LETTERS

Identification of the *sex* genes in an early diverged fungus

Alexander Idnurm[1,2], Felicia J. Walton[1], Anna Floyd[1] & Joseph Heitman[1]

Sex determination in fungi is controlled by a small, specialized region of the genome in contrast to the large sex-specific chromosomes of animals and some plants. Different gene combinations reside at these mating-type (*MAT*) loci and confer sexual identity; invariably they encode homeodomain, α-box, or high mobility group (HMG)-domain transcription factors[1]. So far, *MAT* loci have been characterized from a single monophyletic clade of fungi, the Dikarya (the ascomycetes and basidiomycetes)[2], and the ancestral state and evolutionary history of these loci have remained a mystery. Mating in the basal members of the kingdom has been less well studied, and even their precise taxonomic interrelationships are still obscure[3,4]. Here we apply bioinformatic and genetic mapping to identify the sex-determining (*sex*) region in *Phycomyces blakesleeanus* (Zygomycota), which represents an early branch within the fungi. Each *sex* allele contains a single gene that encodes an HMG-domain protein, implicating the HMG-domain proteins as an earlier form of fungal *MAT* loci. Additionally, one allele also contains a copy of a unique, chromosome-specific repetitive element, suggesting a generalized mechanism for the earliest steps in the evolution of sex determination and sex chromosome structure in eukaryotes.

Ascomycete and basidiomycete fungi exhibit marked variation in the regions of their genomes controlling sexual reproduction. The *MAT* loci comprise two highly dissimilar, idiomorphic alleles. Adjacent conserved genes and conserved intron positions indicate that the loci may derive from a single progenitor. In some (tetrapolar) basidiomycetes, two unlinked *MAT* loci, encoding homeodomain proteins and peptide pheromones/receptors, are involved. In other (bipolar) basidiomycete species, the two loci have coalesced to form a large, single locus[5–7]. Saccharomycotina and Archiascomycete *MAT* loci encode homeodomain and α-box proteins (*Saccharomyces cerevisiae*) or homeodomain, α-box and HMG-domain proteins (*Candida albicans*, *Schizosaccharomyces pombe*)[8]. Some Pezizomycotina (*Neurospora crassa*, *Podospora anserina*) contain a locus in which one allele encodes an HMG-domain protein and the other encodes three genes including another HMG-domain protein that has a less-prominent role in mating[9–12]. In the Microsporidia, which are organisms of ambiguous taxonomic position as fungi or sister-group to the fungi, sex is unknown. However, *Encephalitozoon cuniculi* has a potential *MAT* locus comprising adjacent homeodomain proteins, suggesting that these transcription factors are ancestral[13]. Which sex-determining system (homeodomain, α-box or HMG domain) and arrangement first arose is unknown.

Little is known at the genetic level about how sex determination is controlled in the remainder of the fungal kingdom, even though the abilities to outcross or undergo self-mating were termed heterothallism or homothallism on the basis of studies in one such early diverged lineage, the zygomycetes, a century ago[14,15]. Furthermore, at present only a few species in the order Mucorales of this phylum have a genetically defined sex-determining locus with two alleles that segregate in a 1:1 ratio of (+) to (–) after mating[16–20].

We hypothesized that the types of DNA-binding proteins that function in sex determination in the Dikarya could regulate mating in other fungal lineages. The *Phycomyces blakesleeanus* genome sequence was searched for homologues of Dikarya *MAT*-encoded transcription factors. In reciprocal BLAST searches with the candidates, no predicted *Phycomyces* homeodomain or α-box proteins were closely related to proteins known to regulate fungal mating. However, ten HMG-domain proteins had primary matches to those encoded by ascomycete *MAT* loci.

