Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 1D

Motomichi Matsuzaki1, Osami Misumi2, Tadasu Shin-i3, Shinichiro Maruyama1, Manabu Takahara2, Shin-ya Miyagishima2, Toshiyuki Mori2, Koji Nishida2, Fumi Yagisawa3, Keisuke Nishida2, Yamato Yoshida2, Yoshihiro Nishimura3, Shunsuke Nakao4, Takanori Kobayashi5, Yu Momoyama6, Tetsuya Hisayashi4, Ayumi Minoda4, Masako Sano7, Hisayo Nomoto8, Kazuko Oishi9, Hiroko Hayashi10, Fumiko Ohya11, Satoko Nishizaka12, Shinobu Haga12, Sachiko Miura13, Tomomi Morishita14, Yukihiro Kabeya15, Kimihito Terasawa16, Yutaka Suzuki17, Yasuyuki Ishii18, Shuichi Asakawa1, Hiroyoshi Takano19, Hiroyoshi Oka20, Haruko Kuroiwa21, Kan Tanaka22, Nobuyoshi Shimizu23, Sumio Sugano24, Naoki Sato25, Hisayoshi Nokzaki26, Naotaka Ogasawa21, Yuji Kohara21 & Tatsuya Kuroiwa21

1Department of Biomedical Chemistry, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan
2Department of Life Science, College of Science, Rikkyo (St Paul’s) University, Tokyo 7-3-1 Nishiikebukuro, Toshima-ku, Tokyo 171-8501, Japan
3Department of Genetic Resource Information, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan
4Institute of Molecular and Cellular Biosciences, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
5National Institute of Livestock and Grassland Science, 2 Ikenodai, Tsukuba, Ibaraki 305-0901, Japan
6Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
7Department of Biochemistry and Molecular Biology, Faculty of Science, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama, Saitama 338-8570, Japan
8Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
9Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan
10Department of Biological Sciences, Graduate School of Science and Technology, Kumanoto University, 2-39-1 Kurokami, Kumanoto 860-8555, Japan
11Graduate School of Biological Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan

Small, compact genomes of ultrasmall unicellular algae provide information on the basic and essential genes that support the lives of photosynthetic eukaryotes, including higher plants1,2. Here we report the 16,520,305-base-pair sequence of the 20 chromosomes of the unicellular red alga *Cyanidioschyzon merolae* 1D, as the first complete algal genome. We identified 5,331 genes in total, of which at least 86.3% were expressed. Unique characteristics of this genomic structure include: a lack of introns in all but 26 genes; only three copies of ribosomal DNA units that maintain the nucleolus; and two dynamin genes that are involved only in membrane-bounded compartments; for example, a microbody (peroxisome), single Golgi apparatus with two cisternae, coated vesicles, a single endoplasmic reticulum, and a few lysosome-like structures, as well as a small volume of cytosol (Fig. 1). One of the main points of interest that we discuss regarding this alga focuses on the origin, evolution and fundamental traits (for example, multiplication and differentiation) of single- or double-membrane-bounded organelles in plant cells. *C. merolae*, with its complete genomic information, provides an excellent opportunity for addressing such basic questions using microarray and proteome analyses. In addition, from an evolutionary perspective, *C. merolae* has other noteworthy properties that allow us to study the origin of eukaryotic cells, primary endosymbiosis between cyanobacteria and eukaryotic hosts, and secondary endosymbiosis between red algae and their hosts.

Samples of *C. merolae* 1D were isolated7 from the hot spring algal collection provided by G. Pinto (Naples University). The entire *C. merolae* genome was sequenced using the random sequencing method (see Methods). We obtained 16,520,305 base pairs (bp); approximately 99.98% of the estimated total length) of the nuclear genome sequence (Fig. 2, Table 1, and Supplementary Fig. 1 and Supplementary Table 1) with 46 gaps. The genome is distributed among the 20 chromosomes and ranges in size from approximately 0.42 to 1.62 Mb. No significant deviation in statistical parameters, such as base composition and gene density, were observed among the chromosomes (Supplementary Table 1). The overall G+C composition was 55.0%. The dinucleotide CpG in the *C. merolae* genome was exceptionally over-represented (1.151) compared with the expected value from observations of G+C content; it is generally underrepresented in other eukaryote genomes (Table 1).

