The Plant Cell, Vol. 6, 1467–1475, October 1994 © 1994 American Society of Plant Physiologists

Two Classes of Homeodomain Proteins Specify the Multiple *A* Mating Types of the Mushroom *Coprinus cinereus*

Ursula Kües,[a] Rachel N. Asante-Owusu,[a] Effie S. Mutasa,[b] Anna M. Tymon,[b] Eneida H. Pardo,[a] Suzanne F. O'Shea,[a] Berthold Göttgens,[a] and Lorna A. Casselton [a,1]

[a] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, United Kingdom
[b] School of Biological Sciences, Queen Mary and Westfield College, Mile End Road, London, E1 4NS, United Kingdom

The *A* mating type locus of the mushroom *Coprinus cinereus* regulates essential steps in sexual development. The locus is complex and contains several functionally redundant, multiallelic genes that encode putative transcription factors. Here, we compare four genes from an *A* locus designated *A42*. Overall, the DNA sequences are very different (~50% homology), but two classes of genes can be distinguished on the basis of a conserved homeodomain motif in their predicted proteins (HD1 and HD2). Development is postulated to be triggered by an *HD1* and an *HD2* gene from different *A* loci. Thus, proteins encoded by genes of the same locus must be distinguished from those encoded by another locus. Individual proteins of both classes recognize each other using the region N-terminal to the homeodomain. These N-terminal specificity regions (COP1 and COP2) are predicted to be helical and are potential dimerization interfaces. The amino acid composition of the C-terminal regions of HD1 proteins suggests a role in activation, and gene truncations indicate that this region is essential for function in vivo. A corresponding C-terminal region in HD2 proteins can be dispensed with in vivo. We will discuss these predicted structural features of the *C. cinereus* A proteins, their proposed interactions following a compatible cell fusion, and their similarities to the a1 and α2 mating type proteins of the yeast *Saccharomyces cerevisiae*.

INTRODUCTION

A successful mating in the basidiomycete fungus *Coprinus cinereus* triggers a major switch in mycelial cell type; a sterile monokaryon with uninucleate cells is converted to a fertile binucleate-celled dikaryon that differentiates the mushroom fruit bodies. No special cells are required for mating; hyphal fusion is sufficient, but dikaryon formation only follows if cells have different mating types. *C. cinereus* has multiple mating types determined by multiallelic genes at two unlinked loci that are known as *A* and *B* (Casselton and Kües, 1994). It is the genes at these loci that regulate the developmental pathway that gives rise to the dikaryon (Swiezynski and Day, 1961).

The major phenotypic differences between mated and unmated cells are the abundantly produced uninucleate asexual spores (oidia) on the monokaryon and the characteristic clamp connections between each cell of the dikaryon; both are regulated by the *A* locus (Tymon et al., 1992). Our molecular analysis of three *A* loci has shown that there are several genes separated into two subcomplexes (Kües et al., 1992, 1994a, 1994b) that correspond to the closely linked *Aα* and *Aβ* loci originally defined by classical recombination analysis (Day, 1960) (Figure 1). Two genes (*α-fg* and *β-fg*) mark the boundaries of the *A* locus and appear from DNA gel blot analysis to be present in all strains we have examined. The *A42* locus contains five genes, one in the α complex and four in the β complex. These genes are multiallelic and, remarkably, their different alleles have such dissimilar DNA sequences that they fail to cross-hybridize. One of the genes (*c1-1*) appears to be inactive, but the other four genes determine *A42* mating type specificity. Each of these specificity genes individually promotes *A*-regulated clamp cell development if introduced into a host with an *A5* locus that shares none of these genes with *A42* (Kües et al., 1992).

Preliminary sequence data identified homeodomain motifs in the predicted proteins of all four *A42* genes, indicating that they encode transcription factors (Kües et al., 1992). Two classes of proteins were distinguished on the basis of different amino acid sequences in the homeodomain, and we have called them HD1 and HD2. These two homeodomain motifs are conserved in other fungal mating type proteins. The *Aα* locus of *Schizophyllum commune* contains an *HD1* and an *HD2* gene (Stankis et al., 1992), as does the *b* mating type locus of two *Ustilago* spp (Gillissen et al., 1992; Bakkeren and Kronstad, 1993). The HD1 and HD2 motifs are also related to the different homeodomains present in the α2 and a1 products of the alternative mating-type loci of the ascomycete fungus *Saccharomyces cerevisiae* (Kües and Casselton, 1992a, 1992b; Shepherd et al., 1984). The *S. cerevisiae* proteins are known to heterodimerize following mating to generate a

[1] To whom correspondence should be addressed.

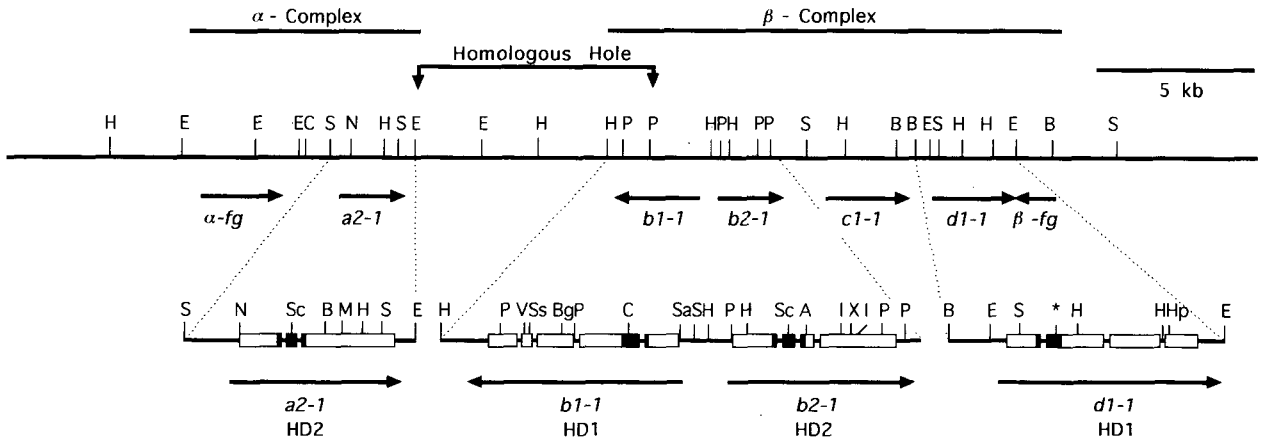


Figure 1. Genes of the *A42* Locus of *C. cinereus*.

