Multiple Versions of the *A* Mating Type Locus of *Coprinus cinereus* are Generated by Three Paralogous Pairs of Multiallelic Homeobox Genes

Eneida H. Pardo, Suzanne F. O'Shea and Lorna A. Casselton

Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, United Kingdom

Manuscript received December 20, 1995
Accepted for publication May 17, 1996

ABSTRACT

The *A* mating type locus of *Coprinus cinereus* determines mating compatibility by regulating essential steps in sexual development. Each *A* locus contains several genes separated into two functionally independent complexes termed *Aα* and *Aβ*, and the multiple alleles of these genes generate an estimated 160 *A* mating specificities. The genes encode two classes of homeodomain-containing proteins designated HD1 and HD2. In this report we describe two newly cloned loci, *A2* and *A5*, and compare them with *A42*, *A43* and *A6* that we have described previously. An *Aβ*-null locus, retaining just a single active *HD1* gene from the *α*-complex, was generated by mutation. Using this as a transformation host, gene combinations that promote *A*-regulated development were identified. We demonstrate that each *A* locus contains members of three paralogous pairs of *HD1* and *HD2* genes. Different allelic versions of gene pairs are compatible but paralogous genes are incompatible. The genes present in four uncloned *A* loci were deduced using Southern analyses and transformations with available cloned genes. The combined analysis of nine *A* factors identifies sufficient *A* gene alleles to generate at least 72 *A* mating specificities.

THE mushroom fungi are remarkable in having large numbers of mating types. In *Coprinus cinereus* these are determined by multiallelic genes at two unlinked loci known historically as the *A* and *B* mating type factors (see RAPER 1966). Based on the number of different specificities found in a random sample of 33 isolates (DAY 1963), RAPER (1966) estimated that there are some 160 versions of the *A* locus and 79 versions of *B*. A compatible mating requires that alleles of genes at both loci are different. The mating type genes determine compatibility of mating cells by regulating initial steps in sexual development. The primary mycelium, the monokaryon, has uninucleate cells and is sexually sterile. Somatic cell fusion between monokaryons with different mating types triggers a developmental pathway that leads to the formation of a fertile mycelium, the dikaryon, that has binucleate cells and characteristic structures known as clamp connections.

A clamp connection is formed each time the apical cell divides. One of the daughter nuclei must pass through it into the subapical cell, a mechanism that ensures that each cell contains a nucleus from each mate. Different *A* genes are required to promote the development of the clamp cell and different *B* genes are required to promote its fusion to the subterminal cell (SWIEZYNSKI and DAY 1960). If only the *A* genes are different, the clamp cell develops but remains unfused and its nucleus is trapped. Introduction of a cloned *A* gene into a host with a different *A* mating specificity elicits the development of unfused clamp cells, providing an assay for testing compatible *A* gene interactions (MUTASA *et al.* 1990).

Classical recombination studies identified two tightly linked genes, *Aα* and *Aβ*, as the determinants of *A* mating specificity (DAY 1960, 1963). Our molecular analysis of three cloned *A* loci, *A42* (KÜES *et al.* 1992), *A43* (KÜES *et al.* 1994c) and *A6* (KÜES *et al.* 1994b), revealed several genes separated by 7 kb of homologous DNA sequence into two subcomplexes that are termed *α* and *β* since they are likely to correspond to the two genes identified by recombination. Two classes of proteins are encoded by the *A* genes, each contains a homeodomain motif, but these are distinctly different and have been designated HD1 and HD2. Similar proteins are encoded by the *Aα* mating type genes of another hymenomycete, *Schizophyllum commune*, and also by genes of the *b* mating type locus of the hemibasidiomycete *Ustilago maydis*. In the *Aα* locus of *S. commune* and in *U. maydis* there are just two genes (*Y* and *Z*, and *bE* and *bW*, respectively; GILLISSEN *et al.* 1992; STANKIS *et al.* 1992) arranged as a divergently transcribed pair. We have suggested that a multiallelic *HD1-HD2* gene pair is the basic functional unit of the *C. cinereus A* locus, and that large numbers of *A* specificities are generated by having several functionally independent (paralogous) gene pairs whose different alleles can be found in different associations. The variation in gene number at any one *A* locus is caused by having just solo representatives of some of these gene pairs (see Figure 1; KÜES and CASSELTON 1993).

C. cinereus, *S. commune* and *U. maydis* have multiple alleles of the *A* or *b* genes, and mating brings together

Corresponding author: Lorna A. Casselton, Department of Plant Sciences, University of Oxford, South Parks Rd., Oxford OX1 3RB, United Kingdom. E-mail: casselton@molbiol.ox.ac.uk

different allelic versions of a particular gene pair. By generating *b*-null or partially deleted *b* loci, GILLISSEN *et al.* (1992) showed that the compatible gene combination that triggers development is a *bE* (*HD1*) allele from one locus and a *bW* (*HD2*) allele from another. A similar interaction between different allelic versions of the *HD1* and *HD2 A* genes of *C. cinereus* and *S. commune* was deduced from transformation studies making use of natural gene deletions (SPECHT *et al.* 1992; KÜES *et al.* 1994c).

In this report we describe the organization of two new *A* loci. We describe how we made a host strain with an *Aβ*-null locus so that we could confirm our predictions as to which gene combinations are compatible by transformation experiments. We then show by a combination of transformation and Southern blot analyses how it is possible to identify specific alleles and predict the number of genes in uncloned loci.

MATERIALS AND METHODS

Strains and culture of *C. cinereus*: The following strains were used as transformation hosts and contained the *trp-1.1;1.6* mutation: LT2, *A6B6;* FA2222, *A5B6;* 218, *A3B1;* LCO2, *A2B6;* LCO7, *A40B40;* LCO5, *A41B41;* LN118, *A42B42;* NL1, *A44B44.* The *A43* host was AT8, *A43 B43 trp-3 ade-8.* NAB2, *ΔAB43m trp-1.1;1.6 pab-1* was generated by crossing the *A43* deletion mutant (*ΔA B43m pab-1*) described in the text to strain 218. The transformation procedure was based on CASSELTON and DE LA FUENTE HERCE (1989). Genes were cotransformed using either pCc1001 containing the *trp-1* gene (BINNINGER *et al.* 1987) or pDB3 containing the *trp-3* (BURROWS 1991) selectable marker. General techniques for using *C. cinereus* were described by LEWIS (1961). Mutagenesis of the *A43mB43m* self-compatible mutant (SWAMY *et al.* 1984) was effected by spreading asexual oidia onto complete medium and UV irradiating with 1 mJ to give 90% kill (Stratelinker, Stratagene). The media were as described in MUTASA *et al.* (1990).

