The Gonium pectorale genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity

Erik R. Hanschen1, Tara N. Marriage2, Patrick J. Ferris1, Takashi Hamaji3, Atsushi Toyoda4,5, Asao Fujiyama4,5, Rafik Neme6, Hideki Noguchi4, Yohei Minakuchi5, Masahiro Suzuki7, Hiroko Kawai-Toyooka7, David R. Smith8, Halle Sparks2, Jaden Anderson2, Robert Bakarić9, Victor Luria10,11, Amir Karger12, Marc W. Kirschner10, Pierre M. Durand1,13,14, Richard E. Michod1,11, Hisayoshi Nozaki7 & Bradley J.S.C. Olson2,11

The transition to multicellularity has occurred numerous times in all domains of life, yet its initial steps are poorly understood. The volvocine green algae are a tractable system for understanding the genetic basis of multicellularity including the initial formation of cooperative cell groups. Here we report the genome sequence of the undifferentiated colonial alga, Gonium pectorale, where group formation evolved by co-option of the retinoblastoma cell cycle regulatory pathway. Significantly, expression of the Gonium retinoblastoma cell cycle regulator in unicellular Chlamydomonas causes it to become colonial. The presence of these changes in undifferentiated Gonium indicates extensive group-level adaptation during the initial step in the evolution of multicellularity. These results emphasize an early and formative step in the evolution of multicellularity, the evolution of cell cycle regulation, one that may shed light on the evolutionary history of other multicellular innovations and evolutionary transitions.
Multicellular organisms have independently evolved numerous times throughout the tree of life including plants, animals, fungi, cyanobacteria, amoeba, brown algae, red algae and green algae. In animals, multicellularity emerged 600–950 million years ago (Myr ago) correlating with a large expansion of genes encoding transcription factors, signalling pathways, and cell adhesion genes that were co-opted from their unicellular ancestors. Similarly, multicellular, terrestrial plants emerged ~750 Myr ago with an expansion of many signalling pathways present in their unicellular relatives. However, because most multicellular lineages have long diverged from their unicellular relatives, the genomic signature of the transition to multicellularity has been obscured, and consequently this evolutionary process remains enigmatic.

The volvocine green algae are a unique model system for the evolution of multicellularity because the unicellular ancestry is clear, the emergence of multicellularity occurred ~230 Myr ago, and species exhibit a stepwise increase in morphological complexity ranging from undifferentiated colonies to differentiated multicellular species. Unicellular Chlamydomonas reinhardtii is thought to resemble the unicellular ancestor of multicellular volvocines, including undifferentiated Gonium pectorale and differentiated Volvox carteri (Fig. 1).

Chlamydomonas undergoes a variant cell cycle (Supplementary Fig. 2), regulated by homologues of the retinoblastoma cell cycle control pathway, termed multiple-fission where germ–soma separation is in extracellular matrix (ECM) inside the spheroid (Fig. 1). Gonium pectorale does not rely upon the evolution of de novo genes. Though gene regulation may be important during multicellular innovation, the diversity and abundance of transcription factors is similar in Chlamydomonas, Gonium and Volvox (Fig. 2b, Supplementary Table 1, Supplementary Fig. 3). Although enrichment of transcription factors can correlate with the evolution of multicellularity (C), expanded ECM (E) and somatic cells (S) are denoted. (b) Micrographs of Chlamydomonas (green; scale bar, 10 μm), Gonium (blue; scale bar, 10 μm) and Volvox (black; scale bar, 25 μm) show morphological differences.

Results

Genomic comparisons of volvocine algae. At the genomic level, the genomes of Chlamydomonas, Gonium and Volvox are similar, though various measures of genome compactness correlate with cell number, consistent with a long-term increase in organisinal size. Chlamydomonas and Gonium have similar GC content near 64%, while Volvox has 56% GC (Table 1). Otherwise, Chlamydomonas, Gonium and Volvox have decreasing gene densities of 159.6, 120.9 and 113.7 genes per megabase, with an increasing average intron length of 279, 349 and 500 base pairs, respectively. Although intron length increases with organisinal complexity, the number of introns per gene (Chlamydomonas, 7.46; Gonium, 6.5; Volvox 6.8, Table 1) does not. GC content, intron length and gene density correlate with morphological complexity.

We next examined genome-wide evolution in all three species to better understand the genetic basis for the evolution of multicellularity. A prediction of lineage-specific genes shows few genes correlate with the evolution of multicellularity in the Volvocales (phylostratum 7; PS7) with a maximum of 180–357 genes (Fig. 2a). This suggests that the evolution of multicellularity does not rely upon the evolution of de novo genes. Though gene regulation may be important during multicellular innovation, the diversity and abundance of transcription factors is similar in Chlamydomonas, Gonium and Volvox (Fig. 2b, Supplementary Table 1, Supplementary Fig. 3). Although enrichment of transcription factors can correlate with the evolution of multicellularity, this is not always the case; we found that Gonium and Volvox have fewer transcription factors than in Chlamydomonas (Fig. 2b). These include PHD domains.