Genes of the α and β complexes are separated by 7 kb of noncoding sequence (homologous hole) (Kües et al., 1992). Gene designations are according to Kües and Casselton (1993). Arrows indicate direction of transcription. Coding regions of the four specificity genes are represented as boxes with homeodomain coding regions shown in black. Abbreviations for restriction sites are A, AsuII; B, BamHI; Bg, BglII; C, ClaI; E, EcoRI; H, HindIII; Hp, HpaI; I, HindII; M, MstI; N, NcoI; P, PstI; S, SalI; Sa, SacII; Sc, ScaI; Ss, SspI; V, PvuII; X, SmaI. The asterisk indicates position of a ClaI site in the *d1-1* gene in the *A6* locus; this gene was used to construct a chimeric *HD1* gene. The dotted lines indicate sequences expanded to show details of restriction sites and introns within the genes relevant to this study.

functional transcription factor that binds target sites upstream of developmentally regulated genes (Dranginis, 1990; Goutte and Johnson, 1993; Mak and Johnson, 1993). Transformation studies with host cells having genes naturally or experimentally deleted leads to the hypothesis that a similar interaction is likely to occur between the HD1 and HD2 mating type proteins of the basidiomycete fungi *U. maydis*, *S. commune*, and *C. cinereus* (Gillissen et al., 1992; Specht et al., 1992; Kües et al., 1994b).

In this study, we examine the complete sequences of the four *A42* specificity genes, predict putative functional domains, and describe experiments that permit us to suggest how HD1 and HD2 proteins might interact to regulate sexual development.

RESULTS

Sequences of the *A42* Specificity Genes Identify Two Classes of Putative Transcription Factors

The position of each of the four functional *A42* specificity genes (*a2-1*, *b1-1*, *b2-1*, and *d1-1*) was mapped previously from data obtained using transformation, RNA gel blot analysis, and cDNA isolation (Kües et al., 1992) (see Figure 1). The complete sequence of *b1-1* was described by Tymon et al. (1992), and we have now obtained the complete genomic and cDNA sequences of the three other genes (the EMBL accession number for CCA21 is X79686; CCB21, X79687; CCD11, X79688). Preliminary sequence data defined *b1-1* and *d1-1* as *HD1* genes and *a2-1* and *b2-1* as *HD2* genes (Kües et al., 1992), but lack of hybridization between genes of the same class suggested that there was little overall conservation of DNA sequence. We now show that *b1-1* and *d1-1* have 51% DNA sequence homology and *a2-1* and *b2-1* have 47%. The genes have two to four introns of ~50 bp each (Figure 1). The most conserved sequences are those encoding the homeodomain motif (65 and 66% homology for *HD1* and *HD2*, respectively); in each class of gene these are interrupted at similar positions by introns having much less similarity (*HD1* introns, 55%, *HD2* introns, 38%, data not shown). It is particularly interesting to note that the intron in the region that encodes the WF.N.R motif of the recognition helix in *HD2* genes (data not shown) is conserved at exactly the same position in the corresponding genes of other basidiomycetes (Gillissen et al., 1992; Stankis et al., 1992; Bakkeren and Kronstad, 1993) and also in the *a1* gene of *S. cerevisiae* (Miller, 1984).

The two HD1 proteins have 632 and 633 amino acids with 42% identity (56% similarity) and the two HD2 proteins 520 amino acids with 35% identity (55% similarity) (Figure 2). Both *HD1* genes have two possible in-frame start codons (Figure 3), and we previously assumed that translation initiated from the first ATG (Tymon et al., 1992). By changing this ATG in *b1-1* to ACG (in pAMT1, Table 1) and deleting the first possible start codon from the cloned *d1-1* gene by cutting with EcoRI (pUK16, Table 1 and Figure 3), we found that both genes are active in our transformation assay. Translation can, therefore, start at the second ATG. Significantly, this would make the length of the region N-terminal to the HD1 homeodomain (108 to 115 amino acids) similar to that in the corresponding HD1 mating

type proteins of other basidiomycetes (AαZ of *Schizophyllum commune*, Stankis et al., 1992, and bE of *Ustilago* spp, Kronstad and Leong, 1990; Schultz et al., 1990; Bakkeren and Kronstad, 1993) and the α2 mating type protein of *S. cerevisiae* (Astell et al., 1984).

The essential promoter elements for *b1-1* and *d1-1* lie within 155 and 177 bp, respectively, of the ATG start codon just defined (Figure 3). This was demonstrated by transformation using a SalI-PstI fragment containing *b1-1* (pAMT3) and an EcoRI fragment containing *d1-1* (pUK16). Complete cDNAs lacking these 5′ sequences and a genomic *b1-1* clone were inactive (pCB1-1, pCD1-1, and pRAO4, Table 1). Similarly for the *HD2* genes, deletion of sequences close to an NcoI site containing the start codon of *a2-1* and at a PstI site just in front of the start codon of *b2-1* (Figure 1) led to gene inactivation (pCA2-1 and pESM7 in Table 2). Comparisons of the 5′ sequences of all four genes reveal no obvious common promoter elements other than TG-rich stretches in front of *HD1* genes (Figure 3) and TA-rich stretches in front of *HD2* genes (data not shown).

Protein Structure

A possible organization of the *b1-1*–encoded HD1 protein with putative domains for DNA binding, dimerization, and