The ten genes encoding HMG-domain proteins were subjected to polymerase chain reaction (PCR) amplification from the sequenced strain NRRL1555 (–), UBC21 (+) of the opposite mating type, and A56 (+)[21], which is isogenic with strain NRRL1555. Nine genes amplified from all strains (Supplementary Fig. 1). The tenth amplified from NRRL1555 (–), but not from the two (+) strains. Furthermore, this gene amplified from four (–) wild-type strains but was absent from six (+) strains (data not shown), suggesting that it is linked to (–) sex specificity.

The DNA region from the opposite (+) sex, corresponding to the (–) HMG-domain protein, was obtained through PCR and sequenced. Strain UBC21 (+) contains a 5,830-base-pair (bp) sex unique region, compared to the 3,494-bp unique region of NRRL1555 (–) (Fig. 1a). The (+) region also encodes an HMG-domain protein, and the genes were named *sexM* (*sex minus*) and *sexP* (*sex plus*). The SexM and SexP proteins share low-level amino acid similarity (Fig. 1b). However, when the SexP protein is used in tBLASTn analysis of the *Phycomyces* genome, the most similar protein is SexM (*E*-value of 2.3×10^{-13}; 48 positive and 29 identical residues of the 81 most conserved), suggesting that the two genes are divergent homologues (Supplementary Fig. 2). The asymmetric positions of the genes within the sex unique regions and their inverted transcriptional orientation suggest that an ancient DNA inversion event may have occurred, driving evolution of the locus.

Genetic manipulation of *Phycomyces* via transformation has remained an elusive procedure[22], and as a consequence gene function for the candidate locus was demonstrated by analysing the properties of strains harbouring both the *sexM* and *sexP* alleles, and establishing genetic linkage between these genes and sex.

Three sex-disomic progeny were identified from genetic crosses that did not mate with either (+) or (–) tester strains. The strains have an odd 'fluffy' colony morphology due to the production of zygophore-like structures (Fig. 2a), suggesting that they are partially self-fertile. One (NRRL1555 × NRRL1554 progeny 54) occasionally even produced zygospores like a normal (+) × (–) cross. A sex-heterozygous strain (B36*A87), previously generated by forced

[1]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina 27710, USA. [2]School of Biological Sciences, University of Missouri-Kansas City, Kansas City, Missouri 64110, USA.

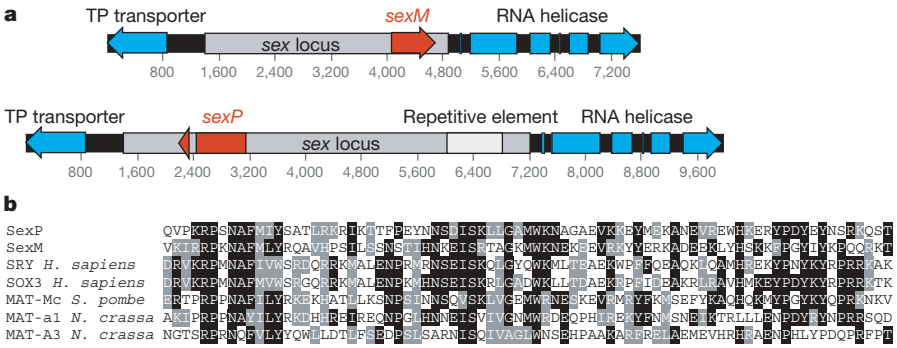


Figure 1 | Structure of the *sex* locus of *Phycomyces blakesleeanus*. **a**, Alignment of the sex unique sequence (shaded grey) and conserved flanking regions (99% identical) encoding a putative triose phosphate (TP) transporter and RNA helicase in strains NRRL1555 (–) and UBC21 (+). The alleles are 3,494 bp (NRRL1555) and 5,830 bp (UBC21), and *sexM* and *sexP* encode HMG-domain proteins. **b**, Alignment of the HMG domain of predicted proteins encoded by *sexM* and *sexP* with human SRY and SOX3, *S. pombe* MAT-Mc, and *N. crassa* MAT-A3 and MAT-a1. Black indicates 100% sequence identity and grey similar amino acid residues.