Table 1 | Summary statistics for genome level analyses for Chlamydomonas, Gonium and Volvox.

| Characteristic                      | Chlamydomonas v5.3 | Gonium v1 | Volvox v2 |
|------------------------------------|--------------------|-----------|-----------|
| Genome size (Mb)                   | 111.1              | 148.8     | 137.8     | 131.1     |
| Scaffold N50 (Mb)                  | 7.78               | 1.27      | 1.49      | 2.6       |
| Number of contigs/scaffolds        | 54                 | 2,373     | 1,265     | 434       |
| % G and C                          | 64.1               | 64.5      | 56.0      | 56.1      |
| Protein coding loci                | 17,737             | 17,984    | 15,669    | 14,971    |
| Gene density (genes/Mb)            | 159.6              | 120.9     | 113.7     | 114.1     |
| Intron/gene                        | 7.46               | 6.50      | 6.78      | 6.29      |
| Average intron length (bp)         | 279.17             | 349.83    | 496.67    | 399.50    |
| % genes w/introns                  | 92.4               | 92.6      | 82.8      | 84.0      |
genes and Pfam domains, suggesting that species-specific domains (Fig. 2d) compared with multicellularity-correlated represented Pfam domains (Supplementary Fig. 4). Interestingly, unicellular green algae, Supplementary Data 1); 394 Pfam domain diversity and abundance (compared with nine Gonium Punctatum, 70% of genes in gene families of size greater than one). Compared with 2,844 net gene families gained, which correlate with the origin of the Chlorophyceae, a phylogenetic analysis of these gene families suggests little protein innovation (110 net gene families) during the evolution of multicellularity (Fig. 2c). These same green algae genomes allowed analysis into Pfam domain innovation, which may correlate with the evolution of multicellularity. We found innovation of only nine Pfam domains correlating with the evolution of multicellularity (Fig. 2d, Supplementary Table 4). Interestingly, there is an excess of species-specific genes (Fig. 2a,c) and Pfam domains (Fig. 2d) compared with multicellularity-correlated genes and Pfam domains, suggesting that species-specific adaptations are more numerous than changes correlating with the evolution of multicellularity. We observe more evidence of species-specific, rather than multicellular-specific, protein innovations, suggesting species-specific adaptation (Fig. 2a,c,d, Supplementary Tables 5–8) rather than genome-wide differences correlating with the evolution of multicellularity. The evolution of multicellularity in the volvocine algae does not require large-scale genomic innovation.

Co-option of cell cycle regulation for multicellularity. Notably, we observe that the genetic innovation correlating with multicellularity, shared between Gonium and Volvox, evolved through co-option of existing developmental programs of cell cycle control. Volvocine algae have a common multiple-fission life cycle, with variation in timing and number of divisions (Supplementary Fig. 1)14. Like most eukaryotes, including plants and animals, their cell cycles are regulated by homologues of the retinoblastoma cell cycle regulatory pathway (Fig. 3). Notably, RB phosphorylate and regulate retinoblastoma (RB or MAT3 in the Volvocales), which in turn de-represses the cell cycle (Fig. 3a).
Although most of these regulators are nearly identical in Chlamydomonas, Gonium and Volvox (Fig. 3b,e), there are two notable differences. First, Volvox has a four gene expansion of cyclin D1 genes (Fig. 3c)\(^{17}\). As Volvox has tissue differentiation, these cyclin D1 genes may have been important for tissue development as is the case in metazoans and land plants\(^{22,23}\), supported by the fact that RB has moved into the mating locus of Gonium and Volvox and is differentially expressed between mating loci (Fig. 3d)\(^{24,25}\). However, the tandem array expansion of the cyclin D1 genes is also found in Gonium (Fig. 3c), where cyclin D genes display elevated dN/dS ratios compared with other cell cycle regulators (Supplementary Fig. 8), suggesting the function of these cyclin Ds may be important for the transition to undifferentiated colonies, rather than tissue differentiation.

Second, there is modification of the RB gene in Gonium and Volvox (Fig. 3d, Supplementary Figs 5–7). Protein dimers of cyclins and CDKs primarily regulate RB by phosphorylating serine or threonine residues, which is thought to regulate RB binding of chromatin via E2F/DP transcription factors\(^{11,26,27}\). Recently it has been shown in human cells that cyclin D and CDK4/6 regulate monophosphorylation of RB proteins for G\(_1\) phase-specific RB functions\(^{25}\). If similar in Gonium, this would suggest a role of the expanded cyclin D1 genes for regulating RB to express multicellularity-related genes during G\(_1\) phase. If the expanded cyclin D1 proteins found in Gonium regulate multicellular cell cycle changes, modification of cyclin D-CDK phosphorylation sites in RB is predicted. Indeed, the linker of the E2F/DP binding pocket of RB is shorter in Gonium and Volvox compared with Chlamydomonas (Fig. 3d, Supplementary Figs 5 and 6), potentially altering how RB binds to chromatin via E2F/DP. In addition, phosphorylation sites between the E2F/DP pocket region (RB-A and RB-B domains) and the conserved carboxy (C)-terminal domain are absent in Gonium and Volvox RB proteins (Fig. 3d). Interestingly, in animals the C terminus of RB is intertwined with E2F/DP and changes in the phosphorylation by cyclin–CDK complexes could also alter E2F/DP binding\(^{28}\). As these phosphorylation sites are absent in Gonium and Volvox RB proteins (Fig. 3d), this suggests that RB co-option for multicellularity may result in differences in locus-specific temporal expression of genes important for multicellularity during G\(_1\), such as cell–cell adhesion genes. Given the role the RB pathway plays in regulating the cell cycle in Chlamydomonas, its early modification found in Gonium, and its co-option for complex morphology in Volvox, RB pathway regulation might be a key step towards multicellularity in the volvocine algae.