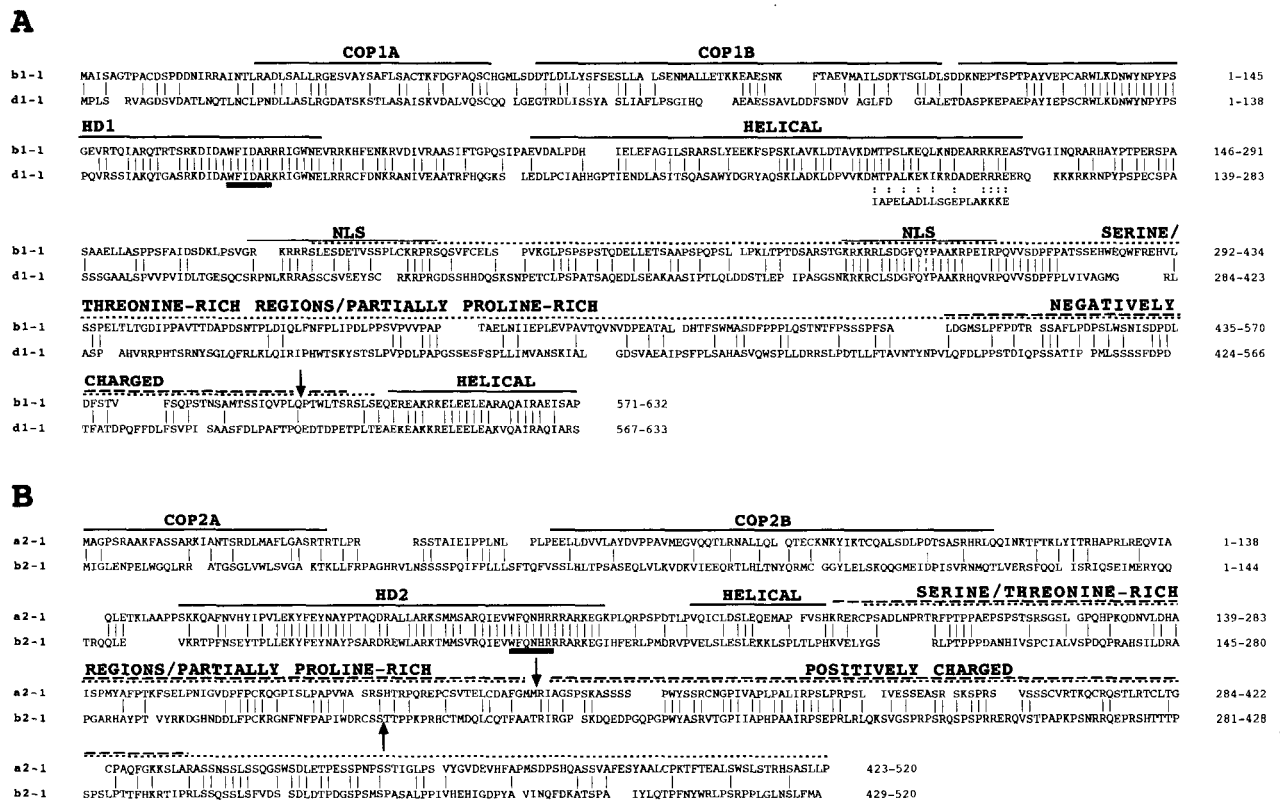


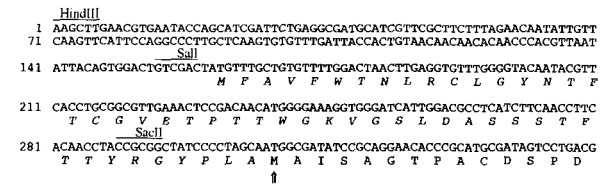
Figure 2. Protein Comparisons.

(A) Alignment of the HD1 proteins b1-1 and d1-1.

(B) Alignment of the HD2 proteins a2-1 and b2-1.

Sequence alignments show the degree of amino acid sequence conservation (vertical lines) and structural features of interest. Relevant features discussed in the text are indicated above the sequence. Continuous bold lines cover predicted helical regions. In the N-terminal regions, these are implicated in specificity and are called COP. The homeodomains are marked by HD1 and HD2; the bar below each indicates amino acids that are highly conserved in all homeodomains (Bürglin, 1994). Homology and similarities to the 19-amino acid tail of the *S. cerevisiae* α2 protein (Astell et al., 1981) is indicated by colons under the proposed helical region C-terminal to the homeodomain of b1-1 and d1-1. Thin lines indicate nuclear localization signals (NLS), and dashed lines cover serine-, threonine-, and proline-rich sequences (24 and 21% S and T in the HD1 proteins and 16 and 13% P between residues 325 to 565; 25 and 18% S and T in the HD2 proteins and 18 and 16% P between residues 240 to 400). Heavily dashed lines indicate a negatively charged region in HD1 proteins (18 and 23% D, E, Q, and N) and a positively charged region in HD2 proteins (15 and 16% R and K). Arrows mark positions at which the C-terminal regions were truncated without loss of ability to elicit *A*-regulated development.

A *b1-1*



B *d1-1*

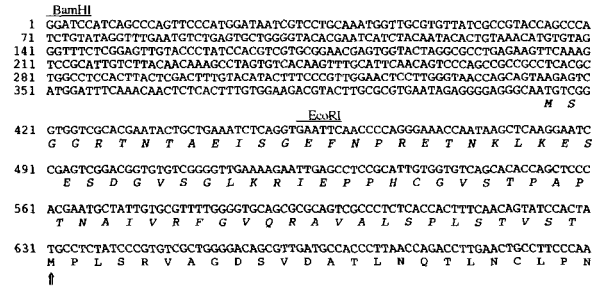


Figure 3. Promoter Sequences and 5′ Ends of the *HD1* Genes from the *A42* Locus.

(A) *b1-1.*

(B) *d1-1.*

Amino acids shown in italics are those that would occur if translation started at the first in-frame ATG. The arrows mark the ATG known to be sufficient for translation of an active gene product.

transactivation was presented by Tymon et al. (1992). Our analysis of the d1-1 protein predicts a similar structure (Figures 2A and 4). The homeodomain, which identifies the proteins as putative transcription factors, lies 108 to 115 amino acids from the N terminus. Two α-helical regions are predicted in the N-terminal region and another just C-terminal to the homeodomain. This latter contains a sequence similar to the short charged C-terminal tail of the α2 mating type protein of *S. cerevisiae* (see Figure 2) that is thought to be involved in dimerization with a1 (Mak and Johnson, 1993). Based on studies of other transcription factors (Sutherland et al., 1992; Tranche and Yaniv, 1992; Leuther et al., 1993; Wegner et al., 1993), we can suggest that possible activation domains are the C-terminal regions (amino acids 310 to 610) that are rich in serine and threonine and partially proline-rich and/or a negatively charged sequence near the C terminus (amino acids 550 to 610) (Figure 2A). Another predicted helical region is found at the extreme C terminus; this region is highly conserved but is known from gene truncations to be dispensable for promoting clamp cell development in vivo (Tymon et al., 1992).

The shorter HD2 proteins encoded by *a2-1* and *b2-1* have 35% identity to each other (Figure 2B). Although HD2 proteins have little sequence similarity to HD1 proteins, computer analysis predicts a similar helical structure on either side of the homeodomain. The C-terminal half of each protein is rich in serine and threonine and partially proline-rich and has an overall positive charge caused by a high lysine and arginine content (Figure 2B).

