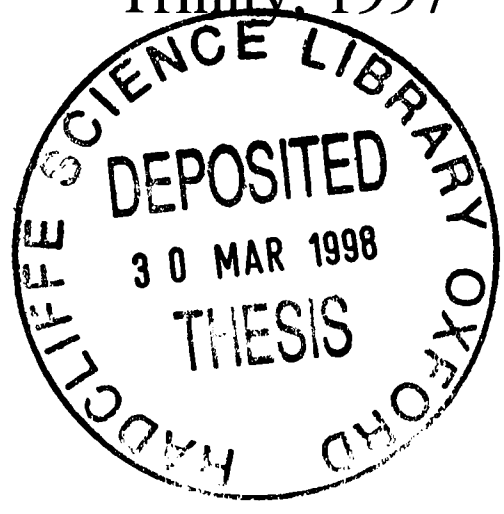
University of Oxford
Department of Plant Sciences
St. Catherine's College

Isolation and characterisation of the *B42* mating type locus of *Coprinus cinereus*

A thesis presented by
John Richard Halsall
for the degree of Doctor of Philosophy

Trinity, 1997



Isolation and characterisation of the *B42* mating type locus of *Coprinus cinereus*

John Richard Halsall
St. Catherine's College

D. Phil. Thesis, Trinity Term 1997

Abstract

The *B* mating type locus of the basidiomycete *Coprinus cinereus* contains a multiallelic gene complex encoding peptide pheromones and transmembrane receptors. There are an estimated 79 *B* mating specificities in *C. cinereus*, any two of which are sufficient to promote *B*-regulated development following cell fusion. The isolation of the *B42* locus is described along with the DNA sequence analysis that identified nine *B* mating type genes within a 27kb *B42*-specific DNA sequence. Six of the genes, with small transcripts of 800-900nt, encode the mating pheromone precursors and the other three, with 1.9 to 2-5kb transcripts, encode the transmembrane pheromone receptors. The genes are arranged in three groups, designated group 1, 2 and 3, each consisting of one receptor gene and two pheromone genes. *B42* and *B6* share the same alleles of the group 1 genes, but not those of groups 2 and 3. This was demonstrated by DNA-sequence analysis and Southern blot analysis. None of the group 1 genes from *B42* were able to activate *B*-regulated development in a *B6* host when introduced by transformation but with one exception, all genes from group 2 and group 3 were able to do so. This analysis led to the recognition that the three genes in any one group are held together in an allele-specific DNA sequence and that Southern blot analysis and transformation can be used to identify shared alleles in uncloned loci. Extensive Southern analyses using cloned genes to probe genomic DNAs from strains having other *B* mating specificites showed that different *B* loci may share identical alleles of two groups of genes. Mating partners thus require different alleles of only one group of genes to generate a compatible *B* mating interaction. Transformation analyses with the same cloned genes confirmed the conclusions derived from the hybridisation data. Multiple *B* mating specificities thus appear to be derived from three groups of multiallelic and functionally redundant genes. A tenth gene located within the *B42*-specific DNA sequence encodes a putative transporter protein belonging to the major facilitator superfamily (MFS). In other genomic backgrounds this gene lies in homologous flanking sequences and its presence within the *B42* locus is unlikely to be related to mating type function.

Acknowledgements

Firstly I would like thank my supervisor Lorna Casselton for her encouragement and enthusiasm throughout my research and for help and support when things were not running quite so smoothly. Thanks also to everyone who has spent time in the lab during the past three years. Tom "Perfect" Brutnell and Miltos for numerous technical discussions, mainly about maize. My second supervisor, Jane Langdale for advice and continuous reminders about how I should have finished already. Thanks to Crawford, Sigga, Suzanne, Eneida and Andreas for help during the 'early years'. More recently, to Gatsby, 'Anal', Nat, 'Rod', Lizzie, Pushpa, Tony and Laura for being understanding while I was slightly less sociable than normal. Last but not least to the fish for keeping me company in the early hours of the morning and not dying until I had finished.

Out of the lab, thanks to all the boys and girls at Catz for making my stay in Oxford very... interesting. For distracting me from my work continuously for eight weeks and then going away again, so I could work all summer. Especially to Topaz and 'Pelvin' for providing me with a floor to sleep on for the last week and to Kenny T. Michael Darren Moore for having a fantastic name.

Finally thanks to John Baker for keeping his cool while being bombarded with requests for 5"x4"s of these and 2"x2$^1/_2$"s of these but not those. Most importantly thanks to BBSRC for the money without which none of this would have been possible (and for the conference in California).

Table of contents

Abbreviations

ABC	ATP-binding cassette
A-on	*A*-mating type regulated development on
ATP	adenosine 5'-triphosphate
B-on	*B*-mating type regulated development on
bp	base pairs
cDNA	complementary deoxyribonucleic acid
CTAB	hexadecyltrimethylammonium bromide
CTRL	control
dCTP	2'-deoxycytosine 5'-triphosphate
DEPC	diethylpyrocarbonate
dNTP	2'-deoxynucleoside 5'-triphosphate
DNA	deoxyribonucleic acid
DNAase	deoxyribonuclease
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	ethylenediamine tetraacetic acid
EtOH	ethanol
h	hour
HD1	*Coprinus cinereus* homeodomain 1
HD2	*Coprinus cinereus* homeodomain 2
λDNA	phage lambda deoxyribonucleic acid
kb	kilobases
M	moles per litre
MFS	major facilitator superfamily
min	minute
mol	mole
MOPS	3-(N-morpholino)propanesulfonic acid
OD	optical density
ORF	open reading frame
pBS	bluescript plasmid
PCR	polymerase chain reaction
pmf	proton-motive force
poly $(A)^+$ RNA	polyadenylated ribonucleic acid
pre	pheromone response element

prf	pheromone response factor
RNA	ribonucleic acid
RNAase	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
s	second
SDS	sodium dodecyl sulphate
Taq	*Thermus aquaticus*
TE	a solution of 10 mM Tris-HCl/1 mM EDTA (pH 8.0)
Tris	Tris(hydroxymethyl)aminoethane
U	units
UTR	untranslated region
UV	ultraviolet
v/v	volume per volume
w/v	weight per volume
xg	relative centrifugal force
X-gal	5-bromo-4-chloro-3-indoyl β-D-galactopyranoside

Chapter 1

Introduction

In most sexually reproducing eukaryotic organisms, mechanisms exist to promote outbreeding and thereby increase the levels of genetic diversity in the population. In animal populations sexual dimorphism predominates and ensures that individuals cannot self fertilise. In plants and fungi, sexual dimorphism is rare. Whilst a flower displays male and female parts, other genetic barriers impose restrictions on fertilisation. Higher plants are generally hermaphrodite but self-incompatibility mechanisms prevent pollen from fertilising egg cells from the same individual. What is particularly interesting is that the best known self-incompatibility systems that have evolved in plants are controlled by a single locus designated S and the genes at this locus are multiallelic (Newbigin *et al.*, 1993; Nasrallah *et al.*, 1994; Golz *et al.*, 1995). The Poaceae have what appears to be the most complex control of pollen compatibility in that there are two unlinked multiallelic loci (Li *et al.*, 1997). From what is known about the molecular basis of self/non-self recognition in higher plants, it is evident that several different mechanisms have evolved and the general genetics of the systems are only superficially similar. In contrast in the fungi, there appears to be a strong conservation in the mechanism that permits compatible mating partners to be recognised, even between such widely diverse species as unicellular ascomycetes and the multicellular basidiomycetes which are the subject of this thesis. Like the higher plants, the higher basidiomycetes have multiallelic genes involved in mating recognition and offer the biologist an opportunity to study complex protein-protein interactions where discrimination is required between members of large families of near identical proteins.

1.1 Mating systems in fungi

In the fungi, the term mating type is used to distinguish individuals that are self-incompatible although morphologically identical. For this reason the genes that control mating compatibility are called the mating type genes. The study of mating type determination has largely been restricted to a few model species originally chosen for classical genetic studies. These include the ascomycetous yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, the filamentous ascomycetes, *Neurospora crassa* and *Podospora anserina* and the basidiomycetes, *Ustilago maydis*, *Coprinus cinereus* and *Schizophyllum commune*.

In ascomycete fungi there are just two mating types, determined by alternative alleles present at a single genetic locus. The two types are termed **a** and α in *S. cerevisiae*, P and M in *S. pombe*, *A* and *a* in *N. crassa* and *mat+* and *mat-* in *P. anserina*. In the basidiomycetes, mating type may be determined by a single genetic locus, in whch case species are said to be bipolar, or by two unlinked loci in which case species are said to be tetrapolar. In the homobasidiomycetes, the so called higher basidiomycetes that include the mushroom fungi, the two mating type loci are termed *A* and *B* whereas in the hemibasidiomycetes, the group to which *U. maydis* belongs, the loci are termed *a* and *b*.

The efficiency of a breeding system is generally measured in terms of how much it restricts inbreeding potential and how much it promotes outbreeding potential. In the ascomycetes, the two mating types segregate equally in the four products of meiosis and half the sibs can mate with the other half. For example, in a cross between **a** and α in *S. cerevisiae,* two of the resulting spores in each ascus have the **a** allele and two the α allele. Inbreeding potential (sib-compatibility) is restricted to 50% but cannot be

eliminated. However, with only two mating types in the population, outbreeding potential is also reduced to 50%. With multiple alleles of the mating type genes in basidiomycetes, inbreeding restriction can be maintained at 50% or reduced to 25% when there are two mating type loci, but outbreeding potential may be very high, depending on the numbers of alleles of the genes as shown in Table 1.1. It requires only 20 alleles of each gene to generate 90% or more outbreeding potential. It is surprising, therefore, that fungi like *C. cinereus* and *S. commune* have quite so many different alleles of the genes. It is estimated that *C. cinereus* has some 160 alleles of the *A* locus and 79 of the *B* locus and thus some 12,000 different mating types. In *S. commune* the numbers are even higher, with a predicted 288 *A* alleles and 81 *B* alleles giving more than 20,000 mating types (Raper, 1966).

There are much fewer mating types in *U. maydis* because only one of the mating type loci has multiallelic genes. The *a* locus has just two alleles whereas the *b* locus has some 25 (Puhalla, 1970). Recent cloning studies with a related bipolar species, *U. hordei* has provided the first insight into how bipolar and tetrapolar species differ (Bakkeren and Kronstad, 1994). *U. hordei* has both the *a* and *b* genes but these are tightly linked within a single complex locus. It seems likely that the many bipolar forms of the homobasidiomycetes, an estimated 35% of all species (Whitehouse, 1949; Quintanilha and Pinto-Lopes, 1950), have equally complex mating type loci.

1.2 The life cycle of *Coprinus cinereus*

This study has been concerned with the molecular organisation and function of genes at the multiallelic *B* mating type locus of *C. cinereus*. To understand the role that these mating type genes play in mating recognition it is first necessary to take a look at the

Table 1.1 Efficiency of bipolar and tetrapolar breeding systems in restricting inbreeding and promoting outbreeding (after Koltin, 1978).

System	Genes	Mating	Meiotic Segregation		Inbreeding restriction (%)	Outbreeding potential† Number of alleles A	B	%
Bipolar	*A1A2....An*	*A1* x *A2*	0.5 *A1*	0.5 *A2*	50	2		50
						5		80
						10		90
						20		95
Tetrapolar	*A1A2....An*	*A1B1* x *A2B2*	0.25 *A1B1*	0.25 *A2B1*	25	2	2	25
	B1B2....Bn		0.25 *A1B2*	0.25 *A2B2*		5	5	64
						10	10	81
						20	20	90

† Outbreeding potential calculated from: 1 gene (n*A*-1)/n*A*; 2 gene (n*A*n*B*-n*A*-n*B*+1)/n*A*n*B*

life cycle of *C. cinereus* (Figure 1.1) and the roles that both classes of mating type genes play in regulating sexual development.

C. cinereus belongs to the hymenomycetes, species that have typical mushroom fruit bodies with characteristic gills covered in the basidia that bear the sexual basidiospores. A single basidiospore germinates to form a sterile filamentous mycelium known as a monokaryon. The cells of the monokaryon are uninucleate, and the aerial mycelium may develop structures termed oidiophores which bud off uninucleate asexual spores called oidia. Somatic cell fusion is sufficient to initiate sexual development provided monokaryons have compatible mating types. Mating gives rise to a fertile mycelium known as the dikaryon that has binucleate cells, one nucleus from each mating partner, and characteristic structures known as clamp connections at each septum. The dikaryon is capable of indefinite growth but given the right light and temperature conditions differentiates into the fruiting bodies. The two dikaryotic nuclei remain discrete until the formation of the basidial cells, in which they fuse and almost immediately undergo meiosis. The four nuclei migrate into the developing spores that are carried on small projections on the surface of the basidium.

1.3 The roles of the *A* and *B* mating type genes in dikaryon formation

The steps between compatible cell fusion and the development of the dikaryon are important to understand in terms of mating recognition. Following cell fusion monokaryons may exchange nuclei. The donor nucleus in each case migrates through the cells of the recipient monokaryon until it reaches a tip cell where the two nuclei divide in synchrony and a complex division is intiated which leads to the development of the clamp connection. The *A* and *B* mating type genes play a critical role in regulating these events as detailed in Figure 1.2. Nuclear migration was recorded by

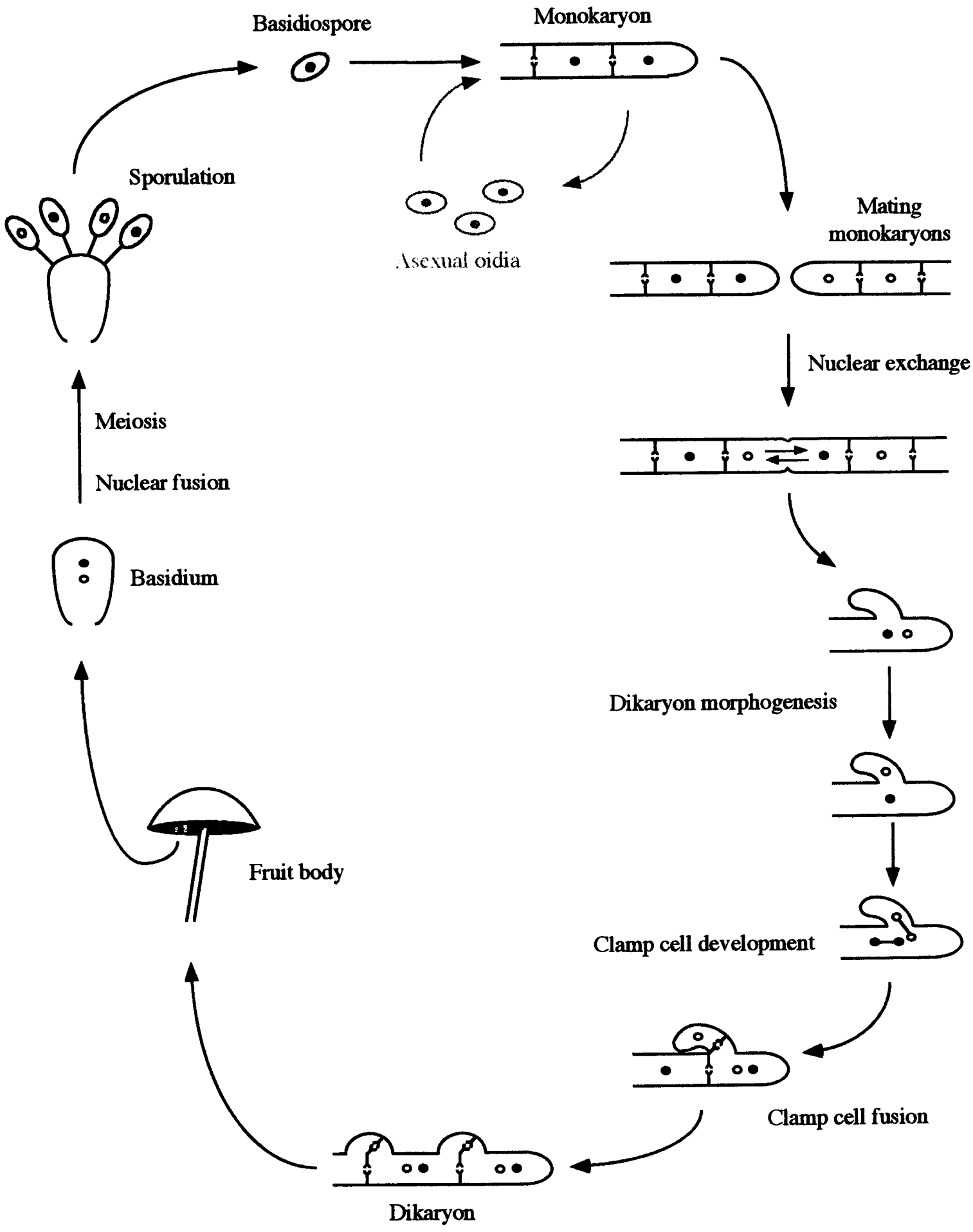


Figure 1.1 Life cycle of *Coprinus cinereus*.

(Buller, 1931) in *Coprinus radiatus*. Estimates of nuclear migration rates range from 2.71mm h^{-1} reported for *S. commune* (Snider and Raper, 1958) to 4cm h^{-1} for *C. congregatus* (Ross, 1976). Nuclear migration requires the disruption of the complex dolipore septa that separate the cells and through which nuclei would normally be unable to pass (Giesy and Day, 1965). Once the apical cells of the mycelium contain both genetically different nuclei the dikaryotic condition is established. The apical cell divides in such a way that the two dikaryotic nuclei always remain together and are distributed equally to each daughter cell. A clamp cell forms on the side of the apical cell, and the leading nucleus passes into this. Both nuclei divide in synchrony and new cell walls are laid down across the mitotic spindles. The apical cell contains both nuclei but one daughter nucleus is cut off in the clamp cell and the other in the newly formed subapical cell. The clamp cell grows backwards and fuses with the subapical cell, releases its nucleus and thereby generates a binucleate subapical cell joined to the apical cell by a clamp connection.

For monokaryons to be compatible and differentiate the dikaryon upon cell fusion, they must have different alleles of both the *A* and *B* mating type genes. Cell fusion is not, however, mating type dependent and fusions between monokarons with identical alleles of the *A* or the *B* genes results in the formation of heterokaryons that express only part of the dikaryon developmental sequence (Swiezynski and Day, 1960a; Swiezynski and Day, 1960b). If only the *B* alleles are different, nuclear migration occurs following cell fusion but nuclear division is unsynchronised and clamp cells do not develop. If only the *A* alleles are different, nuclear migration cannot occur but the two nuclei in the fusion cell undergo a synchronised nuclear division and development proceeds as far as the formation of the unfused clamp cell. This is unable to fuse with the subterminal cell to form the completed clamp connection. It was evident from this study that a compatible *A* gene interaction is required to initiate a synchronised division of the

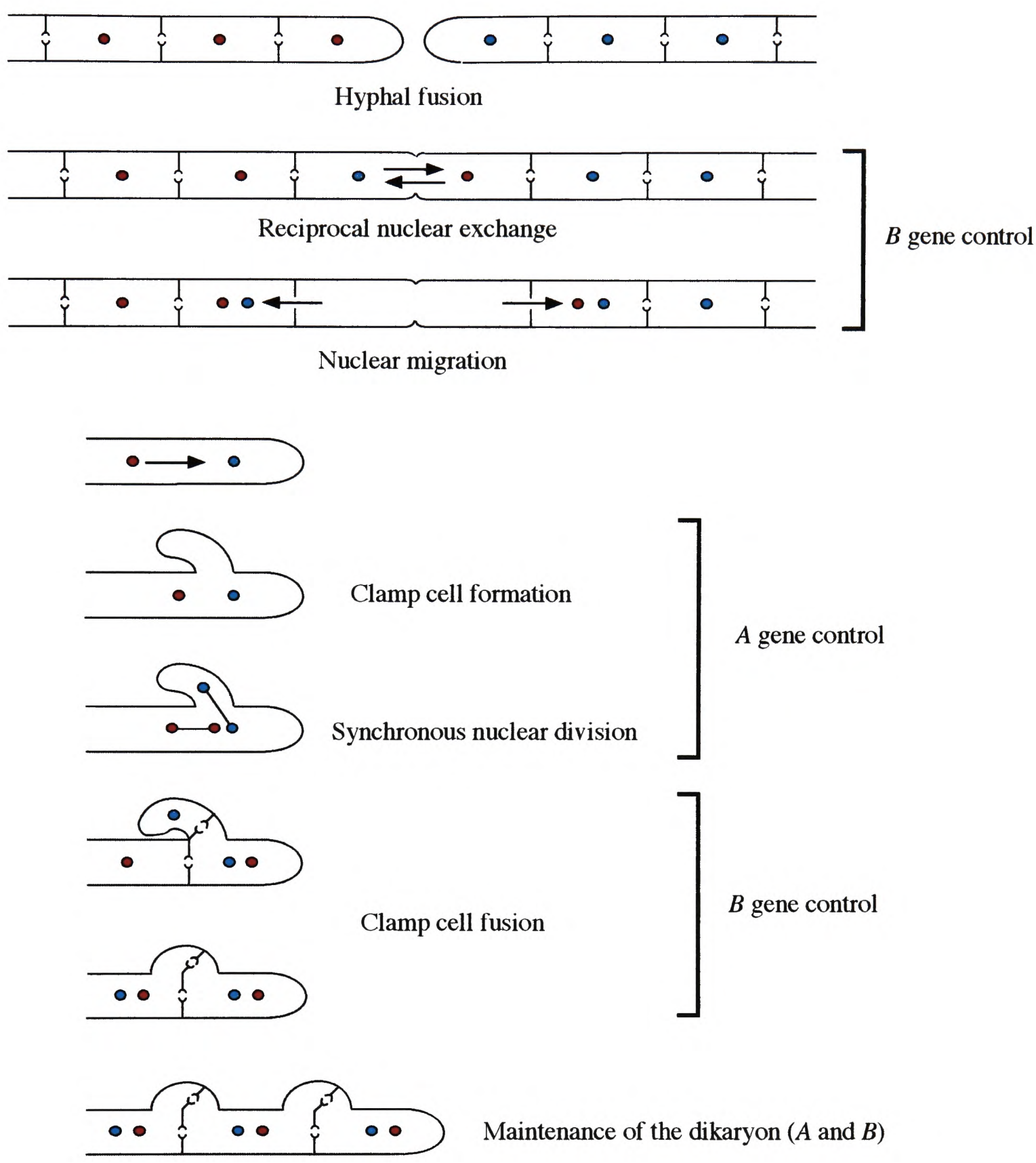


Figure 1.2 The roles of the mating type genes in regulating sexual development in *Coprinus cinereus*.

dikaryotic nuclei and the development of the unfused clamp cell. A compatible *B* gene interaction is required for nuclear migration and septal disruption following the initial cell fusion and completion of the clamp connection in the dikaryon.

1.4 The structure of the *A* and *B* mating type loci in hymenomycetes

Genetic studies in *S. commune* and *C. cinereus* revealed that the *A* and *B* loci were complex long before a molecular analysis was possible. Genetic recombination established that both the *A* and *B* genes in *S. commune* were separated into two closely linked loci which became designated $A\alpha$ and $A\beta$ and $B\alpha$ and $B\beta$ (Papazian, 1951; Raper *et al.*, 1958; Raper *et al.*, 1960). The α and β loci were shown to be multiallelic and functionally redundant. It required only the alleles of the α or the β locus to be different for a compatible *A* or *B* mating interaction. For example, $A\alpha1\beta1$, $A\alpha2\beta2$, $A\alpha1\beta2$ and $A\alpha2\beta1$ are all different *A* mating specificites. From world wide samples it was estimated that there were 9 $A\alpha$ alleles, 32 $A\beta$ alleles, 9 $B\alpha$ and 9$B\beta$ alleles (Raper *et al.*, 1960; Koltin *et al.*, 1972). The distance between the α and β loci varies between strains and ranges from 1-17% for the *A* genes (Raper, 1966) and 0-8% for the *B* genes (Koltin *et al.*, 1972).