Isolation and analysis of nucleic acids: DNA and RNA were isolated according to the methods described by KÜES *et al.* (1992). A genomic library using DNA from the *A2B1* wild-type strain 68 was constructed in λGEM11 (Promega). DNA was partially cut with *Mbo*I and ligated into the *Bam*HI site of the vector. The *A2* mating type locus was isolated on three overlapping clones by screening with pESM2 containing the *d1-1* and *β-fg* genes (KÜES *et al.* 1992), with p43P5 containing the *α-fg,* and with p43E2 containing part of the homologous region between the *α* and *β* complexes (KÜES *et al.* 1994c). A cosmid genomic library was constructed in the *C. cinereus* vector pLLC5200 (PUKKILA and CASSELTON 1991) using DNA from the *A5* strain BM5 partially cut with *Mbo*I and *Sau*3A and ligated into the *Bgl*II site of the vector. The *A5* locus was isolated on a single cosmid clone by screening with pHH5 and pLAC2 containing the *α-fg* and *β-fg,* respectively (KÜES *et al.* 1992). Individual genes were identified by Northern blot analysis and subcloned in pBluescript II KS- (Stratagene). Northern blots and Southern blots were carried out according to KÜES *et al.* (1992).

Plasmids for transformation: Plasmids containing individual *A42* genes were described by KÜES *et al.* (1992): pUK2, *a2-1;* pAMT2, *b1-1;* pESM1, *b2-1;* pUK5, *c1-1;* pESM2, *d1-1.* Plasmids with *A43* genes are described by KÜES *et al* (1994c): pE12P5-3 *a1-2;* p43H4, *a2-2;* p43HB, *b1-2;* p43Bg3, *b2-2;* pE12P2-6, *d1-1.* Plasmids with *A6* genes are described by KÜES *et al.* (1994b): pLAC1, *a2-1;* pA626, *b1-3;* pA625, *b2-3;* pUA61, *d1-1.* The *A2* and *A5* loci were isolated in this study. Individual genes from *A2* were subcloned into pBluescript II KS- to give pA241, *a1-2* on a 3.5-kb *Bam*HI-*Pst*I fragment; pA242, *a2-2* on a 4.0-kb *Sac*II-*Pst*I fragment; pA211, *b1-5* on a 5.0-kb *Hin*dIII-*Xho*I fragment; pA215, *b2-5* on a 4.5-kb *Pst*I-*Hin*dIII fragment; pA25, *d1-1* on a 4.0-kb *Xho*I fragment. pSO1 is a cosmid clone containing the *A5* locus. Subclones of pSO1 used in transformations were as follows: pCR8, a 4.3-kb *Sal*I-*Pst*I fragment with *a1-3* gene; pCR7, a *Bam*HI-*Pst*I fragment with *a2-3* gene; pA5E30, a 4.0-kb *Pst*I-*Eco*RI with *b1-4* gene; pSO12, a 5.0-kb *Hin*dIII-*Bam*HI with *b2-4* gene; pA5E18, a 3.6-kb *Bgl*II-*Hin*dIII containing *c2-2* gene; and pA5E12, a 3.0-kb *Sal*I fragment with *d2-2* gene.

Probes for Southern analyses: *A42: a2-1,* 2.9-kb *Eco*RI-*Hin*dIII; *b2-1, Hpa*I-*Sal*I; *c1-1,* 1.5-kb *Hin*dIII-*Bam*HI; *d1-1,* 0.8-kb *Hin*dIII. *A6: 2-1,* 2.9-kb *Eco*RI-*Hin*dIII; *b1-3,* 0.4-kb *Xho*I-*Eco*RI; *b2-3,* 0.8-kb *Xho*I-*Cla*I. *A43: a1-2,* 1.2-kb *Hin*dIII; *a2-2,* 0.7-kb *Eco*RI-*Hin*dIII; *b1-2,* 1.2-kb *Pst*I-*Xho*I; *b2-2,* 1.4-kb *Bam*HI-*Hin*dIII and *d1-1,* 0.8-kb *Hin*dIII. *A2: a1-2,* 1.0-kb *Hin*dIII; *a2-2,* 1.0-kb *Hin*dIII-*Sal*I; *b1-5,* 1.5-kb *Pst*I-*Bam*HI; *b2-5,* 1.0-kb *Sal*I. *A5: a1-3,* 2.1-kb *Bam*HI-*Eco*RI; *a2-3,* 0.8-kb *Bgl*II-*Eco*RI; *b1-4* and *b2-4,* 3.0-kb *Eco*RI; *c2-2,* 2.4-kb *Bgl*II-*Sal*I and *d2-2,* 3.0-kb *Hin*dIII.

RESULTS

Organization of five wild-type *A* loci: We have previously described the organization of the *A42, A43* and *A6* loci of *C. cinereus* (KÜES *et al.* 1992, 1994b,c). In this study we cloned and mapped genes in two additional loci, *A2* and *A5.* The achetypal *A* locus contains four paralogous pairs of *HD1-HD2* genes designated *a, b, c* and *d* with the *HD1* and *HD2* partners distinguished by the numbers 1 and 2, respectively (FIGURE 1; KÜES and CASSELTON 1993). Transcript analysis (not shown) identified five genes in *A2* and six genes in *A5.* Transcript size (*HD1,* 2.5 kb and *HD2,* 2.1 kb) and relative position within the complex of each gene (determined by reference to two conserved flanking genes and the homologous noncoding sequence that separates the *α* and *β* complexes, KÜES *et al.* 1992) were used to predict the class of each gene and the gene pair to which it belonged. The genes in each locus are aligned below the archetype according to these predictions. Based on this, *A2* and *A5* appeared to contain a complete *a* and *b* gene pair. In addition, *A2* contains a solo *HD1* gene and *A5,* two solo *HD2* genes. The *c* and *d* gene pairs of the *β* complex were predicted on the assumption that *HD2* partners for the solo *HD1 c1-1* and *d1-1* genes in *A42* would be found in other loci (KÜES and CASSELTON, 1993) and the two solo *HD2* genes in *A5* seemed likely candidates for these.