fusion of strains of the opposite sex[23], was obtained for comparison. Analysis of the *sex* gene content in these four fluffy strains showed that they all contain both *sexM* and *sexP* genes (Fig. 2b), confirming their heterozygous nature. Current evidence suggests that fluffy strains derived from crosses are partial diploids containing both *sex* alleles[24]. The three strains identified here behaved similarly to those previously reported[23–25], and exhibited variation in zygophore production, instability after passaging to revert to (+) or (–) sex, and co-segregating *sexP* and *sexM* genotypes, evidence for pheromone production and aneuploidy (Supplementary Fig. 3, Supplementary Data, and data not shown). When *sexM* and *sexP* expression was examined by northern blot analysis, transcript abundance was low in strains grown in isolation. In contrast, transcripts were at higher levels in sex-heterozygous strains and induced during (–) × (+) matings (Fig. 2c). The role of transcript levels in mating is currently unclear: future research in *Phycomyces* and Mucorales species will investigate how the Sex proteins are regulated by environmental conditions, and govern pheromone signalling and other meiotic spore developmental stages.

Genetic segregation data from three crosses of different strains with the wild-type sequenced strain demonstrated that the *sex* genes correspond to the *sex* locus. In the first cross, 69 progeny (one each from 69 zygospores) were obtained from NRRL1555 × NRRL1554. In 68 out of 69 cases, the (–) sex co-segregated with the *sexM* gene and (+) sex co-segregated with the *sexP* gene (χ^2 test, $P < 0.001$, Fig. 3 and Supplementary Information). One notable exception was heterozygous for both *sexM* and *sexP*, as discussed above. Linkage between sex and the HMG-domain genes was confirmed in a second cross previously conducted to establish linkage between the *madA* mutation and loss-of-phototropism in which sex served as an independent genetic marker[26]. Analysis of this cross (*madA* × A56, 63 progeny from 24 zygospores) showed that the (–) sex co-segregated with *sexM* and the (+) sex co-segregated with *sexP* ($P < 0.001$), whereas *madA* segregated independently, confirming linkage in a set of known recombinant progeny.

In a third cross to delimit the *sex* locus and examine the possibility that *Phycomyces* might contain a sex chromosome, we tested recombination in 104 progeny (derived from 104 zygospores, NRRL1555 × UBC21). Again, all (–) progeny contained *sexM* and all (+) progeny contained *sexP* ($P < 0.001$). Nine polymorphisms between UBC21 and NRRL1555 on contigs other than the *sexM*-containing region segregated independently from sex (Supplementary Information). In contrast, markers on the 377-kilobase (kb) sequenced contig containing *sexM* were linked to the sex phenotype. The sex phenotype was then mapped to a small region of this chromosome. Recombination was observed on one side of the *sex* locus but not the other (Fig. 4a). Similar frequencies were observed for the NRRL1554 × NRRL1555 cross, however, 1 out of 68 recombined on the suppressed side of the locus, as had a single progeny from the *madA* × A56 cross, providing recombinant strains for fine mapping. Polymorphisms closer to the locus were identified and, using the rare recombinants,