Corresponding studies in *C. cinereus* revealed that the *A* genes were separated into two loci but these were only 0.07 map units apart (Day, 1960), too close for any general analysis of the numbers of α and β gene alleles that may exist in the population (Day, 1963). There is no evidence for more than a single *B* mating type locus in *C. cinereus* (Haylock *et al.*, 1980). It is interesting to note that recombination between the α and β genes will increase sib-compatibility. With 5% recombination between the *A* loci and the *B* loci sib-compatibility will increase from 25% to 30%. Close linkage will ensure

that sib-compatibility remains low but some level of recombination will serve to maximise the numbers of mating specificities in the population.

1.5 Molecular analysis of the *A* mating type locus of *C. cinereus* and *S. commune*

The complete *A* mating type complex of *C. cinereus* was cloned by (Mutasa *et al.*, 1990) from a cosmid library. A chromosome walk was initiated from the closely linked metabolic gene *pab-1* and this yielded cosmid clones containing the $A\alpha$ and $A\beta$ genes. When transformed into a host with a compatible *A* mating type, the cloned genes elicited the typical *A*-regulated stages in dikaryon development, the formation of unfused clamp cells. Molecular analysis of the *A* locus revealed an unexpected level of complexity. In all, five different *A* gene complexes have been isolated and compared (Kües and Casselton, 1992; Kües and Casselton, 1993; Kües *et al.*, 1994; Pardo *et al.*, 1996). The genetically defined $A\alpha$ locus is separated from the $A\beta$ locus by 7.0kb of DNA sequence, present in all *A* gene complexes, that appears not to contain any genes. The entire complex contains variable numbers of genes, ranging from four to six, depending on the actual *A* mating specificity and all of these are capable of eliciting the *A*-regulated developmental pathway in a host having a different allele of the gene.

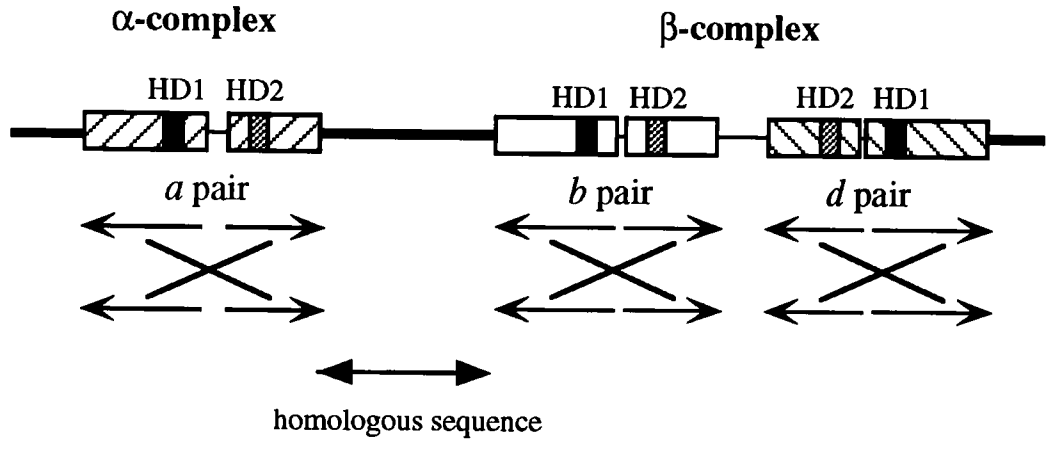
It has been established that genes of the *A* mating type locus encode two families of dissimilar proteins that both contain a characteristic homeodomain DNA binding motif (Scott *et al.*, 1989). These two proteins form the subunits of a heterodimeric protein complex that is assumed to bind DNA to regulate transcription of genes invoved in dikaryon development. In the monokaryon, several versions of these proteins may be present in the cell , but it is only by mating that monokaryons can bring together proteins that can heterodimerise. Compatible proteins have been shown to

heterodimerise (Banham *et al.*, 1995) but the DNA target sites of this complex have yet to be identified.

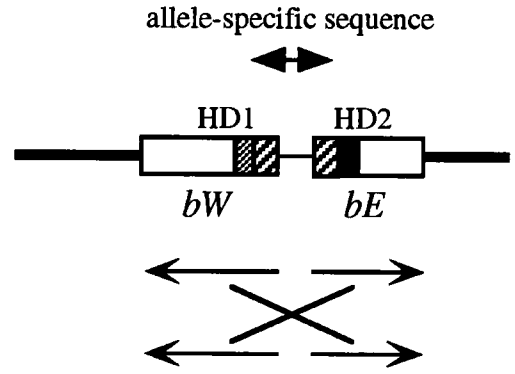
The organisation of the *A* locus of *C. cinereus* is illustrated in Figure 1.3. The figure shows a hypothetical *A* locus in which there are three pairs of divergently transcribed genes. Each pair of genes encodes the two classes of homeodomain proteins which are distinguished as HD1 and HD2 because of the conserved but distinctly different homeodomain sequences. These three sets of genes are multiallelic. Importantly, the DNA sequences of different alleles of the genes are sufficiently dissimilar to not permit recombination, and the gene-pairs are locked into a non-homologous sequence that ensures that the two members of the pair always stay together. The three pairs of genes, indicated as the *a*, *b* and *d* gene pairs are, moreover, functionally independent. For two *A* loci to have different specificities it requires only one pair of genes to have different alleles.

By stripping out most of the genes from a single *A* locus, Pardo *et al.* (1996) were able to demonstrate the compatible gene interaction that promotes *A*-regulated development. A single *HD1* and a single *HD2* gene were sufficient to promote *A*-regulated development provided these two genes came from different allelic versions of the same gene pair; i.e. *a* genes work with *a* genes, *b* genes with *b* genes, and *d* genes with *d* genes. By making use of the non-homology between different alleles, Pardo *et al.* (1996) were able to identify shared alleles of the genes in different uncloned *A* loci because alleles cross-hybridised. A gene that cross-hybridised, moreover, was never able to elicit clamp cell development following transformation, showing that it could not promote a compatible *A* gene interaction.

(A) *A* mating type locus of *Coprinus cinereus*



(B) *b* mating type locus of *Ustilago maydis*



(C) *MAT***a** and *MAT*α loci of *Saccharomyces cerevisiae*

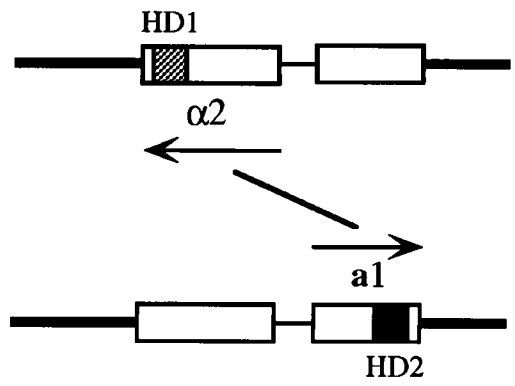


Figure 1.4 Comparison of the *Coprinus cinereus, Ustilago maydis* and *Saccharomyces cerevisiae* mating type loci which encode proteins with homeodomain motifs. HD1 and HD2 refer to the different homeodomain types which are indicated by the black and grey boxes within the genes. Different cross-hatching indicates different gene sub-families in *C. cinereus*. Arrows indicate the direction of gene transcription and diagonal lines the compatible gene combinations that promote sexual development. Thick lines indicate homologous DNA sequences.

Multiple *A* mating type specificities are derived from the many different allele combinations that it is possible to generate with these three pairs of genes. Just 5 alleles of each pair would generate 125 of the predicted 160 classical *A* alleles. Because of the high level of functional redundancy it has been found that *A* loci rarely contain all of the predicted six genes illustrated in Figure 1.3, but at least one gene from each of the pairs is always present.

The *Aα* locus of *S. commune* has been cloned and sequenced and shown to contain just a single pair of *HD1* and *HD2* genes (Specht *et al.*, 1992). These genes clearly act quite independently of the genes within the as yet, uncharacterised *Aβ* locus. Several of the nine alleleic versions of this *Aα* locus have been cloned (Specht *et al.*, 1994) and these are similar to the *A* genes of *C. cinereus* in that different alleles of the genes are very dissimilar in sequence.

Molecular cloning of the mating type genes of *U. maydis* has revealed that the multiallelic *b* locus is equivalent to the *A* loci of the homobasidiomycetes. The *b* locus is much simpler than the *A* loci in that there is just a single pair of *HD1* and *HD2* genes (designated *bE* and *bW*, respectively). As with the *A* genes, the *bE* and *bW* genes are divergently transcribed. Lack of sequence homology in the 5' regions of the genes and the intervening promoter sequence ensures that the two genes in any one version of the *b* locus remain together as a single allelic unit. A compatible *b* gene interaction occurs when mating partners bring together different alleles of the *b* genes (Gillissen *et al.*, 1992) and *b*-regulated development is promoted by heterodimerisation between an HD1 protein from one mate and an HD2 protein from the other (Kämper *et al.*, 1995).

For both the *A* proteins of *C. cinereus* and *S. commune* and the *b* proteins of *U. maydis* it has been shown that heterodimerisation between compatible HD1 and HD2 proteins

is mediated by a domain within the N-terminal sequences of the proteins (Banham *et al.*, 1995; Kämper *et al.*, 1995; Magae *et al.*, 1995) which is not required for DNA binding (Asante-Owusu *et al.*, 1996). This dimerisation domain is sufficient to prevent incompatible gene products activating development. Incompatible proteins encoded either by the same allelic pair of genes in *U. maydis*, *C. cinereus* and *S. commune* or by the functionally independent sets of genes present in the complex *A* locus of *C. cinereus* are unable to heterodimerise and cannot therefore generate the active transcription factor complex.

Mating type compatibility regulated by an interaction between two dissimilar homeodomain proteins is not unique to the basidiomycetes. Exactly the same type of interaction is found in the budding yeast *S. cerevisiae*. *S. cerevisiae* has just a single mating type locus with two alleles, **a** and α (Herskowitz, 1988). The **a** locus contains the **a**1 gene that encodes a protein containing a homeodomain related to the basidiomycete HD2 motif. The α locus contains the α2 gene, that has a homeodomain more related to the basdiomycete HD1 motif. Fusion between compatible cells results in heterodimerisation between the **a**1 and α2 proteins and the formation of a new transcription factor complex not present in unmated cells. As there is only a single version of each of the proteins, the *S. cerevisiae* proteins do not require a dimerisation domain that can discriminate between different versions of the proteins, but interestingly, an N-teminal dimerisation domain that plays no role in DNA-binding is present in these two proteins (Ho *et al.*, 1994).

The functions of the mating type genes in *S. cerevisiae* are very well understood. Each of the two mating type loci contains two divergently transcribed genes, **a**1 and **a**2 in the *MAT***a** locus and α1 and α2 in the *MAT*α locus. In terms of regulating mating recognition, only the functions of the **a**1, α1 and α2 proteins are known (Herskowitz

1988; Johnson, 1995) but these are sufficient together with the activities of several general transcription factors (*e.g.* Mcm1p, Tup1p and Ssn6p), to determine whether or not a cell is expressing **a**-cell specific haploid functions, α-cell specific haploid functions or **a**/α diploid cell-specific functions.

1.6 Signal transduction in *Sacharomyces cerevisiae*

Of relevance to the basidiomycete mating type systems is the fact that haploid cells of *S. cerevisiae* signal to each other via secreted peptide pheromones. Cells of opposite mating type produce mating type-specific **a** or α-pheromones which bind corresponding mating type specific receptors on the cell surface. Pheromone binding elicits a number of responses, which are collectively known as the pheromone response. These include arrest of the cell cycle at the G1 phase, cell elongation to produce shmoos that will fuse at their tips and the production of mating type-specific agglutinins that aid cell adhesion (Kurjan, 1993). The isolation of sterile (*ste*) mutants (MacKay and Manney, 1974; Hartwell, 1980) and the subsequent identification and characterisation of the various gene products has made the *S. cerevisiae* pheromone response one of the best understood signal transduction pathways in eukaryotic cells. The pivotal role this pathway plays in mating is also well understood.

The pheromone secreted by α cells is known as α-factor. Two genes, *MFα1* and *MFα2* encode polypeptide precursors of 165 and 120 amino acids, respectively (Caldwell *et al.*, 1995). Both polypeptide precursors contain tandem repeats of the α-factor peptide separated by proteolytic cleavage sites and a hydrophobic signal peptide leader sequence for targetting to the endoplasmic reticulum. The precursors are then processed and glycosylated before being exported from the cell by the conventional secretory pathway (Julius *et al.* 1984). The mature α-factor pheromone is 13 amino acids long.

The **a** cells secrete **a**-factor. The **a**-factor precursor is also encoded by two genes, *MFa1* and *MFa2*. However, the precursors do not contain tandem repeats of the pheromone or the signal peptide sequences (Brake *et al.*, 1985). *Mfa1* and *Mfa2* encode polypeptides of 36 and 38 amino acids, respectively, and have a single amino acid difference in the region encoding **a**-factor (Table 1.2). Both polypeptides are processed in the same way (reviewed in Chen *et al.*, 1997) (Figure 1.4). The precursors have a carboxy-terminal CaaX motif (C is cysteine, a is aliphatic and X is one of many residues). This motif is present at the carboxy-terminus of a number of eukaryotic proteins including Ras proteins and is a signal for modification of polypeptides by prenylation of the cysteine residue, proteolysis of the carboxy-terminal aaX residues and methylation of the newly exposed cysteine carboxyl group (Omer and Gibbs, 1994).

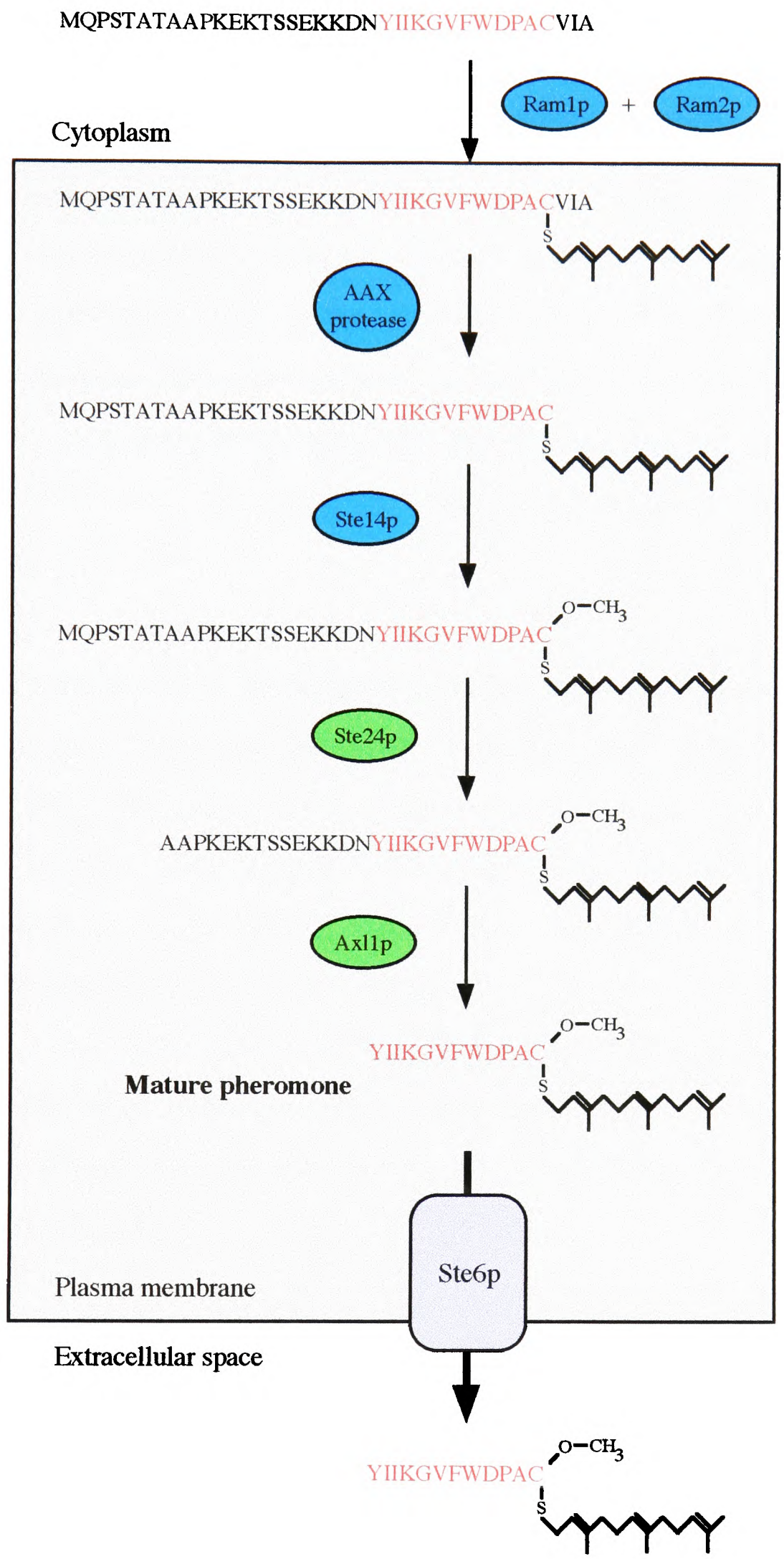
Following carboxy-terminal modification, amino-terminal proteolysis occurs in two distinct steps (Chen *et al.*, 1997). Firstly, the amino-terminal seven amino acids are cleaved by the Ste24p protease (Fujimura-Kamada *et al.*, 1997). The second and final processing event involves Ax11p. The mature **a**-factor is then exported by a non-classical export system instead of the classical secretory pathway used by α-factor (Kuchler *et al.*, 1989). This involves the Ste6p protein which is a member of the ATP-binding cassette (ABC) family of transport proteins which includes the product of the human *MDR1* gene, the multiple drug resistance (MDR) P-glycoprotein (McGrath and Varshavsky, 1989). Indeed, a *STE6* gene mutation has been complemented by expressing a homologue of human MDR1, the mouse *mdr3* gene in *S. cerevisiae* (Raymond *et al.*, 1992). ABC proteins carry out ATP-mediated translocation of various compounds including small peptides and hydrophobic drugs by an uncharacterised mechanism (Gottesman and Pastan, 1993).

Table 1.2 Pheromone precursors of *S. cerevisiae* and *S. pombe*.

Organism	Gene	Amino acid sequence of precursor†
S. cerevisiae	*MFa1*	MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPACVIA
S. cerevisiae	*MFa2*	MQPITTASTQATQKDKSSEKKDNYIIKGLFWDPACVIA
S. pombe	*mfm1*	MDSMANSVSSSSVVNAGNKPAETLNKTVKNYTPKVPYMCVIA
S. pombe	*mfm2*	MDSIATNTHSSSIVNAYNNNPTDVVKTQNIKNYTPKVPYMCVIA
S. pombe	*mfm3*	MDSMANTVSSSVVNTGNKPSETLNKTVKNYTPKVPYMCVIA

†Amino acid sequence of the mature pheromone shown in blue.

Figure 1.4 Model for the pathway of **a**-factor biosynthesis. The components involved in isoprenylation of the cysteine residue, proteolysis of the AAX residues and carboxyl methylation of the cysteine are shown in blue. Proteins involved in N-terminal proteolysis are indicated in green.


MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPACVIA
Ram1p
+
Ram2p
Cytoplasm
MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPACVIA
S
AAX protease
MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPAC
S
Ste14p
MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPAC
O−CH3
S
Ste24p
AAPKEKTSSEKKDNYIIKGVFWDPAC
O−CH3
S
Axl1p
YIIKGVFWDPAC
O−CH3
S
Mature pheromone
Ste6p
Plasma membrane
Extracellular space
YIIKGVFWDPAC
O−CH3
S

The **a**-factor and α-factor are recognised by their cognate receptors, Ste3p (**a**-factor receptor) and Ste2p (α-factor receptor) respectively (Bender and Sprague, 1986; Hagen *et al.*, 1986). These proteins are predicted to have seven transmembrane helices and as such, are members of a family of seven-helix receptors that are linked to heterotrimeric guanine nucleotide-binding proteins (G-proteins) (Dohlman *et al.*, 1991; Baldwin, 1994). Other members of this family include Rhodopsin, β and α-adrenergic receptors and odorant receptors. Mutations in several members of this family are implicated in human diseases (Coughlin, 1994) and in fact, yeast is being used as a model system in which to study β-adrenergic receptor and other 7-transmembrane segment receptor systems (Jeansonne, 1994).

Binding of the pheromones to their cognate receptor protein results in the activation of the mating response pathway common to both **a** and α haploid cells of *S. cerevisiae* (see Kurjan, 1993; Herskowitz, 1995; Leberer *et al.*, 1997). Ligand binding causes the receptor to undergo conformational changes and ubiquitination (Büküsoglu and Jenness, 1996; Roth and Davis, 1996). These changes are then believed to activate the coupled G protein. The heterotrimeric G protein consists of an α, β, and γ subunit. The α subunit has an intrinsic GTPase activity and the inactive form of the heterotrimer contains a GDP-bound α subunit and a βγ subunit. Activation of the receptor causes exchange of GTP for GDP and dissociation of the GTP-bound α subunit from the βγ dimer (Birnbaumer, 1992).

The βγ subunit activates a downstream cascade involving Ste11p and Ste7p (Figure 1.5). Several other proteins interact with the various components of the signalling pathway including Ste5p which associates with all the kinases of the MAP kinase cascade, and Ste20p. The exact functions of these associated proteins is unsure because some of them are not essential for mating. Finally the MAP kinases, Fus3p and Kssp1,

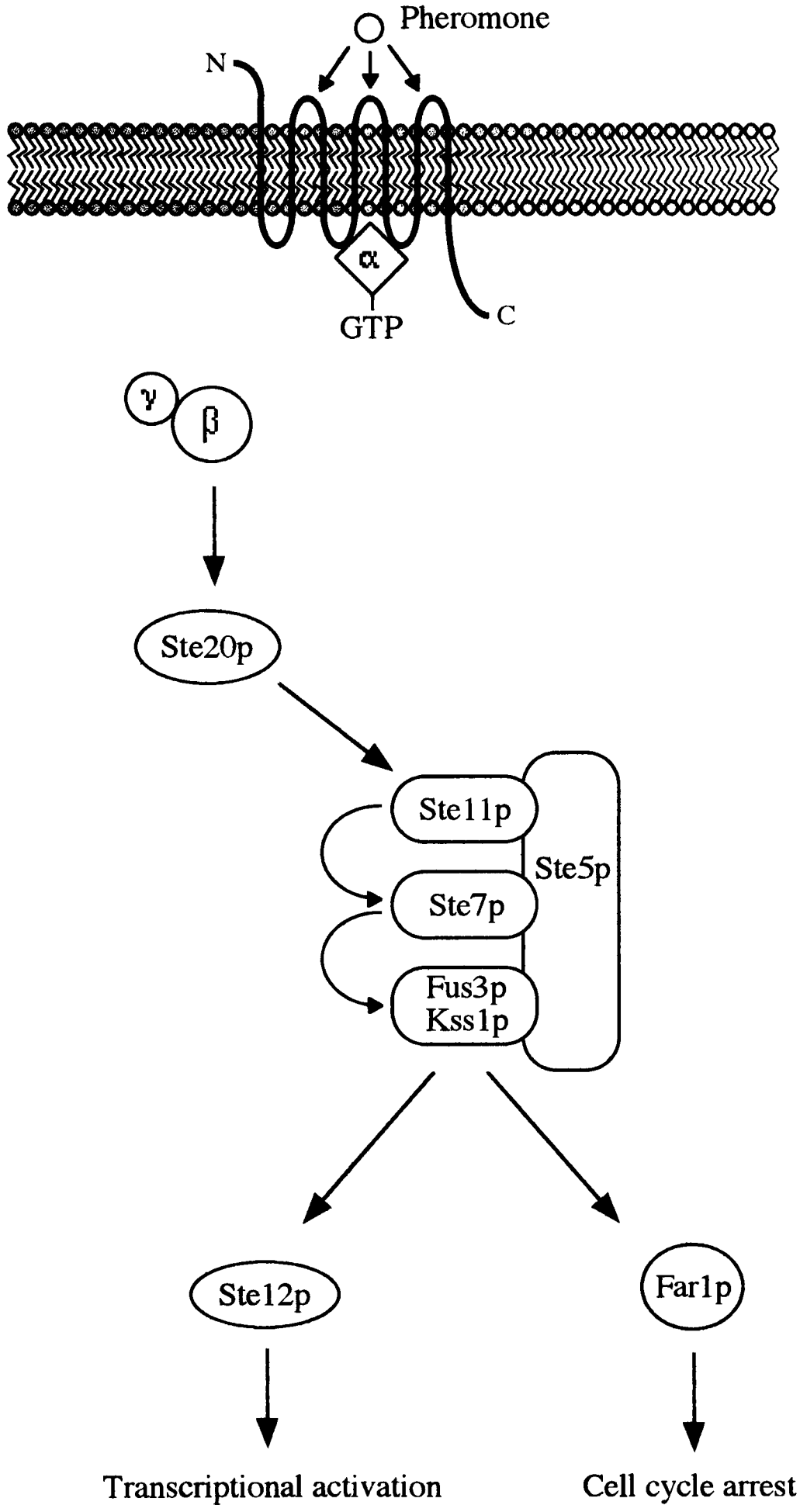


Figure 1.5 Pheromone response signal transduction pathway in *S. cerevisiae*.

activate by phosphorylation, both Far1p which binds and inhibits cyclin-dependent kinases leading to cell cycle arrest, and the transcription factor Ste12p, which is required to activate transcription of many genes that make the haploid *S. cerevisiae* cells competent to mate.

