Cell

# Early Sexual Origins of Homeoprotein Heterodimerization and Evolution of the Plant KNOX/BELL Family

Jae-Hyeok Lee,[1] Huawen Lin,[1] Sunjoo Joo,[1] and Ursula Goodenough[1,*]
[1]Department of Biology, Washington University, St. Louis, MO 63130, USA
*Correspondence: ursula@biology.wustl.edu
DOI 10.1016/j.cell.2008.04.028

## SUMMARY

**Developmental mechanisms that yield multicellular diversity are proving to be well conserved within lineages, generating interest in their origins in unicellular ancestors. We report that molecular regulation of the haploid-diploid transition in *Chlamydomonas*, a unicellular green soil alga, shares common ancestry with differentiation pathways in land plants. Two homeoproteins, Gsp1 and Gsm1, contributed by gametes of *plus* and *minus* mating types respectively, physically interact and translocate from the cytosol to the nucleus upon gametic fusion, initiating zygote development. Their ectopic expression activates zygote development in vegetative cells and, in a diploid background, the resulting zygotes undergo a normal meiosis. Gsm1/Gsp1 dyads share sequence homology with and are functionally related to KNOX/BELL dyads regulating stem-cell (meristem) specification in land plants. We propose that combinatorial homeoprotein-based transcriptional control, a core feature of the fungal/animal radiation, may have originated in a sexual context and enabled the evolution of land-plant body plans.**

## INTRODUCTION

In the eight major eukaryotic radiations (Baldauf, 2003), only the plants, fungi, and animals have yielded widely distributed multicellular descendents. Many key genetic elements responsible for establishing growth and differentiation patterns in multicellular organisms have been found in the genomes of their unicellular or simpler ancestors (Floyd and Bowman, 2007; King et al., 2008; Meyerowitz, 2002), suggesting that the tools to build multicellular forms were "invented" by unicellular forebears, but much remains to be learned about the contexts in which these ancestral tools were used.

The green-plant lineage (Viridiplantae) diverged at least one billion years ago to follow two evolutionary pathways (Figure S1 available online; Lewis and McCourt, 2004): the Chlorophyta, including most of the modern green algae, and the Streptophyta, including the charophyte algae and the land plants (including mosses and ferns). Multicellularity has repeatedly evolved in both lineages, usually as simple colonial or filamentous forms (Graham and Wilcox, 2000). The land-plant form, with multiple-layered upright sporophytes, apparently evolved only once, in charophycean-like ancestors (Karol et al., 2001; Qiu et al., 2006), and presumably played a crucial role in the successful colonization of land.

Since the genetic/genomic resources of the charophytes are limited, study of the evolutionary origins of land plants currently focuses on the Chlorophyta, the assumption being that genetic elements shared by modern chlorophytes and streptophytes were also present in their common unicellular green ancestors (Bowman et al., 2007). The chlorophycean *Chlamydomonas reinhardtii* is particularly informative since its complete genome sequence is available (Merchant et al., 2007) and its sexual life cycle is well characterized (Figure 1; Goodenough et al., 2007). The present report focuses on control of haploid-diploid transitions during this life cycle, which we demonstrate to be mediated by a pair of homeoproteins (homeobox-containing transcription factors).

Genes encoding homeoproteins, identified in all eight eukaryotic radiations except the Alveolata (Derelle et al., 2007), have diversified into two superclasses, TALE (Three Amino acid Length Extension) and non-TALE; animal/plant genomes encode 40–250 members (Nam and Nei, 2005; Shiu et al., 2005). Differential expression of homeoproteins in space and time, and selective heterodimerization of homeoprotein pairs, correlate with numerous differentiation events in plant and animal development (Chan et al., 1998; Mann and Morata, 2000). Our study relates to the KNOX (KNOTTED-like homeobox) and BELL (BEL-Like) class members within the TALE superclass that are expressed in discrete domains of morphogenic tissues, interact via genetic/physical networks, and play key roles in the morphogenesis and developmental transitions of vascular plants (reviewed in Hake et al., 2004; Scofield and Murray, 2006).

Previous studies of *C. reinhardtii* reported that *plus* gametes express a homeoprotein, Gamete-specific *plus*1 (Gsp1), and that when *GSP1* is ectopically expressed in *minus* gametes, the cells transcribe genes that are normally transcribed only after gametes fuse to form diploid zygotes (Kurvari et al., 1998; Zhao et al., 2001). Transcription of these genes is not suppressed by protein-synthesis inhibitors (Ferris and Goodenough, 1987),

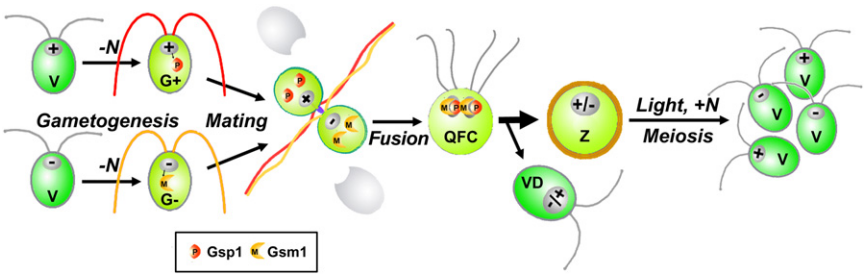


**Figure 1. Sexual Life Cycle of *Chlamydomonas reinhardtii***

(A) Nitrogen starvation (*-N*) induces vegetative (mitotic) cells (V) to differentiate into *plus* and *minus* gametes (G+ and G−), controlled by the two mating-type loci (+ and −).

(B) Upon mixing, agglutinins mediate flagellar adhesion; a rise in intracellular cAMP triggers cell-wall loss and mating-structure activation.

(C) Activated mating structures fuse to form quadriflagellated cells (QFCs). Gsp1 (P) and Gsm1 (M) homeoproteins heterodimerize, translocate into both nuclei, and turn on the zygote program (this report).

(D) Nuclear fusion and secretion of zygote-specific cell-wall proteins results in a thick-walled zygote (Z). A few QFCs resume growth as heterozygous vegetative diploids (VD).

(E) Mature zygotes germinate in response to light and N provision and undergo meiosis.

suggesting that a P factor presynthesized by the *plus* gamete, and an M factor presynthesized by the *minus* gamete, combine when the gametes fuse and initiate the zygote program (Goodenough et al., 1995), with Gsp1 corresponding to the P factor.

Here, we document that the postulated M factor is a second homeoprotein, Gsm1 (Gamete-specific *minus*1), that belongs to the KNOX class, and that Gsp1 is distantly related to the BELL class. We show that Gsm1 and Gsp1 function as a heterodimer, the first demonstration to our knowledge of transcription factor heterodimerization in an algal species. The combinatorial action of Gsm1/Gsp1 resembles the regulatory networks formed by KNOX/BELL heterodimers, suggesting that *Chlamydomonas* may be displaying an ancient usage of this transcription factor family, whose roles appear to be limited to zygotic or postzygotic development in the green-plant lineage.

## RESULTS

### Identification of *GSM1*

Given that interacting networks of homeoprotein subfamilies are widely reported, a likely M factor partner protein for Gsp1 would be a second homeoprotein. Five homeobox-containing genes, including *Gsp1*, were found in the *C. reinhardtii* genome, of which three are TALE-superclass and two are non-TALE superclass members (Table S1). M factor candidates are predicted to be expressed exclusively in *minus* gametes. Figure 2A shows *minus*-gamete-specific expression of a TALE gene, which we have named gamete-specific *minus*1 (*GSM1*). The other three genes are expressed in gametes of both mating types (Figure 2A); their function has not been explored further.

All known *minus*-gamete-specific genes are either encoded in the *MT-* locus (LG VI) (Ferris and Goodenough, 1994) and/or require the dominant transcription factor Mid for expression (Ferris and Goodenough, 1997; Lin and Goodenough, 2007). Since *GSM1* is encoded in LG VIII (Table S1), its Mid dependency was assessed (Figure 2B). *GSM1* expression does not occur in the null *mid-1* mutant, whereas it occurs in an *mt+*/*mt-* diploid strain and in an *mt+* strain harboring a transgenic *MID* gene, both of which differentiate as *minus* gametes (Ebersold, 1967; Ferris and Goodenough, 1997).

The *GSM1* ORF (8 exons, 4615 nt) encodes a predicted 92.8 kD polypeptide of 934 aa with a C-terminal homeobox domain (Figure S2). The predicted protein contains 65% low-complexity regions (LCRs; DePristo et al., 2006), mainly His/Gln- and Ala-rich sequences in the N-terminal region (Figure S4), similar to Gsp1 (69% LCRs; Kurvari et al., 1998).

### Gsm1 Is a KNOX Ortholog; Gsp1 Is BELL-Related

Figure 2C shows a phylogenetic analysis of the TALE homeoproteins from the plant kingdom. Gsm1 and other KNOX-class proteins form a cluster supported by a 61% bootstrap or 98% posterior probability value (Figure 2C, gray); the land-plant BELL-class proteins form a second cluster (Figure 2C, blue), with distantly related algal members; a third algal cluster includes Gsp1 (Figure 2C, pink).

Positions 50 and 54 in the homeobox are critical to DNA recognition (Hanes and Brent, 1989), providing an alternate criterion for homeoprotein classification (Figure 2C). The KNOX proteins all carry 50:I and 54:K and are designated Group I. The cluster including Gsp1 carries 50:V/T and 54:A (Group II). Group III proteins all carry 50:I and 54:V/A, albeit the algal members (Hdg1 and Ot12440) are more distant cladistically. Importantly, Gsp1, Hdg1, and the BELL-class members all share an intron position with one another and with animal PBC/TGIF homeoproteins (Figure S5). Hence we propose that the non-KNOX TALE proteins in the Viridiplantae be designated as three lineages: (1) true BELL (blue); (2) BELL-related1 (pink) with a group II pattern; and 3) BELL-related2 with a group III pattern shared with BELL, possibly indicating that the true BELL class derives from BELL-related2.