The putative repeat unit of telomeres in *C. merolae* is synthetic eukaryotes, and contains a minimal set of small membrane-bounded compartments; for example, a microbody (peroxisome), a single Golgi apparatus with two cisternae, coated vesicles, a single endoplasmic reticulum, and a few lysosome-like structures, as well as a small volume of cytosol (Fig. 1). One of the main points of interest that we discuss regarding this alga focuses on the origin, evolution and fundamental traits (for example, multiplication and differentiation) of single- or double-membrane-bounded organelles in plant cells. *C. merolae*, with its complete genomic information, provides an excellent opportunity for addressing such basic questions using microarray and proteome analyses. In addition, from an evolutionary perspective, *C. merolae* has other noteworthy properties that allow us to study the origin of eukaryotic cells, primary endosymbiosis between cyanobacteria and eukaryotic hosts, and secondary endosymbiosis between red algae and their hosts.

Samples of *C. merolae* 1D were isolated7 from the hot spring algal collection provided by G. Pinto (Naples University). The entire *C. merolae* genome was sequenced using the random sequencing method (see Methods). We obtained 16,520,305 base pairs (bp); approximately 99.98% of the estimated total length) of the nuclear genome sequence (Fig. 2, Table 1, and Supplementary Fig. 1 and Supplementary Table 1) with 46 gaps. The genome is distributed among the 20 chromosomes and ranges in size from approximately 0.42 to 1.62 Mb. No significant deviation in statistical parameters, such as base composition and gene density, were observed among the chromosomes (Supplementary Table 1). The overall G+C composition was 55.0%. The dinucleotide CpG in the *C. merolae* genome was exceptionally over-represented (1.151) compared with the expected value from observations of G+C content; it is generally underrepresented in other eukaryote genomes (Table 1).

The putative repeat unit of telomeres in *C. merolae* is synthetic eukaryotes, and contains a minimal set of small membrane-bounded compartments; for example, a microbody (peroxisome), a single Golgi apparatus with two cisternae, coated vesicles, a single endoplasmic reticulum, and a few lysosome-like structures, as well as a small volume of cytosol (Fig. 1). One of the main points of interest that we discuss regarding this alga focuses on the origin, evolution and fundamental traits (for example, multiplication and differentiation) of single- or double-membrane-bounded organelles in plant cells. *C. merolae*, with its complete genomic information, provides an excellent opportunity for addressing such basic questions using microarray and proteome analyses. In addition, from an evolutionary perspective, *C. merolae* has other noteworthy properties that allow us to study the origin of eukaryotic cells, primary endosymbiosis between cyanobacteria and eukaryotic hosts, and secondary endosymbiosis between red algae and their hosts.
Figure 2. Representation of *C. merolae* chromosomes. a. Chromosome 1 and mapping of the BAC clones (50 kb and 100 kb). The chromosome is represented as a bar with pseudo-colour assignments of local G+C contents. Side rods represent genes and gene-like elements (left for transcription towards the top, and right for the bottom). Telomere and subtelomeric elements are indicated respectively as semicircles and purple rectangles at each end of the chromosome. 

b. A bird's-eye view of 20 chromosomes showing G+C contents, genes, subtelomeric elements (designations on the side: P, H, L, A, E, and so on), and RNA genes. A putative centromeric (A+T-rich) region is located on each chromosome.
GGGGGGAAT, and as far as could be determined experimentally, it is found on both ends of the chromosomes. In addition, several sequence elements up to 20 kilobases (kb) in length were duplicated in 30 of the 40 putative subtelomeric regions (Fig. 2, Supplementary Fig. 1). Each chromosome has, in varying degrees, a single A+T-rich region on its mid-section. As chromosomal centromeric regions generally have a biased base composition, this A+T-rich region possibly defines centromeres (Fig. 2). The centromeres were confirmed via immunological experiments using antibodies against CENP-A, which was identified in the *C. merolae* genome (data not shown). Unlike many other eukaryotes, the *C. merolae* genome does not contain tandem repeated arrays of ribosomal RNA (rRNA) genes (Fig. 2). A single rRNA gene unit (18S-5.8S-28S) was discovered on three separate loci. The three units were virtually identical in sequence. Moreover, *C. merolae* has only three copies of the 5S rRNA gene, the sequences of which are also almost identical. Therefore, *C. merolae* has the smallest set of rRNA genes among all eukaryotes thus far studied. These results might be related to the existence of a single small nucleolus without nucleolus-associated chromatin. Furthermore, they also promote studies on the origin and formation of the nucleolus, because even prokaryotic cells with more than three copies of RNA gene units do not have a nucleolus.