1.7 Mating type determination in *Schizosaccharomyces pombe*

Certain features about mating in the fission yeast, *S. pombe*, are also relevant to mating in basidiomycete fungi. *S. pombe* also has two mating types determined by genes at a single mating type locus *mat1*. M cells have the *mat1-M* locus that contains two genes, *mat1-Mc* and *mat1-Mm*. P cells have the *mat1-P* locus which contains the corresponding *mat1-Pc* and *mat1-Pm* genes. The *mat1-Pc* and *mat1-Mc* genes are responsible for establishing pheromone communication which leads to cell fusion whereas the *mat1-Mm* and *mat1-Pm* genes are necessary for meiosis (Kelly *et al.*, 1988).

M cells produce M-factor which is encoded by three genes, *mfm1*, *mfm2* and *mfm3*. The precursor polypeptides, like the **a**-factor polypeptides of *S. cerevisiae*, have a carboxy terminal CaaX motif (Davey, 1992; Kjærulff *et al.*, 1994) (Table 1.2). Mature M-factor is a nonapeptide with a carboxy terminal cysteine that is carboxymethylated and S-farnesylated. Processing is assumed to be similar to that of **a**-factor and indeed, an ABC transporter, Mam1, which is highly homologous to Ste6p and mediates secretion of M-factor, has been described (Christensen *et al.*, 1997). P-factor is encoded by the *map2* gene (Imai and Yamamoto, 1994). The Map2 polypeptide has a signal sequence and four tandem repeats of the mature P-factor sequence. P-factor is thus analogous to α-factor and is likely to be processed similarly and secreted by the classical secretory pathway.

The Mam2 and Map3 proteins are the receptors that bind P-factor and M-factor, respectively. These are linked to G proteins (Kitamura and Shimoda, 1991; Tanaka *et al.*, 1993). However, it is the α subunit of the G protein that transmits the signal in association with Ras1 to the common phosphorylation cascade (Nielsen *et al.*, 1992). Activation through the α subunit is found more commonly in mammalian signalling pathways (Birnbaumer, 1992). Activation of the MAP kinase cascade leads ultimately to expression of genes that bring about mating competence (Aono *et al.*, 1994).

There appears to be no homologue of the Ste12p transcription factor of *S. cerevisiae*, which is considered to be a degenerate homeodomain protein (Yuan *et al.*, 1993). The Ste11 protein is known to play an essential role in activating transcription of genes that bring about the pheromone response (Aono *et al.*, 1994) but its function has not been shown to be phosphorylation dependent. Ste11 is a member of the HMG (high mobility group) (Grosschedl *et al.*, 1994) class of DNA-binding proteins and binds to a conserved TR-box motif (Sugimoto *et al.*, 1991). This recognition site is found upstream of all the pheromone-controlled genes so far identified including *fus1* which is a gene required for cell fusion (Petersen *et al.*, 1995). Implicating Ste11 as a pheromone response factor has been complicated by the fact that it also mediates transcription of genes in response to nitrogen starvation (Sugimoto, *et al.*, 1991)

1.8 Signalling molecules in mating in basidiomycetes

Sequencing of the second mating type locus, the *a* locus, of *Ustilago maydis* showed that the genes encoding the pheromones and their cognate receptors have become mating type determinants in this group of fungi. The two alleles of the *a* locus, *a1* and *a2*, each contain a gene encoding a peptide pheromone and a 7-segment transmembrane receptor (Bölker *et al.*, 1992). These pheromone and receptor genes have been

designated *mfa1* and *pra1* respectively, in the *a1* locus and *mfa2* and *pra2* respectively, in the *a2* locus.

U. maydis is a plant pathogen which has a dimorphic life cycle. The unmated cells, the sporidia, are unicellular and yeast-like and divide by budding. These cells are saprophytic. Mating gives rise to a filamentous dikaryon that is an obligate pathogen on maize, *Zea mays*. The dikaryon induces tumour development in the plant, and within the tumours the dikaryon differentiates into thick walled diploid cells known as teliospores. These are liberated from the plant, and on germination, meiosis occurs and the four haploid basidiospores are produced.

Different alleles of the *a* locus are essential for mating cell fusion in *U. maydis* and are also required to maintain the filamentous growth of the dikaryon in the host plant. The pheromones encoded by the *mfa1* and *mfa2* genes are secreted and binding of pheromone to the compatible receptor induces the development of long conjugation filaments which fuse at their tips. Immediately after fusion has taken place the filamentous dikaryon starts to develop (reviewed in Banuett and Herskowitz, 1994b).

Both pheromone precursor molecules have a carboxy-terminal CaaX motif indicating that both versions of the pheromones in *U. maydis* are posttranslationally modified by the addition of a farnesyl group. The peptide pheromones have been purified and shown to have 13 (*mfa1*) and 9 (*mfa2*) amino acids (Spellig *et al.*, 1994). Purified pheromone added to a culture of the opposite mating type induces the development of conjugation tubes.

Apart from the signalling molecules themselves, little of the downstream phosphorylation pathway has been determined. An isolated G protein α subunit, Gpa3

has been implicated in the transmisssion of the pheromone signal and in this respect, the *U. maydis* pathway resembles that of *S. pombe* (Regenfelder *et al.*, 1997). A MAP kinase kinase homologue, Fuz7, has also been isolated (Banuett and Herskowitz, 1994a). Of particular interest is the isolation of a gene that encodes a protein which appears to have a similar role to that of Ste11 in *S. pombe* (Hartmann *et al.*, 1996). This protein, Prf1, was so called because it was assumed to be the pheromone response factor activated by the MAP kinase cascade. Like Ste11, it has an HMG DNA-binding domain and binds elements that resemble the TR-box (ACAAAGGGA), found upstream of pheromone-inducible genes (Urban *et al.*, 1996).

The *a* genes play a role in maintaining filamentous growth of the dikaryon which means that pheromone signalling is still important after cell fusion. Prf1 has been shown to play a pivotal role in activating the transcription of the pheromone and receptor genes of the *a* locus. It is itself up-regulated many fold in response to pheromone stimulation and inactivation of the gene leads to loss of ability to transcribe the *a* genes. Significantly, Prf1 is also required to activate transcription of the homeobox genes in the *b* mating type locus. This may explain the role of the pheromone response in maintaining the filamentous growth of the dikaryon. By fusing the *b* genes to constitutive promoters, the need for Prf1 in dikaryotic growth was alleviated. The role of Prf1 seems to parallel that of Ste11 in *S. pombe* in that both are required for sexual development after cell fusion and both lead to up-regulation of mating type genes (Leupold *et al.*, 1989; Sugimoto *et al.*, 1991; Hartmann *et al.*, 1996).

1.9 The *B* mating type loci of the homobasidiomycetes

The fact that the pheromone response continues to play a role after cell fusion in species such as *S. pombe* and *U. maydis* emphasises the fact that signalling is not only required to attract mating cells and promote cell fusion. In the homobasidiomycetes cell fusion is mating type-independent and there is no evidence for pheromone involvement in this. However, a role for pheromone signalling in the maintenance of the dikaryon is now confirmed by the discovery that the *B* mating type loci of the mushroom fungi contain similar signalling molecules to those in the *a* loci of *U. maydis*. However, unlike the *a* genes of *U. maydis*, the *B* mating type genes are multiallelic and there are some 79-81 *B* mating specificities in *S. commune* and *C. cinereus*.

At the start of this study, the *Bα1* locus of *S. commune* had been cloned (Specht, 1995) and shown to contain at least a pheromone receptor gene and a pheromone precursor gene (C. Raper, unpublished). The *B6* locus of *C. cinereus* had been cloned by S. O'Shea in this laboratory and some of the *B6* genes had been identified by transformation and transcript analysis. During the course of the work, the genes of the *Bα1* and *Bβ1* loci of *S. commune* were described (Wendland *et al.*, 1995; Vaillancourt *et al.*, 1997) and provided useful data for the comparative analysis of gene sequences that is made in the final discussion.

1.10 Objectives of this study

The first objective of this study was to isolate a second *B* locus from a cosmid genomic library of *C. cinereus* in order to determine the general organisation of the *B* genes and to establish their likely functions by DNA sequence analysis. Having isolated the *B42* locus, and sequenced the genes, attention was focused on demonstrating how multiple *B* mating specificities might be generated by sets of functionally redundant genes, and comparative sequence analyses that might provide some insight into the basis of allelic variability and specificity.

Chapter 2

Materials and methods

2.1 Strains

2.1.1 *C. cinereus* strains

C. cinereus strains used in this thesis are listed in Table 2.1

2.1.2 Bacterial strains

Escherichia coli strains used in this study are listed in Table 2.2.

2.2 Bacterial cloning vectors

pBluescript II KS^-, pBluescript II SK^-, pBluescript II SK^+ (Stratagene) and pUC18 (Gibco BRL) were used for routine DNA cloning throughout this study. Either pGEM-T or pGEM-T Easy (Promega) were used for cloning PCR amplification products.

2.3 Chemicals and stock solutions

All chemicals used in the preparation of media and buffers were of analytical grade. The following stock solutions were prepared according to Sambrook *et al.* (1989): 50x Denhardt's Reagent; 0.5M EDTA pH 8.0; IPTG (100mM); proteinase K (20mg/ml); salmon sperm DNA (sheared and denatured, 10mg/ml) ; 20% SDS (sodium dodecyl

Table 2.1 Genotype and origin of *C. cinereus* strains used in this study.

Strain	Mating type	Auxotrophic markers	Origin	Source
PR94226	*A6B3*	*ade-5 chol-1*		Day
218	*A3B1*	*trp-1.1, 1.6*	LT2 x C692	Casselton
AT8	*A43B43*	*ade-8, trp-3*	HT29.01 x OK130	Casselton
H9	*A6B6*	wild-type		Lewis
HT8	*A5B5*	*trp-1.2*	JT16	Tilby
FA2222	*A5B6*	*trp-1.1, 1.6*	LT2 x FA222	Casselton
J6,5,5	*A43B42*	wild type	JV6 x OK130	Zolan
JV6	*A42B42*	wild-type		Hedger
LN118	*A42B42*	*ade-2, trp-1.1, 1.6*	LT2 x NGB01	Casselton
LC012	*A2B3*	*trp-3*	94226 x PR2301	Casselton
LT2	*A6B6*	*trp-1.1, 1.6*	HT14.01	Casselton
NL1	*A44B44*	*trp-1.1;1.6*	SF2 x LT2	Casselton
NL1R	*A43B44*	wild type	NL1 x J6,55	Casselton
RSC2	*A6B42*	*met-8*	ZBR9/55 x JV6	Casselton
68	*A2B1*	*wild type*		Lewis

Table 2.2 Genotype of bacterial strains used.

Strain	Genotype
XL1-Blue:	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q}ZΔ*M15* Tn*10* (Tet^{r})] (Bullock *et al.*, 1987)
DH5α:	*supE44* Δ*lacU169* (ϕ80 *lacZ*ΔM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1* (Hanahan 1983)
JM109:	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ(*lac-proAB*) [F' *traD36 proAB lacI*^{q}ZΔM15] (Yanisch-Perron *et al.*, 1985)

sulphate); 3M sodium acetate pH4.8, pH5.2 and pH7.0; 1M sodium phosphate buffer pH7.0; 20x SSC; 20x SSPE; 50x TAE; 10x TE pH8.0; 10x TNE; 1M Tris-HCl buffers; 2% X-gal (5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside).

2.4 Media

All media were sterilized by autoclaving at 121°C for 15 minutes (1.03×10^5Pa).

2.4.1 *C. cinereus* stock solutions and media

Trace elements stock solution: 1.0g calcium chloride (anhydrous), 4.1g magnesium chloride (hydrate), 0.53g iron (III) citrate (hydrate), 0.44g manganous sulphate (hydrate), 0.4g zinc sulphate (hydrate) dissolved in 400 ml water and stored at 4°C.

Complete Medium (Lewis, 1961) 10.0g glucose, 2.0g L-asparagine, 0.75g yeast extract, 0.75g casitone, 1.13g malt extract, 0.1g L-tryptophan, 0.1g adenine sulphate, 0.5g ammonium tartrate, 1.0g potassium dihydrogen orthophosphate, 0.29g disodium sulphate (anhydrous), 2.25g disodium hydrogen phosphate (anhydrous), 40 μg thiamine (aneurine HCl), 0.25g magnesium sulphate dissolved in 1 litre water. 13.0g agar added prior to autoclaving.

Minimal Medium (Shahriari and Casselton, 1974) 2.0g L-asparagine, 10.0g glucose, 0.5g ammonium tartrate, 1.0g potassium dihydrogen orthophosphate, 0.29g disodium sulphate (anhydrous), 2.25g disodium hydrogen phosphate (anhydrous), 40.0μg thiamine (aneurine HCl), 0.25g magnesium sulphate dissolved in 1L water. 14.0g agar added prior to autoclaving.

Yeast, Malt, Glucose Medium (Rao and Niederpruem, 1969) 4.0g yeast extract, 10.0g malt extract, 4.0g glucose dissolved in 1 litre water. 10.0g agar added prior to autoclaving.

Regeneration Medium (Binninger *et al.*, 1987)*:* 5.0g glucose, 2.0g L-asparagine, 0.5g ammonium tartrate, 1.0g potassium dihydrogen orthophosphate, 0.29g disodium sulphate (anhydrous), 2.25g disodium hydrogen phosphate (anyhydrous), 40.0µg thiamine (aneurine HCl), 0.25g magnesium sulphate, 172.0g sucrose, 5.0g soluble starch dissolved in 1 litre water. 14.0g agar added prior to autoclaving.

Liquid media: Liquid minimal medium was prepared as above but with the omission of magnesium sulphate and agar and the addition of 40ml/litre trace elements solution. Liquid yeast, malt, glucose medium was prepared as above but without agar.

Supplemented media: Supplements were added to the media when required at the following concentrations: L-tryptophan (0.1g/litre), adenine sulphate (0.1g/litre), L-histidine (0.1g/litre).

2.4.2 Bacterial Media

LB (Luria-Bertani) Medium: 10.0g tryptone, 5.0g yeast extract, 5.0g sodium chloride, 1.0ml 1M sodium hydroxide solution. 1% agar was added prior to autoclaving for solid medium. 15% glycerol was added prior to autoclaving for LB glycerol medium.

X-gal Medium: 2.5g sodium chloride, 2.5g potassium chloride, 10.0g tryptone, 1.0ml 1M calcium chloride dissolved in 1L water. 10.0g agar added prior to autoclaving.

100mg ampicillin (Sigma A-9518, sodium salt), 2ml 2% X-gal and 2ml 0.1M IPTG were added after cooling.

2xYT Medium: 10.0g NaCl, 10.0g Difco-Bacto yeast extract, 16.0g tryptone dissolved in 1 litre water.

2.5 Antibiotics

Antibiotic stock solutions were prepared as in Table 2.3 and stored at -20°C.

Table 2.3 Stock and working solutions of antibiotics used for selective media.

Antibiotic	Stock solution	Working solution
ampicillin (Sigma A-9518)	50mg/ml in sterile water	100µg/ml
chloramphenicol (Sigma C-0378)	34mg/ml in 100% ethanol	34µg/ml (*C. cinereus*)
kanamycin (Sigma K-4000)	50mg/ml in sterile water	50µg/ml

2.6 *C. cinereus* culturing conditions

Mycelial cultures were grown at 37°C on solid medium in 90mm plastic Petri dishes. Cultures were stored for short periods at 4°C. For long-term storage, 5mm^2 blocks of mycelium were placed in 10% glycerol and stored at -80°C. Small-scale liquid cultures were initiated by inoculating 3-4 5mm^2 blocks of mycelium into 15-20ml of liquid medium in Petri dishes and incubating at 37°C for 5 days. For large-scale cultures, Petri

dish cultures were macerated in 100ml sterile water using an MSE AtoMix blender. 20ml of the macerate were inoculated into 200ml liquid minimal medium containing 34µg/ml chloramphenicol in a 1 litre conical flask and incubated at 37°C in a rotary incubator shaking at 100-125 rpm.

2.7 Bacterial culturing conditions

E. coli strains were grown on solid media (containing the necessary antibiotics to maintain plasmids) in 90mm plastic Petri dishes and incubated overnight at 37°C. Small-scale liquid cultures in boiling tubes were grown by inoculating single colonies into 5ml liquid medium containing the necessary antibiotics and incubating overnight at 37 °C on a roller drum. For long term storage of bacteria, cells were pelleted from overnight cultures and resuspended in 1ml LB containing 15% glycerol and the appropriate antibiotic, and stored at -80°C.

2.8 DNA-mediated transformation

2.8.1 Transformation of *C. cinereus*

modified from Casselton and de la Fuente Herce (1989)

Oidia from a five day old culture were harvested by pouring sterile water onto the surface of the culture and scraping gently with a sterile blade. Mycelial fragments were removed by filtering through glass wool. Cells were washed with MM Buffer (0.5M mannitol, 50mM sodium maleate pH5.5) and resuspended in 2-4ml digestion buffer (30mg/ml Onozuka R10 cellulase, 1mg/ml chitinase (Sigma C-6137) dissolved in sterile MM buffer and sterilised by centrifugation (35,000 xg at 4°C for 30 minutes)) and incubated at 37°C for 3-5 hours. Protoplasts were recovered by centrifugation at 2,500 rpm in a Sorvall Econospin bench top centrifuge at room temperature for 5 minutes. Protoplasts were then washed twice in MM buffer, once in MMC buffer (0.5M mannitol, 50mM sodium maleate pH5.5, 50mM $CaCl_2$) and finally resuspended in an appropriate volume of MMC to give 50µl protoplasts per transformation.

For each transformation, 12.5µl PEG (25% (w/v) PEG 4000, 10mM Tris-HCl pH7.5, 25mM calcium chloride) and 1-5µg DNA were added to 50µl protoplasts. For co-transformations the plasmid to be transformed was combined with 1µg pCc1001 (Binninger *et al.*, 1987) or pDB3 (Burrows, 1991), which contain the *C. cinereus trp-1* gene and *trp-3* gene respectively, prior to mixing with the protoplasts. After gentle mixing the cells were incubated on ice for 20 minutes. 0.5ml PEG was added and following mixing by inversion, the cells were incubated for 5 minutes at room temperature. Following this, 1ml STC (1M Sorbitol, 50mM calcium chloride, 10mM Tris-HCl pH7.5) was added and the cells were plated onto the surface of regeneration agar and incubated at 37 °C for 3-4 days until transformants appeared.

2.8.2 Transformation of bacteria

Preparation of competent *E. Coli* cells

0.2-0.3ml of a fresh overnight culture was inoculated into 250ml liquid LB medium. This was vigorously shaken at 150-200 rpm at 37°C until the cells reached an OD_{600} of 0.2.

The cells were chilled in ice water and harvested by centrifugation at 4,000 xg for 10 minutes at 4°C. The cell pellet was resuspended in (125ml 50mM calcium chloride and 10 mM Tris-HCl pH 8.0) and incubated on ice for a minimum of 2 hours to induce competence. The cells were then harvested by centrifugation at 4,000 xg for 10 minutes at 4°C and resuspended in 20ml ice cold 50mM calcium chloride. After adding 7ml ice cold 50% glycerol, the cells were gently mixed. Aliquots were then dispensed into pre-chilled, sterile 1.5ml Eppendorf tubes and stored at -80°C.

Transformation of competent *E. Coli* cells (Dagert and Ehrlich, 1979)

The frozen competent cells were defrosted on ice. 10-30ng DNA was added to 100µl competent cells and incubated on ice for 20-30 minutes. Following a heat shock at 42°C for 2 minutes (XL1-Blue, DH5α) or 50 seconds (JM109) and a subsequent 2 minutes on ice, 1ml LB pre-warmed to 37°C was added and the cells were incubated at 37°C for 60-90 minutes. Cells were pelleted at 10,000 rpm for 1 minute and 900µl of supernatant removed. Following resuspension in the remaining supernatant, 10-100µl was plated onto selective medium and the plates were incubated overnight at 37°C.

For re-transformations of plasmid DNA, the DNA and cells were mixed, incubated on ice for 5 minutes and plated out onto selective media.

2.9 DNA isolation and manipulation

All micro-centrifugation carried out in a Hereus Sepatech Biofuge 13 unless otherwise specified.

2.9.1 Large-scale *C. cinereus* genomic DNA isolation (Mutasa *et al.*, 1990)

Mycelium from a two-day old 200ml culture was harvested by filtration onto Whatman Number 1 filter paper. The culture was washed twice with sterile water and blotted dry between filter paper and paper towels. 20-30g mycelium was frozen and roughly ground in liquid nitrogen with a pestle and mortar. The frozen mycelium was homogenised with more liquid nitrogen in an MSE Atomix blender for 1 minute at high speed. The mycelium was then ground to a fine powder with the pestle and mortar. 20g powdered mycelium was added to 50 ml digestion buffer (50mM EDTA pH 8.0, 0.5% (w/v) Sarkosyl NL30 (30% (w/v) N-lauroyl sarcosinate; BDH, catalogue number 44275), 100µg/ml proteinase K). The mixture was gently stirred with a glass rod and was incubated at 55°C for 2 hours. 0.1 volume 0.1M EDTA pH 8.0 and an equal volume of phenol:chloroform (1:1) were added for 10 minutes at room temperature. The aqueous and phenol phases were separated by centrifugation at 6,000 rpm for 5 minutes in a Sorvall RC-3B centrifuge using an SS-34 rotor. The aqueous phase was collected and the DNA precipitated at -20°C for at least 2 hours by adding 0.1 volume 3M sodium acetate pH 4.8 and 2 volumes ethanol. After centrifugation at 9,000 rpm for 30 minutes at 4°C, the pellet was resuspended in 4-6ml sterile water. RNA was digested with RNAase A (final concentration 100µg/ml, stock solution 5mg/ml RNAase A type IIIA

from bovine pancreas (Sigma R-4875) in 10mM Tris-HCl pH7.5) at 37°C for 1 hour. Another phenol/chloroform extraction was followed by two SEVAG (24:1 chloroform:isoamyl alcohol) extractions to remove all traces of phenol, centrifuging at 3,500 rpm for 10 minutes at room temperature. Genomic DNA was then purified by caesium chloride density gradient centrifugation. 5g caesium chloride and 0.4ml ethidium bromide (10mg/ml) were added to 5ml DNA solution. The sample was loaded into ultracentrifuge tubes (Quick Seal Bell-top, Beckman) and centrifuged at 100,000 rpm for 16 hours at 20°C in a Beckman TL100 centrifuge using a Beckman TLA100.3 rotor. The genomic band was recovered using a syringe and 19g x 1.5B needle. Ethidium bromide was removed using equal volumes of isopropanol saturated with TE-caesium chloride. The DNA solution was finally dialysed three times against TE pH 8.0 at 4°C.