Southern blot analyses identify shared alleles: The DNA sequence dissimilarity between different alleles has made Southern blot analysis a reliable technique for identifying identical alleles in different *A* loci (KÜES *et al.* 1992). This analysis (Figure 2) predicts that *A2* shares the same *α* complex (*a1-2* and *a2-2*) as *A43* and contains the solo *HD1 d1-1* gene present in the *β* com-

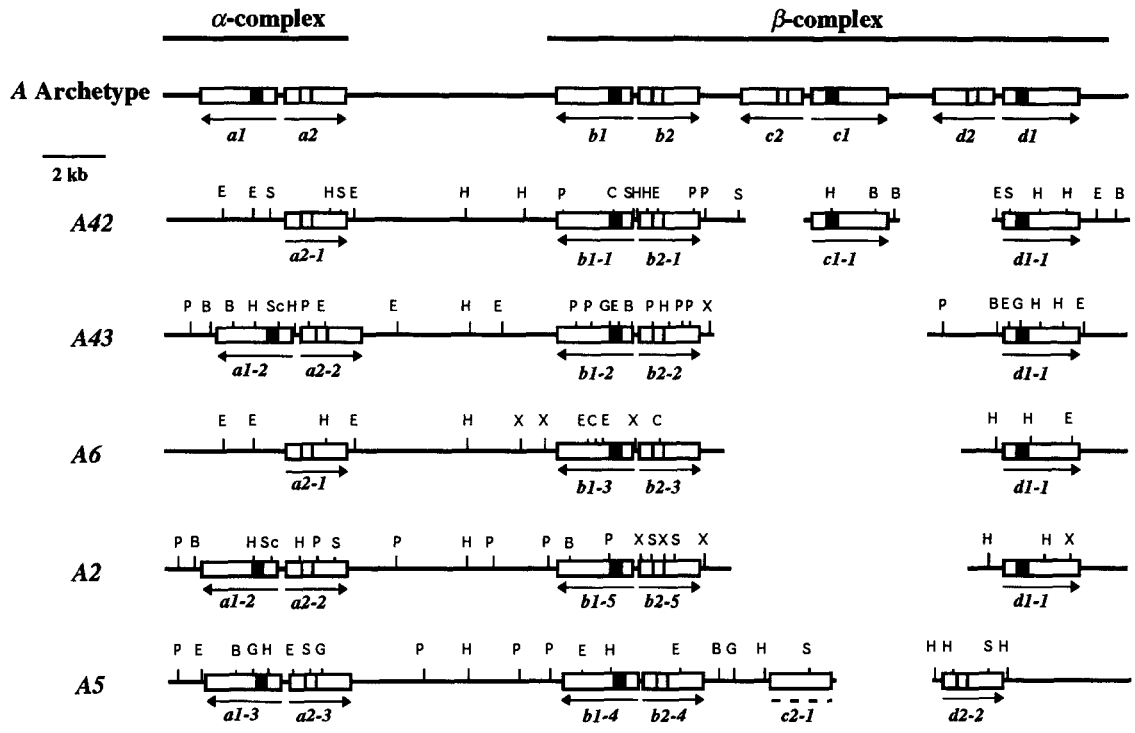

FIGURE 1.—Organization of the archetypal *A* mating type locus predicted by KÜES and CASSELTON (1993) with four pairs of divergently transcribed *HD1-HD2* genes (*a, b, c* and *d*) and maps of five molecularly characterized loci. *HD1* and *HD2* genes are distinguished by the black (*HD1*) and white (*HD2*) boxes marking the homeodomain-encoding regions, and by the numbers 1 or 2 in front of the allele number.

plexes of *A42, A6* and *A43*. The alleles of the *b* gene pair in *A2* and *A5* failed to hybridize to corresponding genes in other loci or to each other, indicating that they are dissimilar. Other genes within the *A5* locus also failed to hybridize to genes in any of the other four loci indicating that their alleles were unique to *A5*.

Transformation analysis: Individual genes from *A5* and *A2* were introduced into host strains having each of the cloned loci to test for compatible gene interactions that would promote clamp cell development (Table 1). Where hybridization predicted that the *A* locus of the host already contained the gene being transformed, no clamp cells developed, as expected. Transformation data confirmed that *A2* and *A43* share the same *a* gene pair. The *a2-2* gene in the *A43* locus is inactive (KÜES *et al* 1994c), and it is interesting to note that *A2* contains a similar inactive allele that is unable to promote clamp cells in the *A5* host that has a potentially compatible *HD1* partner. The solo *HD2* gene that occupies a position corresponding to a predicted *c* gene pair in *A5* failed to promote clamp cells in any background. The other solo *HD2* gene in *A5* found a compatible partner in all other *A* loci and from its position was predicted to be a compatible *HD2* member of the *d* gene pair and is designated *d2-2* in Figure 1.

Generation of an Aβ-null strain: The *A* archetype model assumes that the *a, b, c* and *d* gene pairs are functionally independent. To confirm this, we wanted to generate a host strain with an *A*-null locus into which we could introduce each gene separately and identify precisely its compatible partner. Since we could not predict other consequences of losing the entire complement of *A* genes, we generated an *A* locus lacking only the β gene complex (*Aβ*-null).