To test whether RB modifications present in Gonium and Volvox (compared with Chlamydomonas) are unrelated to, cause or are a consequence of multicellularity, we expressed the Chlamydomonas\(^{13}\) and Gonium RB genes in a Chlamydomonas strain lacking its RB gene (rb, mat3–4 strain, Fig. 4a)\(^{9,11}\), using the promoter and terminator from the Chlamydomonas RB gene to ensure expression near wild-type levels (Fig. 4b,c)\(^{11}\). The Chlamydomonas RB gene rescues the small cell size defect in the rb mutant (HA-CrRB::rb, Fig. 4a), while the Gonium RB gene rescues the cell size defect and causes the Chlamydomonas rb mutant to become non-palmelloid colonial, ranging from 2 to 16 normal-sized cells (HA-GpRB::rb, Fig. 4a). Crossing RB gain-of-function transformed Chlamydomonas strain to a Chlamydomonas strain lacking DP1, a gene that dimerizes with E2F to anchor RB to chromatin\(^ {11}\), results in suppression of the colonial phenotype and large-sized cells (HA-GpRB::rb:dp1, Fig. 4a) consistent with the phenotype of the Chlamydomonas dp1 mutant itself\(^ {12}\). This demonstrates that the Gonium RB gene causes colonial multicellularity (Fig. 4a) through the RB pathway (Fig. 3a) and suggests differences in how RB binds chromatin and
regulates the expression of cell cycle related genes in *Gonium* and *Volvox* are important for co-option of these RB targeted genes for multicellularity (Figs 3 and 4). This gain-of-function demonstrates a causal link between cell cycle regulation and the group level during the evolution of multicellularity, emphasizing that multicellularity can evolve by co-option and modification of regulatory genes rather than extensive genomic differences or innovation.

**Volvox** innovations for morphological complexity. In *Volvox*, somatic differentiation is causally regulated by the *regA* gene cluster, a set of putative DNA-binding transcription factors thought to regulate chloroplast biogenesis\(^29–31\). The *regA* gene cluster is absent in *Chlamydomonas* and *Gonium* (Fig. 5a,b), but is present in diverse *Volvox ferrisii* and *Volvox gigas*\(^15\), suggesting early evolution and co-option of this cluster shortly after the split of *Gonium* and *Volvox* lineages (Fig. 1)\(^16\). *Astrephomene* may determine somatic cell fate through a different pathway than *Volvox* suggesting multiple evolutionary pathways and subsequent evolutionary consequences during the evolution of multicellularity. Indeed, undifferentiated multicellularity evolved once in the *Volvocales*\(^16\), while additional morphological complexity (for example, cellular differentiation and large *Volvox* body size) has repeatedly evolved, suggesting a relative ease to gain and lose additional complexity.

We investigated proteins related to morphological complexity, pherophorins and matrix metalloprotease (MMP) proteins, in the volvocine algae\(^17\). These proteins are hypothesized to produce ECM and break up cell wall components during reproduction in *Gonium* and other *Volvocales*\(^15,33,34\). While *Chlamydomonas* contains no ECM and *Gonium* contains little ECM, a *Volvox* spheroid is largely composed of ECM (Fig. 1). Pherophorin and MMP gene families are expanded in *Volvox* relative to *Chlamydomonas* (Fig. 5c)\(^17\). We found the expansion of pherophorins and MMP genes in *Volvox* (Fig. 5c, Supplementary Data 2 and 3) is not present in *Gonium*, though some species-specific expansion of MMP genes has occurred (Supplementary Data 2 and 3). While some expansion of ECM gene families in *Gonium* was expected\(^14\) to direct the cell wall layer synthesis of a *Gonium* colony, this layer may instead be directed through differential gene expression. Pherophorin, MMP expansion and cellular differentiation correlate with expanded organismal size rather than the origin of multicellularity, suggesting a subsequent step in the evolution of multicellularity.

**Discussion**

We have investigated the evolution of multicellularity in the volvocine algae by sequencing the genome of the undifferentiated *Gonium*. Despite morphological differences, it was known that the *Chlamydomonas* and *Volvox* genomes are strikingly similar, suggesting that multicellularity required few genetic innovations\(^17,35\). However, these two genomes, positioned at the extremes of volvocine morphology, were unable to resolve the tempo and mode\(^36\) of the evolutionary transition to multicellularity.

The evolution of multicellularity in the volvocine algae is thought to involve 12 morphological innovations (Supplementary Fig. 1)\(^15\). Five of these steps correlate with the evolution of cell groups, a period of rapid evolutionary change (tempo). This view emphasizes the importance, and subsequent modification, of innovations correlating with undifferentiated colonies (mode). Finding support for this view, we have generalized these 12 steps into three major phases (Fig. 6): the evolution of cell cycle regulation to form cooperative groups via cell–cell adhesion, the
Figure 5 | Genetic changes present in the Volvox genome. (a) Gene synteny near the regA gene cluster and closely related regA-like genes (bold). Chromosome or scaffold number is indicated. Conserved genes are linked by line segments for regA-like (thick, black) and neighbouring genes (thin, grey). (b) Phylogenetic relationships of regA-like genes. The tree is a midpoint root and bootstrap values above 70% are indicated. (c) Comparison of number of pherophorin and metalloprotease genes in Chlamydomonas, Gonium and Volvox.