A full-length complementary DNA (cDNA) library was used to map expressed genes within the *C. merolae* genome. Fortunately, 99.85% of the expressed sequence tags (ESTs) were mapped on the genome sequence. In addition, many cDNA clones encoded a single open reading frame (ORF) bridging both end sequences. This suggests that most *C. merolae* genes lack introns. The predicted genes were automatically annotated using several databases (see Methods and Supplementary Information). As a result, 5,331 genes were identified, and 86.3% of them had corresponding ESTs (Supplementary Table 2). The number of genes in the *C. merolae* genome is similar to those found in yeasts and malarial parasites, despite the great ecological differences between these species (Table 1). Furthermore, the genes of the *C. merolae* genome are remarkable for their paucity of introns. Only 26 genes (0.5% of the protein genes) contained introns, and all but one of them had only a single intron. These introns had strict consensus sequences (Supplementary Fig. 2 and Supplementary Table 3).

Figure 3 summarizes the repertoire of *C. merolae* proteins on the basis of their assignment to eukaryotic clusters of orthologous groups (KOGs). Of the 4,771 predicted proteins, 2,536 were assigned to KOGs, by emulating the NCBI KOGnitor service (http://www.ncbi.nlm.nih.gov/COG/new/kognitor.html). The distribution of the functional classification of *C. merolae* was compared with those of other free-living unicellular eukaryotes, such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*1,4, and a higher plant *Arabidopsis thaliana*15. The distribution was on the whole similar to both yeasts which have similar genome size although *C. merolae* cells contain plastids. The lowered proportion of genes for ‘secondary metabolites biosynthesis, transport and catabolism’ found in these unicellular organisms, as compared with that of *A. thaliana*, might reflect their simple cellular organizations (Fig. 3, Supplementary Table 4).

In *C. merolae*, the division of double-membrane-bounded mitochondria and plastids involves a dynamic trio: an FtsZ ring of bacterial origin, electron dense mitochondrial/plastid dividing rings (MD and PD rings), and eukaryotic mechanochemical dynamin

| Feature                        | C. merolae | S. pombe | S. cerevisiae | A. thaliana | P. falciparum |
|--------------------------------|------------|----------|---------------|-------------|--------------|
| No. of introns                 | 1          | 7        | 38            | 58          | 3            |
| No. of genes                   | 208        | 174      | 274           | 620         | 43           |
| Genome size (bp)               | 149,987    | 19,431   | 85,779        | 366,924     | 29,422       |
| No. of 18S, 5.8S and 28S rRNA units | 3        | 30       | 100–150       | 700–800     | 3            |
| No. of tRNA genes              | 30         | 174      | 274           | 620         | 43           |
| Genes with introns (%)         | 0.5        | 43.0     | 5.0           | 79.0        | 53.9         |
| Mean gene length† (bp)         | 1,552      | 1,426    | 1,424         | 2,089       | 2,089        |
| Per cent coding                | 44.9       | 57.5     | 70.5          | 28.8        | 52.6         |
| Density (bp per protein genes) | 947        | 1,943    | 2,958         | 6,326       | 1,989        |
| Genes with introns (%)         | 0          | 19.3     | 10.3          | 27.3        | 43.0         |
| No. of chromosomes             | 20         | 3        | 16            | 5           | 14           |
| Sequenced length (bp)          | 16,520,305 | 12,462,637* | 12,496,682*  | 115,409,949* | 22,853,764* |
| G+C content (%)                | 55.0       | 36.0     | 38.3          | 34.9        | 19.4         |
| Genes (Coding sequences)       | 5,331      | 4,929    | 5,770         | 25,498      | 5,268        |
| Mean gene length† (bp)         | 1,552      | 1,426    | 1,424         | 2,089       | 2,089        |
| Per cent coding                | 44.9       | 57.5     | 70.5          | 28.8        | 52.6         |
| Density (bp per protein genes) | 947        | 1,943    | 2,958         | 6,326       | 1,989        |
| Genes with introns (%)         | 0          | 19.3     | 10.3          | 27.3        | 43.0         |
| No. of chromosomes             | 20         | 3        | 16            | 5           | 14           |
| Sequenced length (bp)          | 16,520,305 | 12,462,637* | 12,496,682*  | 115,409,949* | 22,853,764* |
| G+C content (%)                | 55.0       | 36.0     | 38.3          | 34.9        | 19.4         |
| Genes (Coding sequences)       | 5,331      | 4,929    | 5,770         | 25,498      | 5,268        |
| Mean gene length† (bp)         | 1,552      | 1,426    | 1,424         | 2,089       | 2,089        |
| Per cent coding                | 44.9       | 57.5     | 70.5          | 28.8        | 52.6         |
| Density (bp per protein genes) | 947        | 1,943    | 2,958         | 6,326       | 1,989        |
| Genes with introns (%)         | 0          | 19.3     | 10.3          | 27.3        | 43.0         |
| No. of chromosomes             | 20         | 3        | 16            | 5           | 14           |
| Sequenced length (bp)          | 16,520,305 | 12,462,637* | 12,496,682*  | 115,409,949* | 22,853,764* |
| G+C content (%)                | 55.0       | 36.0     | 38.3          | 34.9        | 19.4         |
| Genes (Coding sequences)       | 5,331      | 4,929    | 5,770         | 25,498      | 5,268        |
| Mean gene length† (bp)         | 1,552      | 1,426    | 1,424         | 2,089       | 2,089        |
| Per cent coding                | 44.9       | 57.5     | 70.5          | 28.8        | 52.6         |
| Density (bp per protein genes) | 947        | 1,943    | 2,958         | 6,326       | 1,989        |
| Genes with introns (%)         | 0          | 19.3     | 10.3          | 27.3        | 43.0         |

