Preface

Treasure Hunting in the Chlamydomonas Genome

Olivier Vallon* and Susan Dutcher†,[1]

**Institut de Biologie Physico-Chimique, UMR 7141 CNRS/UPMC, 13 75005 Paris, France and †Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110*

THE draft genome sequence of *Chlamydomonas reinhardtii* attracts an expanding community of researchers to the mysteries hidden in the ~110 million bases that constitute the alga's genome. The Chlamydomonas genome sequence did not materialize overnight—it had been a dream of the community long before it became a reality. After the foundation was laid by insightful forerunners that included Pete Lefebvre and Carolyn Silflow, who constructed a BAC library and found useful polymorphic strains (Gross *et al.* 1988; Kathir *et al.* 2003), the Kazusa Institute, which generated the first set of ESTs for Chlamydomonas (Asamizu *et al.* 1999), and one of us (S.K.D), who trained the gene-finding algorithm Genie for analyzing the first large-scale sequence of the nuclear genome (Li *et al.* 2003), it took the energy of Arthur Grossman to really get things moving. He launched the Chlamydomonas Genome Project with funding from the National Science Foundation that marshaled the enthusiasm of several Chlamydomonas laboratories to produce a large EST collection, the first cDNA microarray, the chloroplast genome sequence, and a more complete set of nuclear markers.

Most importantly, Arthur Grossman also managed to convince Dan Rokhsar at the Department of Energy's Joint Genome Institute (JGI) to sequence the nuclear genome. The sequencing in itself was no small feat; JGI's sequencing team toiled and suffered more than they anticipated in producing the first assembly in 2002. The latest assembly, version 4, was released in spring 2008. This assembly has more than one-half of the genome in seven scaffolds >6.6 Mb. Chlamydomonas has 17 genetic linkage groups (Dutcher *et al.* 1991), and several of them appear to be fully assembled. Further good news is that 92.5% of the sequence has been obtained, bringing the bad news that 7.5% remains undetermined. This results in quite a few holes, with many precious genes likely missing.

[1]*Corresponding author:* Department of Genetics, Washington University School of Medicine, Box 8232, 4566 Scott Ave., St. Louis, MO 63110. E-mail: dutcher@genetics.wustl.edu

Annotation of an intron-rich genome that has many repeats and transposons is a difficult task that demands multiple methods. JGI's annotation team led by Igor Grigoriev generated many gene models that provided the starting point for manual annotation carried out by a team of >100 experts. The result is a catalog of ~15,000 genes, of which 5651 are manually curated, which speaks to the commitment of the Chlamydomonas research community in promoting the best possible use of the genome information. The JGI is also determining the sequence of the genome of a related, multicellular green alga, *Volvox carteri* (Kirk 2005). The annotation of the Chlamydomonas genome has already greatly benefited from comparison of the two genome sequences.

Well before publication of Merchant *et al.*'s (2007) genome article, the genome sequence was put to use. The proteomic studies of flagella (Pazour *et al.* 2005), basal body (Keller *et al.* 2005), eyespot (Schmidt *et al.* 2006), mitochondria (van Lis *et al.* 2003), and thylakoids (Allmer *et al.* 2006) would not have been possible without JGI's early release of the sequence. Comparative genomics for finding genes for flagella and basal bodies exploited the genome sequence (Li *et al.* 2004). Many articles that describe the gene repertoire have appeared, some already grouped in special issues of Plant Physiology in 2004 and Photosynthesis Research in 2005. This special issue of Genetics is another chapter of the encyclopedia of Chlamydomonas.

The lure of a genome sequence is inescapable: this special issue welcomes many newcomers to the Chlamydomonas field. Nine of the 18 laboratories that contributed to it are publishing their first article using Chlamydomonas as a model organism. A genome sequence coupled to powerful experimental tools makes Chlamydomonas even more attractive as an experimental system. Chlamydomonas has a long and proud genetic tradition. Ralph Lewin isolated the first flagellar mutants and Paul Levine isolated the first photosynthesis mutants in Chlamydomonas (Levine 1960; Ebersold *et al.* 1962). Chloroplast inheritance was demonstrated first in Chlamydomonas largely through the isolation of drug resistance mutants and the power of tetrad analysis

(Sager and Ishida 1963; Sager and Ramanis 1970). Chlamydomonas is the first organism in which all three genomes could be easily transformed (Boynton *et al.* 1988; Kindle *et al.* 1989; Randolph-Anderson *et al.* 1993). The Chlamydomonas Center (http://www.chlamy.org/) maintains a large collection of mutants and plasmids and serves as an invaluable source of exchange and information for the community. Cloning genes on the basis of a mutant phenotype is relatively easy using complementation (Purton and Rochaix 1994) or molecular mapping (Bowers *et al.* 2003; Kathir *et al.* 2003; Rymarquis *et al.* 2005). Reverse genetics is becoming routine using efficient RNA interference (Rohr *et al.* 2004) and insertional mutant collections (Tam and Lefebvre 1995; Dent *et al.* 2005).