Figure 6 | Conceptual model for the evolution of multicellularity. Multicellularity hinges on the evolution of cell cycle regulation in a multicellular context with subsequent evolution of cellular differentiation (here, cell size-based) and increased body size.

Evolution of increased organismal size, and the evolution of differentiated germ and soma cells. Having sequenced the genome of Gonium, along with the published genomes of Chlamydomonas and Volvox, we can now identify the genetic pathways associated with each of these steps. The evolution of undifferentiated colonies correlates with RB cell cycle regulatory pathway evolution (Figs 3 and 4), which is further modified as complexity increases in the Volvocales24. Increased organismal size toward Volvox correlates with an expansion of pherophorins and MMPs (Fig. 5c). Finally the evolution of the regA gene cluster underlies somatic differentiation (Fig. 5a,b). Future sequencing of additional Volvocales genomes should clarify the evolutionary steps required for the evolution of germ and soma. Our three-phase model for the emergence of multicellularity, supported by the genetic pathways important for their evolution, changes our understanding of the tempo and mode of multicellular evolution previously obscured in other taxa such as plants, fungi and animals due to genomic divergence (Fig. 6).

Interestingly, an emerging theme throughout the evolution of multicellularity is that the genetic basis for the evolutionary transition emerges much earlier than anticipated6,32. In plants and animals, RB proteins are important for regulating both cell proliferation and differentiation by highly complex locus interactions with chromatin and chromatin remodelling factors37,38. Our finding that the RB pathway was co-opted early for multicellularity in undifferentiated colonies suggests that the template for subsequent evolutionary innovations in developmental programs was laid out during the transition to undifferentiated multicellularity via RB and cell cycle modifications, rather than with emergence of germ and somatic cellular differentiation. Interestingly, RB has been further co-opted for a role in sexual differentiation in Volvox, where there are male- and female-specific isoforms of RB24. This suggests that the evolution of multicellular cell cycle regulation was a critical step for the evolution of multicellularity. By comparing the genomes of these three volvocine green algae, we have determined that the mechanism of multicellular evolution is primarily co-option and regulatory modification of existing genetic pathways39. Gene duplication forms the basis of subsequent multicellular innovations.

The genomic age is illuminating the genetic pathways that are important for the evolution of multicellularity in other organisms where genes such as cadherins and integrins in animals3,4 and cell wall biogenesis genes in plants6. These are roughly analogous to...
metalloproteases and pherophorins in the volvocine algae highlighting convergences on similar genetic innovations for multicellularity. The substantial innovation and expansion of transcription factors and signalling networks found in animals and plants3,6 is not present in the volvocine algae. However, the volvocine algae demonstrate the critical role of transcriptional regulation of the cell cycle by RB for the formation of undifferentiated colonies. RB proteins regulate the cell cycle of most eukaryotes11,29, and are tumour suppressors in humans30, suggesting a broader role for RB and cell cycle regulation during the evolution of multicellularity.

The implications of these findings are greater than simply identifying when genes evolved during the evolution of multicellularity. Theoretical work has emphasized the need for greater understanding of the origin of an accelerated group life cycle during the evolution of multicellularity12,32–34. The field has been concerned with the evolution of germ–soma division of labour as the defining step in the evolution of multicellularity12,32–34. Indeed, a recent review of animal multicellularity45 does not mention the importance of cell cycle regulation and group formation. The Gonium genome reflects the early evolution of cell cycle regulation (Figs 3 and 4) in undifferentiated groups, conserved and modified in differentiated Volvox, that is indicative of the emergence of colony level adaptations. We highlight an early and formative step, the co-option and expansion of cell cycle regulation, as important for the evolution of cooperative groups and impacting the evolution of more complicated body plans; one that may shed light on the evolutionary history of other multicellular innovations and evolutionary transitions.

Methods

Strain and genome sequencing. The Gonium pectorale strain K3-F3-4 (maintained at the Michigan Department of Natural Resources, NIES-2863 from the Microbial Culture Collection at National Institute for Environmental Studies, Tsukuba, Japan, http://mcc.nies.go.jp/) was used for genome sequencing. Gonium was grown in 200–300 ml VTAC media at 20 °C with a 14:10 light–dark cycle using cool-white fluorescent lights (165–175 μmol m−2 s−1).

For next-generation sequencing and construction of a fosmid library, total DNA was extracted. Sequencing libraries were prepared using the GS FLX Titanium Rapid Library Preparation Kit (F. Hoffmann-La Roche, Basel, Switzerland) and the TruSeq DNA Sample Prep Kit ( Illumina Inc., San Diego, CA, USA) were used in both GS FLX (F. Hoffmann-La Roche) and MiSeq (Illumina Inc.) machines. Newbler 2.6 was used to assemble the GS FLX reads. The GC contents of the fosmid library was constructed in-house using vector pKS030. The fosmid library (23,424 clones) and BAC library (18,048 clones, Genome Institute (CUGI), Clemson University, Clemson, SC, USA) were end-sequenced using a BigDye terminator kit v3 (Life Technologies, Carlsbad, CA, USA) analysed on automated ABI3730 capillary sequencers (Life Technologies).

