Additional data file 1. Supplementary Figures:

Figure S1. Cell division concentrations in cultures over the day-night cycle during the period of harvesting. Each symbol type represents a separate culture. The x-axis shows time (in hours) since dawn (lights on) on the first day of harvesting. The dark period (lights off) is shown by the grey box. A) Two 1N cultures grown on 14:10 light:dark over the whole experiment. RNA pools from one of these was used in 1N library construction and samples from both were used in RT-PCR tests. B) Four 2N cultures grown on 14:10 LD. RNA pools from one of these was used in 2N library construction and three were chosen at each time point for RT-PCR testing. C) Two 1N cultures grown on 14:10 LD that experienced a failure of the light regulation one day prior to harvesting, causing them to grown in continuous light. These were not used in library construction but both were used for RT-PCR testing. Only the first 8 time points were used in pools of RNA for library construction and equal quantities of total RNA were mixed to give 100 µg of total RNA in each pool.

Figure S2. Bioanalyzer virtual gel image showing high quality of total RNA. RNA was tested after isolation, DNase treatment, and purification using the Qiagen RNeasy kit RNA clean-up protocol. The distinct 28s, 18s rRNA bands show lack of detectable degradation.

Figure S3. Histogram of EST reads per cluster. Distribution of ESTs across all clusters represented in the 1N library and all clusters represented in the 2N library.

Figure S4. Distribution of clusters by KOG functional class and library for clusters composed of ≥2 EST reads. * indicates KOG classes displaying significant differences between 1N-unique and 2N-unique clusters and @ indicates KOG classes displaying significant differences between 1N-unique and shared clusters.

Figure S5. Taxonomic distribution of clusters “best-hitting” to either Viridiplantae, Stramenopiles, or Metazoa. In each column of numbers, the top row refers to shared clusters, the second row refers to 1N-unique clusters, and the bottom row refers to 2N-unique clusters. Significant differences determined by Fishers Exact Test are marked (*, shared clusters different from both 1N unique and 2N unique; †, 1N-unique clusters different from both shared clusters and 2N-unique clusters; ‰, 2N-unique clusters different from both shared clusters and 1N-unique clusters).

Figure S6. Example images of full gels showing absence of detectable genomic contamination in RT- samples. Top: GS00217 (Elongation factor 1 α), expressed in both cell types, amplified with primer pair e00217F2, e00217R2. Middle: GS09822 (GPA), expressed in both cell types, amplified with primer pair e09822F2/e09822R1. Bottom, GS02894, strongly differential expression in 1N (amplified with primer pair e02904F1/e02993). RT- reactions were tested in parallel with RT+ reactions with a total of 26 different primer combinations and in no case was a signal detected from RT-reactions.
Figure S7. MUSCLE alignment of phototropin homologs over the LOV2 domain
Included also is a sensor hybrid histidine kinase from the planctomycete bacterium
*Gemmata obscuriglobus* identified as the second highest hit to the GS00132 predicted
amino acid sequence in the nr protein database (E-value = 8e-22). The top hit to GS00132
was a PAS domain-containing predicted protein from *Ostreococcus tauri*
(emb|CAL55375.1) (E-value = 4e-23) but it was less conserved over the LOV2 domain.
The following sequences were obtained from Genbank: At-PHOT2
(>gi|145362057|ref|NP_851212.2| PHOT2 (NON PHOTOTROPIC HYPOCOTYL 1-LIKE); kinase [Arabidopsis thaliana]), At-PHOT1
(>gi|15231245|ref|NP_190164.1| PHOT1 (phototropin 1); kinase [Arabidopsis thaliana]),
Gemob (>gi|168702150|ref|ZP_02734427.1| multi-sensor hybrid histidine kinase
[Gemmata obscuriglobus UQM 2246]). Chlre3 (>estExt_fgenesh2_kg.C_190109
[Chlre3:183965]) was obtained from the *C. reinhardtii* protein catalog.

Figure S8. RT-PCR tests of expression pattern of further genes chosen by digital
expression and not shown in Fig. 7-10 of the main test.