These advantages for both fundamental and applied research explain why Chlamydomonas is becoming increasingly more important for systems biology. This special issue contains two articles from the German goFORSYS consortium, which has chosen Chlamydomonas as its model for plant systems biology because of its simplicity relative to Arabidopsis (http://www.goforsys.de/). The first article uses a combination of proteomics, metabolomics, and metabolic modeling that truly improves the description of the genome (May *et al.* 2008, pp. 157–166); the other article describes the suite of transcription factors in Chlamydomonas (Riaño-Pachón *et al.* 2008, pp. 31–39). Another two articles present detailed analyses of Chlamydomonas ESTs (Liang *et al.* 2008, pp. 83–93, and Shen *et al.* 2008, pp. 167–176) that demonstrate how a careful analysis of the sequencing traces can enhance the description of genes and reveal the extent of alternative polyadenylation. These two studies also are part of a broader effort of several laboratories and consortia around the world to take a systems biology approach to the study of Chlamydomonas (http://chlamy.org/abstracts2008/). Chances are that Chlamydomonas research will take a sharp turn toward genomewide and model-building approaches that survey the organism in its entirety.

But for the moment, Chlamydomonas researchers are still busy sorting out individual genes and pathways. A majority of the articles in this special issue focus on descriptions of the Chlamydomonas gene repertoire. They explore important gene families or known biological pathways, revealing interesting variations in Chlamydomonas. The transfer RNA genes are often clustered and cotranscribed (Cognat *et al.* 2008, pp. 113–123), which is unusual for tRNAs. For almost every amino acid, the number of genes for each isoacceptor tRNA parallels codon usage. How does their production match their need in a genome that is 65% GC? Another interesting class of small RNAs is the small nucleolar RNAs that guide methylation or pseudouridylation of ribosomal and spliceosomal RNAs. Chen *et al.* (2008, pp. 21–30) present a detailed analysis of the 322 Chlamydomonas snoRNA genes. Again, clustering predominates, but this time the genes are found within introns of protein-coding genes. It will certainly be interesting to see how their expression correlates with ribosome biogenesis, particularly during nitrogen starvation (Martin *et al.* 1976).

RNA degradation plays a central role in RNA metabolism. Zimmer *et al.* (2008, pp. 125–136) found dozens of RNases and proteins associated with the degradation machinery that localize to all compartments of the cell. This study is complemented by analyses from the Cerutti laboratory on the RNA interference machinery (Casas-Mollano *et al.* 2008, pp. 69–81). Understanding this process is important, not only because it serves as a basis for many reverse genetics studies, but also because miRNA and siRNA abound in Chlamydomonas (Molnar *et al.* 2007; Zhao *et al.* 2007). On the basis of the range of targets, these miRNA are likely to be important regulators of gene expression and of genome stability, as has been discovered in multicellular organisms (Ambros and Chen 2007).

Most chloroplast and mitochondrial proteins are synthesized in the cytosol before being imported into the organelles. Two articles describe the two analogous, but largely nonhomologous, machines that carry out these complex recognition and translocation reactions (Figueroa-Martínez *et al.* 2008, pp. 149–155, and Kalanon *et al.* 2008, pp. 95–112). These studies will help in understanding the signals for sorting a protein to one organelle or the other, or both. Among the most unique plant-specific biochemistry that occurs within the chloroplast is the biosynthesis and degradation of starch. A comparative analysis of several algal genomes suggests a comprehensive model for the evolution of glucan storage in photosynthetic organisms Deschamps *et al.* 2008 (**178:** 2373–2387). Dayer *et al.* (2008, pp. 41–57) have also taken a phylogenetic approach to cataloging the peroxiredoxins and glutathione peroxidases in the Chlamydomonas genome. These enzymes participate in the defense against oxidative stress, which is an unavoidable by-product of photosynthesis. Because iron ions can potentiate oxidative damage, its intracellular concentration and status are tightly controlled. In this issue, the Merchant group expands their description of metal homeostasis in Chlamydomonas by describing two chloroplast-localized ferritin complexes (Long *et al.* 2008, pp. 137–147). Their unusual regulatory properties suggest a buffering function because expression of both genes is induced during Fe starvation, when iron is released by photosystem I degradation. Godman *et al.* (2008, pp. 59–68) examined the multiple pathways that lead to the production of Fe–S clusters, which are essential cofactors in photosynthesis, respiration, and hydrogen production.

