things, allows the scanning of large phylogenetic collections to derive orthology and paralogy relationships. Using such tools, the conflicts between gene trees and species trees in the context of the fungal kingdom have been explored.

This meeting was unique in gathering together a research community involved in the genome-based exploration of eukaryotic diversity. It made an ideal setting to let specialists from distinct fields, studying diverse life forms, share thoughts and data on the unifying theme of comparative and functional genomics and evolution of species. We look forward to the next meeting in this series in 2011.

Published: 22 December 2009
doi:10.1186/gb-2009-10-12-318