Multiple alignments of KNOX-specific domains (Knox1, Knox2, and Elk) document conserved aa positions between algal and land-plant members (Figure 2D). Four intron positions are conserved in land-plant KNOX genes (triangles), of which the Knox1 and Elk introns are present only in subclasses I and II respectively. All four intron positions are conserved in *GSM1*, confirming common ancestry between *GSM1* and the KNOX genes from land plants. Additional phylogenetic analyses are found in Supplemental Data.

### Expression of *GSM1* Parallels Its *plus* Counterpart, *GSP1*

Figure S6 shows Gsm1 and Gsp1 expression profiles during synchronous gametogenesis and zygote development. In *minus*

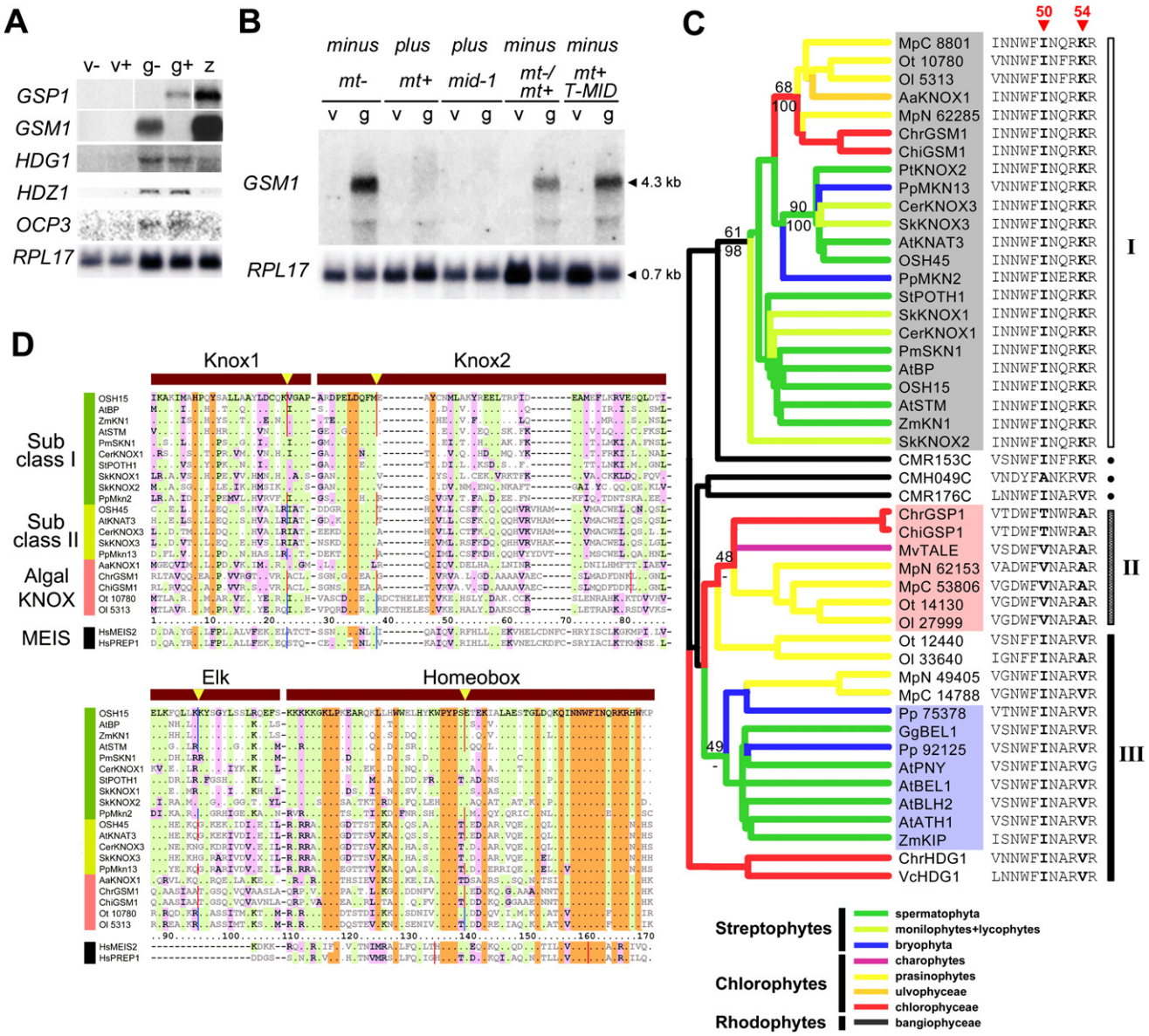


**Figure 2. *GSM1* Is Expressed Only in *minus* Gametes and Encodes a KNOX Homolog**

(A) Homeoprotein gene expression in different life-cycle stages. Northern blots except RT-PCR for *HDZ1*, with *RPL17* as loading control. v−/v+, vegetative cells *minus*/*plus*; g−/g+, gametes *minus*/*plus*; z, 30 min after mixing gametes.

(B) *Minus* gamete-specific expression of *GSM1*. Northern blot (total RNA) from the indicated strains that differ in mating specificity (top row, *plus* or *minus*) and/or carry different alleles of *MT* loci (second row), probed with a *GSM1* cDNA fragment encoding the C terminus. g, gametes; v, vegetative cells.

(C) Maximum Likelihood (ML)-based phylogenetic analysis of the TALE-superclass homeodomains from green plants. 63 aa long alignments of 44 TALE homeodomains from green algae and land plants plus three red algal homeodomains as outgroups were analyzed using the WAG+$\gamma$ aa substitution model and both RAxML-VI-HPC (Stamatakis, 2006; http://phylobench.vital-it.ch/raxml-bb/index.php) and MrBayes3.1 (Ronquist and Huelsenbeck, 2003). The best-scoring tree from RAxML-VI-HPC is shown; numbers on branches are bootstrap support values to clusters (above: ML by RAxML; below: Bayesian probabilities by MrBayes). Information on analyzed sequences is provided in Table S2. Branch colors correspond to the classification key beneath. Sequence alignments of DNA-recognition helices are shown at right, defining groups I–III. Additional information is given in Supplemental Data.

(D) Alignment of conserved KNOX domains. Dots represent identities to rice OSH15. For Knox1, Knox2, and Elk domains identified in algal genes, homology is depicted as orange: 100% identical; green: > 50% identical; pink: > 50% similar. Yellow triangles, intron positions; red vertical lines, intron present; blue lines, intron absent; no line, genomic data unavailable. Animal MEIS and plant KNOX proteins share limited homology in Knox1 and Knox2 domains but do not share intron positions.

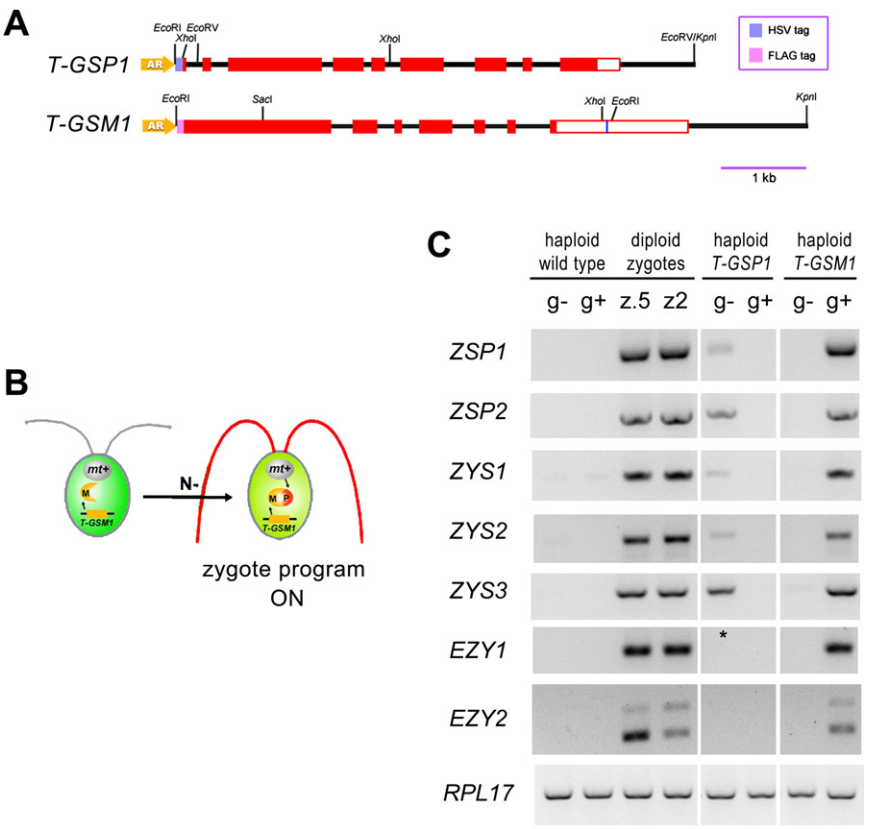


**Figure 3. Ectopic Expression of *GSM1* Induces the Zygote Program**

(A) Diagram of *T-GSP1* and *T-GSM1* constructs. Red filled boxes, exons; red open boxes, 3′UTRs; orange arrows, promoters. Full-length genomic clones of *GSP1* and *GSM1*, from ATG start codons and including 3′UTRs, were in-frame fused to 3XHSV and 3XFLAG tags, respectively, and placed downstream of constitutive AR promoters. Blue line in *GSM1* represents the alternative poly(A) site.

(B) Expected outcome of *GSM1* expression in *plus* gametes. Gsm1 proteins (M) from *T-GSM1* interact with *plus*-specific zygote-program regulator (P) during gametogenesis, activating the zygote program.