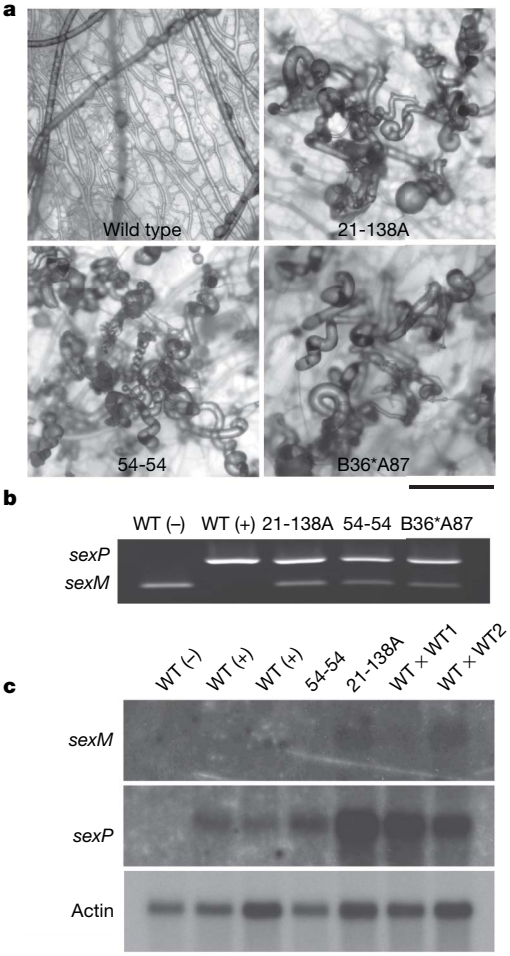
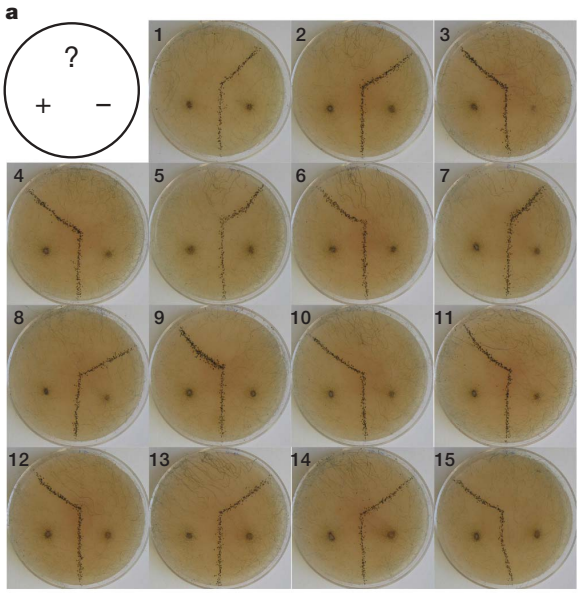


Figure 2 | Heterozygosity at the *sex* locus promotes a partially self-fertile phenotype. **a**, Morphology of wild type and *sexM-sexP* heterozygotes. B36*A87 is a defined *sex*-heterozygous strain. UBC21 × NRRL1555 138A (21-138A) and NRRL1554 × NRRL1555 54 (54-54) are unusual progeny derived from crosses. Scale bar, 200 μm. **b**, *sexM* or *sexP* in strains assessed by multiplex PCR. **c**, Transcript abundance of *sexM* and *sexP* in strains. The blot was sequentially probed with *sexP* (3 days, −80 °C exposure), *sexM* (4 days, −80 °C exposure) and actin (30 min, 22 °C exposure). Lane 1, NRRL1555 (–); lane 2, NRRL1554 (+); lane 3, UBC21 (+); lane 4, NRRL1555 × NRRL1555 54 cross; lane 5, UBC21 × NRRL1555 138A cross; lane 6, NRRL1554 × NRRL1555 cross; lane 7, UBC21 × NRRL1555 cross.



c

Cross	Phenotype and genotype				Probability
	(+) *sexP*	(–) *sexM*	(+) *sexM*	(–) *sexP*	
NRRL1555 × NRRL1554	33	35	0	0	$P < 0.001$
madA × A56	35	28	0	0	$P < 0.001$
NRRL1555 × UBC21	57	45	0	0	$P < 0.001$

Figure 3 | Sex co-segregates with *sexM* and *sexP* genes encoding HMG-domain proteins. **a**, Fifteen progeny of an NRRL1555 × NRRL1554 cross (from 69 analysed) were assessed for sex by crosses against (+) and (–) tester parents. Three strains were co-inoculated on V8 medium; the black lines of zygospores indicate mating, with diagonals to the right and left indicating (+) or (–) progeny, respectively, and a vertical line the positive control interaction between tester strains. **b**, The *sexM* and *sexP* genes were amplified by PCR from genomic DNA from the 15 progeny in panel **a**. **c**, Segregation data for crosses between strains (excluding heterozygous progeny).

sex was mapped to within 10 kb and 18 kb of the sex unique region (Fig. 4a). Thus, there is skewed recombination around the *Phycomyces sex* locus and it maps to less than 34 kb within the 65-megabase (Mb) genome.