Data for the other organisms are from ref. 31. ND, not determined; NA, not applicable.
†Excluding introns.
rings. Four genes representing mitochondrial FtsZ (FtsZ2-1 and FtsZ2-2) and plastid FtsZ (FtsZ1-1 and FtsZ1-2) were identified. A large gene family consisting of more than 10 members encoding functionally diverse dynamins with a wide range of membrane pinching roles have been found in other organisms; however, only two dynamin genes (C. merolae Dnm1 and Dnm2) are found, with a role in the later stages of the mitochondrion and plastid division, respectively. These findings suggest that plastids and mitochondria divide in a similar way, using very common systems consisting of the amalgamation of bacterial and eukaryotic rings. The dynamic trio of plastid division is conserved in lower algae to higher plants. With mitochondrial divisions, however, whilst dynamin rings are retained in higher organisms, FtsZ and MD rings are not clearly observed, and it is possible that they were replaced by other systems during eukaryotic evolution. MD/PD ring genes are yet unknown, although their identification should be accelerated by works such as this. Although the microbody, a single-membrane-bounded organelle, divides by binary fission in C. merolae, it lacks Pex11p, which is a known key regulator of microbody division and proliferation.

The following proteins related to cell motility and cytokinesis were encoded in C. merolae (Supplementary Table 3): one set of tubulin, two actins, five proteins of the kinesin family, and several intermediate filament proteins. However, no genes encoding myosin or proteins containing dynein motor domains were found. The absence of the myosin gene is consistent with the fact that electron microscopy and immuno-detection techniques did not detect microfilaments of actin; cDNA clones for actin genes were also not obtained. In the red alga Cyanidium caldarium RK-1, which is closely related to C. merolae but has a genome double the size, cells divide using a contractile ring of actin filaments. C. merolae cells therefore seem to divide using a system that is simpler than that of actomyosin.

C. merolae has noteworthy properties, which are relevant for examining the origin of eukaryotes, and primary and secondary plastid endosymbiosis. Only 30 transfer RNAs (tRNAs) were detected in the nuclear genome using the program tRNAscan-SE with relaxed parameter settings (Fig. 2). Some of these tRNA genes showed possible archael features, namely, ectopic introns and anticodon GAU for tRNA-Ile (Supplementary Fig. 3). Four of these tRNA genes seemed to have introns in the D-loop region, whereas the introns of eukaryotic tRNA genes are limited to a site 3′ to the anticodon. As ectopic tRNA introns have been reported in some archael genomes, this could explain the paucity of detected tRNA genes in C. merolae. tRNAscan-SE might have overlooked other tRNAs owing to the existence of unknown types of ectopic introns. Another point to note is that C. merolae possesses a single tRNA-Ile with anticodon GAU, which has not been observed in eukaryotes, but only in prokaryotes.