(C) Expression of seven early-zygote genes in diploid wild-type zygotes (lanes 3 and 4), *plus* gametes carrying *T-GSM1* (lane 8) and *minus* gametes carrying *T-GSP1* (lane 6). Genes encode zygotic cell-wall proteins (*ZSP1, ZSP2*), chloroplast-targeted proteins (*EZY1, EZY2*), an ER-localized protein (*ZYS3*), a nuclear-targeted protein (*ZYS1*), and an unknown protein (*ZYS2*). Analysis by semi-quantitative RT-PCR. Asterisk indicates the absence of *EZY1* expression in a *T-GSP1* transgenic line.

cultures, an increased expression of *GSM1* parallels an increase in mating ability during the 7–10 hr after gametogenesis is initiated by transfer to nitrogen (N)-free medium (Figure S6A, left). In *plus* cultures, expression of *GSP1* is also limited to differentiated gametes (Figure S6A, right), confirming Kurvari et al. (1998). Following mating, expression of both genes increases dramatically, followed by a sharp decrease within 1 hr (Figure S6B).

The peak expression of *GSM1* and *GSP1* in mating samples is elicited by an adhesion-induced intracellular cAMP surge during the mating reaction (Pasquale and Goodenough, 1987). Exposure of unmated gametes to dibutyryl cyclic AMP causes a strong accumulation of *GSM1* and *GSP1* mRNA and proteins (Figure S6C), as previously shown for Gsp1 using flagellar-induced activation (Kurvari et al., 1998).

## Activation of the Zygote Program with Ectopic Expression of *GSM1*

Gsp1 was demonstrated to be the P factor by its ability to induce zygote gene expression when ectopically expressed in *minus* gametes (Zhao et al., 2001). To assess the function of Gsm1 as an M factor, *GSM1* was ectopically expressed in *plus* gametes using a constitutive promoter-driven full-length genomic construct (*T-GSM1*, Figure 3A); for comparison, *T-GSP1* transformants were also analyzed (Figure 3A). Wild-type and transgenic strains were subjected to N-starvation to induce endogenous P- or M factor expression, and expression of the zygote program was assessed.

Activation of the zygote program was diagnosed by the formation of zygote-specific cell walls (Figure S7) and by the initiation of zygote-specific gene transcription (Figures 3C and S8) (Ferris et al., 2002; Ferris and Goodenough, 1987; Uchida et al., 1993; Wegener and Beck, 1991). Seven of the probed genes are expressed in *plus* gametes carrying *T-GSM1* but not in *minus* gametes or vegetative cells carrying *T-GSM1*. Reciprocally, *T-GSP1* lines express five of these genes in *minus* gametes, but not in *plus* gametes, consistent with published results (Zhao et al., 2001). The *T-GSP1* lines are not expected to express *EZY2* since *MT-* lacks *EZY2* genes (Ferris et al., 2002), but the differential expression of *EZY1* in the two strains (Figure 3C, asterisk; see also Zhao et al., 2001) is of interest given that *EZY1* is implicated in uniparental transmission of chloroplast DNA (Armbrust et al., 1993). A detailed study of this phenomenon is in progress.

These results document that Gsm1 can drive the zygote program in haploid *plus* gametes just as Gsp1 can drive the zygote program in haploid *minus* gametes.

## Gametogenesis Is Not Influenced by Homeoprotein Expression

RNAi constructs targeted against *GSM1* or *GSP1* generated up to a 5-fold decrease in gametic protein levels with no detectable effects on gametic competency (data not shown); when mated, these RNAi-transformed strains produce normal zygote cell walls, presumably because their RNAi populations are swamped by the rapid cAMP-induced surge of *GSM1* and *GSP1* transcripts

(Figure S6C). Given that both *GSM1* and *GSP1* are expressed after mating competency is achieved (Figure S6A), that transgenic *GSM1* and *GSP1* lines do not change sexual identity, and that transgenic vegetative cells (see below) show no signs of undergoing gametogenesis, it seems unlikely that Gsm1 or Gsp1 play a role in gametogenesis.

## Gsm1 and Gsp1 Form Heterodimers via Domains Also Utilized by KNOX/BELL Heterodimers in Land Plants

KNOX and BELL form both heterodimers and homodimers via α-helical domains N-terminal to the homeobox (Knox1 and Knox2 domains in the KNOX proteins and Sky and Bell domains in the BELL proteins; Bellaoui et al., 2001; Muller et al., 2001; Smith et al., 2002). We therefore used two in vitro approaches—a yeast-two-hybrid assay and a GST pull-down assay—to ask whether Gsm1 and Gsp1 also form heterodimers. Since Gsm1 and Gsp1 possess α-helical domains comparable to the interaction domains of KNOX and BELL (red and navy blue, Figure 4A), truncated constructs containing or removing these domains were also analyzed.

### *GAL4-Based Yeast-Two-Hybrid Assay*

Both full-length Gsm1 and Gsp1 as prey constructs (DNA-binding domain [DBD]-conjugated) were able to activate reporter genes without a bait construct (activation domain [AD]-conjugated) (Figure S9, sets 1 and 5), indicating that they possess autonomous transcriptional activation domains. N-terminal deletions that removed the HQ-rich LCRs eliminated this ability (Figure S9, sets 4 and 8), whereas homeodomain-truncated constructs did not (Figure S9, sets 2 and 6). Given these findings, we used N-terminal truncated Gsm1$^{269–934}$ as prey and full-length Gsp1$^{1–1037}$ as bait and documented their interaction (Figures 4B and 4C). In controls, neither Gsm1$^{269–934}$ with T-antigen nor Gsp1$^{1–1037}$ with lamin activated reporter genes (Figures 4B and 4C, set 2). Using deletion constructs, heterodimeric interaction domains were shown to map to Gsm1$^{512–740}$, containing the Knox1 and Knox2 domains, and Gsp1$^{559–1037}$, containing the homeobox and upstream α-helical domains (Figures 4B and 4C).

To test whether either protein is able to form a homodimer, we again first deleted the self-activating LCRs to create Gsm1$^{269–934}$ and Gsp1$^{356–1037}$ and then prepared bait and prey constructs using these or further deletion constructs (Figure 4D, sets with bars). A quantitative β-galactosidase assay was performed to compare homodimer with heterodimer activity. As shown in Figure 4D, the C-terminal domains of both Gsm1 and Gsp1 are able to form homodimers using the same domains involved in Gsm1/Gsp1 heterodimerization, but the binding affinity is much lower. Notably, longer constructs carrying these C-termini display only background-level activity, suggesting that native protein configurations may prevent homodimerization altogether.

### *In Vitro Pull-Down Assay*

To corroborate that Gsp1 and Gsm1 interact with one another via their C-terminal domains, glutathione-conjugated cellulose was used to trap GST (Glutathione-S-Transferase)-conjugated Gsm1, and its affinity for Gsp1 was assessed.

GST and the N-terminally GST-conjugated C-terminal half of Gsm1$^{512–934}$ (GST-Gsm1Ct) were expressed in *E. coli* and purified using affinity chromatography. As an interacting partner, the Gsp1 C terminus (559–1037) was in vitro translated, as was the C-terminal portion of Hdg1 (546–956), a BELL-related2 homeoprotein (Table S1). The translation products, detected via N-terminal HA-tags, were of the expected size and did not overlap with nonspecific signals present in the no-RNA control (Figure 4E, input lanes). When coincubated, GST-Gsm1Ct, but not GST, could pull down Gsp1$^{559–1037}$ but not Hdg1$^{546–956}$ (Figure 4E, lanes 4–7), reinforcing the conclusion that Gsm1 and Gsp1 are specific interacting partners and demonstrating that this interaction can occur in the absence of DNA.

## Gsm1 and Gsp1 Translocate into the Nuclei of Early Zygotes

Transcription of early zygote-specific genes initiates within 10–20 min of zygote formation, well before the two gametic nuclei fuse at 1–2 hr (Minami and Goodenough, 1978; Uchida et al., 1999), suggesting that heterodimerization of Gsm1 and Gsp1 may occur in the cytosol. We analyzed their localization by immunostaining and by heterologous expression in moss.

In gametes, Gsm1 immunostaining is restricted to *minus* cells, and the signal localizes to the cytosol and does not overlap with nuclear DNA staining (Figure 5A). Reciprocally, Gsp1 is only detected in the cytosol of *plus* gametes (Figure 5B). At 7 min after mixing *plus* and *minus* gametes, both Gsm1 and Gsp1 signals strongly immunolocalize to both nuclei of the newly-formed zygotes (Figures 5A and 5B, bottom panels), where protein that persists in the cytosol is presumably being newly synthesized from transcripts generated by cAMP activation (Figure S6C). By contrast, exclusively cytosolic signals are seen in neighboring cells with a single nucleus, most of which had likely experienced adhesion-induced cAMP elevation but had not yet fused. Hence, cell fusion and not gametic activation is required for nuclear localization of Gsm1 and Gsp1, suggesting that their heterodimerization is necessary for nuclear translocation and/or retention.

To ask whether the Gsm1/Gsp1 interaction domains identified in vitro (Figure 4) are sufficient to drive nuclear localization, sequences encoding the C-termini of each protein (Gsm1$^{512–934}$ and Gsp1$^{559–1037}$) were each conjugated N-terminally to YFP. When efforts to detect the YFP signal in *Chlamydomonas* cells transformed with these constructs were unsuccessful, the constructs, driven by 35S promoters, were adsorbed to gold particles and introduced into cultured moss cells by bombardment. The fluorescent protein products reside in the cytoplasm when each construct is introduced on its own (Figure 5C, upper 2 rows). By contrast, when the constructs are coadsorbed and bombarded, fluorescence localizes to the nucleus (Figure 5C, third row, arrows).