The *plus* sex allele is 2,336 bp larger than the *minus* sex allele for the two strains sequenced. Analysis of the DNA sequence was undertaken to understand the basis for this size difference. A number of different small repetitive elements are present in the *plus* allele. Notably, one element is also present in the genome as two additional copies, both found on the same contig/chromosome as the *sex* locus (Fig. 4b). All three elements are aligned as direct repeats, so mitotic recombination would delete a large region of DNA including *sexP*, leading to sterility or lethality.

Although *Phycomyces* contains a *sex* locus rather than a sex chromosome, its locus structure and repetitive elements may provide insight into the events that shape sex-determining regions more generally. We hypothesize that in both *Phycomyces* and mammals a region of DNA containing a sex-determining HMG-domain gene underwent inversion, thereby blocking recombination and leading to divergence and cell-type specificity (*Phycomyces sexM-sexP*, mammalian *SRY-SOX3*). A similar scenario to the chromosome-specific repeats of *Phycomyces* in an indirect orientation in the early mammalian lineage could have facilitated a larger inversion, leading to the development of a sex chromosome (Supplementary Fig. 4), as is predicted to have occurred for the Y chromosome[27–29]. Simple inversion events probably represent a common mechanism driving the initiation and subsequent expansion of sex-determining regions in diverse lineages.

An early diverged fungus with this gene composition at *sex* suggests that either a locus comprising HMG-domain genes is ancestral to the Dikarya and Zygomycota or sex-determination evolved independently in the two. A similar region is present in another zygomycete genome, *Rhizopus oryzae*, where the candidate *sexP* allele is flanked by the same two genes flanking the *Phycomyces* locus (Supplementary Fig. 5). We currently favour the former hypothesis that the HMG-domain arrangement was ancestral in fungi, and has gradually lost function from some lineages by divergence in sequence or eviction from the locus to become a downstream target of homeodomain *MAT*-encoded proteins (see Supplementary Information for further discussion). However, the *MAT* alleles in the ascomycetes are more different in comparison to *Phycomyces*. This may indicate that functional constraints slow divergence of SexM and SexP or that