We exploited a mutation in the *A43* locus that gives a self-compatible *A* phenotype (SWAMY *et al.* 1984). By comparing the physical maps of the wild-type and mutant *A43* loci, MAY *et al.* (1991) showed that the mutation was the result of a major deletion. Based on the known positions of the genes within *A43* (KÜES *et al.* 1994) and on the transcript analyses shown in Figure 3 we now show that the deletion fused the normally inactive *a2-2 HD2* gene in the α complex to the 3′ end of the *d1-1 HD1* gene in the β complex (Figure 3). The separate 2.1- and 2.5-kb transcripts corresponding to the *a2-2* and *d1-1* genes, respectively, are replaced in the mutant by a 3.6-kb transcript. The residual *HD1* gene, *a1-2*, is detected by a 2.5-kb transcript (Figure 3). The resulting gene fusion is predicted to be analogous to the *A6* mutant fusion gene we have already described (KÜES *et al.* 1994b) that encodes a chimeric HD2:HD1 protein that constitutively promotes *A*-regulated clamp cell development.

We selected a mutation that inactivated the fusion gene thereby generating an *A* locus with a single functional *HD1* gene (*a1-2*). We took the *A43m* strain of SWAMY *et al.* (1984) (*pab-1A43mB43m*) that has self-compatible mutations in both the *A* and the *B* loci and

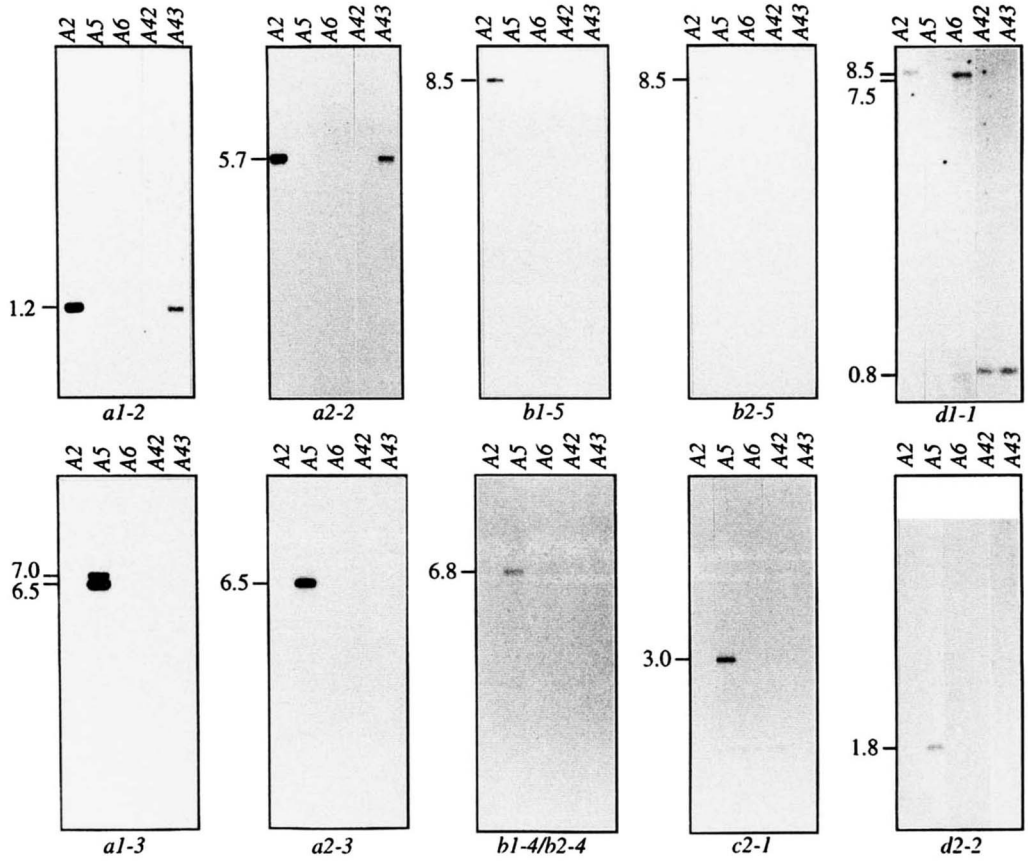

FIGURE 2.—Southern blot analyses identify alleles shared by *A2* or *A5*, and other cloned *A* loci. Genomic DNAs were digested with *Hin*dIII and probed with fragments containing the following: *a1-2* (1.0-kb *Hin*dIII), *a2-2* (1.0-kb *Hin*dIII-*Sal*I), *b1-5* (1.5-kb *Pst*I-*Bam*HI), *b2-5*(1.0-kb *Sal*I), *d1-1* (0.8-kb *Hin*dIII), *a1-3* (2.1-kb *Bam*HI-*Eco*RI), *a2-3* (0.8-kb *Bgl*II-*Eco*RI) *b1-4/b2-4* (3.0-kb *Eco*RI), *c2-1* (3.6-kb *Bgl*II-*Hin*dIII) and *d2-2* (2.4-kb *Bgl*II-*Sal*I).

mimics in appearance a true dikaryon with complete clamp connections. We identified among 12,000 survivors of UV mutagenesis a single colony that no longer produced clamp cells. This mutant had lost the self-compatible *A* phenotype as shown by its inability to form a dikaryon with an *A43* test monokaryon, but retained *A* mating function when mated to monokaryons with an *A* locus having a gene compatible with a residual *a1-2* gene, *i.e.*, *A6, A42* or *A5* (see Figure 1) but not *A2*, which has the same *a1-2* gene. To rule out the possibility of a suppressor mutation masking the self-compatible phenotype, the mutant (*A43m pab-1*, MA220) was crossed to an *A6 pab*$^+$ strain and 140 progeny examined for clamp cell development. The equal segregation of the mutant *A* and wild-type *A6* locus was monitored by looking at the segregation of the *pab-1* gene that maps 0.5 units from *A* (66 *pab-1:* 74 *pab-1*$^+$). None of the progeny developed clamp cells, showing that inability to do so segregated with the A locus.