2.9.2 Small-scale *C. cinereus* genomic DNA isolation (Zolan and Pukkila, 1986)

Mycelium was grown in a 15-20ml Petri dish culture for 5 days at 37°C, then harvested, blot-dried and placed in a 1.5ml Eppendorf tube. The mycelium was rapidly frozen in liquid nitrogen and freeze dried overnight. The freeze dried material (1/3 of the tube volume) was ground with a toothpick in the tube and suspended in 500µl of CTAB extraction buffer (0.7M sodium chloride, 1% CTAB (w/v) (Sigma H-5882), 50mM Tris-HCl pH8.0, 10mM EDTA pH8.0, 1% mercaptoethanol (v/v)). The suspension was incubated for 10 minutes at room temperature and stirred after 5 minutes. SEVAG extraction was carried out with an equal volume of SEVAG (24:1 chloroform:isoamyl alcohol) for 10 minutes. Following centrifugation, the upper aqueous phase was transferred to a fresh tube, mixed with an equal volume of isopropanol and left for 10 minutes at room temperature. The DNA was pelleted by centrifugation for 5-10 minutes at room temperature and resuspended in 300µl 1 x TE. 10µl RNAase A (stock solution

5mg/ml RNAase A type IIIA from bovine pancreas (Sigma R-4875) in 10mM Tris-HCl pH7.5) was added to the DNA suspension which was incubated for 1 hour at 37°C. RNAase A was removed by phenol-chloroform extraction followed by two SEVAG cleaning steps. DNA was precipitated for 10 minutes at -20°C by addition of 0.5 volumes of 7.5M ammonium acetate solution and 2 volumes of cold ethanol, followed by centrifugation for 10 minutes at 13,000 rpm. The pellet was resuspended in 300µl 0.2M ammonium acetate and re-precipitated for 10 minutes at -20°C by the addition of 2 volumes of ethanol. Following centrifugation, the DNA pellet was washed with 70% ethanol, air dried and resuspended in 50µl TE.

2.9.3 Bacterial plasmid DNA isolation (small scale method)

A single colony was inoculated into 5ml LB ampicillin and incubated at 37°C overnight on a rolling drum. 3ml of overnight culture was pelleted by centrifugation at 13,000 rpm for 1 minute. The pellet was resuspended in 200µl solution A (25mM Tris-HCl pH7.5, 50mM glucose, 10mM EDTA pH8.0). 300µl freshly made solution B (0.2M sodium hydroxide, 1% (w/v) SDS) was added, mixed by inversion and the sample incubated on ice for 5 minutes. 300µl solution C (3M potassium acetate, 11.5% (v/v) glacial acetic acid) was added, mixed by inversion and the sample incubated on ice for a further 5 minutes. The precipitate was pelleted by centrifugation at 13,000 rpm for 10 minutes at room temperature. The supernatant was removed to a fresh Eppendorf tube and 3.2µl RNAase A (stock solution 5mg/ml RNAase A type IIIA from bovine pancreas (Sigma R-4875) in 10mM Tris-HCl pH7.5) was added. The tube was incubated at 37°C for 20 minutes. The sample was extracted twice with an equal volume of chloroform for 30 seconds, centrifuging for 1 minute at 13,000 rpm each time. The DNA was precipitated by adding an equal volume of propan-2-ol and then pelleted by centrifuging at 13,000 rpm for 10 minutes at room temperature. The pellet was washed

with 70% ethanol, dried in a Savant DNA110 rotary evaporator and dissolved in 32µl sterile water. The DNA was then reprecipitated by adding 8µl 4M sodium chloride and 40µl 13% (w/v) PEG 8000, mixing well and placing on ice for 20 minutes. The DNA was pelleted by centrifuging at 13,000 rpm for 15 minutes at 4°C. The pellet was washed with 70% ethanol, dried in a Savant DNA110 rotary evaporator and redissolved in 25µl sterile water. The DNA was stored at -20°C.

2.9.4 Restriction enzyme digestion of DNA

Restriction enzyme digests were performed using the conditions recommended by the manufacturers (Boehringer-Mannheim, GibcoBRL Life Technologies Inc., Pharmacia Biotech.).

2.9.5 Estimation of double-stranded DNA concentration

Estimates for DNA concentrations were obtained by electrophoresing the sample through a 1% agarose gel containing ethidium bromide at a final concentration of 0.5µg/ml along with a *Hin*dIII cut λ DNA (GibcoBRL Life Technologies Inc.) standard of known concentration and comparing their fluoresent intensities under short wave UV radiation.

2.9.6 Agarose gel electrophoresis of DNA

Agarose gel electrophoresis was routinely used to separate DNA fragments. Agarose concentrations were varied from 0.8% to 1.2% (w/v) agarose dissolved in 1x TAE. Ethidium bromide was added to the gel to a final concentration of 0.5µg/ml in order to

visualise DNA bands under UV illumination following electrophoresis. Gels were run at 10-75 volts in 1x TAE electrophoresis buffer (Sambrook *et al.*, 1989).

Phage λ DNA digested with *Hin*dIII was used as a size marker for electrophoresed DNA fragments throughout this study. Approximately 1.0μg was run and the fragment sizes produced were 23.13kb, 9.416kb, 6.557kb, 4.361kb, 2.322kb, 2.027kb, 0.564kb, 0.125kb.

2.9.7 Photography of agarose gels containing ethidium bromide stained DNA

Gels were photographed using transmitted short-wave UV radiation (UVP transilluminator) through an orange lens filter with a Polaroid MP-4 camera. Two different types of film were used. For Ilford HP-5 plus film an aperture setting of 8 and exposure time of 3-7 seconds. The film was developed in Kodak X-ray developer (LX-24) for 6 minutes, washed with tap water, fixed in Kodak X-ray fixer (FX-40) for four minutes, then washed thoroughly with tap water before air drying. Alternatively, Polaroid 667 film was used with an aperture setting of 8 and exposure time of 1 second. This film was self-processing (30-60 seconds at room temperature).

2.9.8 Isolation of DNA fragments from agarose gels

Following gel electrophoresis, the gel was illuminated with long wave UV light and the band was excised from the gel using a sterile blade. The DNA was extracted routinely using commercially available kits (Gene-clean II, BIO 101 Inc. and Jetsorb, Genomed Inc.) using the manufacturers instructions. Alternatively the glass wool method of Wang and Rossman (1994) was used.

2.9.9 Cloning Procedures

Cloning of DNA fragments was carried out following standard procedures (Sambrook *et al.*, 1989). The plasmid vector pBluescript II (Stratagene) was used routinely for sub-cloning DNA fragments. The vector fragment was treated with shrimp alkaline phosphatase (Boehringer Mannheim) according to the manufacturers instructions to reduce vector self-ligation. The DNA fragments were purified by agarose gel electrophoresis and band elution prior to ligation.

DNA fragments were ligated using T_4 DNA ligase in 1x ligase buffer (GibcoBRL Life Technologies Inc.) Ligations were carried out for 2 hours or overnight at 16°C and following transformation into a suitable recipient *E. coli* strain (XL1-Blue, DH5α), colonies containing vector with insert were identified using blue/white screening. Successful insertion of a fragment into the polylinker of the vector causes insertional inactivation of the β-galactosidase gene. This produces white colonies on X-gal medium, compared to blue colonies from cells replicating the uninterrupted vector.

2.10 RNA manipulation

2.10.1 Preparation of equipment and solutions

RNAase contamination must be eliminated for the isolation and manipulation of RNA. Equipment such as pipettors, were made RNAase-free by treating with RNase-Away (MβP, Molecular *Bio*-products, San Diego; catalogue number 7000). Disposable plasticware, pipette tips and Eppendorf tubes were autoclaved and baked at 140°C for four hours. Polyallomer ultracentrifuge tubes and 50ml Oakridge centrifuge tubes were

immersed in 10% (v/v) HCl overnight, rinsed in RNAase free water, then autoclaved. Glassware was baked overnight at 180°C.

Solutions were made RNAase free by adding diethylpyrocarbonate (DEPC) to a final concentration of 0.1% (v/v), incubating overnight at 37°C and autoclaving. Tris buffers were made from separate stocks using DEPC treated water.

2.10.2 Total RNA isolation from *C. cinereus* using caesium chloride centrifugation

The caesium step gradient procedure of Glisin *et al.*. (1974) with the modifications of Chirgwin *et al.* (1979) was used for the isolation of total RNA. Unless otherwise stated, all centrifugation was carried out in a Sorvall RC5B or RC5C centrifuge for thirty minutes at 4°C using an SS-34 rotor.

The mycelium from 2-3 200ml liquid cultures was harvested by filtration onto Whatman Number 1 filter paper. The culture was washed twice with sterile water and blotted dry between filter paper and paper towels. The mycelium was immersed in an equal volume/weight of homogenisation buffer (5.5M guanidinium thiocyanate, 25mM tri-sodium citrate pH 7.0, 0.34% (v/v) antifoam A (Sigma A-5758), 0.5% (v/v) Sarkosyl NL30 (30% (w/v) N-lauroyl sarcosinate; BDH, catalogue number 44275), 1.4% (v/v) β-mercaptoethanol) and homogenised for two minutes using a 15mm diameter probe homogeniser (Silverson or Ultra-Turrax). The homogenate was allowed to stand for two minutes before being re-homogenised for a further two minutes. The slurry was centrifuged at 20,000 xg for 20 minutes at room temperature.

2.0ml of the supernatant was overlaid onto 1.0ml 5.5M caesium chloride, 0.1M EDTA in a Beckman thick walled polyallomer centrifuge tube (catalogue number 349623). The

RNA was pelleted in a Beckman TL100 ultracentrifuge, using a TL100.3 rotor, at 65,000 rpm for five hours at 20°C.

The supernatant was carefully removed using a Pasteur pipette and the inside of the tube cleaned with tissue paper without disturbing the pellet. The pellets from six tubes were resuspended in 5.55ml guanidinium hydrochloride solution (7.5M guanidinium HCl, 50mM tri-sodium citrate). 0.3ml 1M acetic acid and 3ml 96% ethanol were added, mixed and the RNA precipitated for 2 hours at -20°C. The RNA was pelleted for 30 minutes at 15,000 rpm for 30 minutes. The supernatant was discarded and the pellets washed with 70% ethanol. The pellets were partially air-dried for 15 minutes and resuspended in 5.4ml DEPC treated water. 0.6ml 1M sodium acetate and 15ml 96% ethanol were added, mixed and the RNA re-precipitated for at least 2 hours at -20°C. The RNA was pelleted at 15,000 rpm. The supernatant was discarded and the pellet thoroughly dried under a stream of nitrogen gas. The RNA pellets were dissolved in 0.5ml DEPC treated water and stored at -80°C.

2.10.3 Rapid total RNA isolation

This method was adapted from Hoge *et al.* (1982). 5-10g of harvested, dried mycelium was ground to a fine powder in liquid nitrogen, then added to 30ml extraction mixture consisting of 15ml phenol:chloroform:isoamyl acohol (25:24:1) and 15 ml extraction buffer (0.2M sodium acetate, 10mM EDTA, 1% (w/v) SDS, 0.5% (v/v) β-mercaptoethanol), that had been pre-heated to 65°C.

The mix was then vortexed vigorously for 2 minutes and centrifuged at 10,000 rpm (12,000 xg). The aqueous phase was carefully removed and the RNA was precipitated by adding 8M lithium chloride to a final concentration of 2M and chilling overnight at

4°C. The RNA was pelleted at 10,000 rpm for 15 minutes and then washed once with 2M lithium chloride and twice with 70% ethanol. The pellet was centrifuged at 10,000 rpm for 15 minutes during each wash. The pellet was air dried, dissolved in 250-500µl DEPC treated water and stored at -80°C.

2.10.4 Isolation of poly $(A)^+$ RNA

Poly $(A)^+$ RNA was purified from total RNA by oligo d(T) cellulose affinity chromatography (Edmonds *et al.*, 1971; Aviv and Leder, 1972). 0.5g of oligo d(T) cellulose type 7 (Pharmacia Biotech) was equilibrated overnight in 1.5-2.0ml 1x loading buffer at 4°C. The column was prepared with a bed volume of 0.5ml in a Gilson 1ml pipette tip plugged with siliconized glass wool.

The column was sterilised by washing with 5ml 0.1M sodium hydroxide, followed by RNAse free water until the eluate pH was below 8. The column was equilibrated with 6ml 1x oligo d(T) column loading buffer (0.5M potassium chloride, 10mM Tris-HCl pH 7.5).

RNA (2-5mg) was denatured at 65°C for fifteen minutes, quenched on ice and applied to the column in a volume of 1ml with 0.25ml 5x oligo d(T) loading buffer (2.5M potassium chloride, 50mM Tris-HCl pH7.5). The eluate was collected in a 15ml Falcon tube. The eluate was heated to 65°C for 15 minutes, quenched on ice for 5 minutes and re-applied to the column. The eluate was collected and re-applied a further three times without denaturing the RNA. The final eluate was poly $(A)^-$ RNA and was saved. The column was washed with 6ml 1x loading buffer to remove unbound poly $(A)^-$ RNA. The poly $(A)^+$ RNA was eluted from the column with 6x 0.5ml fractions of elution buffer (10mM Tris-HCl pH7.5), preheated to 45°C. The poly $(A)^+$ RNA fractions were

identified by UV fluorescence; 3µl of each fraction was mixed with 20µl ethidium bromide (1µg/ml), spotted onto cling-film and viewed on a transilluminator using a control of water and ethidium bromide for comparison. The RNA eluted from this column still contains approximately 50% ribosomal RNA (Sambrook *et al.*, 1989). The fractions containing RNA were pooled and the process repeated using another column to remove poly $(A)^-$ RNA contamination.

Poly $(A)^+$ RNA and poly $(A)^-$ RNA were precipitated overnight at -20°C using 0.1 volume 3M sodium acetate pH5.2 and 2.5 volumes 96% ethanol. The RNA was pelleted by centrifugation at 16,000 rpm (30,500 xg), the supernatant discarded and the pellet washed with 70% ethanol before being dried thoroughly under a stream of nitrogen gas. The poly $(A)^+$ RNA was dissolved in 30µl DEPC treated water and the poly $(A)^-$ RNA in 500µl DEPC treated water. The RNA samples were stored at -80°C. The concentration of the RNA was assessed spectrophotometrically and the samples were size-fractionated on a denaturing gel to check for degradation (See section 2.10.5).

Poly $(A)^+$ RNA was also purified from total RNA using the PolyATtract mRNA Isolation System (Promega) according to the manufacturers instructions.

2.10.5 RNA formaldehyde agarose gel electrophoresis

Standard protocols were used to produce the gel (Sambrook *et al.*, 1989). The gel contained 1% agarose, 0.6M formaldehyde, 1x MOPS (40mM MOPS (3-morpholinopropanesulfonic acid), 10mM sodium acetate, 1mM EDTA pH 7.0).

Typically, 5µg poly $(A)^+$ RNA or 30µg total RNA was loaded onto the gel in the following solution: 22.5% RNA in DEPC treated water; 10% 10x MOPS (0.4M MOPS,

100mM sodium acetate, 10mM EDTA pH 7.0), 17.5% formaldehyde, 50% formamide. To estimate transcript size 5µg 0.24-9.5kb RNA ladder (Gibco-BRL, Life Technologies Inc., catalogue number 15620-016) was also loaded onto the gel. If the RNA was to be visualised by UV fluorescence, it was stained with ethidium bromide prior to loading, using the method of Rosen and Villa-Komaroff (1993).

The RNA was denatured by heating to 55°C for 15 minutes, then quenched on ice for a further five minutes. 0.1 volume 10x RNA loading buffer (50% (v/v) glycerol, 1mM EDTA, 0.4% (w/v) bromophenol blue, 0.4% (w/v) xylene cyanol FF) was added prior to loading. The gel was run at $5Vcm^{-1}$ in 1x MOPS containing 0.6M formaldehyde.

2.10.6 Estimation of RNA concentration

RNA concentration and purity was calculated by measuring the optical density of the solution at 260nm and 280nm. An OD_{260} of 1.0 corresponds to an RNA concentration of 40µg/ml. A pure RNA sample has a $^{260}/_{280}$ ratio of between 1.8 and 2.0.

2.11 Capillary transfer of nucleic acids to nylon membranes

DNA and RNA transfer from agarose gels to nylon membranes was carried out by capillary blotting following standard techniques (Sambrook *et al.*, 1989). Transfer was carried out onto Hybond-N, Hybond-N^+ (Amersham) or Nitran (Schleicher & Schuell). Nucleic acids were fixed to the membrane by baking for 1 hour at 80°C and UV cross-linking using a UV Stratalinker 1800 (Stratagene) on the autocrosslink setting.

2.12 Staining RNA following transfer

If required the RNA was visualised after transfer and fixation to nylon membranes. The membrane was soaked in 0.04% methylene blue for ten minutes, then washed in 25% ethanol for 30 minutes to remove excess stain and air dried.

2.13 ^{32}P labelling of DNA for Southern/northern hybridisation

DNA fragments for northern hybridisations were labelled with [α^{32}-P]dCTP (Amersham, specific activity 3000Ci/mmol) by random priming (Feinberg and Vogelstein, 1983) using the GibcoBRL Life Technologies Inc. Random Primers DNA labelling system (catalogue number 18187-013) or RadPrime DNA labelling system (catalogue number 18428-011) following the manufacturers instructions. DNA fragments for Southern hybridisations were labelled as for northern hybridisations or with [α^{32}-P]dCTP (Amersham, specific activity 3000Ci/mmol) by nick translation using the GibcoBRL, Life Technologies Inc. Nick Translation System (catalogue number 18160-010) following the manufacturers instructions. Unincorporated nucleotides were removed from the radioactively-labelled probe by Sephadex G-50 size exclusion chromatography (Section 2.14)

2.14 Sephadex G-50 size exclusion chromatography

A Sephadex G-50 column was prepared in a 145mm Pasteur pipette plugged with glass wool and packed with Sephadex G-50 (12% (w/v) swollen Sephadex G-50, 1x (v/v) TNE, 0.2% (w/v) SDS). The column was equilibrated with two 150µl washes with 1x

TNE. The sample was applied to the column and 150ml aliquots of 1x TNE added. Separate fractions were collected and monitored for incorporation of the radioactive label using a series 900 Geiger counter. Fractions that contained incoporated label (corresponding to the first radioactive peak) were pooled.

Unincorporated nucleotides were also separated using Quick Spin Sephadex G-50 (fine) columns (Boehringer Mannheim, catalogue number 1 273 973) following the manufacturers instructions.

2.15 Southern hybridisation

Hybridisations were carried out in a Hybaid mini oven rotary incubator. Membranes were prehybridised at 65°C in prehybridisation solution (5x SSC, 0.5% (w/v) SDS, 20mM sodium phosphate buffer pH 7.0, 2x Denhardt's Reagent, 100µg/ml denatured salmon sperm DNA for a minimum of 4 hours. The radioactively labelled probe was denatured by boiling for 10 minutes and then quenched on ice for 5 minutes prior to addition to the pre-hybridisation solution. The membranes were hybridised to the probe overnight at 65°C.

2.16 Southern post-hybridisation washes

Following hybridisation, the membranes were washed at 65°C with increasing stringency in: Wash 1 (3x SSC, 5mM sodium phosphate buffer pH 7.0), 0.1% (w/v) SDS) for a 5 minute wash and a 15 minute wash; Wash 2 (1x SSC, 5mM sodium phosphate buffer pH 7.0, 0.1% (w/v) SDS) for two 45 minute washes; Wash3 (0.2x SSC, 5mM sodium phosphate buffer pH 7.0, 0.1% (w/v) SDS) for a 20-30 minute wash.

2.17 Northern hybridisation

Hybridisation was carried out in a Techne HB-1 rotary incubator. Membranes were prehybridised at 42°C for at least four hours in northern prehybridisation solution (5x SSPE, 50% (v/v) formamide, 5x Denhardt's Reagent, 0.5% (w/v) SDS, 400mg/ml denatured salmon sperm DNA. The radioactively labelled probe was denatured by boiling for 10 minutes and then quenched on ice for 5 minutes prior to addition to the prehybridisation solution. The membranes were hybridised to the probe overnight at 42°C.

2.18 Northern post-hybridisation washes

Following hybridisation, northern membranes were washed at 42°C in: Wash 1 (4x SSC, 25mM sodium phosphate pH 7.0, 0.1% (w/v) SDS) for two washes of 30 minutes; Wash 2 (1x SSPE, 0.1% (w/v) SDS) for 30 minutes; Wash 3 (0.1x SSPE, 0.1% (w/v) SDS) for 15 minutes.

2.19 Autoradiography

Membranes were heat sealed in plastic bags and exposed to Fuji RX X-ray film at -80°C with intensifying screens. The film was developed in Kodak X-ray developer (LX-24) for 30 seconds to 5 minutes depending on the intensity of the signal. The film was rinsed in water to remove developer and fixed in Kodak X-ray fixer (FX-40) for 4 minutes. The film was washed thoroughly under running tap water and air dried.

For northern blots, the film was pre-flashed to increase sensitivity using standard techniques (Laskey and Mills, 1977; Sambrook *et al.*, 1989).

2.20 Re-using Hybond-N membranes

After removal of the DNA probe, the Hybond-N membranes could be re-probed following the same protocol. For Southern blots the probe DNA was removed by washing the membrane in 0.4M sodium hydroxide solution at 65°C for 30 minutes, followed by a further 30 minute wash at 65°C in 0.1x SSC, 0.1% (w/v) SDS, 0.2M Tris-HCl pH 7.5. This procedure was repeated if residual radioactivity was still detected.

Northern blots were stripped by washing the membrane in 5mM Tris-HCl pH 7.5, 2mM EDTA pH 8.0, 0.1x Denhardt's Reagent at 65°C for 3 hours. Alternatively, the membrane was boiled in a 0.1% (w/v) SDS solution until all the radioactivity was removed.

2.21 Maintenance of the Lorist2 cosmid library

The cosmid library generated by Mellon *et al.* (1988) was maintained in *E. Coli* strain 1046 on 4 Pal/Biodyne 1.2µm nylon membranes (Gallenkamp). A second membrane was wetted on the surface of solid LB medium containing kanamycin and 5% glycerol, placed over the master and pressed down firmly with a glass plate to ensure even transfer of cells. This second membrane was the replica. The sandwich was stored as described by Little (1987).

2.22 Screening of the cosmid library

The replica membrane was peeled off the master and placed on the surface of solid LB medium containing kanamycin and incubated at 37°C for 6-8 hours until discrete colonies formed. The membrane was removed and placed colony side up onto Whatman Number 1 filter paper soaked in denaturation solution (1.5M sodium chloride, 0.5M sodium hydroxide) for 7 minutes. The membrane was then transferred onto two further filter circles soaked in neutralisation solution (3M sodium acetate pH5.5), for 3 minutes each. The membrane was then air dried and the DNA fixed to the membrane by baking at 80°C for 1 hour and UV cross-linking in a UV Stratalinker 1800 (Stratagene) on the autocrosslink setting. The membranes were then prehybridised, hybridised with appropriate radioactively labelled probes and washed according to the manufacturer's instructions.

2.23 Small-scale cosmid DNA preparation

A single colony was inoculated into 5ml LB kanamycin and incubated at 37°C overnight on a rolling drum. 3ml of overnight culture was pelleted by centrifugation at 13,000 rpm for 1 minute. The pellet was resuspended in 200µl cold solution I (25% (w/v) glucose, 50mM Tris-HCl pH8.0, 10mM EDTA pH8.0) to which approximately 1mg/ml lysozyme (Sigma L-7001) was added, mixed gently by inversion and left at room temperature for 5 minutes. 400µl freshly made solution II (0.2M sodium hydroxide, 1% (w/v) SDS) was added, mixed by inversion and kept on ice for 5-10 minutes. 200µl cold solution III (3M sodium acetate pH4.8) was added, mixed by inversion and kept on ice for 10 minutes. The precipitate, containing chromosomal DNA, was pelleted by centrifugation at 13,000 rpm for 15 minutes at room temperature and removed with a toothpick. The cosmid DNA was precipitated for 20 minutes at -20°C by adding 480µl propan-2-ol. DNA was pelleted by centrifugation at 13,000 rpm for 10 minutes at room temperature. The DNA pellet was washed with 70% ethanol and

air dried. The DNA pellet was dissolved in 50µl sterile distilled water by incubating at 55°C for 10 minutes.