Standard sets of photosystem genes, including those encoding phycobilisome components, were observed in C. merolae. Many of them (11 PSI genes and 17 PSII genes) are encoded in the plastid genome, while PsbO, P, U, Z as well as a distant PsbQ homologue are encoded in the nuclear genome. Although only PsbU and PsbZ were previously identified in red algal PSI, the localization of PsbP and putative PsbQ in PSII, as recently suggested in Synechocystis sp. PCC6803, is an interesting subject of proteomic study. The genes psaH, N, X, as well as psbS and ndh genes are not encoded in the plastid or the nuclear genomes. Therefore, the photosystems of C. merolae lack various mechanisms for dissipating excessive light energy.

Enzymes of the Calvin cycle in plants are known to be a mosaic of enzymes originating from cyanobacteria-like ancestors of an endosymbiont and its eukaryotic host. Red algal ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is known to be a product of horizontal gene transfer. The origin of other Calvin cycle enzymes is essentially identical in C. merolae and A. thaliana (Supplementary Fig. 4 and Supplementary Table 6). It is highly probable that the complex and mosaic origin of Calvin cycle enzymes derived from common ancestors of green plants and red algae, and no essential changes occurred after the separation of the two lineages. This is strong support for the concept of a single event of primary plastid endosymbiosis. Among the known translocase proteins of plastids, Toc34, Toc75, Tic20, Tic22 and Tic110 were encoded in the C. merolae genome, but other proteins such as Toc159, Tic40 and Tic55 were not found. Results of phylogenetic analysis of the five translocon components (to be published elsewhere) also suggest the concept.

Another aspect of the comparative genomics of the red algal genome is secondary endosymbiosis. Cryptophytes are thought to retain a remnant of the endosymbiotic red algal nucleus, the nucleomorph, in the periplastidic compartment. The sequencing of the cryptophyte alga Guillardia theta nucleomorph genome revealed a number of curious architectural features that might be shared by the genome of red algae. C. merolae chromosomes showed multiple subtelomeric duplications, but did not contain rRNA gene clusters such as those of the nucleomorph genome. This implies that the telomeric rRNA gene clusters observed in the nucleomorph genome, as well as other prominent genome structures such as overlapped genes, appeared after secondary symbiosis. It is also notable that ectopic tRNA introns are also reported in nucleomorph tRNAs. Details of the comparisons with the nucleomorph genome will be presented elsewhere.

Light signal transduction is critical for the growth and differentiation of photosynthetic organisms. As the division of C. merolae cells is synchronized by light, an elaborate mechanism for light signal transduction must exist. Several putative blue light receptor (cryptochrome) genes were found in C. merolae, whereas no genes encoding phytochromes and phototropins were identified. As bacterial phytochrome genes are only found in some species of cyanobacteria with large genomes, the ancestor of plastids might be an ancestral cyanobacterium without phytochromes. This also suggests that the phytochromes of higher plants might not be of cyanobacterial origin. In higher plants, various signalling pathways (such as the two-component system consisting of histidine kinases and response regulators as well as a MAP kinase cascade) are involved in the signal transduction of various hormones, and in the development of organs. In C. merolae, the presence of only a single candidate for histidine kinase and a dozen MAP kinase-related molecules is suggested. However, there are no response regulators other than those that are plastid-encoded, trimeric G protein and adenylate cyclase. Thus, C. merolae appear to use only a limited repertoire of signal transduction mechanisms, which corroborates the lack of cell differentiation in this alga.

C. merolae is an alga in which all of the three genome compartments—nucleus, mitochondrion (32,211 bp) and plastid (149,987 bp)—have been sequenced. Such information is a prerequisite for future studies on proteomics, expression analysis using microarrays, and structural biology with heat-stable proteins that are unique among eukaryotes. All of this information will, in turn, help elucidate the origin, evolution and fundamental mechanisms of the single- as well as double-membrane-bounded organelles, and ultimately all photosynthetic eukaryotes. In addition, this hot spring alga will be useful in analysing the mechanisms of heat and acid tolerance in eukaryotic cells.