TABLE 1

Functional analysis of genes from the *A2* and *A5* loci by testing for compatible gene interactions in host strains having molecularly characterized *A* loci

		Host strains[a]				
Genes	Origin	*A2*	*A5*	*A6*	*A42*	*A43*
a1-2	*A2*	~	+	+	+	−
a2-2	*A2*	~	−	−	−	−
b1-5	*A2*	~	+	+	+	+
b2-5	*A2*	~	+	+	+	+
d1-1	*A2*	~	+	−	−	−
a1-3	*A5*	−	~	+	+	−
a2-3	*A5*	+	~	−	−	+
b1-4	*A5*	+	~	+	+	+
b2-4	*A5*	+	~	+	+	+
c2-1	*A5*	−	~	−	−	−
d2-2	*A5*	+	~	+	+	+

[a] Host strains were as follos: *A2* (LCO2), *A5* (FA2222), *A6* (LT2), *A42* (LN118) and *A43* (AT8). +, presence and −, absence of clamp cells; ~, self host where clamps would not be expected.

Identification of compatible gene combinations: Representative alleles of genes present in *A2, A5, A6, A42* and *A43* were introduced individually into the β-null host strain. Two *HD2* genes, *a2-1* and *a2-3* (from the α-complexes of *A6/A42* and *A5*) promoted clamp cell development showing that each is a compatible partner for the residual *HD1* gene, *a1-2*. Introduction of *a1-3* had no effect since this is a second *HD1* gene. Introduction of any gene from the β-complex, irrespective of

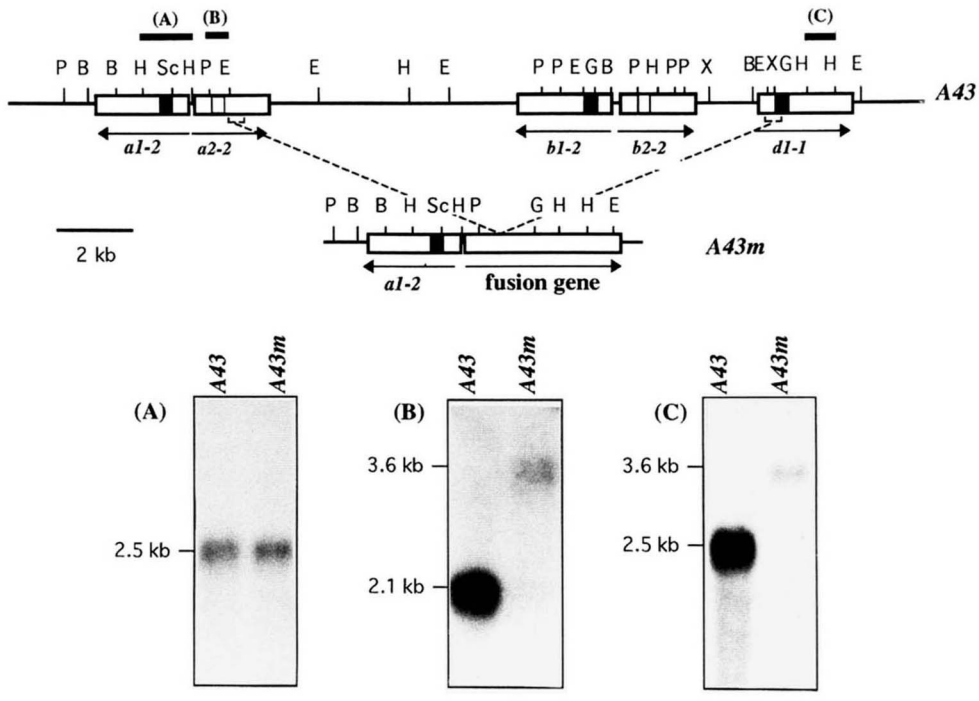

FIGURE 3.—Map of the *A43* locus and derivation of the mutant *A43* locus containing an *a2-2:d1-1* fusion gene that promotes constitutive *A*-regulated development. The approximate fusion point is based on data of MAY *et al.* (1991). Northern blots were probed with three fragments, indicated above the map, containing sequences from the *a1-2* (A), *a2-2* (B) and *d1-1* (C) genes.

which allele, failed to promote clamp cells confirming the conclusions of classical recombination studies that the α and β genes are functionally independent.

Transformants containing genes derived from the β complex were crossed together in all possible combinations to look for a compatible *A* gene interaction. Because the β-null host strain contained the *B43m* mutation, this overcame the need for mating monokaryons to have compatible *B* genes, and development of a dikaryon between two transformants indicated that they contained compatible *A* genes. In all cases the gene combinations that promoted dikaryon formation were an *HD1* gene from one locus and an *HD2* gene from another. All *HD1* × *HD1* and *HD2* × *HD2* gene combinations were incompatible. Specifically, compatible gene combinations were between different alleles of *b1* and *b2* and between *d1-1* and *d2-2* as shown by the data presented in Table 2. This analysis confirmed that the *b* and *d* genes we have identified constitute members of functionally independent gene pairs within the β-complex.

Analysis of four uncloned *A* loci: Four other *A* loci are present in strains in our collection, *A3, A40, A41* and *A44*. Using cloned genes to probe genomic DNA, we identified genes that were likely to be shared between the cloned and uncloned loci. We then introduced all cloned genes into these hosts to test for clamp cell development. The results of these two types of analysis are summarized in Table 3 and lead us to propose the organization illustrated in Figure 4.

The data indicate that *A40, A6* and *A42* share the same α complex as do *A44* and *A5*. None of the cloned *a* genes hybridize to *A3* and *A41* genomic DNA, indicating that these loci contain new alleles. Both members of the *a* gene pair are present in *A3* because a compatible interaction was promoted with either an *HD1* or an *HD2 a* gene, but only *HD1* genes promoted clamp cells in *A41*, indicating that an active *HD1* member of the pair is absent. The only loci seen to share alleles of the *b* gene pair are *A6* and *A3*. Transformation data confirmed this and also demonstrated that both members of the gene pair were present in all four loci. *d1-1* was not present in any other locus, but *d2-2* cross-hybridized to genomic DNA from strains having *A3*,

TABLE 2

Identification of two sets of paralogous genes in the *Aβ*-complex using an *Aβ*-null host to express individual genes and test their function by mating

	HD2 genes					
HD1 genes	*b2-1*	*b2-2*	*b2-3*	*b2-4*	*b2-5*	*d2-2*
b1-1	−	+	+	+	+	−
b1-2	+	−	+	+	+	−
b1-3	+	+	−	+	+	−
b1-4	+	+	+	−	+	−
b1-5	+	+	+	+	−	−
d1-1	−	−	−	−	−	+

The *Aβ*-null host was NAB2 (*ΔA B43mpab-1trp-1.1; 1.6*). +, a successful mating between transformants having the genes indicated as recognized by formation of a dikaryon; −, failure to mate.