2.24 DNA sequencing

2.24.1 DNAase I treatment of DNA fragments

The generation of small, randomly overlapping subclones from large (4-8kb) restriction fragments by partial digestion with DNAase I in the presence of manganese ions is described in Sambrook *et al.* (1989). This method was used to generate subclones of a size that could be sequenced using vector based primers. 5µg target DNA was added to a reaction mixture containing 10x DNAase/Mn^{++} digestion buffer (500mM Tris-HCl pH7.6, 100mM manganese chloride) and made up to 30µl final volume. 0.01-0.1U DNAase I (133.9U/µl, GibcoBRL) diluted in 1x DNAase/Mn^{++} buffer was added to the reaction mixture on ice and mixed. After 30, 60 and 90 seconds, aliquots of 10µl were removed and added to 5µl 50mM EDTA pH8.0 on ice. The digested DNA was size fractionated by electrophoresis on a 0.8% agarose gel. DNA of between 500bp and 1.0kb was excised and eluted from the gel, then resuspended in 23µl sterile water.

2.24.2 Repair and size selection of DNA

To the 23µl fragmented DNA, 3µl 0.5mM dNTPs (Pharmacia Biotech), 50mM magnesium chloride and 10U T4 DNA polymerase (5U/µl, Gibco-BRL) were added. The reaction was incubated for 15 minutes at room temperature, then 5U Klenow fragment of DNA polymerase (3.65U/µl, Gibco-BRL) were added and the reaction was incubated for a further 15 minutes at room temperature. The DNA was then purified by electrophoresis through a 0.8% agarose gel. Target DNA of 500bp-1.0kb was recovered

by elution and redissolved in 20µl sterile water. The DNA was added to Ready-To-Go pUC18 *Sma*I/BAP + Ligase (Pharmacia Biotech, 27-5266-01) and ligated overnight at 15°C according to the manufacturers instructions. 2µl of the ligation mixture was transformed into 100µl competent XL1-Blue cells and 50µl aliquots spread onto X-gal plates. The plates were incubated at 37°C overnight, the white colonies picked and plasmid DNA isolated as in Section 2.9.3.

2.24.3 Automated DNA sequencing

The template DNA for automated sequencing was prepared according to Section 2.7.3. The DNA was labelled using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer P/N402078) which uses the thermally stable Amplitaq DNA polymerase, FS. The cycling reactions were carried out in a Minicycler (Genetic Research Instrumentation Ltd.) according to the manufacturers instructions: rapid thermal ramp to 96°C; 30 seconds denaturing at 96°C; rapid thermal ramp to 50°C; 15 seconds annealing at 50°C; rapid thermal ramp to 60°C; 4 minutes extension at 60°C for 25 cycles. Samples were then purified following the manufacturers instructions and run on an ABI 373A automatic sequencer.

2.24.4 DNA sequencing primers

The oligonucleotides used as primers for DNA sequencing were synthesised by Perkin Elmer, Genosys Ltd and Pharmacia Biotech. The primer sequences are shown in Table 2.4.

2.24.5 DNA sequence analysis

All primary DNA sequence data were analysed using the BESTFIT, GAP, GELASSEMBLE, PILEUP and PRETTYBOX programs of the Wisconsin Package, Version 8.1-UNIX (Genetics Computer Group, Madison, Wisconsin, USA) and the Macintosh compatible software, GeneJockey II (Biosoft, Cambridge, UK). Database similarity searches were carried out using BLAST (Altschul *et al.,* 1990; Altschul *et al.*, 1997) at the URL (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/). Transmembrane helices were predicted from hydrophobicity profiles as calculated by the Kyte Doolittle algorithm (Kyte and Doolittle, 1982), the TMpred program at the URL (http://ulrec3.unil.ch/software/TMPRED_form.html), the Secondary Structure Prediction of Membrane Proteins program (URL http://www.tuat.ac.jp/~mitaku/sosui/), DAS - Transmembrane Prediction Server (http://www.biokemi.su.se/~server/DAS/).

Predicted proteins were searched for known motifs and patterns using the PROSITE database via the URL (http://expasy.hcuge.ch/sprot/prosite.html).

2.25 Amplification of DNA fragments using the Polymerase Chain Reaction (PCR)

Reactions for amplification of *rcb1* were set up using *Taq* DNA polymerase (GibcoBRL Life Technologies Inc.) according to the manufacturer's basic protocol. A reaction volume of 50 µl was used containing 1U of enzyme, 10% (v/v) 10xPCR Buffer (GibcoBRL Life Technologies Inc.) and approximately 500ng of target cosmid DNA. 25pmol of each primer were used. The $MgCl_2$ concentration used was 1.5mM and the dNTPs were used at a concentration of 0.2mM. 50µl of mineral oil was added to prevent evaporation of the components during the PCR cycle. The conditions for the reaction were as follows: 95°C initial denaturation (2 min), followed by 30 cycles of 95°C denaturation (2 min), 55°C annealing (2 min), 72°C extension (3 min). A final 10

minute incubation at 72°C was performed to complete the synthesis. The sequences of the primers, CP13 and REC13, are presented in Table 2.4.

2.25.1 Reverse transcription PCR (RT-PCR)

The cDNA sequences of *rcb1*, *rcb2* and *rcb3* were amplified from total RNA isolated from JH2 (*A42B42/B6*) using the Access PCR system (Promega). Amplification reactions were set up in 50µl according to the manufacturers protocol. Between 500ng and 1µg template RNA was used. The conditions for the reactions were as follows: 48°C reverse transcription (60 min), 94°C denaturation (2 min) followed by 40 cycles of 94°C denaturation (30 sec), 55°C annealing (1 min) and 68°C extension (2 min). A final 10 minute incubation at 68°C was performed to complete the synthesis. The sequences of the primers used are in Table 2.4.

2.25.2 Long PCR of *B3*L and *B3*R

The *B3*L and *B3*R fragments were amplified from LCO12 (*A2B3*) genomic DNA using the GeneAmp XL-PCR Kit (Perkin-Elmer). Reactions were set up according to the manufacturers recommendations. A reaction volume of 100µl was used, containing 2U r*Tth* DNA Polymerase XL, 33% (v/v) 3.3xXLBuffer (Perkin-Elmer) and approximately 50ng target DNA. The $Mg(OAc)_2$ concentration used was 1.3mM, 40pmol primers were used and the dNTPs were used at a concentration of 0.2mM. 50µl of mineral oil was then added to prevent evaporation during cycling. The following conditions for the PCR reactions were used for amplifying *B3*L: 94°C initial denaturation (1 min), followed by 35 cycles of 94°C denaturation (1 min) and 64°C annealing/extension (5min 30s). A final 10 minute incubation at 72°C was performed to complete the synthesis. The

reaction conditions used for amplification of B3R were identical except an annealing/extension temperature of 68°C was used. The sequences of the primers used are presented in Table 2.4.

2.25.3 Cloning of PCR amplification products

The PCR products were purified by agarose gel electrophoresis and band elution prior to ligation. The plasmid vectors pGEM-T and pGEM-T Easy (Promega) were used routinely for cloning PCR products. PCR products were ligated using T_4 DNA ligase in 10x ligase buffer (Promega) according to the manufacturers protocol. Ligations were carried out overnight at 4°C and following transformation into high efficiency JM109 *E. coli* cells (Promega), colonies containing vector with insert were identified using blue/white screening (Section 2.9.9)

Table 2.4 List of oligonucleotide primers used for sequencing and PCR amplification.

Name	Sequence (5' to 3')	Application
P1	GCGTTGATCCGGACAGCCCCTC	Sequencing the *B42* locus
P2	GGCACACTGCACAGCCCGCTG	Sequencing the *B42* locus
P3	GGCAAGTCACTGCCCTTTGGCCACC	Sequencing the *B42* locus
P4	CGATGAAGAGGTCCACAGTATCGAGGG	Sequencing the *B42* locus
P5	GCACGCTATTTAGCTATCGGCTTGAG	Sequencing the *B42* locus
P6	GGCCAGCTCCTTCTACAAGCCAC	Sequencing the *B42* locus
P7	CGATTGATGCGGAATGACTCGATGAC	Sequencing the *B42* locus
P8	GCTACGGGAATACGCGCCACG	Sequencing the *B42* locus
P10	CACTGCCCTTTGGCCACCTGCA	Sequencing the *B42* locus
P11	GTGTAGTGTGATCCCCAACGACAG	Sequencing the *B42* locus
P12	CCCAAAGAATGCGGGGGTAGGTG	Sequencing the *B42* locus
P14	GGAGGCGGCCTCACTTGGTTCT	Sequencing the *B42* locus
P15	CATCGGACGTAGGGGCAGCACAC	Sequencing the *B42* locus
P16	GGCCGAGTACAATAGGTCGGAGG	Sequencing the *B42* locus
P17	GAGTTCGAGCGTGGCCAGAGACAT	Sequencing the *B42* locus
P18	CAACATCGCGGAGTGAACCAGC	Sequencing the *B42* locus
PE8T3	GGCCTACAGGTAACCGACCGC	Sequencing the *B42* locus
PE8T7	GCTGAAGGCAAGTCACTGCCC	Sequencing the *B42* locus
PH48	GGGACGGCGGGTGATTGAATAGAAC	Sequencing the *B42* locus
PJ1	CGATAGGGGAGCGGGTGCAGAG	Sequencing the *B42* locus
PJ2	CTCCGGCGCTAAGCGGTCAGC	Sequencing the *B42* locus
PJ3	GATGATGGGAGATCGAGAGTGGC	Sequencing the *B42* locus
PJ4	CGACCCCTCCCGGGGAGGCG	Sequencing the *B42* locus
PB34	ATCGGGAATCGCCATTTGCGCC	Sequencing the *B3* locus
PB35	TCCAAGAGAGTGGCGAGATCGAC	Sequencing the *B3* locus
PB36	CGATAGGGGAGCGGGTGCAGA	Sequencing the *B3* locus
PJR1	TGCGCTTTCCTCGCTGCCATCCTA	RT-PCR of *rcb2*
PJR3	GACGTTACTCGTGAATGCACTCG	RT-PCR of *rcb2*
PFR3	TCCTCGTCACGAAAGCTGTCCTTC	RT-PCR of *rcb3*
RTREC33	GACTTTTGAGTGTGCGTTAGAGTCG	RT-PCR of *rcb3*
CP13	CCCCGACGGCCTTGTACTGTAGC	Amplification of *rcb1*
REC13	GGTCCGGGAGCAGAGCGAG	Amplification of *rcb1*
PB31	GGCAGGTCTAAAGGTAGCCACG	Amplification of *B3*L
P9	CCACAAGGGCAGTGACTTGCCTTC	Sequencing *B42* locus & PCR of *B3*L
PB32	ATCGTGGGAAACGCAGATGGCG	Amplification of *B3* R
PB33	CGTTCTGTTCGCCCAGTTGTGAC	Amplification of *B3* R
PB3L	GGATAGGATAGGCACAGACGAGG	Recombination experiment
PB3R	GTTGCCAGGTCTTCCACCATTCG	Recombination experiment
PB42L	ACGAGCACAGACCGACCAGTAAG	Recombination experiment
PB42R	CAGCAAGCTATTGGCTCGCGGAA	Recombination experiment
PYAN5	CGTTCTGACGTTCTTTGGTATGTC	RT-PCR of *mfs1*
PYAN3	AGATGAAGAACATGGAGTGATCAGG	RT-PCR of *mfs1*

Chapter 3

Isolation of the *B42* mating type locus of *C. cinereus*

3.1 Introduction

The *B6* mating type locus of *Coprinus cinereus* was isolated by O'Shea (1995). Sequences unique to the genomic DNA of a *B6* strain were recovered following hybridisation to excess genomic DNA of a *B3* strain. Using Southern blot analyses it was possible to define a 23kb DNA sequence present in the *B6* genome but absent from the *B3* genome. The objective of this study was to isolate a second *B* locus having a different specificity from *B6*, in order to learn more about the organisation of the *B* locus and about the basis of allelic variability between different *B* genes. The *B42* locus was chosen since clones containing this were already available in a genomic library constructed from strain JV6 (*A42B42*) in the cosmid vector Lorist2 (Mellon *et al.*, 1988). Lorist2 contains a *Hin*dIII cloning site and the library was generated from genomic DNA partially digested with *Hin*dIII. This is particularly useful when analysing cosmid clones because *Hin*dIII digestion releases the vector and several genomic DNA fragments that have not been altered in size by the cloning procedure.

3.2 Initial screen of the cosmid library

A 5.7kb *Hin*dIII-*Eco*RI fragment containing sequences from a homologous flanking region at the right of the *B6* locus (Figure 3.1) was used to probe filters containing more than 11,000 primary clones of the Lorist2 library. 36 hybridising clones were identified

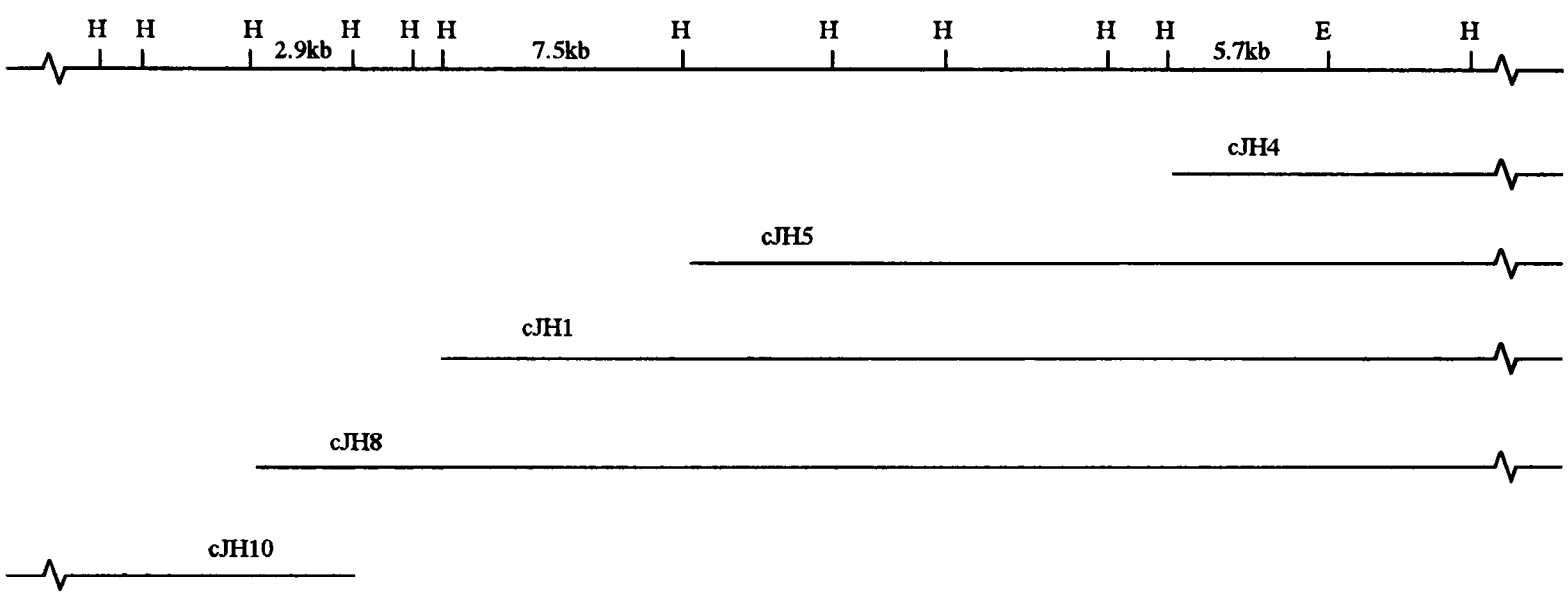


Figure 3.1 Map of the *B42* locus showing the overlapping sequences present in cosmid clones isolated in the genomic library screen. E, *Eco*RI; H, *Hin*dIII.

in the first round screen and of these, seven were purified by a second round screen and five chosen for an initial molecular analysis. Cosmid DNAs were cut with *Hin*dIII, Southern blotted and probed with the 5.7kb *Hin*dIII-*Eco*RI fragment used to screen the library for the clones. The probe was derived from an 11kb *Hin*dIII fragment and this was identified in four of the five cosmid DNAs purified (cJH1, cJH3, cJH4 and cJH5). The DNAs from cJH1, cJH3 and cJH5 produced similar sized fragments when digested with *Hin*dIII indicating that they contained largely the same sequence, whereas cJH4 contained several fragments not present in the other clones (Figure 3.2).

To determine which clones contained DNA sequences that were *B42*-specific, and therefore likely to contain genes of the *B42* locus, cJH1, cJH4 and cJH5 were used to probe *Hin*dIII digested genomic DNAs from a *B42* and a *B6* strain (Figure 3.3). The entire sequence present in cJH4 hybridised to both genomic DNAs indicating that this clone might contain only the homologous region flanking the *B42* locus whereas only five of the fragments derived from cJH1 and four from cJH5 hybridised to *B6* DNA; the rest covering 15kb were *B42*-specific.

Cosmid cJH1 was used to probe a filter containing *Hin*dIII digested cJH1, cJH3, cJH4 and cJH5 DNAs in order to determine whether the similar sized fragments generated were identical. Six of the fragments present in the cJH1 digest were present in other clones, but interestingly, although digestion of all four clones generated a 7.5kb fragment, the fragment in cJH1 failed to hybridise to the fragment present in the other three clones (Figure 3.2B). By looking at the Southern blot analysis of *Hin*dIII cut genomic DNA from a *B42* strain and a *B6* strain presented in Figure 3.3, it can be seen that the fragment present in the cJH1 clone is *B42*-specific (Figure 3.3A). However the fragment present in cJH4 and cJH5 hybridised to a fragment present in both backgrounds (Figure 3.3B and C), and was therefore derived from a homologous sequence flanking the *B* locus. It

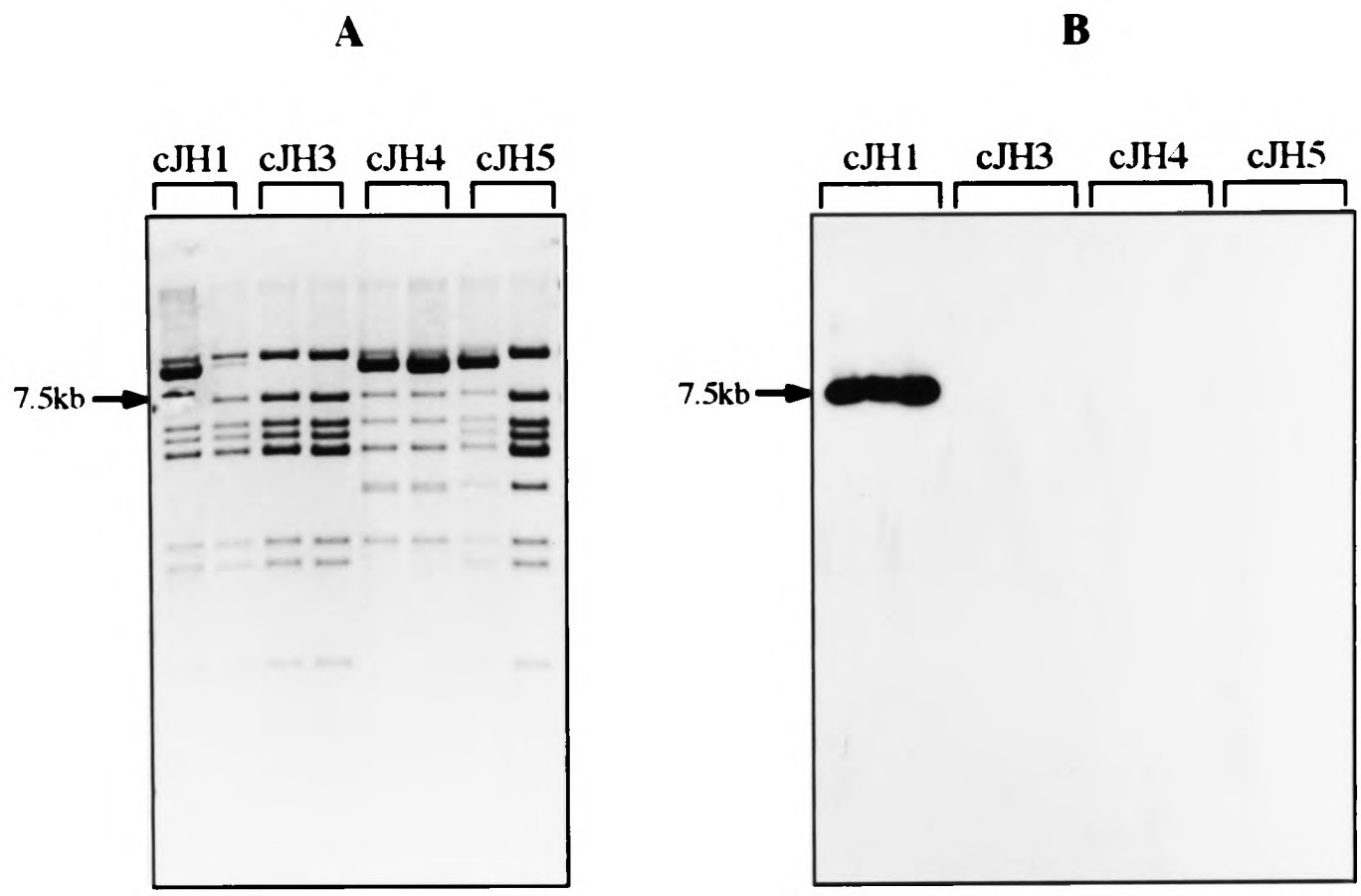


Figure 3.2 (A) Ethidium bromide stained agarose gel of *Hin*dIII digested DNAs from cosmids cJH1, cJH3, cJH4 and cJH5. The two representative DNAs of each cosmid were isolated from a single clone **(B)** Hybridisation of the 7.5kb *Hin*dIII fragment isolated from cJH1, to *Hin*dIII digested DNAs from cosmids cJH1, cJH3, cJH4, cJH5. Each cosmid is represented on the blot by three different DNA preparations from a single clone.