The moss system was also used for bimolecular fluorescence complementation (BiFC) analysis (Hu and Kerppola, 2003), wherein the N-terminal half of a fluorescent tracer is conjugated to one protein and the C-terminal half to a second protein; if the proteins heterodimerize, the halves are brought together to form a functional fluorescent moiety. The C-terminal sequence of Gsm1 was conjugated to the N-terminal sequence of YFP (CIT$^{Nt}$), and the C-terminal sequence of Gsp1 was conjugated to the C-terminal sequence of YFP (CIT$^{Ct}$). No fluorescent signal is detected when the constructs are singly introduced (data not shown), but when the two constructs

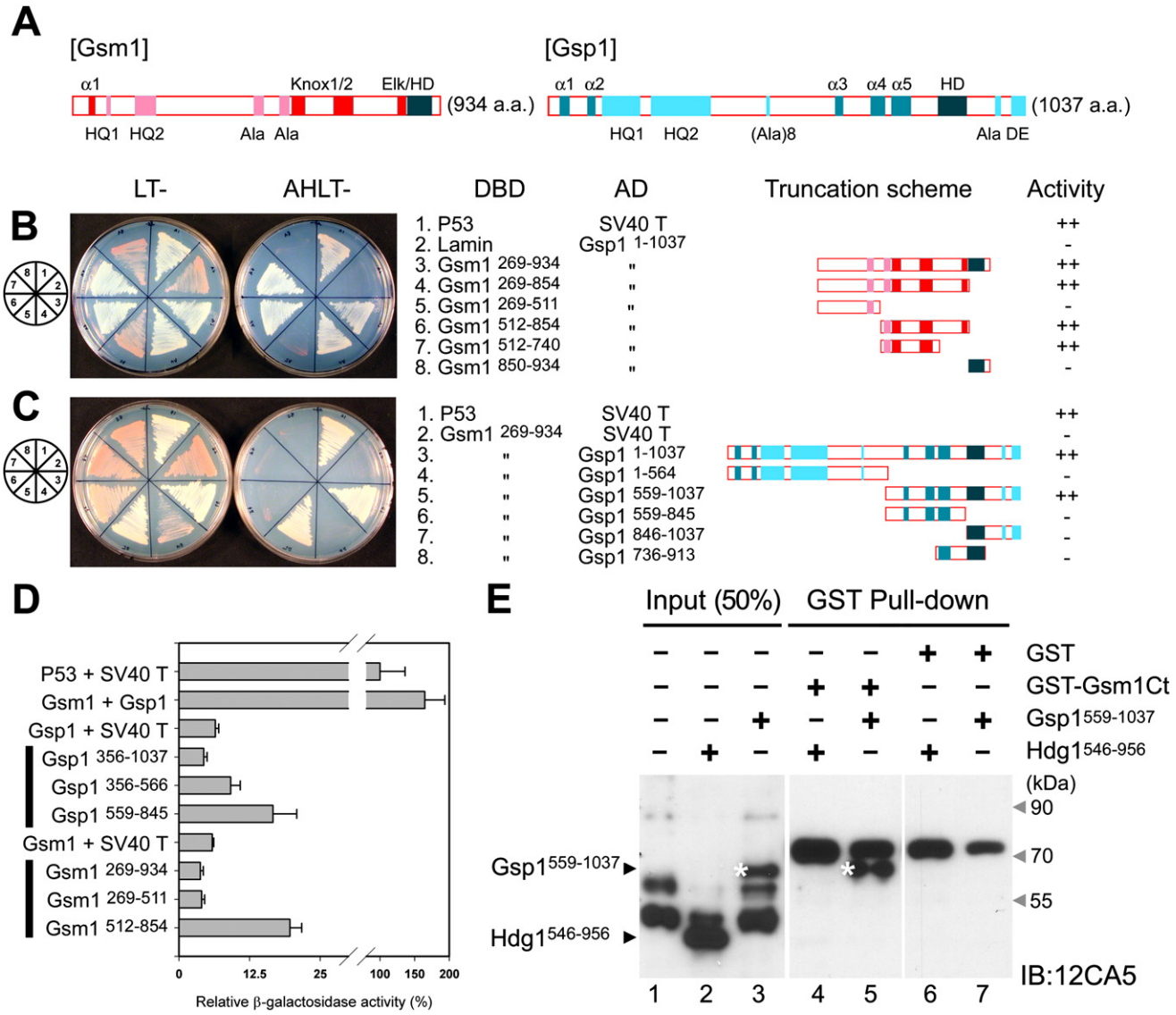


**Figure 4. Gsm1 and Gsp1 Physically Interact via C-Terminal α-Helical Domains**

(A) Domain structures of Gsm1 and Gsp1. Black, homeodomains (HD); red and navy blue, α-helical domains; pink and light blue, low-complexity domains with His/Gln (HQ)-rich, Ala-rich, or acidic (DE)-rich residues.

(B and C) Mapping Gsm1 and Gsp1 domains that participate in heterodimerization. White (but not pink) growth on Leu-/Trp- (LT-) and growth on Ade-/His-/Leu-/Trp- (AHLT-) media indicates interaction of the introduced proteins. Gsm1-based constructs served as prey conjugated with the GAL4 DNA-binding domain (DBD), and Gsp1-based constructs as bait conjugated with the GAL4 transcriptional activation domain (AD). Set 1 utilizes P53 and SV-40 T, known to be interacting partners. Set 2 represents negative controls without a homeoprotein partner.

(D) Weak homodimerization of Gsm1 and Gsp1. N-terminal truncated proteins were conjugated with DBD and AD as full-length, N-terminal half, and C-terminal half (without homeodomain) constructs (sets with bars). Relative affinity of homodimers was quantified by β-galactosidase activity in comparison to a positive control (P53 + SV40 T), set at 100%, and a heterodimer set (Gsm1 + Gsp1). Error bars represent the SD.

(E) Coprecipitation of Gsp1 with Gsm1. In vitro-expressed Gsm1$^{512–934}$ attached to glutathione S-transferase (GST-Gsm1Ct) was used to pull down in vitro translated HA-tagged Hdg1$^{546–956}$ or Gsp1$^{559–1037}$. Translated and pulled-down products were analyzed by western blotting with anti-HA monoclonal Ab (12CA5). Asterisks indicate Gsp1-specific signals from in vitro translation and pull-down products. 12CA5 also interacts with three proteins present in the wheat-germ extract (lanes 1–3) and with one protein present in the *E. coli* extract (lanes 4–7).

are cobombarded, a fluorescent signal localizes to the nucleus (Figure 5C, bottom row, arrows). Hence the C-terminal domains of Gsm1 and Gsp1 heterodimerize in vivo, and their interaction is sufficient to drive nuclear localization in a heterologous system.

## Dual Ectopic Expression of *GSM1* and *GSP1* Activates the Zygote Program without Gametogenesis

Transformants carrying either *T-GSM1* or *T-GSP1* alone require gametogenesis to activate the zygote-specific genes, presumably because gametogenesis triggers the synthesis of partner

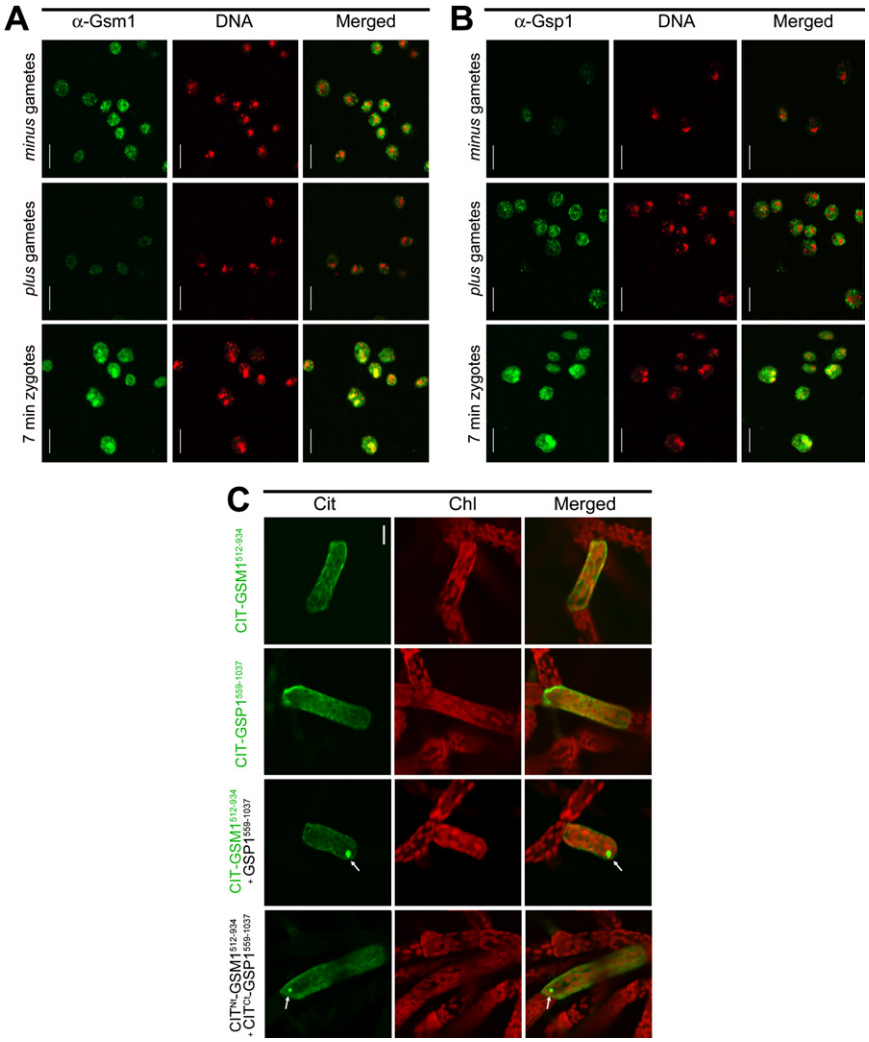


**Figure 5. Gsm1 and Gsp1 Translocate from the Cytosol to the Nucleus upon Heterodimerization**

(A and B) Immunostaining of Gsm1 (A) or Gsp1 (B) using anti-Gsm1 or anti-Gsp1 in gametes and 7 min zygotes. Left panels: FITC-conjugated secondary antibody staining; Center panels: propidium iodide staining for nuclei.