TABLE 3

Identification of three paralogous pairs of genes in four uncloned *A* loci using a combination of genomic DNA hybridization and transformation with characterized genes

		Host strains[a]							
		A3		*A40*		*A41*		*A4*	
Genes	Origin	H	C	H	C	H	C	H	C
a2-1	*A42*	−	+	+	−	−	−	−	+
a1-2	*A43*	−	+	−	+	−	+	−	+
a2-2	*A43*	−	−	−	−	−	−	−	−
a1-3	*A5*	−	+	−	+	−	+	+	−
a2-3	*A5*	−	+	−	−	−	−	+	−
b1-1	*A42*	−	+	−	+	−	+	−	+
b2-1	*A42*	−	+	−	+	−	+	−	+
b1-2	*A43*	−	+	−	+	−	+	−	+
b2-2	*A43*	−	+	−	+	−	+	−	+
b1-3	*A6*	+	−	−	+	−	+	−	+
b2-3	*A6*	+	−	−	+	−	+	−	+
b1-4	*A5*	−	+	−	+	−	+	−	+
b2-4	*A5*	−	+	−	+	−	+	−	+
b1-5	*A2*	−	+	−	+	−	+	−	+
b2-5	*A2*	−	+	−	+	−	+	−	+
c1-1	*A42*	−	−	−	−	−	−	−	−
c2-1	*A5*	−	−	−	−	−	−	−	−
d1-1	*A42*	−	+	−	+	−	+	−	+
d2-2	*A5*	+	−	−	−	+	−	+	+
Genes present		*a1, a2, b1, b2, d2*		*a2, b1, b2, d2*		*a2, b1, b2, d2*		*a1, a2, b1, b2, d1, d2*	
New alleles		*a1, a2*		*b1, b2, d2*		*a2*		*b1, b2, d1, d2*	

[a] *A3* (218), *A40* (LCO7), *A41* (LCO5) and *A44* (NL1). H, Southern analysis; C, clamp cell development in transformed host; +, DNA hybridization or presence of clamp cells; −, absence of either DNA hybridization or clamp cells.

A41, and *A44*. Based on all our other analyses, we expected all three loci to contain *d2-2*, however, transformation data showed this to be true only for *A3* and *A41* where *d2-2* fails to elicit clamp cells. *A44* contains genes that are compatible partners for both *d1-1* and *d2-2* and is the first locus we have identified with a complete *d* gene pair. The *HD2* gene cannot be *d2-2* but had sufficient DNA homology to cross-hybridize. *A40* lacks either of the identified *d* genes, as shown by Southern analysis, but the compatible interaction with *d1-1*, but not *d2-2*, indicates that it contains only a new allele of *d2*. *c1* and *c2* failed to elicit clamp cells in any host.

For the most part, there is a good correlation between cross-hybridization to a hosts' genomic DNA and failure to elicit clamp cell development. Where a gene fails to both elicit clamp cells or cross-hybridize, this may be due to the absence of a compatible partner, as occurs for the *a2-3* gene in the *A40* or *A41* hosts (Table 3). While there is a general lack of sequence similarity between the different allelic and paralogous versions of the genes, cross-hybridization can be misleading as shown by our data for the *d2-2* gene.

DISCUSSION

It is estimated that there are 160 versions of the *A* locus (RAPER 1966). Our analysis of nine *A* mating type loci of *C. cinereus* allowed us to demonstrate that these

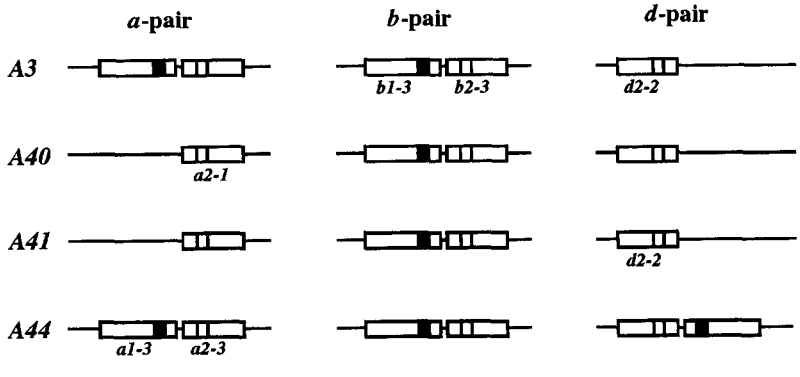

FIGURE 4.—Predicted *a*, *b* and *d* genes present in four uncloned *A* loci. *HD1* and *HD2* genes are distinguished by the black and white boxes, respectively.

large numbers could be accounted for by a small number of multiallelic forms of three paralogous pairs of genes. With the small sample of *A* loci that we have analyzed, we have identified at least four alleles of the *a* genes, six alleles of the *b* genes and three alleles of the *d* genes.

It is essential that the two members of a gene pair, when present, are always kept together since recombination would generate a compatible combination capable of promoting *A*-regulated development. However, if paralogous gene pairs can be recombined in all possible combinations, we have identified sufficient alleles to generate at least 72 different *A* loci.

MAY and MATZKE (1995) have addressed the problem of *A* locus variation and used Southern blotting to identify three variable regions within the *A* locus that correspond to those containing the *a*, *b* and *d* gene pairs and some homologous flanking sequence. These authors analyzed 16 strains all with different *A* loci and taken from locations around the world. They confirmed that the members of a particular gene pair are always found together but that the three blocks of sequence, representing paralogous pairs, were in linkage equilibrium indicating that all three could be recombined despite their physical proximity. Our molecular analysis has shown that members of a gene pair are separated by nonhomologous sequence that should prevent recombination between them (KÜES *et al.* 1992, 1994 and present study) but that the *a* and *b* gene pairs are separated by 7.0 kb of homologous sequence that should permit recombination. The *b* and *d* gene pairs in the loci we have examined are not separated by homologous sequence, and it is unlikely that the genes in these loci can be recombined.