Evidence-based gene prediction. Introns hint file generation was done through a two-step, iterative mapping approach using Bowtie/Tophat command-line and custom Perl scripts written by Mario Stanke as part of AUGUSTUS36, (available at: http://bowtie.bioinf.uni-greifswald.de/bowtie2/wiki/index.php?n=IncorpSangRNAseq.Tophat). AUGUSTUS version 2.6.1 was selected because its algorithm has been successfully tuned to predict genes in Chlamydomonas and Volvox genomes, which contain high GC content17. Reads were first mapped to the genome assembly with Tophat version 2.0.2 (ref. 47) and the raw alignments were filtered to create an initial (introns) hint file, which was subsequently provided to AUGUSTUS during gene prediction. An exon–exon junction database was generated from the initial AUGUSTUS prediction via a Perl script. The twice-mapped reads (once to the genome and once to the exon–exon sequences) were then merged, filtered and a final intron hint file was created. From this, the final gene prediction with AUGUSTUS was performed.

Pfam domain analysis. Diversity and abundance of Pfam domains were determined for all published green algae genomes. Chlorophyte genomes including Bathycoccus prasinus48, Chlamydomonas reinhardtii35, Chlorella variabilis35, Coccomyxa subellipsoidea C-169 (ref. 30), Micromonas pusilla CCMP1545 (ref. 51), Micromonas pusilla RCC299 (ref. 51), Ostreococcus lucimarinus53, Ostreococcus sp. RCC809 (available on the DOE Phytapase website, version 10.1) and Volvox carteri RCC299 (ref. 51) were searched using direct Pfam domain analysis. Pfam domain prediction with AUGUSTUS was performed. Results were then merged, filtered and a final intron hints file was created. From this, the final gene prediction with AUGUSTUS was performed.

Analysis of transcription-associated proteins. Transcription-associated proteins (TAPs) include transcription factors (enhance or repress transcription) and transcription regulators (proteins which indirectly regulate transcription such as scaffold proteins, histone modification or DNA methylation). We combined three TAP classification rules for plants; PlantTFDB32, PlantTFDB35 and PlantAPADB36 to make a set of classification rules for 96 TAP families. Conflicts between the three sets of rules were manually resolved using the rule that included more genes as transcription-associated proteins.

Each transcription family includes at least one, up to three, mandatory domains. Families may include up to six forbidden domains (that is, a gene G cannot be in family F if domain D is present); not all families have defined forbidden domains. All mandatory and forbidden domains were represented by a full-length, global, Hidden Markov Model (HMM) database52. All HMMs were not extracted from the Pfam database52,57,59. When HMMs were not available from the Pfam database, custom HMMs were made using multiple sequence alignments from PfTFDB32 and the HMM was calculated using HMMER version 3.0 (ref. 59) using ‘hmmbuild’ with default parameters and ‘hmmcalibrate—seed 0’. This approach was used to create a HMM database52,59 for both Volvox and 16S rRNA genes. Intron HMMs were set as the lowest score of a true positive hit using a ‘hmmscan’ search against several complete Chlorophyte genomes. Chlorophyte genomes including Bathycoccus prasinus48, Chlamydomonas reinhardtii35, Chlorella variabilis35, Coccomyxa subellipsoidea C-169 (ref. 30), Micromonas pusilla CCMP1545 (ref. 51), Micromonas pusilla RCC299 (ref. 51), Ostreococcus lucimarinus53, Ostreococcus sp. RCC809 (available on the DOE Phytapase website, version 10.1) and Volvox carteri RCC299 (ref. 51) were searched using ‘hmmscan’ to search the library of 103 domains against the predicted protein sequences. Analyses were replicated with both Volvox version 1 and version 2; however, both version 1 and version 2 domains were not qualitatively different and results from version 1 are provided (Supplementary Table S3). Support values were classified into a TAP family. Conflicts between multiple TAP families were resolved by assigning the gene to the TAP family with the highest score (Supplementary Table 1).

Construction of protein families. Protein families were created using OrthoMCL60 with a variety of inflation values ranging from 1.2 to 4.0 in steps of 0.1 (Supplementary Figs 16–17). This analysis was performed using Chlorophyte genomes available on the DOE JGI Phytapase website, version 10.1 including Bathycoccus prasinus48, Chlamydomonas reinhardtii35, Chlorella variabilis35, Coccomyxa subellipsoidea C-169 (ref. 30), Micromonas pusilla CCMP1545 (ref. 51), Micromonas pusilla RCC299 (ref. 51), Ostreococcus tauri52,53, Ostreococcus lucimarinus53, Ostreococcus sp. RCC809 (available on the DOE JGI Genome Initiative website) and Volvox carteri version 2. These analyses were repeated for both Volvox version 1 and Volvox version 2. The inflation value of 1.9 was used for both analyses for consistency and was chosen to have relatively large, coarser grained clusters that were robust to higher inflation values (Supplementary Figs 16–19). To avoid bias introduced by not including all genes for each species, genes not assigned to a gene family (singletons) were assigned to singleton gene families and included in all subsequent phylogenetic gene family analyses.

A species tree was calculated by extracting OrthoMCL gene families containing only one copy in each species, for a total of 1,457 genes. The OrthoMCL run with an inflation value of 1.9 was chosen to use larger, coarser grained clusters, thus increasing the likelihood of capturing clusters that are true 1:1:1 orthologues. Given the relatively high gene content in any species) resulting in 1:1:1 orthologues. Given the relatively high gene duplication rates in volvocine algae (data not shown), these strict criteria support
an interpretation of 1:1 orthology. Genome-wide pairwise comparisons of dN, dS and dN/dS were calculated (Supplementary Fig. 21; Supplementary Table 11) using PAML and codeml analysis48 based on nucleotide translation based alignments (proteins were aligned using MUSCLE49).