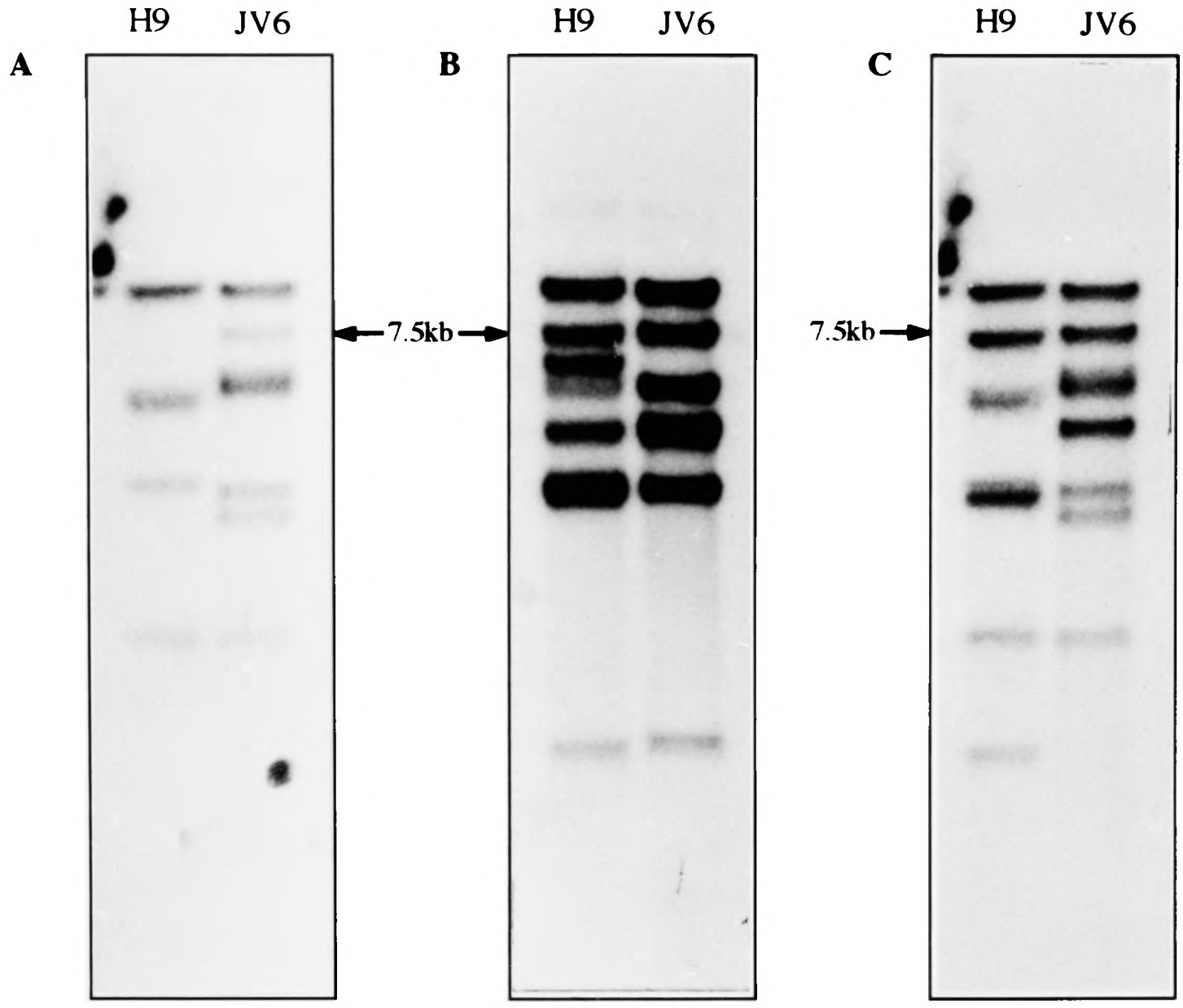


Figure 3.3 Southern blots of *Hin*dIII digested genomic DNA from H9 (*B6*) and JV6 (*B42*) strains probed with (**A**) cJH1, (**B**) cJH4 and (**C**) cJH5.

was evident that the 7.5kb fragment from cJH1 represented the sequence furthest from the homologous flanking region of the *B* locus and could therefore be used to probe the genomic library for new clones, containing sequences extending further into the *B42* locus.

3.3 Isolation of overlapping clones covering the entire *B42* locus

Screening the library with the 7.5kb *Hin*dIII fragment from cJH1 identified several new clones two of which, cJH7 and cJH8, were purified and shown to contain the 7.5kb *Hin*dIII fragment. *Hin*dIII digestion and Southern blotting showed that cJH7 contained exactly the same sequences as cJH1, but that cJH8 contained three new fragments of 2.9kb, 1.8kb and 0.8kb thus extending the isolated sequence by 5.5kb (Figure 3.4). The unique 2.9kb fragment from cJH8 was used to screen the genomic library a further time and a new clone, cJH10, was obtained that contained the 2.9kb probe sequence but shared no other sequences with the cosmids isolated previously (Figure 3.4). cJH10 provided a further 40kb of sequence to the 50kb covered by cJH4, cJH1 and cJH8 giving in total some 90kb of contiguous sequence, more than sufficient to contain the entire *B42* locus (Figure 3.1).

3.4 Testing for *B42* gene activity by transformation

Cosmids that were isolated during library screening were assayed for *B42* gene activity by transformation. Lorist2 vector does not contain a marker than can be selected in *C. cinereus* and cosmid DNA sequences were co-transformed into host strains using the strategy illustrated in Figure 3.5. For hosts like LT2, which have a *trp-1* (tryptophan synthetase defective) mutation , co-transformation was carried out with plasmid pCc1001 containing the *C. cinereus trp-1* gene, but for the *B3* host LCO12, which is a *trp-3*

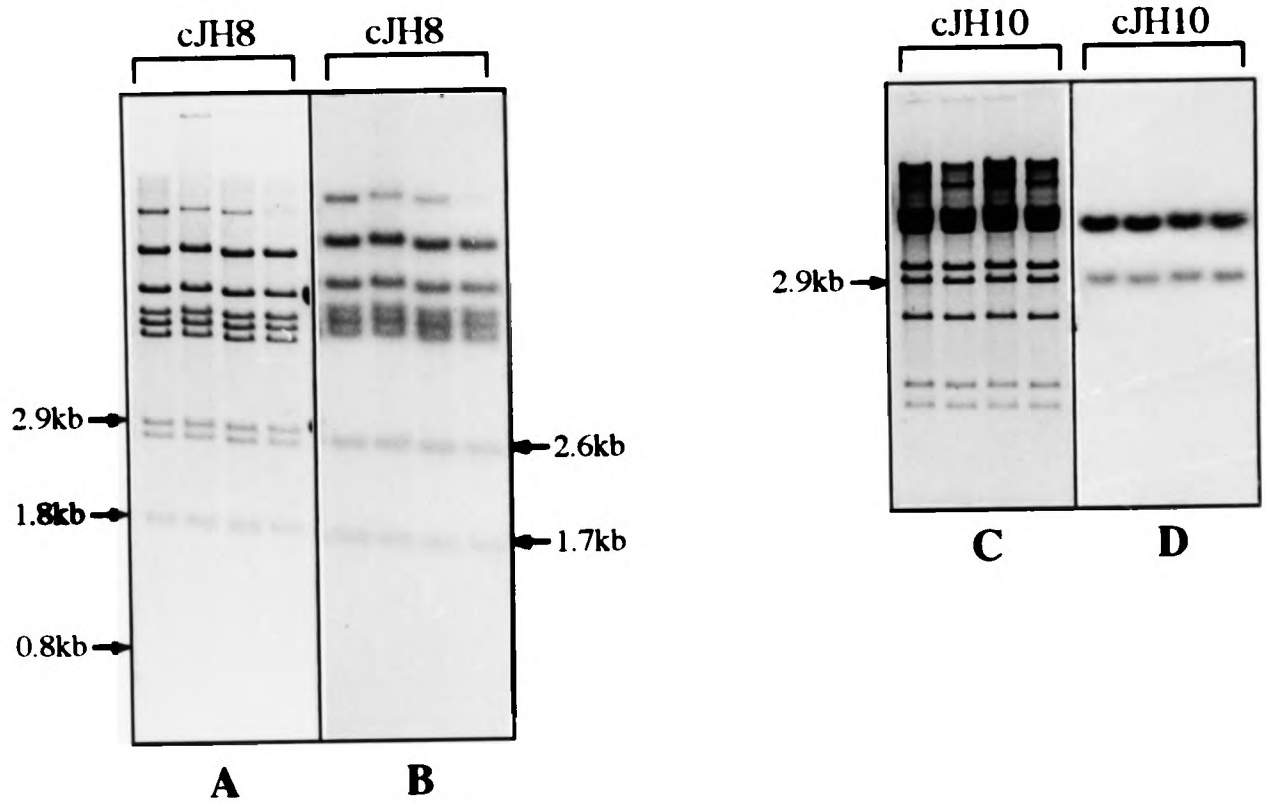


Figure 3.4 (**A**) Ethidium bromide stained agarose gel of *Hin*dIII digested DNA from cosmid cJH8 and (**B**) the corresponding Southern blot probed with cosmid cJH1; (**C**) ethidium bromide stained agarose gel of *Hin*dIII digested DNA from cosmid cJH10 and (**D**) the corresponding Southern blot probed with cosmid cJH8. In both cases the cosmids were represented on the blot by four different DNA preparations from a single clone.

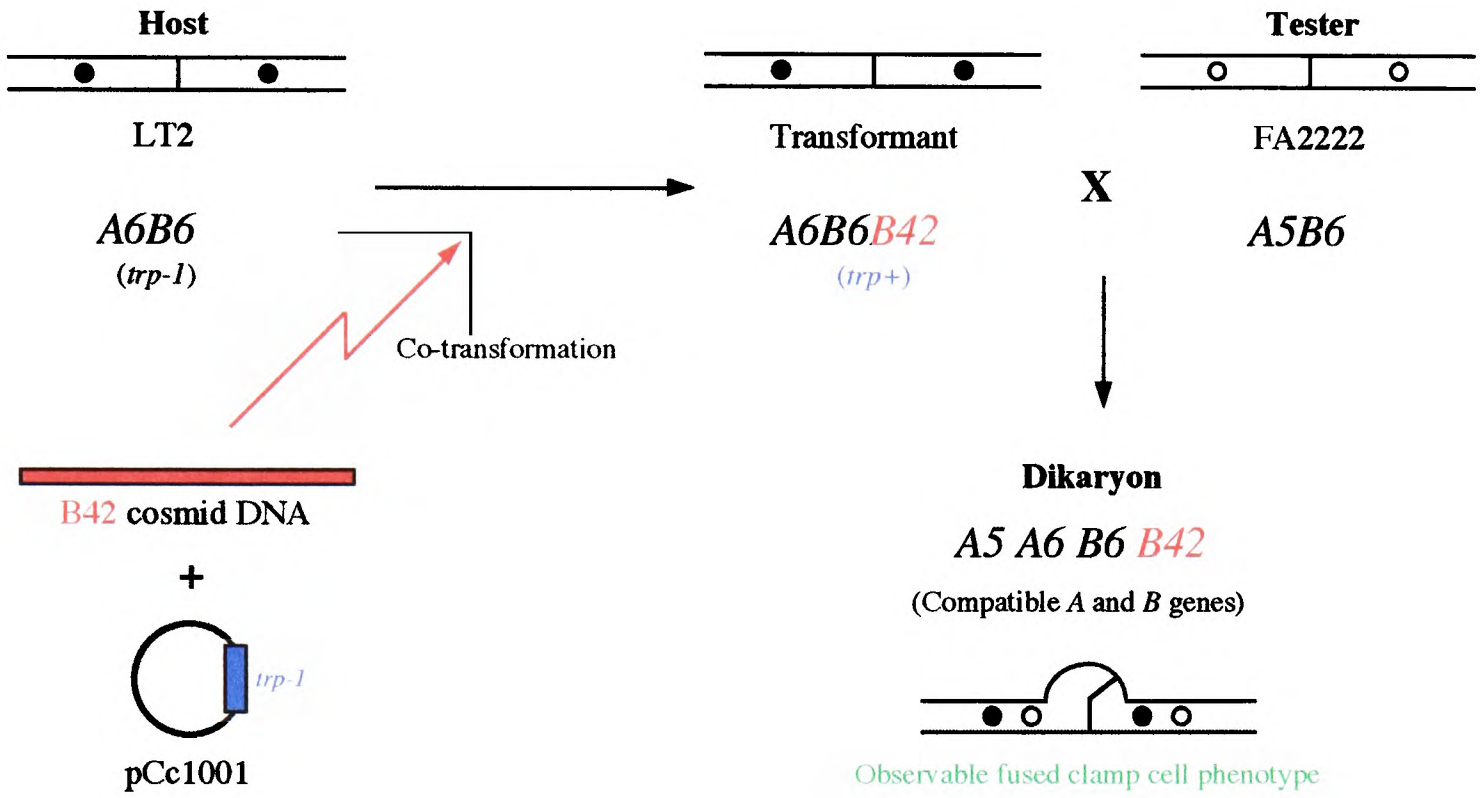


Figure 3.5 Schematic diagram illustrating the co-transformation technique used to identify DNA carrying *B42* specificity genes.

(anthramlate synthetase) mutant, pDB6 containing the *C. cinereus trp-3* gene was used. There is no observable phenotype for a compatible *B* gene interaction within the transformed cells, and a two-step assay was used to identify those transformants that expressed an introduced *B* gene.

Introduction of a *B42* gene into a host strain having *B6* genes will change the mating specificity of the host and confer the ability to form a dikaryon with a tester strain having the same *B6* genes, *i.e.* an *A6B6* untransformed host will be unable to form a dikaryon with a tester strain that is *A5B6* and the two monokaryons will remain discrete when grown together (Figure 3.6A). If the host is transformed with *B42* genes to give the genotype *A6B6/B42*, transformants will be compatible with the *A5B6* tester strain and be able to form a dikaryon (Figure 3.6B), which appears as faster and less dense growth at the margin of the mated colonies. Under the microscope, the dikaryon is distinguished by the fused clamp connections at each cross wall (Figure 3.6C and D)

Five of the cosmid clones, containing sequences spanning the entire region isolated, were introduced into host strains having *B6*, *B3* and *B42* genes. The results are recorded in Table 3.1. Since the genes were introduced by co-transformation, not all transformants tested were expected to express the cloned genes. Co-transformation frequencies vary, but at least 10% of transformants would be expected to express the unselected genes (Casselton and de la Fuente Herce, 1989). Up to 50 transformants were tested for *B42* gene activity. If none of these formed dikaryons, it was concluded that the introduced sequences could not confer a change in mating specificity.

None of the cosmid DNAs altered the mating specificity of the B42 strain which was included as the negative control. cJH1, cJH5 and cJH8 contained sequences that altered the mating specifcity of both the *B6* and the *B3* hosts and these positive mating reactions

Table 3.1 Results of co-transformation experiments with cosmids cJH1, cJH4, cJH5, cJH8 and cJH10. (+) represents a change in *B* mating specificity of the host strain in 10-80% of transformants, (-) represents no change in host *B* mating specificity.

	Host strain specificity		
Cosmid	*B6*	*B3*	*B42*
cJH1	+	+	-
cJH4	-	-	-
cJH5	+	+	-
cJH8	+	+	-
cJH10	-	-	-

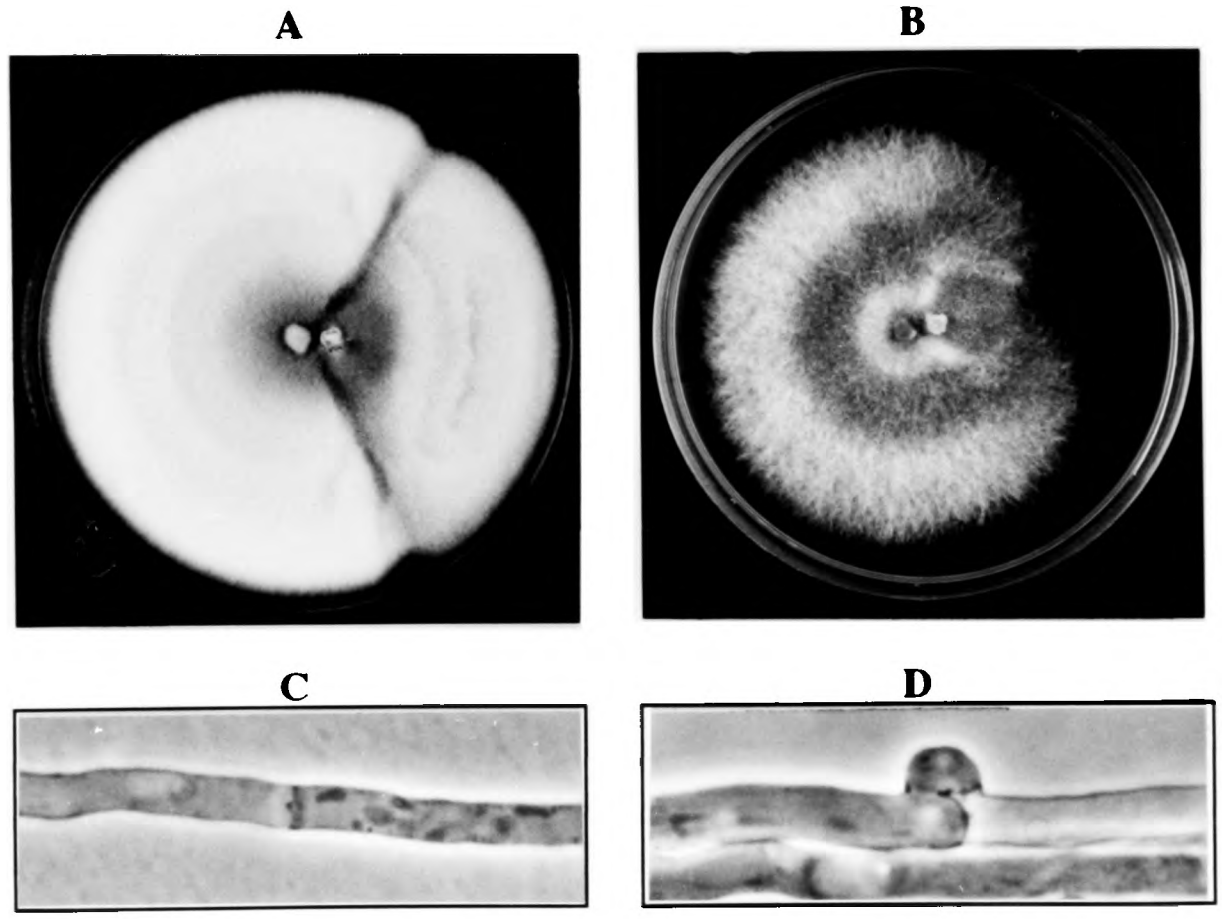


Figure 3.6 Photographs showing (**A**) an incompatible mating between the untransformed host and tester strain; (**C**) the corresponding lack of clamped septa in the hyphae; (**B**) a compatible mating reaction between the transformed host and tester strain; (**D**) the characteristic fused clamp cell of the dikaryotic hypha.

confirmed that sequences from the *B42* locus had been cloned. The sequences present in both cJH4 and cJH10 failed to change the mating specificity of either the *B6* or the *B3* host and either contained genes that were inactive in these two backgrounds or lacked *B42* genes altogether.

3.5 Mapping the cloned sequences and defining the borders of the *B42* locus

At this stage it was necessary to construct a complete physical map of the *B42* sequences present in the isolated cosmids and to define the borders of the *B42*-unique DNA which constitutes the *B42* locus. The position of the homologous flanking sequences to the left and right of the *B42* locus were established by Southern blot analyses which are described below. The restriction map of *B42* presented in Figure 3.7 is based on the restriction enzymes *Hin*dIII, *Bam*HI, *Eco*RI, *Pst*I and *Sal*I. The entire sequence to be mapped was contained in two cosmid clones, cJH8 and cJH10. Single and double digests of these two clones were probed with labelled *Hin*dIII fragments derived from both cosmid sequences and indicated below the map. The results of hybridisations with cJH8 fragments as probes are presented in Figure 3.8 and those with cJH10 fragments as probes in Figure 3.9.

Comparison of *B6* and *B3* genomic DNAs made it possible to identify a region covering 17kb that was unique to *B6* and within which all the *B6* genes were located (O'Shea, 1995; O'Shea *et al.*, 1998). A similar hybridisation analysis was made using *Hin*dIII fragments from the clones isolated in this study to probe genomic DNAs from *B6*, *B3* and *B42* strains. The results of this analysis are summarised in Figure 3.10.

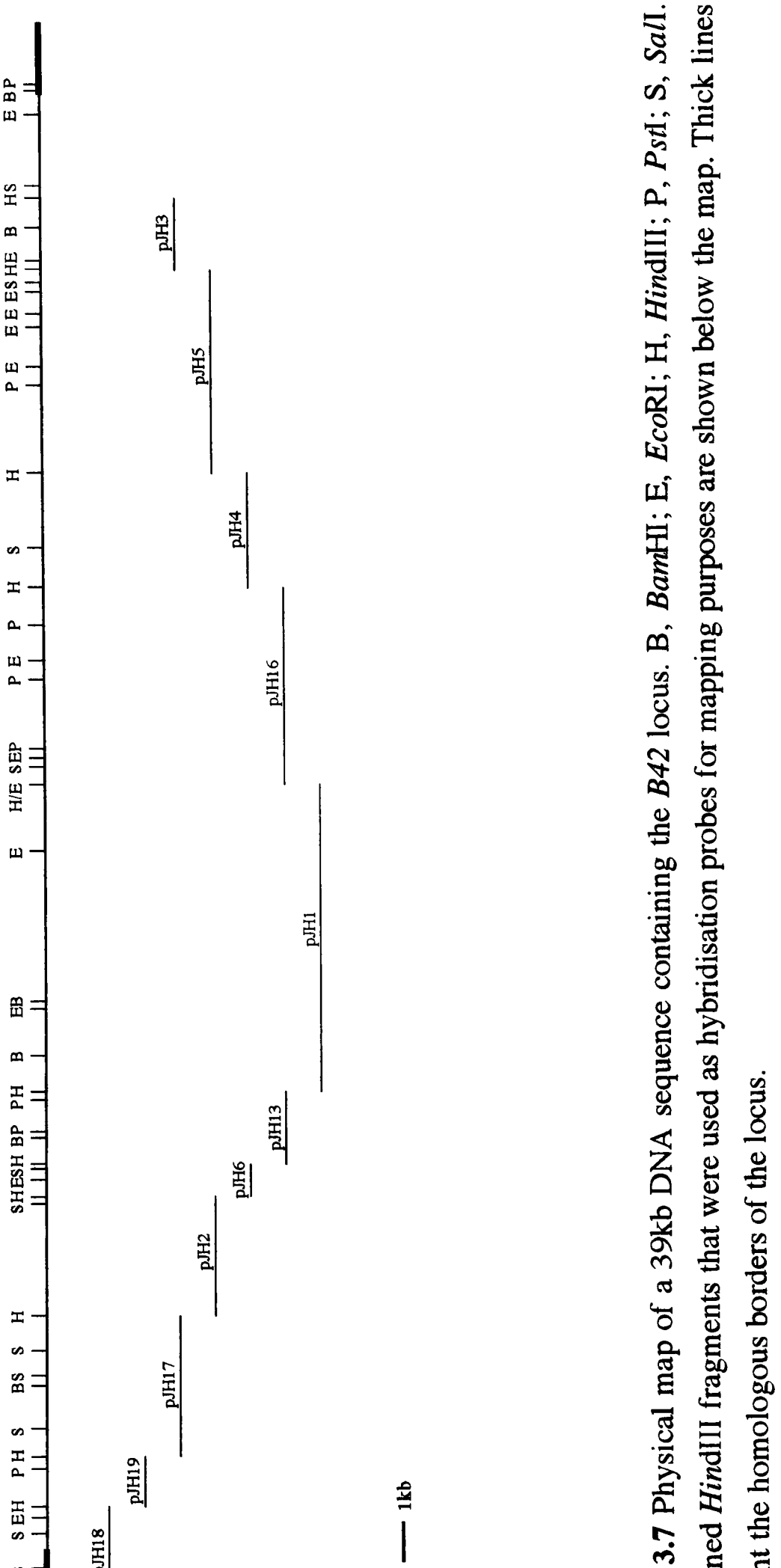
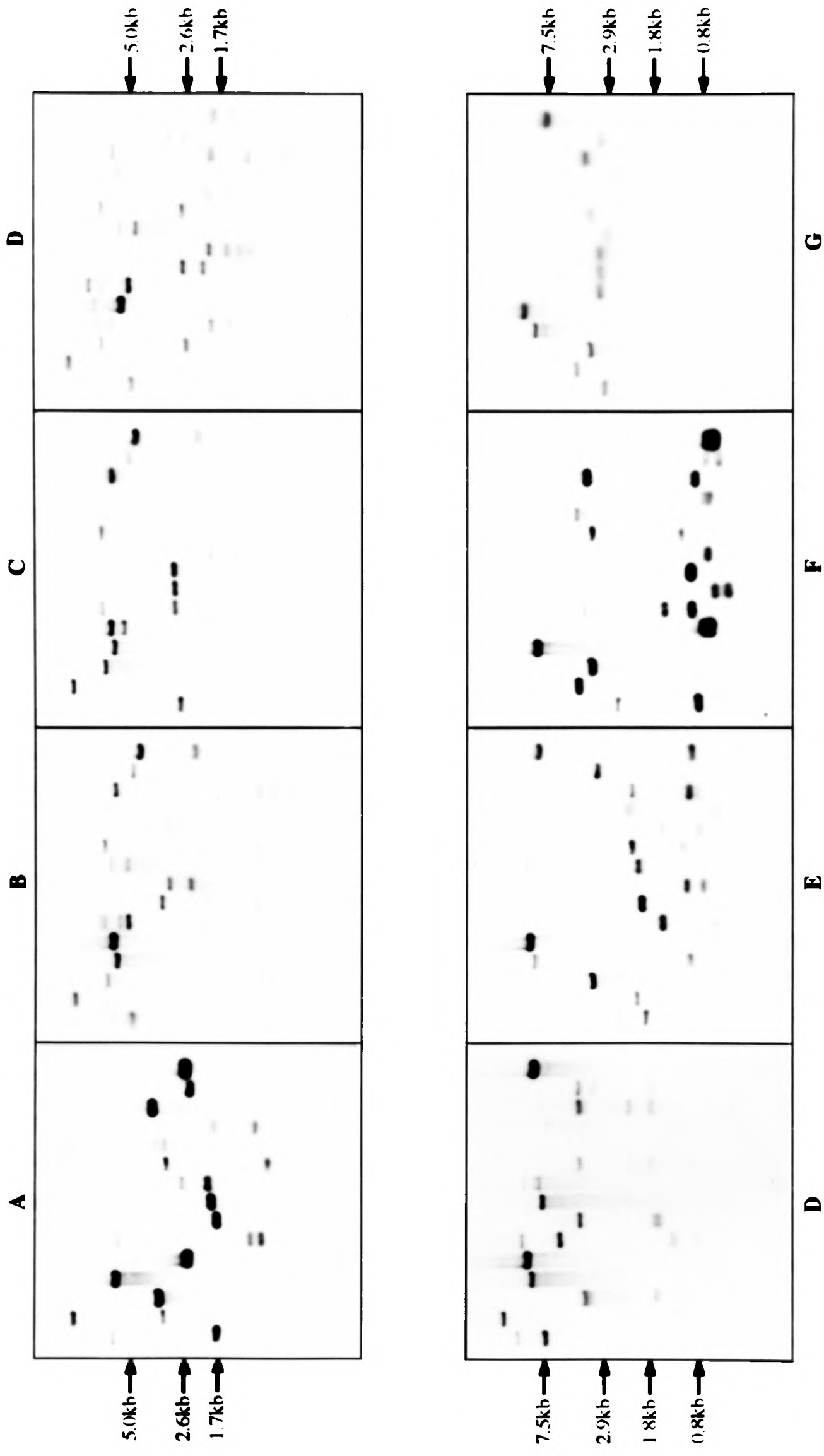


Figure 3.7 Physical map of a 39kb DNA sequence containing the *B42* locus. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I. The cloned *Hin*dIII fragments that were used as hybridisation probes for mapping purposes are shown below the map. Thick lines represent the homologous borders of the locus.