Several signaling pathways have been discovered. Many histidine kinases are found in the genome, as shown by Wheeler *et al.* (2008, pp. 193–197). Their study also reveals a large number of proteins with C-type lectin or

scavenger receptor cysteine-rich domains, two motifs often associated with the innate immune system of metazoans. Signaling events can also be mediated by small intracellular peptides, which include the ubiquitin-like SUMO proteins. Bailey *et al.* (2008, pp. 177–192) present a detailed analysis of six SUMO and SUMO-like proteins in Chlamydomonas that are encoded at three different loci. They show that heat or osmotic stresses induce SUMOylation of a number of target proteins, an important first step toward the elucidation of SUMO function in Chlamydomonas. Finally, the Keller group has studied genes expressed during flagellar assembly and disassembly (Chamberlain *et al.* 2008, pp. 7–19), which contributes to our understanding of this exquisitely complex organelle.

Since the beginning of its career as an experimental organism, the "green yeast" (Goodenough 1992) has strived to become the model organism of choice for the study of fundamental biological processes that *Saccharomyces cerevisiae* lacks. In these days of global warming and rapid depletion of oil reserves, Chlamydomonas and related species hold technological promises that have not escaped the notice of researchers, funding agencies, or venture capitalists. Potential industrial applications include hydrogen production, bio-diesel and bio-remediation, as well as production of high-value compounds and pharmaceuticals in the chloroplast. Another area where Chlamydomonas will continue to benefit human health is through the study of ciliopathies. Chlamydomonas flagella, which have both sensory and motility functions, are very similar to animal cilia. Their intensive study has revealed scores of genes whose human homologs are involved in a wide range of genetic diseases.

This special issue provides only a brief preview of areas in which the Chlamydomonas genome sequence will propel research. There are sure to be a lot of interesting clues about how organisms (including us) perform the various functions that keep them alive and evolving. But as the cover of this issue of Genetics illustrates, the ultimate object of study could well be this little green organism itself. Chlamydomonas is not just a model, not just a bag of genes. It is a beautiful creature with a wonderful story to tell—a story far from finished.

LITERATURE CITED

Allmer, J., B. Naumann, C. Markert, M. Zhang and M. Hippler, 2006 Mass spectrometric genomic data mining: novel insights into bioenergetic pathways in Chlamydomonas reinhardtii. Proteomics **6:** 6207–6220.

Ambros, V., and X. Chen, 2007 The regulation of genes and genomes by small RNAs. Development **134:** 1635–1641.

Asamizu, E., Y. Nakamura, S. Sato, H. Fukuzawa and S. Tabata, 1999 A large scale structural analysis of cDNAs in a unicellular green alga, Chlamydomonas reinhardtii. I. Generation of 3433 non-redundant expressed sequence tags. DNA Res. **6:** 369–373.

Bowers, A. K., J. A. Keller and S. K. Dutcher, 2003 Molecular markers for rapidly identifying candidate genes in *Chlamydomonas reinhardtii*: Ery1 and ery2 encode chloroplast ribosomal proteins. Genetics **164:** 1345–1353.

Boynton, J. E., N. W. Gillham, E. H. Harris, J. P. Hosler, A. M. Johnson *et al.*, 1988 Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. Science **240:** 1534–1538.

Casas-Mollano, J. A., J. Rohr, E.-J. Kim, E. Balassa, K. van Dijk *et al.*, 2008 Diversification of the core RNA interference machinery in *Chlamydomonas reinhardtii* and the role of DCL1 in transposon silencing. Genetics **179:** 69–81.

Chamberlain, K. L., S. H. Miller and L. R. Keller, 2008 Gene expression profiling of flagellar disassembly in *Chlamydomonas reinhardtii*. Genetics **179:** 7–19.