The BLAST searching program clearly identifies a relationship between the A proteins of *C. cinereus* and *S. commune* and the corresponding bE and bW proteins of *U. maydis.* The regions most similar in the *C. cinereus* and *S. commune* proteins are the homeodomains, the putative helical regions that flank them, and the conserved C terminus of the HD1 proteins. The *C. cinereus* A proteins and *U. maydis* b proteins are most similar in the homeodomains and adjacent helical regions and in the putative bipartite nuclear localization signals (see Figure 2B) (Kronstad and Leong, 1990; Tymon et al., 1992). The computer search aligns the HD2 homeodomain with its conserved WF.N.R in the recognition helix to the classical homeodomain motif found in a range of animal transcription factors (e.g., human Hox2F, frog Hox3, *Caenorhabditis* Ceh7, and Drosophila Ant; for compilation, see Bürglin, 1994), whereas the HD1 homeodomains with WF.D.R in the recognition helix and extra amino acids between helix I and II (Bürglin, 1994) only align to atypical motifs, such as those found in Knotted-1 of maize (Vollbrecht et al., 1991).

3′ Gene Truncations Identify Nonessential C-Terminal Regions of the Proteins

Transforming DNA generally integrates ectopically in *C. cinereus* (Binninger et al., 1987; Mellon et al., 1987). When an *A*

Table 1. Effect of 3′ Truncations on the Ability of the *HD1* Genes *b1-1* and *d1-1* to Promote Clamp Cell Development and Repress Asexual Sporulation in an *A5* Host

Designations	Fragment	HD1 Protein (aa)	Extra (aa)[a]	Clamp Cells	Oidia
b1-1					
pAMT1	2.9-kb H, H	632	–	+	–
pAMT3	2.1-kb S, P, P	594	20	+	–
pRAO4	2.1-kb Sc, P, P[b]	594	20		
pRAO3	1.8-kb S, S	501	25	–	+
pRAO2	1.7-kb S, Ss	488	25	–	+
pRAO1	1.4-kb S, Bg	385	17	–	+
pAMT6	1.4-kb H, P	335	83	–	+
pCB1-1	2.1-kb cDNA[b]	632	–	–	+
d1-1					
pESM2	4.5-kb B, B	632	–	+	–
pUK16	2.5-kb E, E	632	–	+	–
pUK17	2.4-kb B, Hp	529	78	–	+
pESM4	1.4-kb B, H	233	26	–	+
pCD1-1	2.1-kb cDNA[b]	632	–	–	+

[a] Possible additional amino acids (aa) provided by pBluescript KS – sequences.

[b] Fragments with only 19 and 5 (*b1-1*) and 21 (*d1-1*) bp in front of the start codon.

Restriction sites are as given in Figure 1. +, presence; –, absence.

Table 2. Effect of 3′ Truncations on the Ability of the *HD2* Genes *a2-1* and *b2-1* to Promote Clamp Cell Development and Repress Asexual Sporulation in an *A5* Host

Designation	Fragment	HD2 Protein (aa)	Extra (aa)[a]	Clamp Cells	Oidia
a2-1					
pLAC1	9.0-kb	520	–	+	–
pUK2	2.1-kb S, S	471	31	+	–
pUK7	2.9-kb E, H	392	21	+	–
pUK11	1.7-kb S, M	344	10	+	–
pUK10	2.0-kb C, B	268	16	–	+
pCA2-1	1.8-kb cDNA[b]	520	–	–	+
b2-1					
pUK4	3.4-kb S, S	520	–	+	–
pESM1	3.2-kb H, S	520	–	+	–
pESM7	1.3-kb P, P[b]	374	82	–	+
pUK12	1.7-kb S, I, I	354	87	+	–
pUK13	1.7-kb S, X	353	80	+	–
pUK14	1.6-kb S, I	320	9	+	–
pUK15	1.3-kb S, A	211	25	–	+

[a] Possible additional amino acids (aa) provided by pBluescript KS– sequences.

[b] Fragments with only 13 (*a2-1*) and 8 (*b2-1*) bp in front of the start codon.

Restriction sites are as given in Figure 1. +, presence; –, absence.

mating type gene is introduced that is not present in the host *A* locus, it is unlikely that this can integrate homologously because of the lack of DNA sequence similarity between different allelic forms of the genes. Transformation has the effect of introducing a second *A* mating type into the cell and a compatible mating response results; in the case of *b1-1*, we have shown that here is repression of asexual sporulation and promotion of clamp cell development (Tymon et al., 1992). We now confirm that all four *A42* genes, present on the plasmids pAMT1 (*b1-1*), pESM2 (*d1-1*), pLAC1 (*a2-1*), and pUK4 (*b2-1*), can each individually promote this phenotypic change (Tables 1 and 2).

We have identified nonessential C-terminal regions of the A proteins by constructing a series of pBluescript KS– clones of each gene truncated at available restriction sites (Tables 1 and 2). We monitored transformants for the development of clamp cells and asexual sporulation, but in no case could we separate the regulation of these two functions; loss of ability to promote clamp cell development simultaneously restored asexual sporulation. The analysis suggests that much of the 3′ end of *HD2* genes is dispensable (e.g., *a2-1*, pUK11; *b2-1*, pUK14, Table 2), but regions close to the 3′ end of the *HD1* genes are not (e.g., *b1-1*, pRAO3; *d1-1*, pUK17, Table 1). The minimal HD1 protein sequence found to be functional contained 594 amino acids of b1-1 and was missing just the conserved C terminus. The minimal functional HD2 protein sequences contained 320 amino acids of b2-1 and 344 amino acids of a2-1. These truncations deleted most of the serine-, threonine-, and proline-rich regions but left all the predicted helical regions (Figures 2 and 4).

The truncated proteins will have gained amino acids at their C termini because of the fusion of the coding sequences to pBluescript KS– (Tables 1 and 2). These extra amino acids only prevented the expression of the mated phenotype in one case, pUK7, which contains a truncated *a2-1* gene. This plasmid had to be linearized at the HindIII fusion point to give an active gene product. Linearization of other nonfunctional constructs had no effect.