Figure S9. MUSCLE alignment of GS00273 and Myb transcription factor family
members. The following sequences were obtained from Swissprot:
>gi|127591|sp|P01103.1|MYB_CHICK, >gi|1709195|sp|P52550.1|MYBA_CHICK ,
>gi|417333|sp|Q03237.1|MYBB_CHICK , >gi|75336839|sp|Q9S7L2.1|MYB98_ARATH,
>gi|75336831|sp|Q9S7G7.1|MB3R1_ARATH, >gi|127590|sp|P01104.2|MYB_AVIMB.
Conserved regularly repeating tryptophan residues are marked with arrows. The fourth
tryptophan (the first in the R3 domain) has been substituted conservatively by
phenylalanine in GS00273. This tryptophan is often substituted by other aromatic
or hydrophobic residues in other plants (Martin and Paz-Ares, 1997).

Figure S10. CLUSTAL alignments of predicted amino acid sequences of clusters
with GPA homology against GPA. A) Frame 3 ORF in GS09822 against GPA. B)
Longest ORF of GS02894 translated frame 1 against GPA. C) A very strong homology to
GPA is found over a 30 amino acid section of translated frame 2 of GS02894 against
GPA, yet this frame is interrupted by multiple stop codons (marked in red), including two
surrounding a methionine (potential start, marked in green). • marks perfect match, :
marks strong conservation, and . marks weak conservation. The predicted EF-hand motif
in GPA is underlined. The GPA sequence used was obtained from Genbank
(>gi|4102565|gb|AAD01505.1| putative calcium binding protein [Emiliania huxleyi]).

Figure S11. MUSCLE alignment of VCX1 homologs. The following sequences were
obtained from Swissprot: VCX1_SCHPO (>gi|74582235|sp|O59768.1|VCX1_SCHPO
RecName: Full=Vacuolar calcium ion transporter), VCX1_YEAST
(>gi|74623660|sp|Q99385.1|VCX1_YEAST RecName: Full=Vacuolar calcium ion
transporter; AltName: Full=Vacuolar Ca(2+)/H(+) exchanger), CAX2_ARATH
(>gi|122056116|sp|Q39254.2|CAX2_ARATH RecName: Full=Vacuolar cation/proton
exchanger 2), CAX5_ARATH (>gi|75154113|sp|Q8L783.1|CAX5_ARATH RecName:
Full=Vacuolar cation/proton exchanger 5).
Figure S12. MUSCLE alignment of partial sequences of V-type ATPase V0a subunits. The highly conserved arginine (R735 in VPH1_YEAST) is marked with an arrow. Two other clusters, GS08326 and GS12017, had hits against the KOG database to V-type ATPase V0a subunits but did not have strong homologs in the Uniprot or Swissprot databases and did not align over the conserved region shown here. The following sequences were obtained from Swissprot:
gi|1711568|sp|P37296.2|STV1_YEAST, gi|418296|sp|P32563.3|VPH1_YEAST, gi|3929395|sp|Q01290.1|VPH1_NEUCR, gi|3929385|sp|O13742.1|VPH1_SCHPO, gi|182702220|sp|Q54E04.2|VATM_DICI, gi|59803038|sp|Q93050.3|VPP1_HUMAN, gi|172046607|sp|Q9Y487.2|VPP2_HUMAN, gi|12643719|sp|Q13488.2|VPP3_HUMAN, gi|38372616|sp|Q9HBG4.1|VPP4_HUMAN, gi|15226542|ref|NP_179736.1|VHA-A2 [Arabidopsis thaliana], gi|30683925|ref|NP_850122.1|VHA-A1 [Arabidopsis thaliana], gi|18420373|ref|NP_568051.1|VHA-A3 [Arabidopsis thaliana], gi|74502607|sp|Q5JDS2.1|VATI_PYRKO [Thermococcus kodakarensis].

Figure S13. MUSCLE alignment of histone H4 homologs. Possible initiator methionines of GS02435 are marked with arrows. The following sequences were obtained from Swissprot: gi|28202123|sp|P59259.2|H4_ARATH [Arabidopsis thaliana], gi|51317339|sp|P62805.2|H4_HUMAN, >gi|74752149|sp|Q99525.1|H4G_HUMA, gi|122107|sp|P02309.2|H4_YEAST, gi|59799579|sp|P69152.2|H42_TETTH [Tetrahymena thermophila]. To increase phylogenetic coverage the following sequences were obtained from the Genbank nr protein dataset: gi|72389584|ref|XP_845087.1|histone H4 [Trypanosoma brucei TREU927], gi|219116983|ref|XP_002179286.1|histone H4 isoform 1b [Phaeodactylum tricornutum CCAP 1055/1], gi|159464912|ref|XP_001690685.1|histone H4 [Chlamydomonas reinhardtii], gi|159480048|ref|XP_001698098.1|histone H4 variant [Chlamydomonas reinhardtii].