The active gene combination that triggers development is an *HD1* gene from one mate and an *HD2* gene from the other. This we have previously deduced from transformation studies (KÜES *et al.* 1994c), but here we provide direct evidence by generating a transformation host containing an *Aβ*-null locus and into which we could introduce single genes to test-compatible combinations. Where both members of a gene pair are present, and this is always true for the *b* pair, two equally effective *HD1-HD2* combinations result on mating. For the *a* and *d* pairs many loci have just a solo *HD1* or *HD2* gene, but this is sufficient to generate a single compatible interaction if the appropriate partner is found in another locus.

The existence of two classes of mating type proteins with dissimilar homeodomains is not unique to the basidiomycete fungi. The a1 and α2 proteins encoded by the alternative forms of the mating type locus of the budding yeast *Saccharomyces cerevisiae* (MAT**a** and MATα) have homeodomains that are analogous to the HD2 and HD1 domains, respectively. a1 and α2 proteins heterodimerize on mating and generate a new transcription factor complex that binds unique operator sites that cannot be recognized in unmated cells. There is now evidence from studies in both *U. maydis* (KÄMPER *et al.* 1995) and *C. cinereus* (BANHAM *et al.* 1995) that development is triggered by a similar heterodimerization between bE and bW proteins and between HD1 and HD2 A proteins, respectively. However, unlike *S. cerevisiae,* the basidiomycete mating type loci contain genes that encode both classes of proteins and, moreover, there are multiple versions of these proteins. Proteins already present in the cell must be distinguished from many nonself proteins that are potentially compatible. In the cells of *C. cinereus,* homologous gene products must also be distinguished from paralogous gene products. In a mating between monokaryons having *A42* and *A5* loci, there are five HD1 proteins and five HD2 proteins resulting in four compatible protein combinations and 21 incompatible ones. In both *U. maydis* and *C. cinereus* discrimination has been shown to be effected by an N-terminal domain in both proteins classes that permits compatible but not incompatible proteins to heterodimerize (BANHAM *et al.* 1995; KÄMPER *et al.* 1995).

By generating mutations in a *bE* gene, KÄMPER *et al.* (1995) were able to introduce amino acid changes that allowed the protein to dimerize with its normally incompatible partner encoded by the *bW* gene in the same *b* locus. The substitutions conferred changes in charge or hydrophobicity and led the authors to propose a model to account for the compatible and incompatible interactions of proteins encoded by at least 25 different versions of the *b* locus based on four polar or hydrophobic amino acid contacts.

Coiled coil motifs have been implicated in dimerization of several proteins including the a1 and α2 mating type proteins of *S. cerevisiae* (HO *et al.* 1994). Changes in charge are likely to affect the stability of dimerization via coiled coils where these affect the *e* and *g* positions, flanking the hydrophobic *a* and *d* positions of the 3,4 heptad repeat (O'SHEA *et al.* 1992). Significantly, GIESER and MAY (1994) and BANHAM *et al.* (1995) predicted the presence of two coiled coil motifs within the N-terminal domains of the HD1 A proteins of *C. cinereus* and suggested that these were likely to be involved in dimerization. Particularly interesting was the prediction that the positions of these motifs were exactly the same in the proteins encoded by three different alleles of *b1* (amino acid positions 15–33 and 73–99) but that they were displaced in the protein encoded by the paralogous *d1-1* gene (37–59 and 79–100) (BANHAM *et al.* 1995). We have now extended this analysis to the HD1 product of another paralogous gene, a1-3, where the predicted amino acid positions are 3–22 and 59–85. This analysis of a1-3 lends further support to our hypothesis that paralogous gene products are distinguished by the position of the dimerization interface.

The odd inactive *c* genes that we found in the *A42* and *A5* loci may represent the remains of a gene pair

that has largely been eliminated, or alternatively these genes may have been introduced by illegitimate recombination events. DNA sequence analysis of the *A5 c2-1* gene has failed to identify a homeodomain encoding region and this gene can have no functional significance. Upon subjecting the c1-1 protein sequence to the coil analysis, we concluded that this is likely to be a pseudogene rather than a member of a fourth gene pair. Only a single coil could be predicted at position 9-28 that more nearly corresponds to the first coil position in a1-3. *c1-1* may have been derived from an *a1* gene that has been accidently inserted into the *β* complex and inactivated by losing part of its dimerization domain.

The *a2-2* gene was inactive in all host strains used. Interestingly, the constitutive *A43* fusion gene that we inactivated to generate a partial null-*A* host contains the *a2-2* gene as its *HD2* component. The regulatory activity of the fused protein is clearly unimpaired, and it seems likely that the defect in a2-2 is in its N-terminal dimerization domain. We have shown for the analogous A6 fusion protein that this domain is only essential for recognition of a compatible dimerization partner and is dispensible if this function is no longer required (BANHAM *et al.* 1995).

We thank MARKUS AEBI for the mutagenesis screen that yielded the *Aβ*-null mutant, URSULA KÜES and CORNELIA RINDERLE for help with mapping the *A5* locus, WENDY RICHARDSON for the Northern analysis of the *A43* constitutive mutant, CRAWFORD KINGSNORTH for providing Northern blots for probing with *A5* genes and ALISON BANHAM for making the computer analysis described by LUPAS *et al.* (1991) to predict coiled coils within a1-3 and c1-1. We are grateful to VEREL SHULL for critical reading of the manuscript. This work was supported by Biotechnology and Biological Sciences Research Council grants 43/P00389 and 43/P01791 to L.A.C. and a Brazilian Government studentship (CNPq) to E.H.P. L.A.C. is a BBSRC Senior Research Fellow.