(C) Transient expression of C-terminal-truncated Gsm1 or Gsp1 conjugated to Citrine (CIT-GSM1$^{512–934}$, CIT-GSP1$^{559–1037}$) monitored in moss cells by confocal microscopy (the scale bar represents 20 μm). Left panels: Cit signals (488 nm excitation); center panels: chlorophyll (Chl) signals (568 nm excitation); right panels: merged images. Upper two rows: Cit-Gsm1 or Cit-Gsp1 localizes to the cytosol. Third row: Cit-Gsm1 localizes both to the cytosol and the nucleus (arrows) when cobombarded with non-conjugated Gsp1. Bottom row: In vivo interaction and nuclear localization (arrows) of Gsm1 and Gsp1 C-terminal domains demonstrated by bi-fluorescent complementation (CIT$^{Nt}$ + CIT$^{Ct}$). Each set of moss-bombardment experiments was repeated five times with identical localization patterns.

homeoproteins in the transgene recipients. Left unaddressed is whether additional gamete-specific proteins are also required to initiate the zygotic pathway or whether Gsp1/Gsm1 is sufficient. We therefore analyzed the consequence of expressing both *T-GSM1* and *T-GSP1* in vegetative cells that have not undergone gametic differentiation.

*T-GSM1* and *T-GSP1* single transformants were crossed to generate double transformants expressing both *T-GSM1* and *T-GSP1*. Plate-grown cultures deplete available nitrogen after 4–5 days of growth and display mating behavior after suspension in N-free medium (Minami and Goodenough, 1978); therefore, to characterize the double transformants as vegetative cells, 3-day-old (mitotic) agar cultures were suspended in N-containing (N+) liquid medium. After 12–24 hr, both *plus* and *minus* strains of the double transformants form zygote-specific cell walls and express zygote-specific genes whereas single-transformant controls do not (Figure 6). Interestingly, the double transformants grow as well as wild-type cells on N+ agar plates, leading us to analyze zygote-gene expression under these conditions. We ascertained that *T-GSP1* expression fails to occur in 3 day N+ plate cells but is detected 12 hr after N+ liquid suspension (data not shown). The factors exerting *T-GSP1* inhibition under agar growth conditions are currently unidentified.

## Functional Characterization of Transgene-Induced Zygotes

Natural zygotes differentiate directly into dormant spores that resist environmental insult (freezing, desiccation, and exposure to chloroform). Under laboratory conditions, zygote maturation typically takes 2–5 days in the dark, after which the spores are competent to undergo meiotic germination when provided with an N source and a light signal (Figure 1; van Winkle-Swift, 1977). We went on to ask whether *T-GSM1* and/or *T-GSP1*-induced zygotes can form resistant spores that undergo meiosis.

Since only diploid nuclei can undergo meiosis (Dutcher, 1988), we constructed diploid strains transformed with *T-GSM1* and/or *T-GSP1*; since heterozygous *mt+/mt-* diploids differentiate as *minus* (Ebersold, 1967), we generated *plus* diploids using a novel strategy (details in Experimental Procedures). Six diploid strains were then generated carrying *T-GSP1* and/or *T-GSM1* (Table S3), all of which are capable of activating the zygote program without mating, and two of which—the double *T-GSM1 T-GSP1*—are capable of activating the zygote program without N-starvation and gametogenesis. In addition, all six strains were made heterozygous in *trans* for *ac17* and *nit2*, two linked auxotrophic markers on LG III, allowing us to characterize segregation of homologous chromosomes via meiosis and to detect recombination between *ac17* and *nit2* by tetrad analysis.

Early efforts to detect meiotic germination from the transgene-induced zygotes were unsuccessful: when the standard

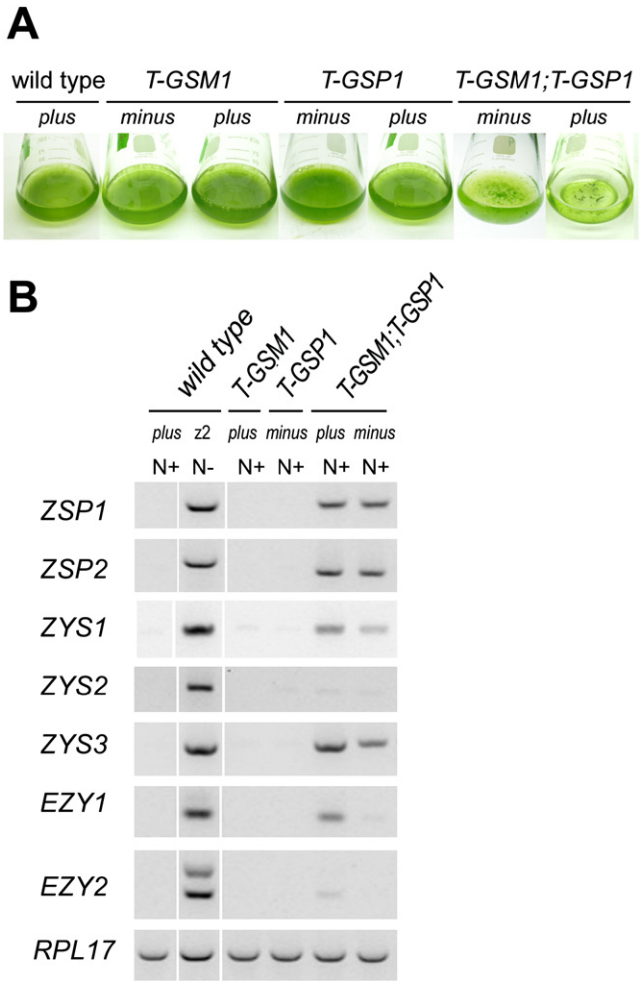


**Figure 6. Coexpression of *GSM1* and *GSP1* Turns on Zygote Development without Gametogenesis**

(A) Cellular phenotypes of single- and double-transgenics resuspended in N+ media from 3-day-old plates and maintained for 24 hr. Adhesive zygotic cell walls generate flocculant aggregates.

(B) Zygote-specific gene expression in vegetative cells of single- and double-transgenics by semiquantitative RT-PCR. Vegetative (N+) samples were harvested at 12 hr after resuspension in N+ media from 3-day-old plates. z2 indicates 2 hr after mixing wild-type *plus* and *minus* gametes.

laboratory protocol for zygote maturation was followed—1 day in the light, 5 days in the dark, return to N+ medium and light, expose to chloroform vapor to kill nonzygotic cells (Harris, 1989)—most products from the chloroform-resistant cells yielded prototrophic clones and gave no evidence of having segregated heterozygous markers, indicating that they had not undergone meiosis and were instead dividing by mitosis as diploids. Pre-exposure to dibutyryl-cAMP to mimic the mating reaction did not alter this outcome.

For meiotic germination to occur, it proved necessary to keep the transgene-induced zygotes in the dark for >3 weeks. Segregation of auxotrophic markers was then observed in 60%–100% of the germinated zygotes, depending on the strains. Complete tetrads consistently displayed 2:2 marker segregation, indicating the occurrence of normal meiosis and not random chromosome loss (Table 1; see also Figure S10). Importantly, no large variations in meiotic germination rates were found among the six strains, indicating that the decision to enter meiosis during germination is dependent on the Gsm1/Gsp1-activated zygote program but not on the heterozygosity of *MT* (shown also in Ferris et al., 2002; Ferris and Goodenough, 1997) nor on prior gametic differentiation.

**Table 1. Zygotes Induced by *T-GSM1* and/or *T-GSP1* Undergo Meiosis during Germination**

| Cell Line | Maturation Time | Meiosis Rate[a] | PD:NPD:T[b] | MAP Distance (cM)[c] |
|---|---|---|---|---|
| P1 | 21 days | 97.7% (84/86) | 54:00:12 | 9.1 |
| P2 | 27 days | 59.0% (26/49) | 13:00:09 | 20.5 |
| M1-homo | 29 days | 100% (65/65) | 51:00:11 | 8.9 |
| M1-hetero | 27 days | 98.1% (53/54) | 11:00:04 | 13.3 |
| MP-minus | 27 days | 89.4% (34/38) | 18:00:06 | 12.5 |
| MP-plus | 29 days | 53.7% (22/41) | 9:00:00 | 0.0 |
| | | | combined | 10.6 |
| | | | published[d] | 10.6 |

See legend to Figure S10 for details.

[a] In parenthesis: (number of tetrads with marker segregation) / (number of isolated tetrads).

[b] Based on full tetrads (with >3 genotypes).

[c] Calculated as described in Experimental Procedures.

[d] (Fernandez and Matagne, 1984)

Recombination rates between *ac17* and *nit2* were measured in all germination products that yielded full tetrads (Table 1). The combined average map distance between these markers, 10.6 cM, proved to be identical to the map distance reported for natural zygotes (Fernandez and Matagne, 1984), indicating that meiosis in the transgene-induced zygotes occurs normally, with a proper configuration of recombination machinery.

## DISCUSSION

We show that TALE-superclass homeoproteins Gsm1 and Gsp1, when localized to the nucleus as heterodimers, are sufficient to initiate the diploid phase of the *Chlamydomonas* sexual cycle, with the resultant zygotes competent to form spores that undergo meiosis. Here, we discuss Gsm1/Gsp1 heterodimerization in the context of current understandings of combinatorial transcriptional control, and we then explore the evolutionary implications of our findings for the origins of sexual development and the emergence of the land-plant lineage.