5′ Ends of the Genes Determine Specificity

Proteins encoded by genes within the same *A* locus cannot promote sexual development; this requires different allelic forms of the proteins to be brought together by mating or transformation (Kües et al., 1992; Kües and Casselton, 1993). Clearly, there are compatible and incompatible protein associations, and there must be domains within both classes of A proteins that are responsible for discriminating between these. Because HD2 proteins truncated at the C terminus can still discriminate, it seems likely that specificity is imposed by the

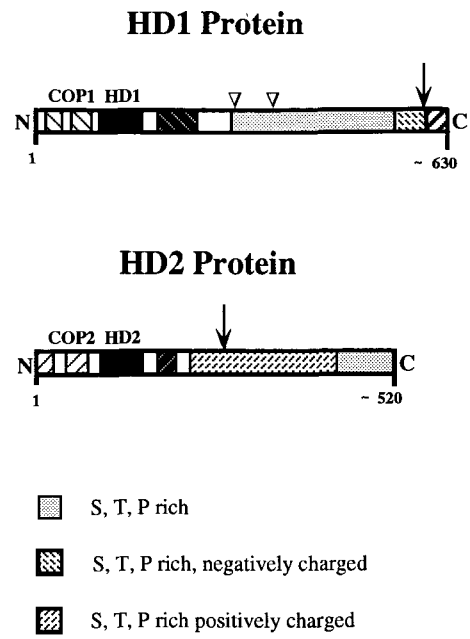


Figure 4. Predicted Organization of HD1 and HD2 Proteins.

Homeodomains are shown as black boxes (HD1 and HD2) and other helical regions by striped boxes. COP1 and COP2 indicate the regions implicated in specificity. Patterns in the figure refer to characteristics of putative transactivation domains. Triangles mark positions of putative bipartite nuclear localization sites. C-terminal regions of the proteins were truncated as far as the sites indicated by arrows without loss of function. N, N terminus; C, C terminus. The numbers indicated at left and right refer to amino acid residues.

Table 3. Transformation with Wild-Type and Chimeric *A42* Genes to Show That the 5′ Ends of the Genes Determine Specificity

			A42 Genes Introduced by Transformation					
	Genotype		*HD1* Genes			*HD2* Genes		
Host *A* Locus	*HD1* Genes	*HD2* Genes	*b1-1*	*d1-1*	*b1-1/d1-1*[a]	*a2-1*	*b2-1*	*b2-1/a2-1*[a]
A5[b]			+	+	+	+	+	+
A42	*b1-1 d1-1*	*a2-1 b2-1*	–	–	–	–	–	–
A6	*b1-3 d1-1*	*a2-1 b2-3*	+	–	+	–	+	+

[a] The chimeric gene *b1-1/d1-1* has the 5′ end of *b1-1* and the 3′ end of *d1-1*; the chimeric gene *b2-1/a2-1* has the 5′ end of *b2-1* and the 3′ end of *a2-1*.

[b] The genotype of the *A5* locus has not been determined. The genotype of the *A6* locus is described by Kües et al. (1994a).

+ indicates clamp cell development by the host; – indicates no clamp cell development.

N-terminal regions. Significantly, these regions have only low homology (16% identity, 41% similarity) when compared to the strongly conserved homeodomains (79% similarity) and subsequent sequences containing predicted helices (66% similarity). In the comparable *U. maydis* bE and bW proteins, the N-terminal regions are the only sequences that are variable between alleles (Kronstad and Leong, 1990; Schulz et al., 1990; Gillissen et al., 1992) and have actually been shown to determine allelic specificity of bE proteins (Yee and Kronstad, 1993).

We tested whether the N-terminal regions of the *C. cinereus* A proteins confer specificity by exchanging the 5′ ends of the *A42* genes using conserved ClaI or ScaI restriction sites within the HD1 and HD2 homeodomain coding regions, respectively (see Figure 1). Two chimeric genes were generated, an *HD2* gene with the 5′ end of *b2-1* and the 3′ end of *a2-1* (*b2-1/a2-1*) and an *HD1* gene with the 5′ end of *b1-1* and the 3′ end of *d1-1* (*b1-1/d1-1*). These chimeric genes were introduced into hosts with three different *A* loci, *A42*, *A6*, and *A5* (Table 3). *A42* and *A6* share *a2-1* and *d1-1* but have different alleles of *b1* and *b2*; thus, only the *b1-1* and *b2-1* genes can promote clamp cells in an *A6* host. *A5* and *A42* do not share any specificity genes, and all four *A42* genes promote clamp cells in an *A5* host (Kües et al., 1992) (Tables 1 and 2). Both chimeric genes promoted clamp cell development in the *A5* host, showing that they were functional but failed to do so in the *A42* host, showing that no new specificity had been generated. The critical test is their ability to promote clamp cell development in the *A6* host. The chimeric genes are behaving as *b* gene alleles; these are the genes from which they derived their 5′ ends and thus the N-terminal regions of their proteins.

DISCUSSION

Sexual development in several fungi has been shown to be regulated by two classes of homeodomain proteins, HD1 and HD2. The HD2 homeodomain has a typical length of 60 amino acids and the WF.N.R motif in the recognition helix, whereas the HD1 homeodomain is classed as atypical with extra amino acids between its helices and different amino acids in the recognition helix (Bürglin, 1994; Casselton and Kües, 1994). The best characterized of these homeodomain proteins are the a1 and α2 proteins of *S. cerevisiae* that are encoded by the alternative alleles of the mating-type locus, MATa and MATα. Following mating between a and α cells, these proteins dimerize to generate a transcription factor complex that binds unique operator sites upstream of haploid cell-specific genes (for review, see Dolan and Fields, 1991). A similar interaction between HD1 and HD2 homeodomain proteins has been postulated to regulate sexual development in basidiomycetes (Gillissen et al., 1992; Specht et al., 1992; Kües et al., 1994a, 1994b). The situation is, however, more complex because the basidiomycetes have evolved multiple mating types, and there are large numbers of functionally equivalent proteins encoded by multiallelic genes. Every cell contains genes for both classes of proteins, but it is only when different allelic forms of the genes are brought together by mating or by transformation that sexual development can be promoted. Compared with the simple a1–α2 interaction in *S. cerevisiae*, in the basidiomycetes there is a need for discrimination between incompatible and compatible proteins (Kües and Casselton, 1993).