**Prediction of lineage-specific genes.** The phylostratigraphy method20 assumes Dollo’s parsimony (that is, it is more likely that a gene observed in two distant clades was present in the common ancestor and multiple independent gains are not possible). This provides an entry point for testing evolutionary hypotheses related to the age of genes and to quantify how much gene-level innovation has occurred along each phylogenetic branch. Old genes are classified in low phylostrata (present in distant species, PS1–PS7) and young genes are classified in higher phylostrata (for example, genus- or species-specific genes, PS8–PS9). The resolution of each phylostratum strictly depends on the availability of reliable outgroups (the availability of reliable genomic outgroups is relatively low in Chlorophyte algae).

The phylostrata for each species were defined from those in each NCBI Taxonomy entry for Chlamydomonas, Gonium and Volvox, resulting in nine expected phylostrata for each species. All proteins were subjected to a BLASTP search with an E-value threshold of 0.001 against the NCBI nr database. Placement in phylostrata was determined using a simple threshold of 0.95 against the NCBI nr database. The phylogenetic classes were defined from those in each NCBI Taxonomy entry for Chlamydomonas, Gonium and Volvox, resulting in nine expected phylostrata for each species. All proteins were subjected to a BLASTP search with an E-value threshold of 0.001 against the NCBI nr database. Placement in phylostrata was determined using a simple threshold of 0.95 against the NCBI nr database.

**Chlamydomonas strains culture conditions.** Wild-type Chlamydomonas reinhardtii 6145 and 21gr, and HA-CrRB (HA-MAT3=tet3-4, here referred to as HA-CrRB-tet), mat3-4 (here referred to as tet), and dp1 have been previously described50, 51. Wild-type strains 6145, 21gr (mt1), and 21gr (mt2) are mating pairs that have been back crossed to eliminate the y1 mutation in 6145 (ref. 10). The BB knockout strain has been previously characterized as a null allele, and the knockout mutation is the rr allele10, 11. The rb mutation can be complemented by a amino (N)-terminally tagged version of the gene that behaves identical to wild type. Previously, a knockout mutation in the Chlamydomonas Dpi gene, dp1, was identified and characterized10, 11. All the strains were maintained on TAP plates.

**Cloning of Gonium pectorale RB and transformation into rb.** A 3X haemagglutinin (HA) tagged copy of the Gonium pectorale RB gene was cloned using InFusion Cloning (Clontech) to be driven by the Chlamydomonas RB promoter and terminator that includes a AphVIII selectable marker for Chlamydomonas transformation (Fig. 4, ref. 11). Gonium pectorale genomic DNA from KF43 was used as a template and the genomic region of RB was amplified with its ATG start codon using the primers 5'-CAGATTACGCTACTAGATCTGCCGAAGCTG-3' and 5'-CAGATTACGCTACTAGATCTGCCGAAGCTG-3'. These primers overlap with the 3X HA tag and 3X terminator from the previously created HA-CrRB transformation clone that complements the rb mutation10. The HA-CrRB plasmid was amplified by inverse PCR with 5'-TCTAGATTACGCTACTAGATCTGCCGAAGCTG-3' and 5'-CAGATTACGCTACTAGATCTGCCGAAGCTG-3' primers. PCR products were gel purified with a QiAquick gel extraction kit (Qiagen). Purified PCR fragments were fused by InFusion (Clontech) cloning based on overlaps in the amplified sequences and transformed into chemically competent DH5-alpha cells, after which the clone was confirmed by sequencing.

**Transformation of Chlamydomonas reinhardtii.** The rb strain was transformed with glass beads29, with the HA-GpRB clone (above) and as a control with HA-CrRB-pS1015 (AphVIII selectable marker only) and selected on TAP plates supplemented with 20 μg ml⁻¹ paromycin51. Candidate strains were screened by growth morphology51, 52 and then screened for expression by immunoblotting with an anti-HA antibody (Roche 3F10, high affinity53). Four independent strains expressing the HA-GpRB, and five independent strains expressing HA-CrRB were created. Control complementation of the rb mutation with HA-CrRB occurred at rates similar to previous results10. The presence of the rb mutation was confirmed by replica plating on TAP plates supplemented with 10 μg ml⁻¹ emetine51, 53.

**Genetic analysis of HA-GpRB-expressing strains.** Two lines expressing HA-GpRB were crossed to a dp1 null mutation10. Because both the HA-GpRB and dp1 mutations are linked to AphVIII, single tetras were dissected. HA-GpRB was genotyped with primers in the XHA tag 5'-AGGTCTAACAGCATGTTACTC-3' and 3'-CGGCCAACAGCGGGCTGAGA-3'. The dp1 mutation was genotyped as previously described10.

**Immunoblotting HA-GpRB and HA-CrRB strains complementing rb.** Whole-cell lysates from strains were prepared, separated and immunoblotted51. Briefly, the anti-HA antibody used for detection of HA-GpRB and HA-CrRB was an anti-HA high affinity monoclonal antibody (clone 3F10, Roche) and anti-alpha-tubulin monoclonal antibody (Sigma), as previously described11. The expression levels of RB in HA-CrRB strains have been previously shown to be similar to wild-type Chlamydomonas11. The expression levels of RB in HA-GpRB are similar, if not slightly below, the expression levels of HA-CrRB, suggesting that overexpression of RB is not causing the observed colonial phenotype, but rather modification to the Gonium RB gene.