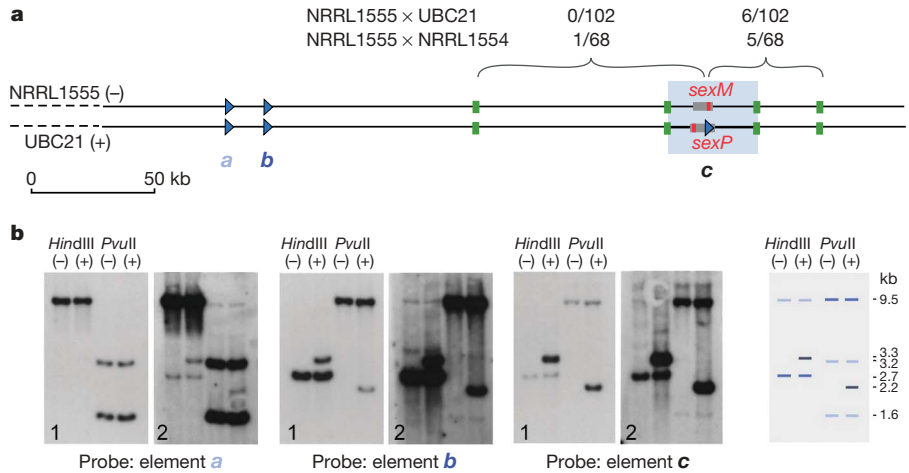


Figure 4 | Recombination around the *sex* locus and presence of chromosome-specific repetitive elements. **a**, Polymorphic markers approximately 50 and 87 kb from *sexM*/*sexP* were scored in progeny from crosses between NRRL1555 with either UBC21 or NRRL1554. In the subset exhibiting recombination, closer markers (blue box) map the locus within 10 kb and 18 kb of *sexM*/*sexP*. *a*, *b* and *c* represent repetitive elements. **b**, Southern blot hybridizations of chromosome-specific repetitive elements. Strain UBC21 contains a third element (*c*) within the *sex* locus. DNA from NRRL1555 or UBC21 was probed with each element (*a*, *b*, or *c*). Two autoradiograph exposures (1 and 2) illustrate cross-hybridization. Restriction maps for elements are summarized on the right.

an inter-allelic gene conversion event reset the locus. Only a better understanding of sex in fungi, particularly among the diverse non-Dikarya lineages, and identification of sex/mating-type loci will fully reveal the evolutionary trajectory of this sex-determining region in the fungal kingdom.

METHODS SUMMARY

Genes present in characterized fungal *MAT* loci as well as the human *SRY* HMG-domain male determinant were used to search the *Phycomyces* genome database. Candidate genes were amplified by PCR, with one gene encoding an HMG-domain protein amplifying only from (–) strains. The equivalent region in a (+) strain was obtained by inverse and regular PCR, and sequenced. Genetic crosses and segregation analyses were used to demonstrate that these genes correspond to the *sex* locus. The two genes were amplified from DNA from these progeny strains and compared with their sex type. To identify molecular genetic markers, ~160 kb of DNA derived from genomic libraries and PCR products of genes or regions of interest was sequenced from strain UBC21, as well as regions flanking *sex* in NRRL1554, and the sequence compared to the NRRL1555 genome database. Genetic markers linked and unlinked to the putative *sex* locus enabled mapping to a small genomic region. Three strains with unusual morphology were identified from crosses and their content and expression of the *sexM* and *sexP* genes determined by PCR and northern blot analyses, respectively. The low copy number of a repetitive element in the *plus* allele, as suggested by the sequencing project, was confirmed by Southern blot analysis.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 August; accepted 5 November 2007.