Figure 3.8 Southern blot of cJH8 DNA probed with the cloned sequences in (**A**) pJH3, (**B**) pJH5, (**C**) pJH4, (**D**) pJH16, (**E**) pJH1, (**F**) pJH13, (**G**) pJH6 and (**H**) pJH2. The cJH8 DNA is digested with (from left to right) *Hin*dIII (H), *Bam*HI (B), *Eco*RI (E), *Pst*I (P), *Sal*I (S), H/B, H/E, H/P, H/S, B/E, B/P, B/S, E/P, E/S and P/S.


5.0kb
2.6kb
1.7kb
7.5kb
2.9kb
1.8kb
0.8kb
A
B
C
D
E
F
G

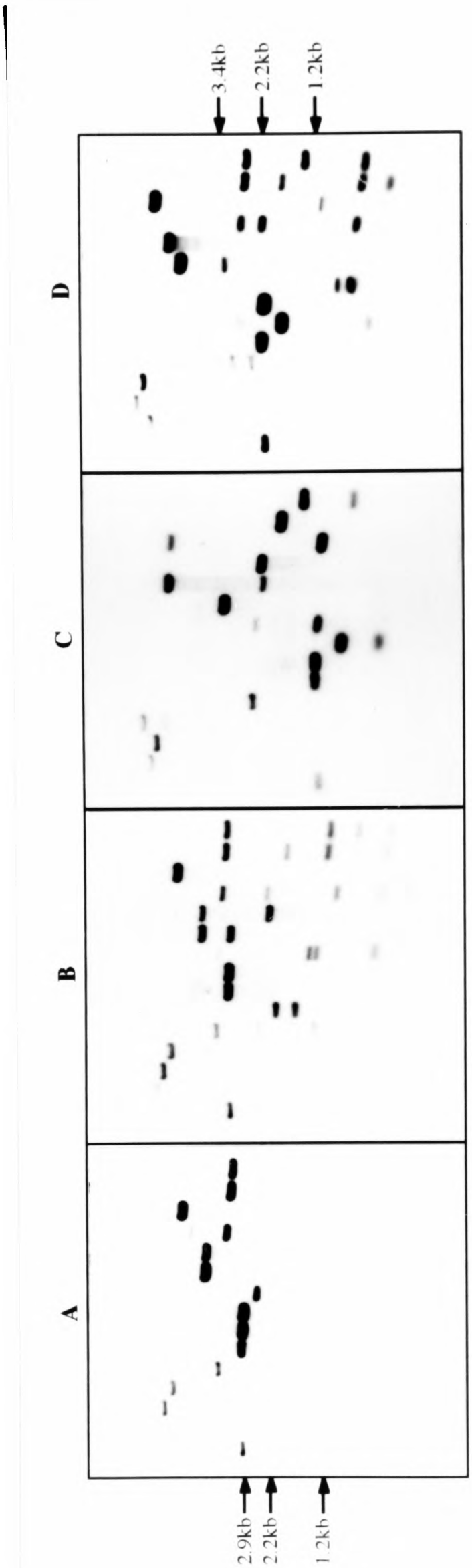


Figure 3.9 Southern blot of cJH10 DNA probed with the cloned sequences in (**A**) pJH2, (**B**) pJH17, (**C**) pJH19 and (**D**) pJH18. The cJH10 DNA is digested with (from left to right) *Hin*dIII (H), *Bam*HI (B), *Eco*RI (E), *Pst*I (P), *Sal*I (S), H/B, H/E, H/P, H/S, B/E, B/P, B/S, E/P, E/S and P/S.

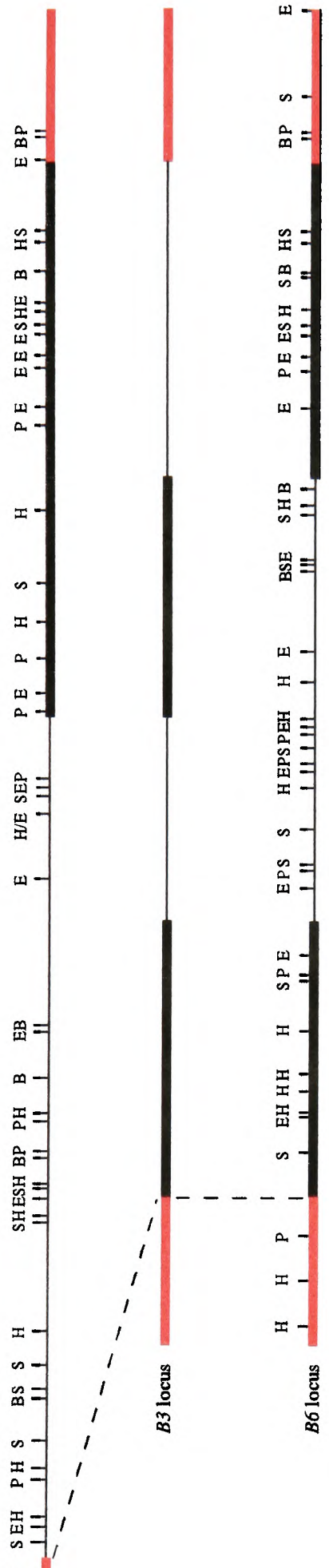


Figure 3.10 Diagram to show DNA sequences that are shared in genomic DNAs from *B42*, *B6* and *B3* strains. Restriction maps of the regions containing the *B42* and *B6* loci are included. Restriction enzymes are as shown in Figure 3.9. Thin lines represent sequences unique to a particular *B* locus, thick lines represent sequences shared by two *B* loci. Thick lines shown in red correspond to the homologous flanking sequences shared by all three genomic DNAs.

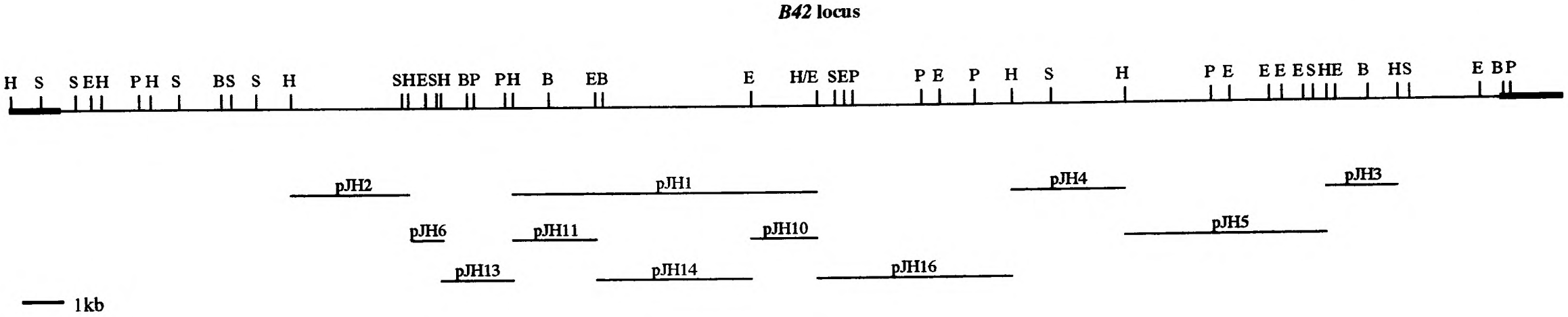
The pattern of cross-hybridisation indicates that different *B* loci may share some sequences but not others. *B6* and *B3* share no sequences and the homologous borders of the *B* locus were defined by comparing these sequences (shown in red in Figure 3.10). When *B42* was compared to *B6* and *B3*, it was found that the 7kb identified as the right end of the *B6* locus (thick black line) was also present in *B42* whereas the 7kb of DNA in the middle of the *B42* locus was present in *B3* (thick black line). The region constituting the left of the *B6* locus was different in all three genomic DNAs. The *B3* and *B6* loci share a common sequence at the right of the map that contains no genes with *B* mating specificity (O'Shea *et al.*, 1998). There is *B42*-specific sequence at the left of the locus which covers some 17kb. Transformation analysis with cosmid cJH10 suggests that this region contains no genes conferring *B42* mating specificity.

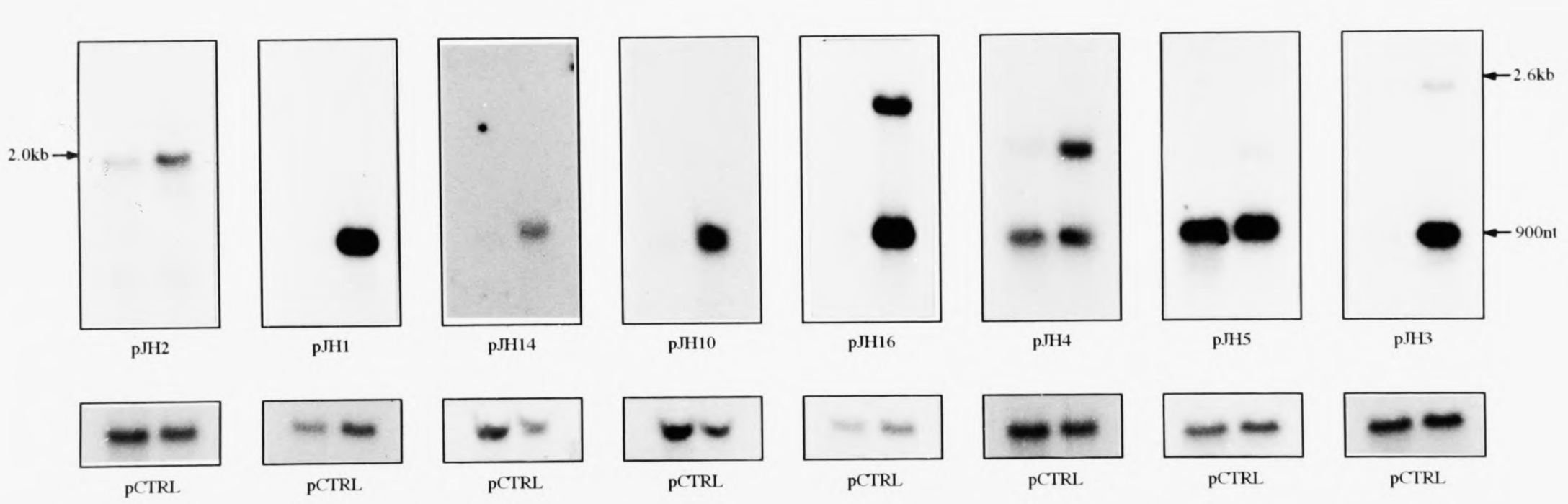
The fact that the *B42* locus shares sequences with both *B6* and *B3* suggested that the different *B* mating specificities might be derived from several functionally redundant genes. In the *A* locus there are three sets of multiallelic and functionally redundant genes which can be arranged in different combinations to generate large numbers of *A* mating specificities (Pardo *et al.*, 1996). A similar organisation for the *B* locus would permit the 79 predicted versions (Raper, 1966) to be generated in a similar fashion. The next objective was to identify the individual genes within *B42* and compare this organisation with genes present in the *B6* locus.

3.6 Preliminary transcript analysis to identify *B42* genes

A first step towards locating the *B42* mating type genes was to look for transcripts derived from the *B42*-specific sequence. Previous work by Kingsnorth (1996) established that the transcripts of most of the *B* genes are barely detectable in mRNA extracted from monokaryotic cells but are induced to relatively high levels in cells which

Figure 3.11 Transcript analysis of the *B42* mating type locus. Cloned sequences used as probes for hybridisation are shown below the map. The autoradiographs are northern blots of poly $(A)^+$ RNA subjected to formaldehyde gel electrophoresis. The left lane is RNA isolated from a monokaryotic strain, LN118 (*A42B42*). The right lane is RNA isolated from an *A*-on strain, JH1 (*A5/A42B42*), in which the *A*-regulated steps of development had been activated by the introduction of a compatible *A* locus. The same blots probed with a constitutively expressed gene cloned in plasmid pCTRL (Kingsnorth, 1996) are also shown. Restriction enzymes are detailed as in Figure 3.9.


B42 locus
H S SEH PH S BS S H SHESH BP PH B EB E H/E SEP P E P H S H P E E E S H E B HS E B P
pJH2
pJH1
pJH4
pJH3
pJH6
pJH11
pJH10
pJH5
pJH13
pJH14
pJH16
1kb


2.0kb
2.6kb
900nt
pJH2
pJH1
pJH14
pJH10
pJH16
pJH4
pJH5
pJH3
pCTRL
pCTRL
pCTRL
pCTRL
pCTRL
pCTRL
pCTRL
pCTRL

have the *A* gene pathway activated (*A*-on). In order to look for *B42* gene transcripts, RNA was isolated from a *B42* monokaryon (LN118, *A42B42*) and the same strain into which a cosmid containing the entire *A5* mating type locus had been introduced (JH1, *A42/A5B42*).

Northern blots containing RNA from both strains was probed with a series of fragments covering most of the *B42*-specific sequence. The map and fragments chosen are illustrated in Figure 3.11. Transcripts detected were either between 2.1 and 2.5kb or 800-900nt. Most of the transcripts could only be detected well in RNA derived from *A*-on mycelium though some of the small transcripts were detectable in both RNAs.

Without knowing the functions of the genes, this preliminary analysis was used to detect the location of active genes and not to study in detail how they might be regulated. This is considered in much more detail in Chapter 5. As can be seen from the northern blots illustrated in Figure 3.11, at least three large transcripts are derived from the *B42*-specific sequence and four small transcripts. Interestingly, one of the large transcripts derives from a region covered in cJH10 that does not confer *B42* mating specificity and which is separated at some distance from sequences giving rise to the other transcripts. This analysis served to demonstrate that genes that confer *B42* mating type specificity are concentrated towards the right side of the *B42*-specific DNA sequence and that this would be where to initiate the DNA sequence analysis.

Chapter 4

Sequence analysis of the *B42* mating type locus

4.1 Introduction

This chapter describes the identification by DNA-sequence analysis of nine genes which determine *B42* mating specificity. As described in the introduction, sequence analysis of the *Bα1* locus of *Schizophyllum commune* provided the first evidence that the multiallelic *B* genes of the mushroom fungi encode peptide pheromones and transmembrane receptors. These are characteristic elements of a signalling pathway essential for mating cell recognition and cell fusion in the ascomycetous yeasts, *Saccharomyces cerevisiae* (Kurjan, 1993) and *Schizosaccharomyces pombe* (Nielsen and Davey, 1995) and the yeast-like basidiomycete, *Ustilago maydis* (Bölker et al., 1992). In the mushroom fungi this mating type function had not previously been predicted because mating cell fusion is mating type-independent.

It is now evident that the *a* locus of *U. maydis* is equivalent to the *B* mating type loci of the mushroom fungi *C. cinereus* and *S. commune*. However, the *a* locus has only two alleles, *a1* and *a2*, and each contains just a single gene encoding a pheromone and a single gene encoding its corresponding transmembrane receptor. In contrast, the *Bα1* locus of *S. commune* (Wendland et al., 1995) and the more recently described *Bβ1* locus, (Vaillancourt et al., 1997) each contain a gene encoding a receptor protein and three genes encoding functionally redundant pheromones. There are moreover, nine alleles of each *B* locus. Little is yet known about the basis for allelic variability between receptor and pheromone genes and the mushroom fungi may well provide clues as to

how large receptor and pheromone families have evolved in mammals (Dulac and Axel, 1995; Troemel et al., 1995).

In parallel with the sequence analysis of the *B42* locus, the sequence of the *B6* locus of *C. cinereus* was determined by (O'Shea et al., 1998). This has made it possible to compare the organisation of genes within the two loci and to make predictions about which genes are functional alleles and how multiple *B* mating types may be generated by three sets of functionally independent genes.

4.2 Sequencing strategy

Based on the transformation data and preliminary trasncript analysis described in the previous chapter, it seemed likely that the *B42* genes would be clustered towards the right of the *B42* specific sequence, as illustrated in Figure 3.11. The aim was to identify genes as far as possible by DNA sequence analysis.In order to do this, a series of overlapping fragments covering the predicted region of function were chosen as illustrated in Figure 4.1. The majority of the sequence was obtained by generating a series of random overlapping clones by partial DNAase I digestion. Fragments of 500bp to 1.0kb were size selected and cloned into *Sma*I cut pUC18. Cloned fragments were then sequenced using two general primers that anneal to the flanking vector sequences. The random overlapping sequences obtained were aligned using the GELASSEMBLE program (GCG, version 8.1). Any gaps in the sequence were covered using custom made oligonucleotide primers.

The sequencing program was initiated by taking the 7.5kb *Eco*RI fragment present in pJH24 followed by the 4.8kb *Hin*dIII fragment in pJH16 (Figure 4.1). In total, data from 156 clones provided some 11kb of DNA sequence. The region to the right of that

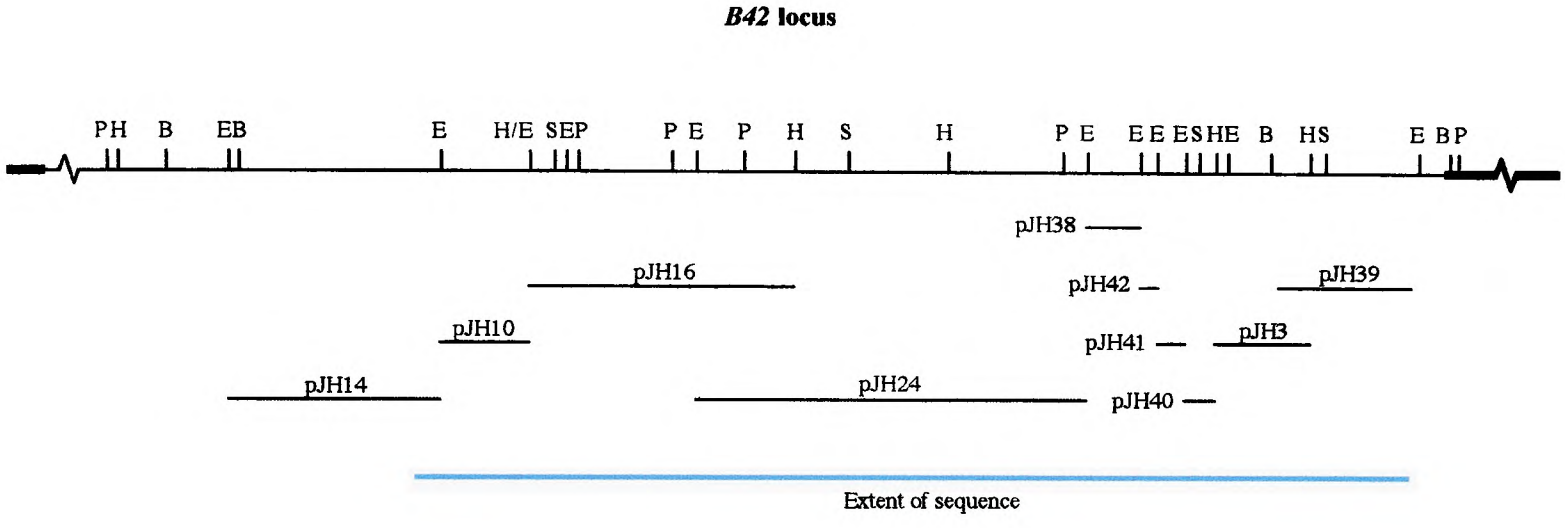


Figure 4.1 Strategy for sequencing the *B42* locus. The clones used for sequencing are shown below the restriction map. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I. The homologous borders of the locus are represented by thick lines. The blue line indicates the region for which the complete DNA sequence was obtained.

present in pJH24 was shown by hybridisation to be shared by the *B6* locus. This region was sequenced using a combination of smaller fragments and primers already available from sequencing the *B6* locus (O'Shea, 1995). A region to the left of the map was covered by sequencing a 1.4kb *Hin*dIII-*Eco*RI fragment in pJH10 and part of a 3.9kb *Eco*RI fragment from pJH14. A contiguous sequence covering in total 18kb of genomic DNA was obtained and individual genes within this were identified using a combination of BLAST database searches (Altschul et al., 1990), transcript location and transformation. A comparative map of the *B42* and *B6* loci is presented in Figure 4.2.

4.3 Identification of six putative pheromone precursor genes

An important observation derived from sequence analysis of the *U. maydis* and *S. commune a* and *B* loci was that the pheromone precursor genes encoded proteins with a carboxy-terminal CaaX motif (Bölker et al., 1992; Wendland et al., 1995; Vaillancourt et al., 1997). A search for putative open reading frames that might encode pheromone precursors in the *B42* locus identified six. Each gene is predicted to encode a short polypeptide ranging in size from 56-75 amino acids and terminating in the CaaX sequence, CVIA, CVVS or CIIA. The genes are distributed throughout the region sequenced and their positions correspond to sequences from which small 900nt transcripts are derived (Figure 3.11). To conform with the gene designations adopted for the *B6* locus, which will be discussed later in this chapter, the six pheromone genes are are listed as *phb1.1*, *phb1.2*, *phb2.1*, *phb2.2*, *phb3.1* and *phb3.2*.