Methods

Whole genome shotgun sequencing

We sequenced the C. merolae genome by the whole genome random sequencing method (see Supplementary Information for details). About 335,000 insert ends were sequenced, which covered the genome 11 times. BAC libraries with two subsets were constructed and a large-scale full-length cDNA library from cells cultured under various growth conditions was prepared. The sequences were assembled using Phrap, further examined by referring to another assembly using ARACHNE, and edited using CONSED. The scaffolds were built within the hybridization groups using read-pair information from the BAC, shotgun and cDNA clones. The gaps between the contigs were closed by primer walking PCR, and mate-pair clone and BAC clone sequencing.
Supplementary Information

1. Kuroiwa, T. The primitive red alga: Cyanidioschyzon merolae and Cyanidiouchromis merolae as model system for investigating the dividing apparatus of mitochondria and plastids. Bioessays 26, 344–354 (1998).

2. Kuroiwa, T. et al. The division apparatus of plastids and mitochondria. Int. Rev. Cytol. 181, 1–41 (1998).

3. McFadden, G. I. & Ralph, S. A. Dynamin: the endosymbiosis ring of power. Proc. Natl Acad. Sci. USA 100, 3557–3559 (2003).

4. Surridge, D. Ancient rings. Nature 422, 275 (2003).

5. Terui, S., Suzuki, K., Takahashi, H., Itoh, R. & Kuroiwa, T. High synchronization of chloroplast division in the ultramicro-alga Cyanidiouchromis merolae by treatment with both light and aphidicolin. J. Phycol. 31, 958–961 (1995).

6. Kuroiwa, T. et al. Comparison of ultrastructures between the ultra-small eukaryote Cyanidiouchromis merolae and Cyanidium caldarium. Cytologia (Tokyo) 59, 129–138 (1994).

7. Toda, K., Takahashi, H., Itoh, R. & Kuroiwa, T. DNA contents of cell nuclei in two Cyanidiouchromis merolae and Cyanidium caldarium. Forma A. Cytologia (Tokyo) 60, 183–188 (1995).

8. Tatsutou, R. L. et al. The COG database: an updated version includes eukaryotes. BMC Bioinform. 4, 41 (2003).

9. Goffeau, A. et al. Life with 6000 genes. Science 274, 563–567 (1996).

10. Wood, V. et al. The genome sequence of Schizosaccharomyces pombe. Nature 415, 871–880 (2002).

11. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815 (2000).

12. Takahara, M. et al. A putative mitochondrial fH2 gene is present in the unicellular primitive red alga Cyanidiouchromis merolae. Mol. Gen. Genet. 264, 452–460 (2000).

13. Miyagishima, S., Nishida, K. & Kuroiwa, T. An evolutionary puzzle: chloroplast and mitochondrial division rings. Trends Plant Sci. 8, 432–438 (2003).

14. Nishida, K. et al. Dynamic recruitment of dynamin for final mitochondrial segregation in a primitive red alga. Proc. Natl Acad. Sci. USA 100, 2146–2151 (2003).

15. Kuroiwa, T., Mext, T., Takahara, M., Miyagishima, S. & Kuroiwa, T. Chloroplast division machinery as revealed by immunofluorescence and electron microscopy. Planta 215, 185–190 (2002).

16. Gao, H., Kadirjan-Kalbach, D., Freihiel, J. I. & Osterryoung, K. W. ARCs, a cytosolic dynamin-like protein from plants, is part of the chloroplast division machinery. Proc. Natl Acad. Sci. USA 100, 4328–4333 (2003).

17. Miyagishima, S. et al. Microbody proliferation and segregation cycle in the single-microbody alga Cyanidiouchromis merolae. Planta 208, 326–336 (1999).

18. Marshall, P. A. et al. Pmp27 promotes peroxisomal proliferation. J. Cell Biol. 129, 345–355 (1995).

19. Takahashi, H. et al. A possible role of actin dots in the formation of the contractile ring in the ultramicro-alga Cyanidium caldarium RR-1. Protoplasma 201, 115–119 (1998).

20. Nozaki, H. et al. The phylegetic position of red alga revealed by multiple nuclear genes from mitochondria-containing eukaryotes and an alternative hypothesis on the origins of plastids. J. Mol. Evol. 56, 485–497 (2003).

21. March, C. & Grosjean, H. tRNNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticonvulsant-sparring strategies and domain-specific features. RNA 8, 1189–1232 (2002).

22. Ohta, N. et al. Complete sequence and analysis of the plastid genome of the unicellular red alga Cyanidiouchromis merolae. DNA Res. 10, 67–77 (2003).