1. Fraser, J. A. & Heitman, J. Evolution of fungal sex chromosomes. *Mol. Microbiol.* **51,** 299–306 (2004).
2. Heitman, J., Kronstad, J. W., Taylor, J. W. & Casselton, L. A. (eds) *Sex in Fungi: Molecular Determination and Evolutionary Implications* (ASM Press, Washington, DC, 2007).
3. James, T. Y. *et al.* Reconstructing the early evolution of fungi using a six-gene phylogeny. *Nature* **443,** 818–822 (2006).
4. White, M. M. *et al.* Phylogeny of the Zygomycota based on nuclear ribosomal sequence data. *Mycologia* **98,** 872–884 (2006).
5. Bakkeren, G. *et al.* Mating factor linkage and genome evolution in basidiomycetous pathogens of cereals. *Fungal Genet. Biol.* **43,** 655–666 (2006).
6. Lee, N., Bakkeren, G., Wong, K., Sherwood, J. E. & Kronstad, J. W. The mating-type and pathogenicity locus of the fungus *Ustilago hordei* spans a 500-kb region. *Proc. Natl Acad. Sci. USA* **96,** 15026–15031 (1999).
7. Fraser, J. A. *et al.* Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biol.* **2,** e384 (2004).
8. Butler, G. *et al.* Evolution of the *MAT* locus and its Ho endonuclease in yeast species. *Proc. Natl Acad. Sci. USA* **101,** 1632–1637 (2004).
9. Staben, C. & Yanofsky, C. *Neurospora crassa a* mating-type region. *Proc. Natl Acad. Sci. USA* **87,** 4917–4921 (1990).
10. Glass, N. L., Grotelueschen, J. & Metzenberg, R. L. *Neurospora crassa A* mating-type region. *Proc. Natl Acad. Sci. USA* **87,** 4912–4916 (1990).
11. Ferreira, A. V.-B., An, Z., Metzenberg, R. L. & Glass, N. L. Characterization of *mat A-2, mat A-3* and Δ*matA* mating-type mutants of *Neurospora crassa. Genetics* **148,** 1069–1079 (1998).
12. Arnaise, S., Zickler, D., Le Bilcot, S., Poisier, C. & Debuchy, R. Mutations in mating-type genes of the heterothallic fungus *Podospora anserina* lead to self-fertility. *Genetics* **159,** 545–556 (2001).
13. Bürglin, T. R. The homeobox genes of *Encephalitozoon cuniculi* (Microsporidia) reveal a putative mating-type locus. *Dev. Genes Evol.* **213,** 50–52 (2003).
14. Blakeslee, A. F. Zygospore formation is a sexual process. *Science* **19,** 864–866 (1904).
15. Blakeslee, A. F. Sexual reproduction in the Mucorineae. *Proc. Am. Acad. Arts Sci.* **40,** 205–319 (1904).
16. Cerdá-Olmedo, E. The genetics of *Phycomyces blakesleeanus. Genet. Res.* **25,** 285–296 (1975).
17. Alvarez, M. I., Peláez, M. I. & Eslava, A. P. Recombination between ten markers in *Phycomyces. Mol. Gen. Genet.* **179,** 447–452 (1980).
18. Eslava, A. P., Alvarez, M. I., Burke, P. V. & Delbrück, M. Genetic recombination in sexual crosses of *Phycomyces. Genetics* **80,** 445–462 (1975).
19. Eslava, A. P., Alvarez, M. I. & Delbrück, M. Meiosis in *Phycomyces. Proc. Natl Acad. Sci. USA* **72,** 4076–4080 (1975).
20. Gauger, W. L. Meiotic gene segregation in *Rhizopus stolonifer. J. Gen. Microbiol.* **101,** 211–217 (1977).
21. Alvarez, M. I. & Eslava, A. P. Isogenic strains of *Phycomyces blakesleeanus* suitable for genetic analysis. *Genetics* **105,** 873–879 (1983).
22. Obraztsova, I. N., Prados, N., Holzmann, K., Avalos, J. & Cerdá-Olmedo, E. Genetic damage following introduction of DNA in *Phycomyces. Fungal Genet. Biol.* **41,** 168–180 (2004).
23. Govind, N. S. & Cerdá-Olmedo, E. Sexual activation of carotenogenesis in *Phycomyces blakesleeanus. J. Gen. Microbiol.* **132,** 2775–2780 (1986).
24. Mehta, B. J. & Cerdá-Olmedo, E. Intersexual partial diploids of *Phycomyces. Genetics* **158,** 635–641 (2001).
25. Burgeff, H. Untersuchungen über Variabilität, Sexualität und Erblichkeit bei Phycomyces nitens Kuntze. *Flora* **107,** 259–316 (1914).
26. Idnurm, A. *et al.* The *Phycomyces madA* gene encodes a blue-light photoreceptor for phototropism and other light responses. *Proc. Natl Acad. Sci. USA* **103,** 4546–4551 (2006).
27. Lahn, B. T. & Page, D. C. Four evolutionary strata on the human X chromosome. *Science* **286,** 964–967 (1999).
28. Skaletsky, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423,** 825–837 (2003).
29. Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95,** 118–128 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge access to the *Phycomyces* genome project: these sequence data were produced by the US Department of Energy Joint Genome Institute. We thank L. Corrochano and A. Eslava for encouragement and providing *Phycomyces* strains, and L. Corrochano and X. Lin for comments on the manuscript. This research was supported by National Institutes of Health grants to J.H.

Author Information DNA sequences for the reported genes have been deposited at GenBank under accessions EU009461 and EU009462. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.H. (heitm001@mc.duke.edu).