**Measurement of cell or colony size distribution.** The size of cells and groups of cells was measured with a Moxi Z automated cell sizer/counter using type ‘S’ cassettes (ORFLO Technologies). Sizing is based on the Coulter principle used previously with Chlamydomonas reinhardtii10, 11.

**References**

1. Bonner, J. T. The origins of multicellularity. *Integr. Biol.* 1, 27–36 (1998).

2. Grosberg, R. K. & Strathmann, R. R. The evolution of multicellularity: a major evolutionary transition? *Annu. Rev. Ecol. Evol. Syst.* 38, 621–654 (2007).

3. Suga, H. et al. The Capsaspora genome reveals a complex unicellular prehistory of animals. *Nat. Commun.* 4, 2325–2333 (2013).

4. King, N. et al. The genome of the chonooangellate *Monosiga brevicollis* and the origin of metazoa. *Nature* 451, 783–788 (2008).

5. Parfrey, L., Lahr, D. J. G., Knoll, A. H. & Katz, L. A. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl Acad. Sci. USA* 108, 13624–13629 (2011).

6. Hori, K. et al. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* 5, 1–9 (2014).

7. Herron, M. D., Hackett, J. D., Aylward, F. O. & Michod, R. E. Triassic origin and early radiation of multicellular volvocine algae. *Proc. Natl Acad. Sci. USA* 106, 3254–3258 (2009).

8. Kirk, D. L. *Volvox*: Molecular-Genetic Origins of Multicellularity and Cellular Differentiation (Cambridge Univ. Press, 1998).

9. Umen, J. G. & Goodenough, U. W. Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. *Genes Dev.* 15, 1652–1661 (2001).

10. Fang, S.-C., de los Reyes, C. & Umen, J. G. Cell size checkpoint control by the retinoblastoma tumor suppressor pathway. *PLoS Genet.* 2, e167 (2006).

11. Olson, B. J. S. C. Regulation of the *Chlamydomonas* cell cycle by a stable, chromatin-associated retinoblastoma tumor suppressor complex. *Plant Cell* 22, 3331–3347 (2010).

12. Shelton, D. E. & Michod, R. E. Group selection and group adaptation during a major evolutionary transition: insights from the evolution of multicellularity in the volvocine algae. *Biol. Theory* 9, 452–469 (2014).

13. Harper, J. D. I. & John, P. C. L. Coordination of division events in the volvocine algae. *Annu. Rev. Ecol. Syst.* 45, 106, 3254–3258 (2011).

14. Lynch, M. & Conery, J. S. The origins of genome complexity. *Science* 302, 1401–1404 (2003).

15. Smith, D. R. et al. *Organelle genome complexity scales positively with organism size in volvocine green algae*. *Mol. Biol. Evol.* 30, 793–797 (2013).

16. Domazet-Loso, T., Brjakovic, J. & Tautz, D. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet.* 23, 533–539 (2007).

17. Lang, D. et al. Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: a timeline of loss, gain, expansion, and correlation with complexity. *Genome Biol. Evol.* 2, 488–503 (2010).

18. Beumer, T. L., Roepers-Gajadien, H. L., Gademann, I. S., Kal, H. B. & de Rooij, D. G. Involvement of the D-type cyclins in germ cell proliferation and differentiation in the mouse. *Biochim. Biophys. Acta Reprod. 63, 1893–1898 (2000).