Chen, C.-L., C.-J. Chen, O. Vallon, Z.-P. Huang, H. Zhou *et al.*, 2008 Genomewide analysis of box C/D and box H/ACA snoRNAs in *Chlamydomonas reinhardtii* reveals an extensive organization into intronic gene clusters. Genetics **179:** 21–30.

Cognat, V., J.-M. Deragon, E. Vinogradova, T. Salinas, C. Remacle *et al.*, 2008 On the evolution and expression of *Chlamydomonas reinhardtii* nucleus-encoded transfer RNA genes. Genetics **179:** 113–123.

Dayer, R., B. B. Fischer, R. I. L. Eggen and S. D. Lemaire, 2008 The peroxiredoxin and glutathione peroxidase families in *Chlamydomonas reinhardtii*. Genetics **179:** 41–57.

Dent, R. M., C. M. Haglund, B. L. Chin, M. C. Kobayashi and K. K. Niyogi, 2005 Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of Chlamydomonas reinhardtii. Plant Physiol. **137:** 545–556.

Deschamps, P., H. Moreau, A. Z. Worden, D. Dauvillée and S. G. Ball, 2008 Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. Genetics **178:** 2373–2387.

Dutcher, S. K., J. Power, R. E. Galloway and M. E. Porter, 1991 Reappraisal of the genetic map of Chlamydomonas reinhardtii. J. Hered. **82:** 295–301.

Ebersold, W. T., R. P. Levine, E. E. Levine and M. A. Olmsted, 1962 Linkage maps in *Chlamydomonas reinhardi*. Genetics **47:** 531–543.

Figueroa-Martínez, F., S. Funes, L.-G. Franzén and D. González-Halphen, 2008 Reconstructing the mitochondrial protein import machinery of *Chlamydomonas reinhardtii*. Genetics **179:** 149–155.

Godman, J., and J. Balk, 2008 Genome analysis of *Chlamydomonas reinhardtii* reveals the existence of multiple, compartmentalized iron–sulfur protein assembly machineries of different evolutionary origins. Genetics **179:** 59–68.

Goodenough, U. W., 1992 Green yeast. Cell **70:** 533–538.

Gross, C. H., L. P. Ranum and P. A. Lefebvre, 1988 Extensive restriction fragment length polymorphisms in a new isolate of Chlamydomonas reinhardtii. Curr. Genet. **13:** 503–508.

Kalanon, M., and G. I. McFadden, 2008 The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. Genetics **179:** 95–112.

Kathir, P., M. LaVoie, W. J. Brazelton, N. A. Haas, P. A. Lefebvre *et al.*, 2003 Molecular map of the Chlamydomonas reinhardtii nuclear genome. Eukaryot. Cell **2:** 362–379.

Keller, L. C., E. P. Romijn, I. Zamora, J. R. Yates, III and W. F. Marshall, 2005 Proteomic analysis of isolated Chlamydomonas centrioles reveals orthologs of ciliary-disease genes. Curr. Biol. **15:** 1090–1098.

Kindle, K. L., R. A. Schnell, E. Fernandez and P. A. Lefebvre, 1989 Stable nuclear transformation of Chlamydomonas using the Chlamydomonas gene for nitrate reductase. J. Cell Biol. **109:** 2589–2601.

Kirk, D. L., 2005 A twelve-step program for evolving multicellularity and a division of labor. BioEssays **27:** 299–310.

Levine, R. P., 1960 Genetic control of photosynthesis in *Chlamydomonas reinhardi*. Proc. Natl. Acad. Sci. USA **46:** 972–977.

Li, J. B., S. Lin, H. Jia, H. Wu, B. A. Roe *et al.*, 2003 Analysis of Chlamydomonas reinhardtii genome structure using large-scale sequencing of regions on linkage groups I and III. J. Eukaryot. Microbiol. **50:** 145–155.

Li, J. B., J. M. Gerdes, C. J. Haycraft, Y. Fan, T. M. Teslovich *et al.*, 2004 Comparative genomics identifies a flagellar and basal body proteome that includes the *BBS5* human disease gene. Cell **117:** 541–552.

Liang, C., Y. Liu, L. Liu, A. C. Davis, Y. Shen *et al.,* 2008 Expressed sequence tags with cDNA termini: previously overlooked resources for gene annotation and transcriptome exploration in *Chlamydomonas reinhardtii.* Genetics **179:** 83–93.