### Homeoproteins and Combinatorial Control

Homeoprotein-based combinatorial control is widely used throughout the eukaryotes. Animals and plants utilize homeoprotein-based networks to specify core domains of the body plan: PBC/MEIS + HOX (TALE + non-TALE) specify the anterior-posterior axis of animals, and KNOX + BELL (TALE + TALE) specify the shoot apical meristems (reviewed in Hake et al., 2004; Mann and Morata, 2000).

We show that Gsm1 + Gsp1 and KNOX + BELL use the same interaction domains. The Gsm1 interaction region includes the

Knox 1 and 2 α-helical domains but not the homeobox; land-plant KNOX proteins use these same domains to interact with partner proteins (Bellaoui et al., 2001). The Gsp1 interaction requires both homeobox and α-helical domains N-terminal to the homeobox; the same is true for true BELL proteins (Muller et al., 2001).

### *DNA-Binding Specificity of Gsm1/Gsp1*

The third α-helix of the homeobox (residues 42–58 in a 60-residue non-TALE homeobox) engages in DNA recognition: residues at positions 47, 50, 51, and 54 interact with nucleotides in the major groove (Gehring et al., 1994); residue 51 is highly invariable; and residues 50 and 54 determine DNA-binding specificity (Damante et al., 1996; Viola and Gonzalez, 2006). The Gsm1 homeobox contains I and K at residues 50 and 54 in its DNA recognition helix, as do MEIS- and KNOX-class homeoboxes, suggesting that they all recognize a similar DNA target. The Gsp1 homeobox, by contrast, contains T and A residues at these positions (Figure 2C), meaning that the Gsm1/Gsp1 combination may identify new *cis* elements.

In a preliminary study, ~2000 bp upstream of seven early-zygotic ORFs were searched for consensus motifs using MEME and ACEalign (Bailey and Elkan, 1994; Roth et al., 1998). Both algorithms found CGtGACATGaCC sequences within 400 bp of the transcriptional start site of 6 of the 7 ORFs; these contain two TGAC motifs, a motif repeatedly identified as the binding site for TALE-class homeodomains with "$WFI_{50}N$" cores (Knoepfler et al., 1997; Krusell et al., 1997). Upstream sequences from 28 nonzygotic genes failed to include this sequence. Future studies will investigate whether these regions represent bona fide DNA binding elements for Gsm1/Gsp1 and whether heterodimerization enhances affinity for the zygotic-cis-elements directly (by contributing to DNA-binding specificity) and/or indirectly (by inducing allosteric changes).

### *Nuclear Localization of Gsm1/Gsp1*

Gsp1 and Gsm1 localize to the cytoplasm in *plus* and *minus* gametes and to the nucleus in zygotes, and their heterodimerization is sufficient to drive nuclear localization in a heterologous moss system (Figure 5). Hence, nuclear translocation/retention is solely dependent on the interaction of these two proteins, as is the case for *Arabidopsis* KNOX/BELL heterodimers (Bhatt et al., 2004; Cole et al., 2006).

Neither Gsp1 nor Gsm1 displays candidate NLS motifs (searched at http://cubic.bioc.columbia.edu/predictNLS/), although such signals often prove to be embedded in the basic amino-acid clusters of homeoproteins (Meisel and Lam, 1996). While Gsp1 carries no candidate NES sequences, two are predicted for Gsm1, at amino acid positions 24–28 and 707–709 (searched at http://www.cbs.dtu.dk/services/NetNES/), where the latter is embedded in the interaction domain of Gsm1 (512–740; Figure 4C). By analogy with Exd/Hth (Stevens and Mann, 2007), the second NES may be involved in the cytosolic retention of Gsm1 in *minus* gametes, and heterodimerization with Gsp1 may mask this motif, allowing nuclear retention.

## Homeoprotein Combinatorial Control and Eukaryotic Sexual Cycles

Representatives of each of the major eukaryotic groups have been observed to engage in a sexual cycle, either obligately or facultatively (Dacks and Roger, 1999), or else to possess a highly conserved set of genes that are selectively expressed during meiosis (Ramesh et al., 2005), suggesting that the common ancestor to all modern eukaryotes engaged in haploid-diploid transitions that were followed by meiotic reduction.

Our molecular understanding of haploid-diploid transitions in modern single-celled or simple multicellular eukaryotes has to date been limited to fungi where, except for the filamentous ascomycetes (e.g., *Neurospora crassa*; Saupe, 2000), haploid-diploid transitions are mediated by homeoproteins. (1) Pairs of homeoproteins heterodimerize to initiate zygote-specific gene expression in *S. cerevisiae* (Goutte and Johnson, 1988) and *C. neoformans* (Hull et al., 2005), and to initiate fruiting body development after dikaryon formation in several basidiomycetes (Gillissen et al., 1992; Kues et al., 1992); (2) Coexpression of the homeoprotein mat1-Pm with the novel short ORF mat1-Mm activates entry into meiosis in *S. pombe* (Willer et al., 1995). Where analyzed, moreover, the homeoporoteins are presynthesized in gametes of opposite type (Hull et al., 2005; Urban et al., 1996).

This report provides an example of haploid → diploid homeoprotein heterodimerization in a second radiation, the green-plant lineage, where the interacting proteins are also presynthesized in gametes of opposite type. The plant and fungal/animal lineages are thought to have radiated independently for at least a billion years; they also share deep eukaryotic ancestral roots (Baldauf, 2003). The finding that both lineages use homeoproteins in the same life-cycle context suggests that the homeoprotein family may have served as components of an ancient—perhaps the pioneering—sexual strategy in deep eukaryotic ancestors. Key to evaluating this hypothesis will be the analysis of homeoprotein expression in organisms from additional radiations (e.g., diatoms and slime molds), where a fruitful approach, suggested by our studies, would be to monitor the phenotypic effects of driving the constitutive expression of their homeoprotein-encoding genes. Indeed, this approach might reveal sexual diploid phases that have been refractory to laboratory induction.

## Homeoprotein Combinatorial Control and Plant Evolution

Our studies document that KNOX/BELL-related (TALE/TALE) heterodimers are used by *Chlamydomonas* and land plants in a homologous fashion. This suggests a testable scenario for the participation of homeoprotein combinatorial control in the evolutionary transition from green algae to land plants in the Viridiplantae radiation,

In most green algae, the diploid phase is limited to a unicellular zygospore that undergoes meiosis. The haploid meiotic products then go on to develop either as haploid unicellular organisms, e.g., *Chlamydomonas* (chlorophyceae), *Acetabularia* (ulvophyceae), and *Closterium* (charophyta), or as haploid multicellular organisms, e.g., *Volvox* (chlorophyceae), *Ulva* (ulvophyceae), and *Chara* (charophyta). Exceptions include the unusual mitotic diploids of *Chlamydomonas* (Figure 1) and the diploid stage of certain ulvophytes (Graham and Wilcox, 2000), but in these cases the haploids and diploids are isomorphic and hence presumably specified by the same genetic networks.

Land plants differ from this algal pattern in three important respects. (1) Their diploid (sporophytic) phase is morphologically distinctive from their haploid (gametophytic) phase, indicating that the two forms are generated via distinctive genetic networks (Kenrick and Crane, 1997). (2) The sporophyte is not committed to transition directly into meiosis, but instead differentiates into novel cell types and tissues, postponing the onset of meiosis to a more terminal stage in its developmental program wherein (3) a subset of sporophytic cells undergoes meiosis to generate and release haploid, rather than diploid, spores or gametes.

We show here that in *Chlamydomonas*, the algal pattern is governed by Gsp1/Gsm1 combinatorial control. When this kernel-like network (Davidson and Erwin, 2006) is disrupted or delayed, the alternative is to resume a haploid-like/vegetative growth pattern, displayed both by heterozygous diploids and by the transgenic diploids, described in this report, that resume mitotic growth if germination is prematurely induced. The isomorphic transition from haploid to diploid ulvophytes may prove to occur by a similar process.

Hence, core innovations were needed to break out of this "algal rut"—adaptive as it clearly is for the algae—and establish the land-plant pattern with distinctive haploid and diploid bauplans. We document that the Gsp1-like class of BELL-related1 genes is present in the genomes of all examined green algae and absent from the genomes of all examined land plants; hence the loss of BELL-related1 genes presumably participated in escaping from the algal pattern. This loss was accompanied by (1) the diversification of KNOX-class gene family members (e.g., into subclasses I and II) and (2) a dramatic expansion of the true BELL-class, possibly derived from the BELL-related2 class, leading to the establishment of new KNOX/BELL kernel networks. These then (3) recruited novel inputs (negative regulators) to establish such body-plan innovations as lateral branches and marginally expanded leaves (megaphyll), analogous to the inputs on the HOX-based kernels in animals specifying bilateral symmetry and AP-axis formation in animals (Pearson et al., 2005).

Such a scenario predicts that in a larger survey, green organisms with diploid premeiotic spores will carry BELL-related genes and those with post-meiotic haploid spores will carry true BELL genes. It also lifts up the importance of analyzing the homeoprotein endowment and function of the charophyte algae, which retain the diploid-spore phenotype but display other land-plant-related features (e.g., matrotrophy and axial growth) as haploids, and which share more recent common ancestry with the land plants than does *Chlamydomonas* (Lewis and McCourt, 2004).

The thesis that novel KNOX/BELL networks "took over" as the core kernels for elaborating a sporophytic body plan in land plants predicts that KNOX/BELL function may be limited to sporophytic development. Evidence for such a pattern is emerging in the "lower" land plants, whose extensive gametophytic phase involves simple filamentous or thalloidal forms similar to green algae. (1) In moss, expression of its five homeobox-containing genes is predominantly restricted to the sporophytic phase of the life cycle, and the knockout of three of these genes causes developmental defects exclusively in sporophytes (Singer and Ashton, 2007; Sakakibara et al., 2008). (2) In ferns, the 3 KNOX genes are exclusively expressed in the sporophyte meristem (Sano et al., 2005).