19. Dewitte, W. et al. Altered cell cycle distribution, hyperplasia, and inhibited differentiation in *Arabidopsis* caused by the D-type cyclin *CYCD3*. *Plant Cell* 15, 79–92 (2003).
24. Ferris, P. J. et al. Evolution of an expanded sex-determining locus in Volvox. Science 328, 351–354 (2010).
25. Hanschen, E. et al. Evolution of male-female sexual dimorphism predates the gender-based divergence of the mating locus gene MAT3/BR. Mol. Biol. Evol. 30, 1038–1040 (2013).
26. Khidr, L. & Chen, P.-L. RB, the conductor that orchestrates life, death and differentiation. Oncogene 25, 5210–5219 (2006).
27. Narasimha, A. M. et al. Cyclin D activates the Rb tumor suppressor by monophosphorylation. EMBO J. 3, 1–21 (1984).
28. Rubín, S. M., Gall, A. L., Zheng, N. & Pavletich, N. P. Structure of the Rb C-terminal domain bound to E2F1-DR1: a mechanism for phosphorylation-induced E2F release. Cell 123, 1093–1106 (2005).
29. Meissner, M., Stark, K., Cresnar, B., Kirk, D. L. & Schmitt, R. Volvox germline-specific genes that are putative targets of RegA repression encode chloroplast genes. Genome Res. 16, 363–370 (1999).
30. Kirk, M. M. et al. regA, a Volvox gene that plays a central role in germl-cell differentiation, encodes a novel regulatory protein. Development 126, 639–647 (1999).
31. Duncan, L. et al. The VARG family gene and the evolutionary origins of the master cell-type regulatory gene, regA, in Volvox carteri. J. Mol. Evol. 65, 1–11 (2007).
32. Hanschen, E. R., Ferris, P. J. & Michod, R. E. Early evolution of the genetic basis for soma in the Volvocacae. Evolution 68, 2014–2025 (2014).
33. Hallmann, A. Extracellular matrix and sex-inducing pheromone in Volvox. Int. Rev. Cytol. 227, 131–182 (2003).
34. Hallmann, A. The pherophorins: common, versatile building blocks in the evolution of extracellular matrix architecture in Volvocales. Plant J. 45, 292–307 (2006).
35. Merchant, S. S. et al. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318, 245–250 (2007).
36. Simpson, G. G. Tempo and Mode in Evolution (Columbia Univ. Press, 1944).
37. Sage, J. The retinoblastoma tumor suppressor and stem cell biology. Genes Dev. 26, 1409–1420 (2012).
38. Kwabara, A. & Gruissem, W. Arabidopsis Retinoblastoma-related and Polycomb group proteins: Cooperation during plant cell differentiation and development. J. Exp. Bot. 65, 2667–2676 (2014).
39. Kapheim, K. M. et al. Genomic signature of evolutionary transitions from solitary to group living. Science 348, 1139–1144 (2015).
40. Hanschen, E. R., Shelton, D. E. & Michod, R. E. in Evolutionary transitions to multicellular life: principles and mechanisms. (eds Nedelcu, A. M. & Ruiz-Trillo, I.) 165–188 (Springer, 2015).
41. Libby, E. & Rainey, P. B. A conceptual framework for the evolutionary origins of multicellularity. Phys. Biol. 10, 035001 (2013).
42. Rainey, P. B. & Kerr, B. Cheats as first propagules: a new hypothesis for the evolutionary origin of multicellularity. Proc. Natl Acad. Sci. USA 9113–9117 (2006).
43. Michod, R. E. The group covariance effect and fitness trade-offs during evolutionary transitions in individuality. Proc. Natl Acad. Sci. USA 103, 159–205 (Wiley, 2008).
44. Nozaki, H. Origin and evolution of the genera Pleodorina and Volvox (Volvocales). Biologia 58, 425–431 (2003).
45. Leliäert, F. et al. Phylogeny and molecular evolution of the green algae. CRC Crit. Rev. Plant Sci. 31, 1–46 (2012).
46. Nozaki, H. & Itoh, M. Phylogenetic relationships within the colonial Volvocales (Chlorophyta) inferred from clades analysis based on morphological data. J. Phycol. 30, 353–365 (1994).
47. Csurös, M. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. Bioinformatics 26, 1910–1912 (2010).
48. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591 (2007).
49. Liu, K., Baghavan, S., Nelsen, S., Linder, C. R. & Warnow, T. Rapid and accurate large-scale coestimation of sequence alignments and phylogenetic trees. Science 324, 1561–1564 (2009).

Acknowledgements
We acknowledge the support of the National Aeronautics and Space Administration (grant number NNX13AH14G to R.E.M., P.J.F. and P.M.D.), the National Institute of Health (grant number GM084905 to the University of Arizona, P20GM103638 to B.J.S.C.O., P20GM103418 to B.J.S.C.O. and T.N.N., GM103785-02 to M.W.K.), the National Science Foundation (grant number MCB-1412395 to B.J.S.C.O. and R.E.M.), grant number DGE-0664435 to the University of Arizona and PHY11-29515 to U.C.S.B., K.I.T.P.), the KSU Johnson Cancer Center (Innovative Research Award to B.J.S.C.O.), MEXT/JSPS KAKENHI Scientific Research on Innovative Areas ‘Genome Science’ (grant number 22150002 to A.T.), MEXT/JSPS KAKENHI Scientific Research (A) (grant number 24247042 to H.N.), JSPS Fellows (19-7661 and 23-5499 to T.H.) and the JSPS Postdoctoral Fellowship for Research Abroad (26-495 to T.H.). Part of this work was undertaken under the U.S.C.B. K.I.T.P. ‘Cooperation and the Evolution of Multicellularity’ program. Part of the computing for this project was performed on the Beocat Research Cluster at KSU, which is funded in part by NSF grants CNS-1006860, EPS-1006860 and EPS-0919443. We thank M. Barker and M. Sanderson for technical assistance; A. M. Nedelcu for comments; three anonymous reviewers for comments and all the technical staff of the Comparative Genomics lab at the National Institute of Genetics for their assistance. Publication of this article was funded in part by the Kansas State University Open Access Publishing Fund.

Author contributions
E.R.H. conceived the project, performed the analyses and wrote the manuscript; T.N.M. did gene model prediction and annotation and prepared the RNA samples; T.H. prepared gDNA; A.T, A.E., H.N., Y.M., M.S. and H.K.T. did assemblies; P.J.F., D.R.S., H.S. and E.R.M. did gene model prediction and annotation and prepared the RNA samples; T.S. and H.K.-T. did assemblies; P.J.F., D.R.S., H.S. and E.R.M. performed transformations and wrote the manuscript; all the authors contributed to preparation of the manuscript and the text.

Additional information
Accession codes: This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession LSV90008000. The version described in this paper is version LSV90008000. Data are also available via JGI Phytozome.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.
Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Hanschen, E. R. et al. The Gonium pectorale genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. Nat. Commun. 7:11370 doi: 10.1038/ncomms11370 (2016).