Long, J. C., F. Sommer, M. D. Allen, S.-F. Lu and S. S. Merchant, 2008 *FER1* and *FER2* encoding two ferritin complexes in *Chlamydomonas reinhardtii* chloroplasts are regulated by iron. Genetics **179:** 137–147.

Martin, N. C., K. S. Chiang and U. W. Goodenough, 1976 Turnover of chloroplast and cytoplasmic ribosomes during gametogenesis in Chlamydomonas reinhardi. Dev. Biol. **51:** 190–201.

May, P., S. Wienkoop, S. Kempa, B. Usadel, N. Christian *et al.*, 2008 Metabolomics- and proteomics-assisted genome annotation and analysis of the draft metabolic network of *Chlamydomonas reinhardtii.* Genetics **179:** 157–166.

Merchant, S. S., S. E. Prochnik, O. Vallon, E. H. Harris, S. J. Karpowicz *et al.*, 2007 The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science **318:** 245–250.

Molnar, A., F. Schwach, D. J. Studholme, E. C. Thuenemann, and D. C. Baulcombe, 2007 miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature **447:** 1126–1129.

Pazour, G. J., N. Agrin, J. Leszyk and G. B. Witman, 2005 Proteomic analysis of a eukaryotic cilium. J. Cell Biol. **170:** 103–113.

Purton, S., and J. D. Rochaix, 1994 Complementation of a Chlamydomonas reinhardtii mutant using a genomic cosmid library. Plant Mol. Biol. **24:** 533–537.

Randolph-Anderson, B. L., J. E. Boynton, N. W. Gillham, E. H. Harris, A. M. Johnson *et al.*, 1993 Further characterization of the respiratory deficient dum-1 mutation of Chlamydomonas reinhardtii and its use as a recipient for mitochondrial transformation. Mol. Gen. Genet. **236:** 235–244.

Riaño-Pachón, D. M., L. G. G. Corrêa, R. Trejos-Espinosa and B. Mueller-Roeber, 2008 Green transcription factors: a Chlamydomonas overview. Genetics **179:** 31–39.

Rohr, J., N. Sarkar, S. Balenger, B. R. Jeong and H. Cerutti, 2004 Tandem inverted repeat system for selection of effective transgenic RNAi strains in Chlamydomonas. Plant J. **40:** 611–621.

Rymarquis, L. A., J. M. Handley, M. Thomas and D. B. Stern, 2005 Beyond complementation: map-based cloning in Chlamydomonas reinhardtii. Plant Physiol. **137:** 557–566.

Sager, R., and M. R. Ishida, 1963 Chloroplast DNA in Chlamydomonas. Proc. Natl. Acad. Sci. USA **50:** 725–730.

Sager, R., and Z. Ramanis, 1970 Genetic studies of chloroplast DNA in Chlamydomonas. Symp. Soc. Exp. Biol. **24:** 401–417.

Schmidt, M., G. Gessner, M. Luff, I. Heiland, V. Wagner *et al.*, 2006 Proteomic analysis of the eyespot of Chlamydomonas reinhardtii provides novel insights into its components and tactic movements. Plant Cell **18:** 1908–1930.

Shen, Y., Y. Liu, L. Liu, C. Liang and Q. Q. Li, 2008 Unique features of nuclear mRNA poly(A) signals and alternative polyadenylation in *Chlamydomonas reinhardtii.* Genetics **179:** 167–176.

Tam, L. W., and P. A. Lefebvre, 1995 Insertional mutagenesis and isolation of tagged genes in Chlamydomonas. Methods Cell Biol. **47:** 519–523.

van Lis, R., A. Atteia, G. Mendoza-Hernandez and D. Gonzalez-Halphen, 2003 Identification of novel mitochondrial protein components of Chlamydomonas reinhardtii: a proteomic approach. Plant Physiol. **132:** 318–330.

Wheeler, G. L., D. Miranda-Saavedra and G. J. Barton, 2008 Genome analysis of the unicellular green alga *Chlamydomonas reinhardtii* indicates an ancient evolutionary origin for key pattern recognition and cell-signaling protein families. Genetics **179:** 193–197.

Zhao, T., G. Li, S. Mi, S. Li, G. J. Hannon *et al.*, 2007 A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. Genes Dev. **21:** 1190–1203.

Zimmer, S. L., Z. Fei and D. B. Stern, 2008 Genome-based analysis of *Chlamydomonas reinhardtii* exoribonucleases and poly(A) polymerases predicts unexpected organellar and exosomal features. Genetics **179:** 125–136.