In conclusion, our studies suggest that homeoprotein combinatorial control was instantiated in a deep eukaryotic ancestor in the context of establishing the meiotic sexual cycle. In the green lineage, the resultant novel diploid phase, initially committed only to fostering dormancy and meiosis, represented "virgin territory" for generating novel bauplans; the genetic networks that availed themselves of this opportunity, we propose, were regulated by novel combinations of an expanded/modified set of homeoprotein family members. Possibly a similar sequence of events accompanied the adoption of haploid/diploid alternations in the basidiomycetes and the brown-algal radiations. The recent sequencing of the choanoflagellate (*Monosiga*) genome suggests that the expansion of homeoprotein networks also accompanied metazoan evolution (King et al., 2008), in which the alternation-of-generations idea was largely abandoned in favor of diploid-bauplan dominance, often featuring a succession of variant diploid larval forms.

## EXPERIMENTAL PROCEDURES

Procedures developed for this study are presented below; routine protocols are found in Supplemental Experimental Procedures.

### Novel Diploid Generation Protocols

Homozygous diploid strains were generated using the *iso1* mutant, previously described as a recessive allele (Campbell et al., 1995) but since shown to be dominant. *ISO1/iso1;mt+/mt−* heterozygous diploids isoagglutinate and fuse among themselves, generating diploid *mt+/mt+* and *mt−/mt−* meiotic progeny. These were subjected to tetraploid genetics (Galloway and Goodenough, 1985) using *mt+/mt−* or *mt−/mt−* as *minus* gametes and *mt+/mt+* as *plus* gametes. Backcrosses to *mt+/mt-* diploids allowed the *iso1* mutation to be discarded as confirmed by PCR. *mt+/mt+* diploids carrying *T-GSM1* (M1-homo in Table 1) were generated by mating *mt+/mt+* with *mt+/mt−;T-GSM1* diploids. To generate *plus* sexuality in a *mt+/mt−* background, [*ac17 NIT2*];*mt−* was mated to [*AC17 nit2*];*mid-1;T-FUS* and [*ac17 NIT2*]/[*AC17 nit2*];*mt-/mid-1*;*T-FUS*/- diploids were selected; these were mated with *mt+/mt+*;*T-GSM1* to generate *mt+/mid-1*;*T-GSM1* (M1-hetero in Table 1).

### Germination of Transgene-Induced Zygotes

*T-GSM1* or *T-GSP1*-transformed diploids were N-starved, and the doubly transformed diploids with *T-GSM1* and *T-GSP1* were resuspended from 3-day-old N+ plates to induce the zygote program. Gametic/vegetative cell walls were removed with gametolysin to facilitate detection of zygote cell walls (Figure S7). When zygote walls were observed, 1–5 × $10^6$ cells were plated on N+ high-salt-minimal plates, exposed to light overnight, and maintained in darkness for 5–30 days.

### In Vitro GST Pull-Down Experiments

GST (29 kDa) and GST-Gsm1Ct (512–934, 90.5 kDa) were expressed in *E. coli* (BL21[DE3]pLysS) at 25°C. GST and GST-Gsm1Ct fusion proteins were purified by Glutathione Sepharose 4B (GE Healthcare) and desalted against 1× PBS (140 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO4$, 1.8 mM $KH_2PO4$, [pH 7.3]). HindIII-digested pAD-Gsp1$^{559–1037}$ and pAD-Hdg1$^{546–956}$ plasmids, with HA N-terminally tagged to inserts, were used for in vitro transcription, then purified and translated in vitro using Wheat Germ Extract Plus (Promega) with T7 RNA polymerase and biotin-labeled tRNA (Transcend, Promega) to label translated products. Translated products were analyzed by immunoblotting using Streptavidin-conjugated HRP (Invitrogen), showing HA-Gsp1$^{559–1035}$ (est. 47.7 kDa) migrating as ~65 kDa and HA-Hdg1$^{546–956}$ (est. 40.3 kDa) as ~48 kDa. Pull-down experiments were performed in NP40-binding buffer (1× PBS, 1% NP-40, 0.1% BSA, and a protease inhibitor cocktail) as described in Bellaoui et al. (2001) (see Supplemental Data for details).

### Transient Gene Expression in Moss

Four-day-old lawn-grown protonemal tissue of *Physcomitrella patens* cultures was used for transient transformation by particle bombardment following published protocols (Bezanilla et al., 2003). 2 μg of each DNA construct was used to coat gold particles for two shots. For cobombardment, two constructs were mixed before coating the particles. The bombarded moss tissue was incubated for 24–48 hr at room temperature before imaging. Images were captured on a glass slide with the tissue in water under a coverslip. YFP signals were visualized using a laser excitation line at 488 nm and chlorophyll fluorescence by a 568 nm laser excitation (see Supplemental Data for details).

## ACCESSION NUMBERS

Sequence data has been filed in the EMBL/GenBank data libraries under accession number(s) EU029996, EU271678, EU271676, and EU271677, for *C. reinhardtii GSM1, HDG1, HDZ1, and OCP3*; AY898810 and DQ989631 for *C. incerta GSP1 and GSM1*.

## SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, Supplemental Results, Supplemental References, ten figures, and four tables and can be found with this article online at http://www.cell.com/cgi/content/full/133/5/829/DC1/.

## ACKNOWLEDGMENTS

We thank William Snell (UT Southwestern) for his generous gifts of Gsp1-related molecular probes, Patrick Ferris (Salk Institute) for *C. incerta GSP1* cloning and pNic7.9, Jean-David Rochaix (U Geneva) for pT7-ARG7, and Pierre-François Perroud and Lauren Gunther for help with the moss experiments. This work is supported by grant MCB 0326829 from the National Science Foundation (NSF).

Received: July 26, 2007
Revised: April 3, 2008
Accepted: April 21, 2008
Published: May 29, 2008

## REFERENCES

Armbrust, E.V., Ferris, P.J., and Goodenough, U.W. (1993). A mating type-linked gene cluster expressed in Chlamydomonas zygotes participates in the uniparental inheritance of the chloroplast genome. Cell *74*, 801–811.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In Proceedings Second International Conference on Intelligent Systems for Molecular Biology (Menlo Park, CA, AAAI Press), pp. 28–36.

Baldauf, S.L. (2003). The deep roots of eukaryotes. Science *300*, 1703–1706.

Bellaoui, M., Pidkowich, M.S., Samach, A., Kushalappa, K., Kohalmi, S.E., Modrusan, Z., Crosby, W.L., and Haughn, G.W. (2001). The *Arabidopsis* BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. Plant Cell *13*, 2455–2470.

Bezanilla, M., Pan, A., and Quatrano, R.S. (2003). RNA interference in the moss *Physcomitrella patens*. Plant Physiol. *133*, 470–474.

Bhatt, A.M., Etchells, J.P., Canales, C., Lagodienko, A., and Dickinson, H. (2004). VAAMANA–a BEL1-like homeodomain protein, interacts with KNOX proteins BP and STM and regulates inflorescence stem growth in Arabidopsis. Gene *328*, 103–111.

Bowman, J.L., Floyd, S.K., and Sakakibara, K. (2007). Green genes–Comparative genomics of the green branch of life. Cell *129*, 229–234.

Campbell, A.M., Rayala, H.J., and Goodenough, U.W. (1995). The *iso1* gene of *Chlamydomonas* is involved in sex determination. Mol. Biol. Cell *6*, 87–95.

Chan, R.L., Gago, G.M., Palena, C.M., and Gonzalez, D.H. (1998). Homeoboxes in plant development. Biochim. Biophys. Acta *1442*, 1–19.

Cole, M., Nolte, C., and Werr, W. (2006). Nuclear import of the transcription factor SHOOT MERISTEMLESS depends on heterodimerization with BLH proteins expressed in discrete sub-domains of the shoot apical meristem of *Arabidopsis thaliana*. Nucleic Acids Res. *34*, 1281–1292.

Dacks, J., and Roger, A.J. (1999). The first sexual lineage and the relevance of facultative sex. J. Mol. Evol. *48*, 779–783.

Damante, G., Pellizzari, L., Esposito, G., Fogolari, F., Viglino, P., Fabbro, D., Tell, G., Formisano, S., and Di Lauro, R. (1996). A molecular code dictates sequence-specific DNA recognition by homeodomains. EMBO J. *15*, 4992–5000.

Davidson, E.H., and Erwin, D.H. (2006). Gene regulatory networks and the evolution of animal body plans. Science *311*, 796–800.

DePristo, M.A., Zilversmit, M.M., and Hartl, D.L. (2006). On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. Gene *378*, 19–30.

Derelle, R., Lopez, P., Guyader, H.L., and Manuel, M. (2007). Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. Evol. Dev. *9*, 212–219.

Dutcher, S.K. (1988). Nuclear fusion-defective phenocopies in *Chlamydomonas reinhardtii*: mating-type functions for meiosis can act through the cytoplasm. Proc. Natl. Acad. Sci. USA *85*, 3946–3950.

Ebersold, W.T. (1967). *Chlamydomonas reinhardi*: heterozygous diploid strains. Science *157*, 447–449.

Fernandez, E., and Matagne, R.F. (1984). Genetic analysis of nitrate reductase-deficient mutants in *Chlamydomonas reinhardii*. Curr. Genet. *8*, 635–640.

Ferris, P.J., Armbrust, E.V., and Goodenough, U.W. (2002). Genetic structure of the mating-type locus of *Chlamydomonas reinhardtii*. Genetics *160*, 181–200.

Ferris, P.J., and Goodenough, U.W. (1987). Transcription of novel genes, including a gene linked to the mating-type locus, induced by *Chlamydomonas* fertilization. Mol. Cell. Biol. *7*, 2360–2366.

Ferris, P.J., and Goodenough, U.W. (1994). The mating-type locus of *Chlamydomonas reinhardtii* contains highly rearranged DNA sequences. Cell *76*, 1135–1145.