23. Kashino, Y. et al. Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium Synechocystis sp. PCC 6803 reveals the presence of novel polypeptides. Biochemistry 41, 8004–8012 (2002).

24. Martin, W. & Schnaremberger, C. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr. Genet. 32, 1–18 (1997).

25. Ohta, N., Sato, N., Ueda, K. & Kuroiwa, T. Analysis of a plastid gene cluster reveals a close relationship between Cyanidiouchromis and Cyanidium. J. Plant Res. 110, 235–245 (1997).

26. Cavalier-Smith, T. Genetic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimeras (meta-algae). Phil. Trans. R. Soc. Lond. B 358, 109–134 (2003).

27. Douglas, S. et al. The highly reduced genome of an enslaved algal nucleus. Nature 40, 1091–1096 (2001).

28. Zauner, S. et al. Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unassigned contigs. Sequences and annotation are available at http://merola.biol.s.u-toyko.ac.jp/ or http://dolphin.lab.nig.ac.jp/publish/.

Acknowledgements We thank many colleagues for discussions, especially K. Kita, Y. Watanabe, H. Fujisawa and T. Q. Ueda. We also thank Trans New Technology, Inc. for providing computational resources. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas "Genome" from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and a Grant-in-Aid from the Promotion of Basic Research Activities for Innovative Biosciences (ProBRAIN).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.K. (tsune@rikkyo.ne.jp) or M.M. (mazaki@biol.s.u-toyko.ac.jp). Chromosome sequences were submitted to DDBJ with accession numbers AP006483–AP006502 (chromosome 1–20) and AP006600–AP006614 (unassigned contigs). Sequences and annotation are available at http://merola.biol.s.u-toyko.ac.jp/ or http://dolphin.lab.nig.ac.jp/publish/.

Long-lasting sensitization to a given colour after visual search

Chia-huei Tseng1*, Joetta L. Gobell1* & George Sperling1,2

1Department of Cognitive Sciences, and 2Neurobiology and Behavior, and Institute of Mathematical Behavioral Sciences, University of California, Irvine, California 92697, USA

* Present addresses: Center of Cognitive Science, Rutgers University, New Brunswick, New Jersey 08854, USA (C.-h.T.); Department of Psychology and Center for Neural Science, New York University, New York, New York 10003, USA (J.L.G.)

Visual attention enables an observer to select specific visual information for processing. In an ambiguous motion task in which a coloured grating can be perceived as moving in either of two opposite directions depending on the relative salience of two colours in the display, attending to one of the colours influences the direction in which the grating appears to move1, Here, we use this secondary effect of attention in a motion task to measure the effect of attending to a specific colour in a search task. Observers performed a search task in which they searched for a target letter in a 4 × 4 coloured matrix. Each of the 16 squares within a matrix was assigned one of four colours, and observers knew that the target letter would appear on only one of these colours throughout the experiment. Observers performed the ambiguous motion task before the search task profoundly influenced the perceived direction of motion. This effect lasted for up to one month and in some cases had to be reversed by practising searches for the complementary colour, indicating a much longer-persisting effect of attention than has been observed previously.

To investigate the consequences of attending to a particular colour, we designed a search task that requires observers to attend to a particular colour and to ignore all other colours. The task is to report the location of a target letter among other ‘distractor’ characters. Each trial consists of ten consecutive 4 × 4 matrices displayed rapidly (Fig. 1a). Each of the 16 squares of a matrix is randomly assigned one of four colours (red, green, yellow or blue). The initial frame rate was two matrices per second and it was increased as observers’ performance improved to as high as 19 matrices s–1 for the fastest observer. At the average final speed of 10 frames s–1, each colour display was shown for 50 ms and then the to-be-searched letters and numbers appeared superimposed on the coloured squares for an additional 50 ms. The purpose of advancing the colour matrix relative to the target was to induce the observers to use the colour to find the target. A target colour, either red or green, was assigned to each observer for the duration of the first phase of the experiment.

At the beginning of each trial, observers were shown a randomly

1*Department of Cognitive Sciences, and 2Neurobiology and Behavior, and Institute of Mathematical Behavioral Sciences, University of California, Irvine, California 92697, USA

* Present addresses: Center of Cognitive Science, Rutgers University, New Brunswick, New Jersey 08854, USA (C.-h.T.); Department of Psychology and Center for Neural Science, New York University, New York, New York 10003, USA (J.L.G.)