Figure S14. Partial MUSCLE alignment of histone H2A homologs. The C terminals, which are extremely variable in length and sequence, are not included. Two mini-clusters that composed cluster GS07501 differed slightly in predicted aa sequence. These are shown by e07501.1 and e07501.2. The following Arabidopsis thaliana, Homo sapiens, C. reinhardtii, and Tetrahymena thermophila sequences were obtained from Swissprot: gi|75306451|sp|Q94F49.1|H2A5_ARATH, gi|75276926|sp|O04848.1|H2AXA_ARATH, gi|75313113|sp|Q9S9K7.1|H2AXB_ARATH, gi|75308805|sp|Q9C681.1|H2A1_ARATH, gi|75311179|sp|Q9LHQ5.1|H2A2_ARATH, gi|75311051|sp|Q9LD28.1|H2A6_ARATH, gi|75279005|sp|O81826.1|H2A3_ARATH, gi|75309136|sp|Q9FJE8.1|H2A7_ARATH, gi|75311717|sp|Q9LZ46.1|H2A4_ARATH, gi|75313476|sp|Q9SI0.1|H2AV2_ARATH, gi|75277395|sp|O23628.1|H2AV1_ARATH, gi|75308904|sp|Q9C944.1|H2AV3, gi|75314165|sp|Q9T0H7.1|H2A8_ARATH, gi|74752099|sp|Q96QV6.3|H2A1A_HUMAN, gi|74750623|sp|P8IUE6.3|H2A2B_HUMAN, gi|121992|sp|P16104.2|H2AX_HUMAN, gi|12643341|sp|Q93077.3|H2A1C_HUMAN, gi|47117890|sp|Q16777.4|H2A2C_HUMAN, gi|74733131|sp|Q9BM1.1|H2AJ_HUMAN, gi|74751984|sp|Q96KK5.3|H2A1H_HUMAN.
Four diatom sequences were obtained from *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* via the Genbank nr protein database. H2A_Tpseud (gi|209585850|gb|ACI64535.1), H2A-1_Ptric (gi|219122004|ref|XP_002181345.1| histone H2A isoform 1), H2A-2_Ptric (gi|219119185|ref|XP_002180358.1| histone H2A isoform 2), H2A-3_Ptric (gi|219116815|ref|XP_002179202.1| histone H2A isoform 3a).

**Figure S15. Phylogenetic tree classifying different H2A homologs.** MUSCLE-aligned sequences were curated with GBLOCKS and aligned using the PhyML method and the Approximate Likelihood Ratio Test. Branch nodes with support <50% have been collapsed.
Supplementary Figure S1.
Supplementary Figure S2. Bioanalyzer analysis of total purified RNA samples pooled for construction of 1N and 2N libraries.
Supplementary Figure S3.
Posttranslational modification, protein turnover, chaperones
General function prediction only
Signal transduction mechanisms
Function unknown
Translation, ribosomal structure and biogenesis
Carbohydrate transport and metabolism
Energy production and conversion
Intracellular traffic, secretion and vesicular transport
Amino acid transport and metabolism
Lipid transport and metabolism
RNA processing and modification
Transcription
Inorganic ion transport and metabolism
Cytoskeleton
Secondary metabolites biosynthesis, transport, catab.
Replication, recombination, and repair
Coenzyme transport and metabolism
Chromatin structure and dynamics
Nucleotide transport and metabolism
Cell cycle control, division and chromosome partition
Cell wall/membrane/envelope biogenesis
Defense mechanisms
Nuclear structure
Extracellular structures
Cell motility

Supplementary Figure S4.
Supplementary Figure S5.
Supplementary Figure S6.
Supplementary Figure S7.
Supplementary Figure S8.
Supplementary Figure S9.
Supplementary Figure S10.
Supplementary Figure S10 (continued).
Supplementary Figure S11.
Supplementary Figure S13.
Supplementary Figure S14.
Supplementary Figure S15.