In *U. maydis*, the multiallelic *HD1* and *HD2* genes are designated *bE* and *bW*, respectively. The compatible combination that triggers development is a *bE* and a *bW* allele from different *b* loci (Gillissen et al., 1992). In the *Aα* locus of *S. commune* (Stankis et al., 1992; Specht et al., 1992) and in the equivalent *Aα* complex of *C. cinereus* (Kües et al., 1994b), there is a pair of divergently transcribed *HD1* and *HD2* genes (*Y* and *Z* in *S. commune* and *a1* and *a2* in *C. cinereus*), and development is similarly triggered by an *HD1* and *HD2* gene from different loci. If we look at the *A42* locus of *C. cinereus*, only the *HD2* gene (*a2-1*) of this pair is present; the *HD1* gene is missing (Figure 1). This is also true of the *A6* locus, which shares the same *α* complex as *A42* (Kües et al., 1992, 1994a). This is fortunate because it allows us to show that the 5′ ends of the genes play a critical role in determining the specificity of a compatible interaction. Genes of the *α* and *β* complexes are functionally independent and can be recombined into different *α*/*β* combinations without promoting sexual development (Day, 1960). For *a2-1* or any other allele of the *a2* gene to promote clamp

cell development, it must be introduced into a host with a compatible *a1 HD1* gene (Kües et al., 1994b). By replacing the 5′ end of *a2-1* with that of *b2-1*, we generated a chimeric gene that promoted clamp cell development in an *A6* host that has no *a1* gene. We concluded that the chimeric gene is recognized as a *b2* gene and not a new *a2* allele. Furthermore, the transformation into the *A42* host strain showed that no new *b2* allele was generated. We concluded that the N-terminal region discriminates between genes, but our experiments did not show whether this region also discriminates between alleles.

The same argument can be put forward to explain why the chimeric *b1-1/d1-1* gene has *b1-1* specificity and not a new *d1* allele specificity. In the *β* complex of *A42*, there is only one obvious pair of divergently transcribed *HD1* and *HD2* genes (*b1-1* and *b2-1*). The *d1-1* gene has been found associated with different alleles of the *b* genes in the *A42*, *A43*, and *A6* loci (Kües et al., 1992; 1994a, 1994b), and it seems likely that it is the *HD1* half of another independently acting *HD1–HD2* pair of genes (Kües and Casselton, 1993). Because the *A6* locus lacks a *d2* gene, the chimeric *b1-1/d1-1* gene has only the *b2* gene as a possible *HD2* partner and therefore has the specificity of a *b1* gene.

Our transformation experiments suggest that there is an interaction between an HD1 and an HD2 protein following a compatible *A* mating in *C. cinereus* (Kües et al., 1994b). This interaction could be analogous to dimerization between the a1 and α2 mating type proteins of *S. cerevisiae.* Two regions have been implicated in a1–α2 dimerization. One is in the C-terminal 89 and 82 amino acids, which includes the homeodomains of both proteins and a short-charged tail of 19 amino acids from α2 (Goutte and Johnson, 1993; Mak and Johnson, 1993). The second is in the N-terminal regions of both proteins and contains 3,4-hydrophobic heptad repeat motifs, which are thought to mediate dimerization by two leucine zipper-like coiled-coil motifs (Ho et al., 1994). Analogous features in HD1 and HD2 proteins suggest that these also have two regions that could be involved in dimerization. There is a predicted helical region just C-terminal to the homeodomain in the *C. cinereus* HD1 proteins; this region is highly conserved and has some sequence similarity to the charged tail of α2 (Figure 2), and there are corresponding helical regions in the relatively unconserved N-terminal regions of the proteins. These N-terminal regions, which have a similar length to those of a1 and α2 are particularly important in the *C. cinereus* A proteins because we have shown with our chimeric genes that they determine the specificity between compatible and incompatible gene products. To emphasize the importance of these regions, we have termed them COP1 and COP2 (Figures 2 and 4).

A major function of an N-terminal dimerization domain is likely to be discrimination between compatible and incompatible proteins, whereas a second C-terminal domain could simply stabilize the association. Alteration of amino acids in the N-terminal heptad repeats of α2 weakens or destroys the affinity for dimerization with a1 (Ho et al., 1994), showing that quite subtle specificity changes can be brought about by individual amino acids. This could be a key to discrimination between compatible and incompatible HD1 and HD2 proteins. The amino acid sequence in the N-terminal regions is highly variable, and the dimerization potential of variant forms of the proteins might be quite differently influenced by single amino acids. We have previously suggested that the N-terminal region of the b1-1 protein might have analogy to the POU_S domain of POU homeodomain transcription factors that has both DNA binding and dimerization functions (Tymon et al., 1992). Mutations in the POU_S-like dimerization domain of the mammalian liver-enriched factor (LFB1) interfere with DNA binding of the wild-type protein by sequestering it in heterodimers that are defective in DNA recognition (Nicosia et al., 1992). Variations in the N-terminal regions of HD1 and HD2 proteins might have analogous effects and prevent DNA binding of incompatible pairs.

The HD1 and HD2 proteins of *C. cinereus* are longer than the a1 and α2 proteins of *S. cerevisiae* (125 and 210 amino acids, respectively, Astell et al., 1981; Miller, 1984) and may have other functional domains. Our gene truncation experiments indicate that most of the C-terminal region of the HD1 protein is essential, and based on its amino acid composition, we suggest that this region contains a potential transactivation domain. The only transformation assay available to us, detection of clamp cells, is not sensitive enough to distinguish between different levels of efficiency, and it is likely that sequences that appear to be dispensable, such as the C-terminal region of HD2 proteins, normally contribute to optimal protein activity.

METHODS

Coprinus cinereus Strains

Tryptophan auxotrophs used as host strains for transformation include the following: LN118, *A42B42 ade-2 trp-1.1,1.6*; LT2, *A6B6 trp1-1.1,1.6*; FA2222, *A5B5 trp-1.1,1.6.* Plasmids containing mating type genes were cotransformed with plasmid pCc1001 (Binninger et al., 1987) containing the *C. cinereus trp-1* gene. Fifty trp^+ transformants from each experiment were examined microscopically for presence of clamp cells, and surface scrapes were taken to look for oidia. Media and general methods of culture were those described by Lewis (1961) and Mutasa et al. (1990). Transformation procedures were based on Binninger et al. (1987) and Casselton and de la Fuente Herce (1989).