METHODS

Genetic crosses. Strains were inoculated ~6 cm apart on V8 juice agar media and incubated until zygospores were produced. Zygospores were removed to wet filter paper and allowed to germinate (3–4 months dormancy). Individual meiotic spores were isolated on V8 juice agar as serial dilutions. The sex of each progeny was assessed by crossing to NRRL1555 (–), A56 (+), UBC21 (+) or NRRL1554 (+) strains on V8 juice agar or potato dextrose agar. To assess pheromone production, strains were inoculated ~4 cm apart on 1-mm thick Sutter I, V8 or potato dextrose agar medium and examined 2–4 days later for the confrontation reactions.

Statistical analyses. χ^2 tests were used with the null hypothesis that genetic markers were unlinked and that any difference from the 1:1:1:1 expected distribution was due to chance.

Nucleic acid extraction and manipulations. Genomic DNA was extracted from mycelium with a CTAB extraction buffer. For Southern blots, genomic DNA was digested with restriction enzymes and transferred to Zeta-Probe membrane (Bio-Rad). For transcript analysis, strains were grown for 5 days on potato dextrose agar and in Fig. 2 the plates overlaid with dialysis membrane (Spectra/Por, 3,500-Mw cutoff, Spectrum Medical Instruments) to aid removal of mycelium without agar contamination. RNA was extracted with TRIzol reagent (Invitrogen) using the manufacturer's instructions, and 120 μg was further purified with the polyATtract system (Promega). RNA was resolved on 1.4% agarose/formaldehyde gels and transferred to Zeta-Probe membrane. Membranes were probed with [^{32}P]dCTP-labelled DNA fragments. To identify the ends of the *sexM* and *sexP* genes, 5′ and 3′ RACE was preformed using cDNA templates prepared from RNA of the heterozygous strain UBC21 × NRRL155 progeny 138A, with the GeneRacer kit (Invitrogen).

PCR reactions. Primers to amplify *sexM* are JOHE18353 (5′-CCTAGCATTCTATCAAGC-3′) and JOHE18354 (5′-GTAGTCTTGTGAGGTAGC-3′), and *sexP* are JOHE18168 (5′-GCAAATGTGTACCGAAAGC-3′) and JOHE18169 (5′-ATTATTACAAAAGTCCTGCC-3′). Additional primers and those used for RFLP mapping are listed in Supplementary Table 2.

DNA and protein sequence analyses. Genome databases were searched with BLASTp, BLASTn or tBLASTn programs. Gene predictions from unannotated genomic DNA were made with FGENESH software using the settings for *Schizosaccharomyces pombe.* HMG domains were aligned in ClustalW, in MacClade 4.08, and by eye. Neighbour-joining was performed with ClustalW; maximum likelihood and parsimony, with 100–500 bootstrap re-samplings, were performed in PHYLIP 3.67, using default settings unless specified.

Generation of sequencing libraries and sequencing *sexM* and *sexP* alleles. Sequencing libraries were generated to identify polymorphisms between (+) strains and the NRRL1555 (–) strain sequenced by the US Department of Energy. Genomic DNA of strains UBC21 and NRRL1554 was digested with *Hin*dIII restriction enzyme. Cut DNA was resolved on 0.8% agarose gels, and two gel slices of sizes between 4–6 kb and 6–10 kb were excised. Purified cut DNA was ligated into the *Hin*dIII site of plasmid pBluescript II KS (–). Plasmid DNA was purified with QIAprep spin miniprep kits (QIAGEN Sciences) and sequenced with the T3 and T7 universal primers. Additional sequence was obtained by PCR amplification and sequencing of regions flanking *sex* in strains NRRL1555, UBC21, NRRL1554 and A56. Sequence chromatograms were compared to the NRRL1555 database to identify polymorphisms. Fragments of the *sexM* and *sexP sex* locus alleles from strains NRRL1555 and UBC21, respectively, were amplified by PCR and sequenced to a minimal double-strand coverage, with the chromatograms assembled in Sequencher. The sequences of the two strains were submitted to GenBank and assigned accession numbers EU009462 for NRRL1555 (–) and EU009461 for UBC21 (+).