LITERATURE CITED

BANHAM, A. H., R. N. ASANTE-OWUSU, B. GÖTTGENS, S. A. J. THOMPSON, C. S. KINGSNORTH *et al.*, 1995 An N-terminal dimerization domain permits homeodomain proteins to choose compatible partners and initiate sexual development in the mushroom *Coprinus cinereus*. Plant Cell **7:** 773–783.

BINNINGER, D. M., C. SKRZYNIA, P. J. PUKKILA and L. A. CASSELTON, 1987 DNA-mediated transformation of the basidiomycete *Coprinus cinereus*. EMBO J. **6:** 835–840.

BURROWS, D. M. 1991 Transformation studies in the basidiomycete fungi *Coprinus cinereus* and *Coprinus bilanatus*. PhD. Thesis, University of London.

CASSELTON, L. A., and A. DE LA FUENTE HERCE, 1989 Heterologous gene expression in the basidiomycete fungus *Coprinus cinereus*. Curr. Genet. **16:** 35–40.

DAY, P. R., 1960 The structure of the *A* mating-type factor in *Coprinus lagopus*. Genetics **45:** 641–650.

DAY, P. R.. 1963 The structure of the *A* mating type factor in *Coprinus lagopus*. Wild alleles. Genet. Res. Camb. **4:** 323–325.

GIESER, P. T., and G. MAY, 1994 Comparison of two *b1* alleles from within the *A* mating-type of the basidiomycete *Coprinus cinereus*. Gene **146:** 167–176.

GILLISSEN, B., J. BERGEMANN, C. SANDMANN, B. SCHROER, M. BÖLKER *et al.* 1992 A two-component system for self-non-self recognition in *Ustilago maydis*. Cell **68:** 647-657.

HO, C.-Y., J. G. ADAMSON, R. S. HODGES and M. SMITH, 1994 Heterodimerization of the yeast MATa1 and MAT*α*2 proteins is mediated by two leucine zipper-like coiled-coil motifs. EMBO J. **13:** 1403–1413.

KÄMPER, J., M. REICHMANN, T. ROMEIS, M. BÖLKER and R. KAHMANN, 1995 Multiallelic recognition: nonself-dependent dimerization of the bE and bW homeodomain proteins in *Ustilago maydis*. Cell **81:** 73–83.

KÜES, U., and L. A. CASSELTON, 1993 The origin of multiple mating types in mushrooms. J. Cell Sci. **104:** 227–230.

KÜES, U., W. V. J. RICHARDSON, A. M. TYMON, E. S. MUTASA, B. GÖTTGENS *et al.*, 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. **6:** 568–577.

KÜES, U., R. N. ASANTE-OWUSU, E. S. MUTASA, A. M. TYMON, E. H. PARDO *et al.*, 1994a Two classes of homeodomain proteins specify the multiple *A* mating types of the mushroom *Coprinus cinereus*. Plant Cell **6:** 1467-1475.

KÜES, U., B. GÖTTGENS, R. STRATMANN, W. V. J. RICHARDSON, S. F. O'SHEA *et al.*, 1994b A chimeric homeodomain protein causes self-compatibility and constitutive sexual development in the mushroom *Coprinus cinereus*. EMBO J. **13:** 4054–4059.

KÜES, U., A. M. TYMON, W. V. J. RICHARDSON, G. MAY, P. T. GIESER *et al.*, 1994c *A* mating-type factors of *Coprinus cinereus* have variable numbers of specificity genes encoding two classes of homeodomain proteins. Mol. Gen. Genet. **245:** 45–52.

LEWIS, D., 1954 Comparative incompatibility in angiosperms and fungi. Adv. Genet. **6:** 235-285.

LEWIS, D., 1961 Genetical analysis of methionine suppressors in *Coprinus*. Genet. Res. Camb. **2:** 141–155.

LUPAS, A., M. VAN DYKE and J. STOCK 1991 Predicting coiled coils from protein sequences. Science **252:** 1162–1164.

MAY, G., and E. MATZKE, 1995 Recombination and variation at the *A* mating-type of *Coprinus cinereus*. Mol. Biol. Evol. **12:** 794–802.

MAY, G., L. LE CHEVANTON and P. J. PUKKILA, 1991 Molecular analysis of the *Coprinus cinereus* mating type *A* factor demonstrates an unexpectedly complex structure. Genetics **128:** 529–538.

MUTASA, E. S., A. M. TYMON, B. GÖTTGENS, F. M. MELLON, P. F. R. LITTLE *et al.*, 1990 Molecular organization of an *A* mating type factor of the basidiomycete fungus *Coprinus cinereus*. Curr. Genet. **18:** 223–229.

O'SHEA, E. K., R. RUTKOWSKI and P. S. KIM, 1992 Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell **68:** 699–708.

PUKKILA, P. J., and L. A. CASSELTON, 1991 Molecular genetics of the Agaric *Coprinus cinereus*, pp. 249–264 in *More Gene Manipulations in Fungi*, edited by J. W. BENNETT and L. L. LASURE. Academic Press, San Diego.

RAPER, J. R., 1966 *Genetics of Sexuality in Higher Fungi*. The Ronald Press, New York.

SPECHT, C. A., M. M. STANKIS, L. GIASSON, C. P. NOVOTNY and R. C. ULLRICH, 1992 Functional analysis of the *Aα* homeotic mating-type genes of *Schizophyllum commune*. Proc. Natl. Acad. Sci. USA **81:** 7174–7178.

STANKIS, M. M., C. A. SPECHT, H. YANG, L. GIASSON, R. C. ULLRICH *et al.*, 1992 The *Aα* mating type locus of *Schizophyllum commune* encodes two dissimiar multiallelic homeodomain proteins. Proc. Natl. Acad. Sci. USA **89:** 7169–7173.

SWAMY, S., I. UNO and T. ISHIKAWA, 1984 Morphogenetic effects of mutations at the *A* and *B* incompatibility factors in *Coprinus cinereus*. J. Gen. Microbiol. **130:** 3219–3224.

SWIEZYNSKI, K. M., and P. R. DAY, 1960 heterokaryon formation in *Coprinus lagopus*. Genet. Res. Camb. **1:** 129–139.

Communicating editor: P. J. PUKKILA