DNA Manipulations

Routine cloning and plasmid amplification was in *Escherichia coli* XL-1 Blue (*recA1 lac⁻ endA1 gyrA96 thi hsdR17 supE44 relA1* [F′ *proAB* $lacI^q$ *lacZΔM15* Tn*10*]) (Stratagene) or DH5α (F⁻, *endA1 hsdR17* r_k^- m_k^+ *supE44 thi⁻ recA1 gyrA96 relA1* Φ80Δ[*lacZM15*]) (Bethesda Research Laboratories) using standard DNA techniques (Sambrook et al., 1989). Site-directed mutagenesis at the first putative ATG start codon of *b1-1* was performed according to the LP-USE strategy (Ray and Nickoloff,

1992) using the Clontech (Palo Alto, CA) transformer site-directed mutagenesis kit with *E. coli* BMH71-18 (*mutS thi, supE Δ[lac-proAB] [mutS::Tn10] F' proAB, lacI^q ZΔM15*), and *E. coli* HB101 (*sup44 hsdS20 [rB⁻mB⁻] recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1*) (Sambrook et al., 1989) for the secondary transformation, with GTCGACTAGGTTTGCTG as the mutagenic primer and GGAGCTCCACCGCGGTGGCGGCCGATCGAGAACTAGTGGATCCCCCG as the selection primer. Underlining indicates base changes in the *b1-1* and pBluescript KS− sequences. The base change introduced into pAMT1 (to give pRAO5) was confirmed by sequencing.

Plasmid Constructs

pLAC1 was isolated from a genomic library generated in YRp12 (Yashar and Pukkila, 1985) and contains the *A6* α-complex, which has the same *α-fg* and *a2-1* alleles as the *A42* α complex (Kües et al., 1992). All other genomic fragments used for fungal transformations (see Table 1) originated from *A42* and were subcloned in pBluescript KS− (Stratagene). Full-length cDNAs were isolated from a λgt10 cDNA library (Kües et al., 1992) and were subcloned into pBluescript KS− using either the BamHI or EcoRI sites in the λgt10 adapter to give pCA2-1 (*a2-1* cDNA), pCB1-1 (*b1-1* cDNA), pCB2-1 (*b2-1* cDNA), and pCD1-1 (*d1-1* cDNA). pUA61 contains a 4.9-kb BamHI fragment with a full-length *d1-1* gene copy from the *A6* factor. This sequence was obtained from pLAC3, a YRp12 clone having a 7.7-kb insert containing the *d1-1* and *β-fg* genes. The *A6 d1-1* gene is distinguished from the *A42 d1-1* gene by a T residue in place of a C (at position 1245 in sequence CCD11 [X79688, EMBL data bank]) in the homeodomain-encoding region (Figure 1), which generates a recognition site for the restriction enzyme ClaI. This site was used to fuse the 5′ end of *b1-1* to the 3′ end of *d1-1* as follows: a ScaI-ClaI fragment of pUA61 with the 5′ end of *b1-1*, the f1 filamentous phage *ori*, and the 5′ end of the ampicillin resistance gene was fused to a ScaI-ClaI fragment of pAMT3 containing the ColE1 *ori* and the 3′ end of the ampicillin resistance gene. This generated plasmid pB1D1, which has a restored ampicillin resistance gene function. Similarly, the 5′ end of *a2-1* from pUK2 was fused to the 3′ end of *b2-1* from pUK4 using conserved ScaI sites within the sequence encoding the HD2 motif (Figure 1) and the ampicillin resistance gene of pBluescript KS− to give plasmid pB2A2.

Sequencing of Genomic and cDNA Clones

Sequencing was performed mainly with double-stranded DNA from pBluescript KS− or pUC13 subclones, ^{35}S-labeled ATP, the T7 polymerase kit of Pharmacia, and either the M13 universal or reverse primers. Where no suitable restriction sites were available for generating subclones, custom-made oligonucleotides were used as primers. DNA was sequenced in both directions. Sequence analysis was performed using the GCG-Package, version 7 (Genetics Computer Group, Madison, University of Wisconsin); data base searches were conducted using the BLAST program at the National Center for Biotechnology Information (NCBI) computer (Altshul et al., 1990; Gish and States, 1993). Helical regions were predicted according to Chou and Fasman (1978).

ACKNOWLEDGMENTS

We thank Alison Banham for critical reading of the manuscript. This work was supported by Science and Engineering Research Council (SERC) Grants No. GR/E90175 and No. GR/H45421 and Agricultural and Food Research Council (AFRC) Grant No. PG43/548 to L.A.C.; SERC studentships to A.M.T. and B.G.; a Gatsby Charitable Foundation grant to B.G.; an AFRC studentship to R.N.A.-O.; and a Brazilian government studentship to E.H.P. L.A.C. is an AFRC Postdoctoral Fellow.

Received June 8, 1994; accepted August 16, 1994.

REFERENCES

Altshul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215,** 403–410.

Astell, C.R., Ahlstrom-Jonasson, L., Smith, M., Tatchell, K., Nasmyth, K.A., and Hall. B.J. (1981). The sequence of the DNAs coding for the mating-type loci of *Saccharomyces cerevisiae.* Cell **27,** 15–23.

Bakkeren, G., and Kronstad, J.W. (1993). Conservation of the *b* mating-type gene complex among bipolar and tetrapolar smut fungi. Plant Cell **5,** 123–136.

Binninger, D.M., Skrzynia, C., Pukkila, P.J., and Casselton, L.A. (1987). DNA mediated transformation of the basidiomycete *Coprinus cinereus.* EMBO J. **6,** 835–840.

Bürglin, T.R. (1994). A comprehensive classification of homeobox genes. In Guidebook to the Homeobox Genes, D. Deboule, ed (Oxford, U.K.: Oxford University Press), pp. 27–71.

Casselton, L.A., and de la Fuente Herce, A. (1989). Heterologous gene expression in the basidiomycete fungus *Coprinus cinereus.* Curr. Genet. **16,** 35–40.

Casselton, L.A., and Kües, U. (1994). The multiallelic mating type genes of the Hymenomycetes. In The Mycota, K. Esser and P.A. Lemke, eds, Vol. I: Growth, Differentiation and Sexuality, J.G.H. Wessels and F. Meinhardt, vol. eds (Berlin: Springer-Verlag), pp. 307–321.

Chou, P.Y., and Fasman, G.D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47,** 45–147.

Day, P.R. (1960). The structure of the *A* mating type locus in *Coprinus lagopus.* Genetics **45,** 641–650.

Dolan, J.W., and Fields, S. (1991). Cell-type specific transcription in yeast. Biochim. Biophys. Acta **1088,** 155–169.

Dranginis, A.M. (1990). Binding of yeast a1 and α2 as a heterodimer to the operator DNA of a haploid-specific gene. Nature **347,** 682–685.