Ferris, P.J., and Goodenough, U.W. (1997). Mating type in Chlamydomonas is specified by *mid*, the minus-dominance gene. Genetics *146*, 859–869.

Floyd, S.K., and Bowman, J.L. (2007). The ancestral developmental tool kit of land plants. Int. J. Plant Sci. *168*, 1–35.

Galloway, R.E., and Goodenough, U.W. (1985). Genetic analysis of mating locus linked mutations in *Chlamydomonas reinhardii*. Genetics *111*, 447–461.

Gehring, W.J., Qian, Y.Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A.F., Resendez-Perez, D., Affolter, M., Otting, G., and Wuthrich, K. (1994). Homeodomain-DNA recognition. Cell *78*, 211–223.

Gillissen, B., Bergemann, J., Sandmann, C., Schroeer, B., Bolker, M., and Kahmann, R. (1992). A two-component regulatory system for self/non-self recognition in *Ustilago maydis*. Cell *68*, 647–657.

Goodenough, U.W., Armbrust, E.V., Campbell, A.M., and Ferris, P.J. (1995). Molecular genetics of sexuality in *Chlamydomonas*. Annu. Rev. Plant Physiol. Plant Mol. Biol. *46*, 21–44.

Goodenough, U., Lin, H., and Lee, J.H. (2007). Sex determination in *Chlamydomonas*. Semin. Cell Dev. Biol. *18*, 350–361.

Goutte, C., and Johnson, A.D. (1988). a1 Protein alters the DNA binding specificity of α2 repressor. Cell *52*, 875–882.

Graham, L.E., and Wilcox, L.W. (2000). Algae (Upper Saddle River, NJ: Prentice Hall).

Hake, S., Smith, H.M., Holtan, H., Magnani, E., Mele, G., and Ramirez, J. (2004). The role of KNOX genes in plant development. Annu. Rev. Cell Dev. Biol. *20*, 125–151.

Hanes, S.D., and Brent, R. (1989). DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. Cell *57*, 1275–1283.

Harris, E.H. (1989). The *Chlamydomonas* Sourcebook (San Diego, CA: Academic Press).

Hu, C.D., and Kerppola, T.K. (2003). Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. Nat. Biotechnol. *21*, 539–545.

Hull, C.M., Boily, M.-J., and Heitman, J. (2005). Sex-specific homeodomain proteins Sxi1α and Sxi2a coordinately regulate sexual development in *Cryptococcus neoformans*. Eukaryot. Cell *4*, 526–535.

Karol, K.G., McCourt, R.M., Cimino, M.T., and Delwiche, C.F. (2001). The closest living relatives of land plants. Science *294*, 2351–2353.

Kenrick, P., and Crane, P.R. (1997). The Origin and Early Diversification of Land Plants: A Cladistic Study (Washington, DC: Smithsonian Institution Press).

King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., et al. (2008). The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature *451*, 783–788.

Knoepfler, P.S., Calvo, K.R., Chen, H., Antonarakis, S.E., and Kamps, M.P. (1997). Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. Proc. Natl. Acad. Sci. USA *94*, 14553–14558.

Krusell, L., Rasmussen, I., and Gausing, K. (1997). DNA binding sites recognised in vitro by a knotted class 1 homeodomain protein encoded by the hooded gene, k, in barley (*Hordeum vulgare*). FEBS Lett. *408*, 25–29.

Kues, U., Richardson, W.V., Tymon, A.M., Mutasa, E.S., Gottgens, B., Gaubatz, S., Gregoriades, A., and Casselton, L.A. (1992). The combination of dissimilar alleles of the Aα and Aβ gene complexes, whose proteins contain homeo domain motifs, determines sexual development in the mushroom *Coprinus cinereus*. Genes Dev. *6*, 568–577.

Kurvari, V., Grishin, N.V., and Snell, W.J. (1998). A gamete-specific, sex-limited homeodomain protein in *Chlamydomonas*. J. Cell Biol. *143*, 1971–1980.

Lewis, L.A., and McCourt, R.M. (2004). Green algae and the origin of land plants. Am. J. Bot. *91*, 1535–1556.

Lin, H., and Goodenough, U.W. (2007). Gametogenesis in the *Chlamydomonas reinhardtii* minus mating type is controlled by two genes, *MID* and *MTD1*. Genetics *176*, 913–925.

Mann, R.S., and Morata, G. (2000). The developmental and molecular biology of genes that subdivide the body of *Drosophila*. Annu. Rev. Cell Dev. Biol. *16*, 243–271.

Meisel, L., and Lam, E. (1996). The conserved ELK-homeodomain of KNOTTED-1 contains two regions that signal nuclear localization. Plant Mol. Biol. *30*, 1–14.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., et al. (2007). The *Chlamydomonas* Genome Reveals the Evolution of Key Animal and Plant Functions. Science *318*, 245–250.

Meyerowitz, E.M. (2002). Plants compared to animals: The broadest comparative study of development. Science *295*, 1482–1485.

Minami, S.A., and Goodenough, U.W. (1978). Novel glycopolypeptide synthesis induced by gametic cell fusion in *Chlamydomonas reinhardtii*. J. Cell Biol. *77*, 165–181.

Muller, J., Wang, Y., Franzen, R., Santi, L., Salamini, F., and Rohde, W. (2001). In vitro interactions between barley TALE homeodomain proteins suggest a role for protein-protein associations in the regulation of Knox gene function. Plant J. *27*, 13–23.

Nam, J., and Nei, M. (2005). Evolutionary change of the numbers of homeobox genes in bilateral animals. Mol. Biol. Evol. *22*, 2386–2394.

Pasquale, S.M., and Goodenough, U.W. (1987). Cyclic AMP functions as a primary sexual signal in gametes of *Chlamydomonas reinhardtii*. J. Cell Biol. *105*, 2279–2292.

Pearson, J.C., Lemons, D., and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet. *6*, 893–904.

Qiu, Y.L., Li, L., Wang, B., Chen, Z., Knoop, V., Groth-Malonek, M., Dombrovska, O., Lee, J., Kent, L., Rest, J., et al. (2006). The deepest divergences in land plants inferred from phylogenomic evidence. Proc. Natl. Acad. Sci. USA *103*, 15511–15516.

Ramesh, M.A., Malik, S.-B., and Logsdon, J.M. (2005). A phylogenomic inventory of meiotic genes: evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. Curr. Biol. *15*, 185–191.

Ronquist, F., and Huelsenbeck, J.P. (2003). MrBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics *19*, 1572–1574.

Roth, F.P., Hughes, J.D., Estep, P.W., and Church, G.M. (1998). Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. Nat. Biotechnol. *16*, 939–945.

Sakakibara, K., Nishiyama, T., Deguchi, H., and Hasebe, M. (2008). Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development. Evol. Dev. *10*, in press.

Sano, R., Juarez, C.M., Hass, B., Sakakibara, K., Ito, M., Banks, J.A., and Hasebe, M. (2005). KNOX homeobox genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. Evol. Dev. *7*, 69–78.

Saupe, S.J. (2000). Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. Microbiol. Mol. Biol. Rev. *64*, 489–502.

Scofield, S., and Murray, J.A. (2006). KNOX gene function in plant stem cell niches. Plant Mol. Biol. *60*, 929–946.

Shiu, S.H., Shih, M.C., and Li, W.H. (2005). Transcription factor families have much higher expansion rates in plants than in animals. Plant Physiol. *139*, 18–26.

Singer, S.D., and Ashton, N.W. (2007). Revelation of ancestral roles of KNOX genes by a functional analysis of *Physcomitrella* homologues. Plant Cell Rep. *28*, 28.

Smith, H.M., Boschke, I., and Hake, S. (2002). Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. Proc. Natl. Acad. Sci. USA *99*, 9579–9584.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688–2690.

Stevens, K.E., and Mann, R.S. (2007). A balance between two nuclear localization sequences and a nuclear export sequence governs Extradenticle subcellular localization. Genetics *175*, 1625–1636.

Uchida, H., Kawano, S., Sato, N., and Kuroiwa, T. (1993). Isolation and characterization of novel genes which are expressed during the very early stage of zygote formation in *Chlamydomonas reinhardtii*. Curr. Genet. *24*, 296–300.

Uchida, H., Suzuki, L., Anai, T., Doi, K., Takano, H., Yamashita, H., Oka, T., Kawano, S., Tomizawa, K.I., Kawazu, T., et al. (1999). A pair of invertedly repeated genes in *Chlamydomonas reinhardtii* encodes a zygote-specific protein whose expression is UV-sensitive. Curr. Genet. *36*, 232–240.

Urban, M., Kahmann, R., and Bolker, M. (1996). Identification of the pheromone response element in *Ustilago maydis*. Mol. Gen. Genet. *251*, 31–37.

van Winkle-Swift, K.P. (1977). Maturation of algal zygotes: Alternative experimental approaches for *Chlamydomonas reinhardtii* (Chlorophyceae). J. Phycol. *13*, 225–231.

Viola, I.L., and Gonzalez, D.H. (2006). Interaction of the BELL-like protein ATH1 with DNA: role of homeodomain residue 54 in specifying the different binding properties of BELL and KNOX proteins. Biol. Chem. *387*, 31–40.

Wegener, D., and Beck, C.F. (1991). Identification of novel genes specifically expressed in *Chlamydomonas reinhardtii* zygotes. Plant Mol. Biol. *16*, 937–946.

Willer, M., Hoffmann, L., Styrkarsdottir, U., Egel, R., Davey, J., and Nielsen, O. (1995). Two-step activation of meiosis by the mat1 locus in *Schizosaccharomyces pombe*. Mol. Cell. Biol. *15*, 4964–4970.

Zhao, H., Lu, M., Singh, R., and Snell, W.J. (2001). Ectopic expression of a *Chlamydomonas* mt+-specific homeodomain protein in mt- gametes initiates zygote development without gamete fusion. Genes Dev. *15*, 2767–2777.