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

Gish, W., and States, D.J. (1993). Identification of protein coding regions by database similarity search. Nature Genetics **3,** 266–272.

Goutte, C., and Johnson, A.D. (1993). Yeast a1 and α2 homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein. J. Mol. Biol. **233,** 359–371.

Ho, C.Y., Adamson, J.G., Hodges, R.S., and Smith, M. (1994). Heterodimerization of the yeast MATa1 and MATα2 proteins is mediated by two leucine zipper-like coiled-coil motifs. EMBO J. **13,** 1403–1413.

Kronstad, J.W., and Leong, S.A. (1990). The *b* mating type locus of *Ustilago maydis* contains variable and constant regions. Genes Dev. **4,** 1384–1395.

Kües, U., and Casselton, L.A. (1992a). Homeodomains and regulation of sexual development in basidiomycetes. Trends Genet. **8,** 154–156.

Kües, U., and Casselton, L.A. (1992b). Regulation of fungal development by mating type genes. In The Eukaryotic Genome—Organization and Regulation, Vol. 50, P.M.A. Broda, S.G. Oliver, and P.F.G. Sims, eds (Cambridge, U.K. : Cambridge University Press), pp. 185–210.

Kües, U., and Casselton, L.A. (1993). The origin of multiple mating types in mushrooms. J. Cell Sci. **104,** 227–230.

Kües, U., Richardson, W.V.J., Tymon, A.M., Mutasa, E.S., Göttgens, 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 homeodomain motifs, determines sexual development in the mushroom *Coprinus cinereus.* Genes Dev. **4,** 568–577.

Kües, U., Göttgens, B., Stratmann, R., Richardson, W.V.J., O'Shea, S.F., and Casselton, L.A. (1994a). A chimeric homeodomain protein causes self-compatibility and constitutive sexual development in the mushroom *Coprinus cinereus.* EMBO J. **13,** 4054–4059.

Kües, U., Tymon, A.M., Richardson, W.V.J., May, G., Gieser, P.T., and Casselton, L.A. (1994b). *A* factors of *Coprinus cinereus* have variable numbers of specificity genes encoding two classes of homeodomain proteins. Mol. Gen. Genet. **243,** in press.

Leuther, K.K., Salmeron, J.M., and Johnston, S.A. (1993). Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. Cell **72,** 575–585.

Lewis, D. (1961). Genetical analysis of methionine suppressors in *Coprinus.* Genet. Res. **2,** 141–155.

Mak, A., and Johnson, A.D. (1993). The carboxy-terminal tail of the homeodomain protein α2 is required for function with a second homeodomain protein. Genes Dev. **7,** 1862–1870.

Mellon, F.M., Little, P.F.R., and Casselton, L.A. (1987). Gene cloning and transformation in the basidiomycete fungus *Coprinus cinereus*: Isolation and expression of the isocitrate lyase gene (*acu-7*). Mol. Gen. Genet. **210,** 352–357.

Miller, A.M. (1984). The yeast *MATα1* gene contains two introns. EMBO J. **3,** 1061–1065.

Mutasa, E.S., Tymon, A.M., Göttgens, B., Mellon, F.M., Little, P.F.R., and Casselton, L.A. (1990). Molecular organization of an *A* mating type factor of the basidiomycete fungus *Coprinus cinereus.* Curr. Genet. **18,** 223–229.

Nicosia, A., Tafi, R., and Monaci, P. (1992). *trans*-Dominant inhibition of transcription activator LFB1. Nucl. Acids Res. **20,** 5321–5328.

Ray, F.A., and Nickoloff, J.A. (1992). Site-specific mutagenesis of almost any plasmid using a PCR-based version of unique site elimination. BioTechniques **13,** 342–346.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Shepherd, J.C.W., McGinnis, W., Carrasco, A.E., DeRobertis, E.M., and Gehring, W.J. (1984). Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. Nature **310,** 70–71.

Schulz, B., Banuett, F., Dahl, M., Schlesinger, R., Schäfer, W., Martin, T., Herskowitz, I., and Kahmann, R. (1990). The *b* alleles of *U. maydis*, whose combinations program pathogenic development code for polypeptides containing a homeodomain-related motif. Cell **60,** 295–306.

Specht, C.A., Stankis, M.M., Giasson, L., Novotny, C.P., and Ullrich, R.C. (1992). Functional analysis of the homeodomain-related proteins of the *Aα* locus of *Schizophyllum commune.* Proc. Natl. Acad. Sci. USA **89,** 7174–7178.

Stankis, M.M., Specht, C.A., Yang, H., Giasson, L., Ullrich, R.C., and Novotny, C.P. (1992). The *Aα* mating locus of *Schizophyllum commune* encodes two dissimilar multiallelic homeodomain proteins. Proc. Natl. Acad. Sci. USA **89,** 7169–7173.

Sutherland, J.A., Cook, A., Bannister, A.J., and Kouzarides, T. (1992). Conserved motifs in *Fos* and *Jun* define a new class of activation domain. Genes Dev. **6,** 1810–1819.

Swiezynski, K.M., and Day, P.R. (1961). Heterokaryon formation in *Coprinus lagopus.* Genet. Res. **1,** 114–128.

Tranche, F., and Yaniv, M. (1992). HNF1, a homeoprotein member of the hepatic transcription regulatory network. BioEssays **14,** 579–587.

Tymon, A.M., Kües, U., Richardson, W.V.J., and Casselton, L.A. (1992). A mushroom mating-type protein that regulates sexual and asexual development contains a POU-related domain. EMBO J. **11,** 1805–1816.

Wegner, M., Drolet, D.W., and Rosenfeld, M.G. (1993). POU-domain proteins: Structure and function of developmental regulators. Curr. Opin. Cell Biol. **5,** 488–498.

Vollbrecht, E., Veit, B., Sinha, N., and Hake, S. (1991). The developmental gene *knotted-1* is a member of a maize homeobox gene family. Nature **350,** 241–243.

Yashar, B.M., and Pukkila, P.J. (1985). Changes in polyadenylated RNA sequences associated with fruiting body morphogenesis in *Coprinus cinereus.* Trans. Br. Mycol. Soc. **84,** 215–226.

Yee, A.R., and Kronstad, J.W. (1993). Construction of chimeric alleles with altered specificity at the *b* incompatibility locus of *Ustilago maydis.* Proc. Natl. Acad. Sci. USA **90,** 664–668.