The predicted amino acid sequences of the six polypeptides is given in Table 4.1 and the full DNA sequences in Figure 4.3. A BLAST search revealed no sequence similarities to known proteins in the databases. This was not unexpected since the sequences are very different from each other. They display no sequence similarity to

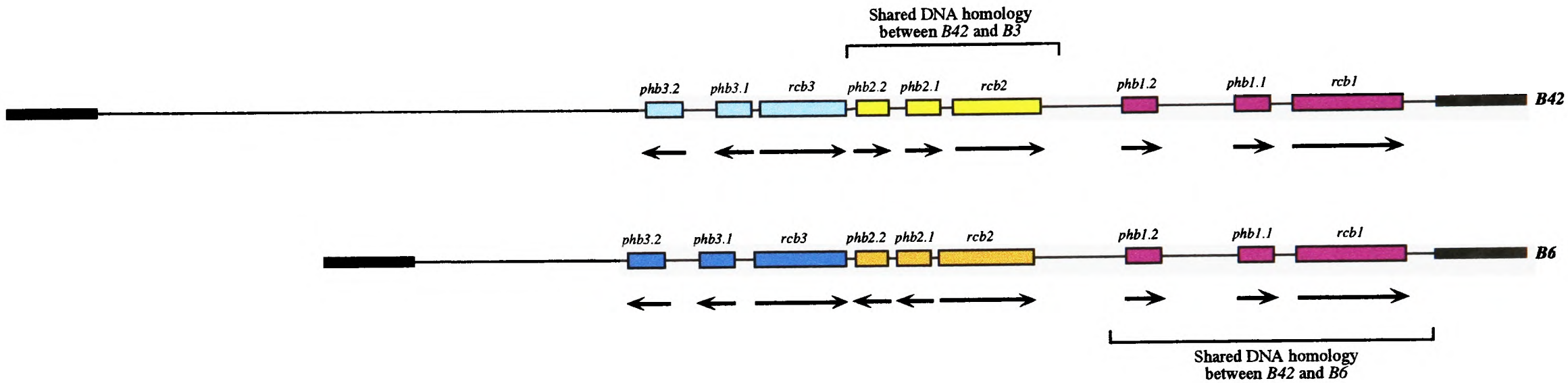


Figure 4.2 Comparison of the *C. cinereus B42* and *B6* loci. Thick lines represent the homologous borders to the loci. Arrows indicate the direction of transcription of the genes. Different colours are used to identify the three predicted groups of genes and illustrate DNA homology between the *B42* and *B6* loci.

Table 4.1 Comparison of the amino acid sequences of basidiomycete pheromone precursors. The predicted sequences of (**A**) the six pheromone precursors encoded by genes at the *B42* mating type locus of *C. cinereus*, (**B**) the six pheromone precursors encoded by genes at the *Bα1* (*bap1(1)*, *bap1(2)* and *bap1(3)*; Wendland *et al*, 1995)) and *Bβ1* (*bbp1(1)*, *bbp1(2)* and *bbp1(3)*; Vaillancourt *et al*, 1997)) loci of *S. commune* and (C) the pheromone precursors encoded by the alternative *a1* (*mfa1*) and *a2* (*mfa2*) loci of *U. maydis* (Bölker *et al*, 1992). The mature pheromone sequence is shown in red (Spellig *et al*, 1994).

	Organism	Gene	Amino acid sequence of precursor
A			
	C. cinereus	*phb1.1*	MDSFDSLDSLNLSVEETTLQTLLESMDTTDAASESELDAILVNSERDPGFTSKGFCVIA
	C. cinereus	*phb1.2*	MDSFQQLNLFVEETIQQSLPEAIPSSDSTDTGASERDATPVNTERHLGFTTKGFCVVS
	C. cinereus	*phb2.1*	MDSFSTLSLPATLGNEEAQQTTVIATVAQPQESPSSSGTPVDSERPGAGKVRAFCIIA
	C. cinereus	*phb2.2*	MDNFTVDLATLFEEFPELQEIQATASEHCSQDQYGSCEGPPINQERPGSGVNRAFCVIA
	C. cinereus	*phb3.1*	MDSGNTVDLAELCDMDPNIGFTPDSSAPTEDNVAKQLVDSDQRLPGGYYGGQCIIA
	C. cinereus	*phb3.2*	MSDAFTTLDTVDLFIEENEQEVVEVPSCPPPRRPSFSSADAESIFLTVEEVNDLPVDYERRTQGGGGLTWFCVIA
B			
	S. commune	*bap1(1)*	MDGEGHDINIRGARHSPSPAAAPVSATRGAPWSGCEGCPSRAADRRCVCH
	S. commune	*bap1(2)*	MRSRASAEGIAVLGLRRRGESPVCRRRNVVVCFWGDRSCVEREGCVRG
	S. commune	*bap1(3)*	MDDFAEFFPTLVLDEPEVARRPARDAEVLAILADAERPGGSNCTAWCVVA
	S. commune	*bbp1(1)*	MDAFTAMFPELFPIEEGLEDALVGSLSDTSAASASATHTSPASTDTFDDADILAILADAEHWRGGNTTAHGWCVVA
	S. commune	*bbp1(2)*	MDAFTDFSILADGLASLGDESSHTILAEFSPSILDGPFVADSAPLTEAPCNHDQIADYGSYCVVA
	S. commune	*bbp1(3)*	MASSVLARPGPSTVLPAMTRPPPPMAHRAAATPSFARSAQPQTDDAVLALLANAEHTEAGEETTARGWCVVA
C			
	U. maydis	*mfa1*	MLSIFAQTTQTSASEPQQSPTAPQGRDNGSPIGYSSCVVA
	U. maydis	*mfa2*	MLSIFETVAAAAPVTVAETQQASNNENRGQPGYYCLIA

Figure 4.3 DNA sequences of the *C. cinereus B42* pheromone precursor genes, (**A**) *phb1.1*, (**B**) *phb1.2*, (**C**) *phb2.1*, (**D**) *phb2.2*, (**E**) *phb3.1* and (**F**) *phb 3.2*. The predicted protein sequence is given in one letter code. The introns are indicated in lower case.

A (*phb1.1*)

```
GGTCGACAGCGACATTTCACACCCCAACATCACTCTACCCTACGGAGTGTCGGCCTAGGCTTGAAGGAAGCTTTGCAT  78

TCACGTCGCGCTTGGTACATTGGGTGACTGTCGCACCCGCAGAAGCGAGTTCAGGATGAAACGCCGCTCTAATTGAGC  156

TCGGTGTCAATGGCCCCTATTGAATTCCGCTTGTTTGGTCAACATTGCGCTTAGTGTTTGGCGAATAAAACAATTCTC  234

TAGAACGCTAGGTTTACCTGGCACAATGCCACGACCATTCTATTGTTCCTATTGATAGTCAGTGTGCACGCAAATGAG  312

GTATAAAATCTCTGTGTGATGTACACGCGACCCCTATATCTATCGCACTCACTGGCACGACTCACGACGATCCCAAGC  390

CGTTCCAAGTTTCCAAGCTCATTACCACCTTCCTTTACGCTCTCACAAGCTAACCAGCTTCACACTCACTCAAACCAT  468

CTGCAACACTCAACGCCACACACCAATCGCCATGGATTCATTCGATTCACTCGATTCACTCAACCTCTCTGTCGAAGA  546
                                M  D  S  F  D  S  L  D  S  L  N  L  S  V  E  E

GACCACCCTCCAAACCTTGCTCGAGTCCATGGATACCACAGACGCTGCGAGCGAATCTGAGCTCGATGCAATCCTCGT  624
 T  T  L  Q  T  L  L  E  S  M  D  T  T  D  A  A  S  E  S  E  L  D  A  I  L  V

CAACTCTGAGCGGGACCCAGGCTTTACTAGCAAAGGCTTTTGCGTCATTGCTTGAGATTGTATTCGTCTTCTCGACCG  702
 N  S  E  R  D  P  G  F  T  S  K  G  F  C  V  I  A  ·

TTATCATTATTGTTTGTTTATGGAATGAAGCAACAGTGAGCTCGCGATATCTCTTCCCACCTTTGGGCCATGCTAATG  780

GAGGCGTTTCTCCAGACTGTATCAATAGCCATCAGGATCCCGTCTTTTGCGCTCTGGAAGCGTTTGTATCACTATCTA  858

TTACTCCCCCGACAAGGGACTCCTCACTCATTCGTTCGTACGAAACCCCGACGGCCTTGTACTGTAGCATACTGTCCA  930

CATACACCTACTTGATTGAATAGCTAATTGAGTCGTGGCCCTT  979
```

B (*phb1.2*)

```
ATGGCATCCTAGCCGTTTAGTGTGCTTATTCTCTCCTACAAATTTTCGGTTCGTTATGACACGAGGTCCTTCTTAGAA  78

CGCCCTCTCACAAATATCGTAGGAAATGTCCACAGAGATTGGGACGGAACAACGAGGAGAACACGCGGGCGTTCTCCG  156

AGTCACGGAGGCGGACAGGGTCTCTTGTCTCCGACATAGCTCCGCTTGGCTGGGACCCCATGCGCAATACTTTTTATG  234

CTTGCTAATGCTTCGAATTCACTAGTCGAACAAAGTCATTTATAGTATTCTTGATCACGGGTATATAAGGCGACAGCT  312

CGAGTTCCTTGGGTTGGTCCCGTCATCACCACCATCCGATAGCTTCGTGAATCGCCCGTCGCCGTTGAACTAAACTTG  390

CTTCTGTCTTCCTGCTCGTCCGTGCACTCCTTCTTTTCGAGCTTCACTCAACAACCCACCTTACCGACACTCCTTCCT  468

CGTCACAACCGATCCACCTTTCTCCAATAGCAATGGATTCCTTCCAGCAACTCAACCTCTTCGTCGAGGAGACCATCC  546
                                  M  D  S  F  Q  Q  L  N  L  F  V  E  E  T  I  Q

AACAGTCCCTCCCCGAGGCGATCCCCTCCTCTGACTCCACCGACACCGGTGCCTCAGAGCGAGACGCCACCCCGGTCA  624
  Q  S  L  P  E  A  I  P  S  S  D  S  T  D  T  G  A  S  E  R  D  A  T  P  V  N

ACACCGAACGCCACCTCGGCTTCACCACCAAGGGCTTTTGCGTTGTCTCGTGAAGCGTGAAGCTCTATCCGATCCCTT  702
  T  E  R  H  L  G  F  T  T  K  G  F  C  V  V  S  ·

GCGTCGTCCCTGCGGAGAAGAAGAATTGAAGCAATAGCCAgtacgtcccgggtcgctcttgccactccatctgaatgt  780

ggttctaacgatcgctctgccgcagATGTACCACTATATTTATTATCTCTCAGCCCACCAAGTAGGCCTCAATGCCTC  858

ATCGAGCGGGACCTCAGCGCCATGGTGTTTTATTCTGTTCACTATTTCGACTCCCCTTCATTATCATTCCTGTTGTTT  936

TAGTACCATTTTATCGTTGACAACACAATTATGACCCGTTA  977
```

C (*phb2.1*)

```
TATCCGGAATAGATTCTCTGTCTGGAACACATTGCACCCAGCTCCGCCACCTGAGAGAGCTATGGTCCACCCAGCTAC  78

CTGCATAAACCCAAAGAATCGCTGGCCAGACATAGCCGGTCTGCGCAACCGAGGATATGAAATCGGTGCACGGATCTG 156

GATCGGGACTGGCGCTCAGATGACTGATATGCAGAATCAAGAAAGAAAGATAAGTAAAAAAAGAACTGAGTAAGAACA 234

AACCACTTTCAAAAACCCCGAGGCATGCGCGCGCCGCATTCCCTGGAAAGAGGGGGTCGGAACGGACAGATGTTCCGA 312

TGTAAATCCAGCGCCTCGTCGGAGTCTACCGATTGTATTTCTCACGCTTGGAAGGGCACAAAAGCTATCCGCGGTCAG 390

CTGGAGTCACAACAGAACGTTCTATTCAATCACCCGCCGTCCCTACGTCTTTGCGACTCGCTACAAGCTTTTCTCTTC 468

TGCCCAACCCCAAGCTAACAGCCCGATCGATAATGGATTCCTTCTCCACCCTAAGCCTCCCCGCCACCCTCGGAAACG 546
                                M  D  S  F  S  T  L  S  L  P  A  T  L  G  N  E

AGGAGGCGCAACAAACCACTGTGATCGCCACTGTCGCTCAACCTCAGGAGTCGCCCTCTTCATCGGGCACCCCAGTTG 624
  E  A  Q  Q  T  T  V  I  A  T  V  A  Q  P  Q  E  S  P  S  S  S  G  T  P  V  D

ACTCCGAACGACCAGGCGCAGGCAAGGTTCGCGCTTTCTGTATTATCGCATAGCCTCTAAGgtacgtctgttaatcgc 702
  S  E  R  P  G  A  G  K  V  R  A  F  C  I  I  A  ·

cacacttctaaccttccgcgctcaatcccctttgcagCTATCCCCAAGGCTCTTTGACCGCACTCCTTTTCCAATCAAC 780

CAACATCTATCGCGCCCATCAACAACACCTAGAGGCATTACAACTCAACAAGTCTACCATTACGATGTGGCGACCCCC 858

TTGCCACATTGTTCGGCATCGCGCTATTCGTTGCTCATTGCTTGTAATTTATTTAAATTCGTTCGTTAAGAAGGTGTT 936

ATGCCTTAACCGCCATTTTGGCATTCTTTAACCATGCTCTT  977
```

D (*phb2.2*)

```
ATTTCCTTGCCCTCGACATCTTCGTCGGTACTTCTCTCGTAGGCTCAATGTCACAGCGTACTGCTATGTAGCGGTAAT  78

GTGAGATTTATATAATTGCTGTCTATGGCTACCGAAGTTGCAGGTTGCCAGTGCAATCGTCCGCGCCTGACACTCCCA 156

CATGTGGGCTCGGTCCCTGATTCACGAGTTTTTCGGCAATATAAGTAGAGGTAGCTTCAGACAGAGCGTCATATGGCC 234

TCTCAATACGCGCTCCTTTGCGGCTGAAATCGACATATTCCCGCCCAGGGGGACATAATGGGTTTCGGAGGACAATGG 312

CTCCCTCTTGCATCAGTTTTCAAACTCCGATCATTGTCCCTCACTGCCCATTGTTCGCATTGCATATAACGACGCCGT 390

TCAACGAGCACAGACCGACCAGTAAGCCATCGTTAACTCGCGCGTTAGCCACCGTTTCTTCTGCTTAACATATCCACG 468

CTCTCTACGAACATCCGTTCTTTCTTTCCGCCATGGATAATTTCACCGTCGATCTTGCCACCCTCTTCGAAGAGTTTC 546
                                M  D  N  F  T  V  D  L  A  T  L  F  E  E  F  P

CTGAACTTCAAGAGATCCAGGCCACTGCCTCAGAGCACTGCTCACAGGATCAGTACGGCTCTTGCGAGGGCCCTCCAA 624
  E  L  Q  E  I  Q  A  T  A  S  E  H  C  S  Q  D  Q  Y  G  S  C  E  G  P  P  I

TAAATCAGGAACGACCGGGAAGTGGCGTGAATCGCGCATTTTGTGTGATCGCCTAACCTGCTGAAGGCAAGTCACTGC 702
  N  Q  E  R  P  G  S  G  V  N  R  A  F  C  V  I  A  ·

CCTTGTGGCCACCTGCAGTCGATAAACTTACAAGGCTCTTACCGTCTTTAGGTTGTTCCGACCTACCCTTCAGTGTTC 780

AAAGCGGTCCGAAACCCTTTCACTTGCTCTTTTCACTTGCTCTCGTGGTAAGGCTCCACTTTCGCTGACTCGTATAAG 858

GAACTAAGGGAGAGTTTTTTCCGCACAGTCCGATCAACAGCAACTTTTTACGCCGGCGTCTTCTCATTGCATGTCTCT 936

CTTCCTTTCTCGTTTCTCCCTTCGTCCTTCTTTGCGGTACTGCT  980
```

E (*phb3.1*)

```
GCAGCCATCTCGTGAAATCTCATCGCTGGGGTTTACAGGATATCGCATCGGTTAGGGCAGTAGGCCCTCATCTTACAG  78

ACTTTTCTCCTCCTTACTCCTGAATGCCCGGCTTGATAACCAATAATTAATAAACCTTCATGGAAAGAGGGCGCCTGT 156

CGCTTATTGACACTTATAGTAAAGAGGATTCACGTAGATGCTCAGAATTTACCCGAAGATACACTGAGATACATAACC 234

CTTTTGACCTAGCCCCCGCGCCTCCCCGGGAGGGGTCGTCGTATATATTGTTTGGTGACAATGCGACAATGGAGGACA 312

AAGTTCGCACGGTACCATTGTCCCATAAATACTGTGGTATTGAGGATTCGCCGCAACGTTAACCTTGTCATCCATCCA 390

GCGAATTCAGACGACGCTCATCGATCAAGTGCCTTCCCTCGTTCCTGATGCTTTATCCACCGCGCCAAGAACCAACGA 468

CTACCACAACTTCCTCCGTATCATCTCCAACTATGGATTCCGGCAACACTGTCGACCTCGCAGAGCTCTGCGACATGG 546
                                M  D  S  G  N  T  V  D  L  A  E  L  C  D  M  D

ACCCAAATATTGGCTTTACCCCCGACTCTTCAGCACCCACGGAGGACAATGTCGCCAAACAACTCGTGGACTCCGACC 624
  P  N  I  G  F  T  P  D  S  S  A  P  T  E  D  N  V  A  K  Q  L  V  D  S  D  Q

AGAGACTTCCGGGGGGTTATTACGGAGGCCAATGCATCATCGCATAACGTTTGGCACAATGCATGTATCTTCCCCCGC 702
  R  L  P  G  G  Y  Y  G  G  Q  C  I  I  A  ·

TTCTTCCTGCTCTTGTGCTGATAGCTCAACAGGTGCACAGTCGCAGTCATTACTCTGCCTGGCTCATTATGAGCCCTT 780

CTTTAATCGTCATTTTTTCGTTCATCAATGGCCACTTCACTTTGACTTCACTTTCGTTCTGTTGTTTCTTTCAATCCG 858

GGTGCTTGACTTTGCCATCATGTAACAAGAATTCATGACCACCACGACTGTTGATATGTATGTAGTAGTAACTTATTT 936

GCCAATCGCTGCGAGACTGTAGTAGTGTAGTGTGA  971
```

F (*phb3.2*)

```
AGACCGCAGCTGAACCATCACCTCGTGTTCGCGGAGAACGCCGATGCTCCGTATGGGCATCCCTGATTCAGTTCGCCC  78

AAAACGGYTCACAATGGACGACAATGGTTCAGACTCCGAACTCAGGGAGCAAATCGGGTAATCAGCCCCTATTTCGAG 156

AACAGGGCCCTGCTCCCGATGTTCCATTCGCCTTGAAAAGGTGTTGTCCCCCCAGAGAATGGTACTATAATTATTTTT 234

ACGAAATTATCCGGCTTTGTAGAACATTAGAGCGATAACATAAGGCCACGTTGCTCATTATTGATTCTTTAAACGCTT 312

TCCACAATGGCAATTCTCATTATTTGTACATATTTGTACAGTAAATTCTGAGATATAAGCAAGTTGGATCGCTCCCGA 390

TTTCTCACACATACCAGCACCATTGTCAATCGTCTTCCATCCCCCCACCTCCAGTTAACCAAGCTAATCAACCATCAA 468

GCCACTCTCGATCTCCCATCATCATCGCCGACATGTCTGACGCATTCACTACCCTCGATACTGTGGACCTCTTCATCG 546
                                M  S  D  A  F  T  T  L  D  T  V  D  L  F  I  E

AGGAGAACGAACAGGAAGTCGTCGAGGTGCCATCCTGTCCGCCTCCCCGTCGCCCCTCCTTCTCTTCTGCCGACGCAG 624
  E  N  E  Q  E  V  V  E  V  P  S  C  P  P  P  R  R  P  S  F  S  S  A  D  A  E

AATCTATCTTTCTGACTGTGGAGGAAGTTAATGACCTTCCCGTTGACTACGAGCGCCGGACTCAAGGTGGAGGCGGCC 702
  S  I  F  L  T  V  E  E  V  N  D  L  P  V  D  Y  E  R  R  T  Q  G  G  G  G  L

TCACTTGGTTCTGCGTTATCGCATGAACGCGGCCGAGGTTTACCTCGCCTCGCTCGCGAATTCTTCGGGCTGACTCCT 780
  T  W  F  C  V  I  A  ·

TAATCCGTACGTTATTTTTTACTTCCTTTCGACCTTGGACGTGACTGATGGTGATTAACAGTTCGGGGAGGTCAACCA 858

TCTTGGAAATCCATCGCCAATCAAGGACCACAAGTTCATCTCCCCGACCTCTACATCTTTCACAATCCATCCCCACAT 936

CCGTTCTGTTTCTTTCCTCAATAAGGAGCTCATATTCGGCTATGCTCGTGCTCCGTATTAGCAACCTTCGTTGCGATC 1014

AACGCTCATTCATG  1028
```

pheromone precursors encoded by the *B* genes of *S. commune*, which like the *C. cinereus* polypeptides, vary considerably in predicted amino acid sequence though they are known to be functionally equivalent (Wendland *et al.*, 1995; Vaillancourt *et al.*, 1997). The sequences of the *Bα1* and *Bβ1* pheromone precursors are shown in Table 4.1 (B) together with those of the *U. maydis* Mfa1 and Mfa2 polypeptides (C). Only in *U. maydis* is the mature peptide pheromone sequence known (shown in red), because these are secreted pheromones which have been relatively easy to purify (Spellig *et al.*, 1994).

C. cinereus genes sequenced to date contain short introns of some 50bp, a size characteristic of introns in genes of other filamentous fungi (Ballance, 1990). The 5' and 3' splice sites of the introns are highly conserved. The consensus 5' site is GTANGT and the 3' site is YAG, where Y=C or T (Gurr *et al.*, 1987). Potential splice sites were not detected within the predicted coding regions of each of the pheromone precursor genes. The *U. maydis mfa1* amd *mfa2* genes do however, have an intron downstream of the translational stop site in the 3' untranslated region (UTR). Predicted 5' and 3' splice borders are found in the 3' UTRs of the *phb1.2* and *phb2.1* genes so these may have similarly located introns.

The most obviously conserved feature of the pheromone precursor polypeptides is the CaaX motif at the carboxy terminus which is either CVIA, CVVS or CIIA. The aliphatic amino acids are thus valine or isoleucine and the X residue either alanine or serine. The carboxy-terminal X residue largely determines which type of prenyl group is added to the cysteine residue, a farnesyl group or a geranylgeranyl group (Caldwell et al., 1995). X in the *B42* pheromone precusors is A or S which indicates that the cysteine will be farnesylated like **a**-factor of *S. cerevisiae* and M-factor of *S. pombe*.

4.4 Identification of genes encoding pheromone receptors

Three genes were found that corresponded in position to sequences that identified transcripts of 2.1-2.5kb in size. These three genes were present within the same region encoding the pheromone genes. A BLAST search revealed that the translation products showed significant identity (BLAST P(N) < $5e^{-5}$)to the pheromone receptor proteins of other fungi, namely Ste3p of *S. cerevisiae* which encodes the **a**-factor receptor, Map3 of *S. pombe* which encodes the M-factor receptor, and the Pra1 and Pra2 pheromone receptors encoded within the *a* loci of *U. maydis*. The three genes were designated *rcb1*, *rcb2* and *rcb3* to conform with the designations of equivalent genes within the *B6* locus.

It was clear from the homology searches that there must be several introns present in each gene and sequence analysis identified several consensus splice sites. To confirm the actual coding sequence, cDNA was synthesised by Reverse Transcription PCR (RT-PCR). The cDNAs were amplified from the RNA, cloned into the vector pGEM-T (Promega) and sequenced. The rcb1 and rcb2 genes were each shown to contain four introns whereas rcb3 has five introns (Figure 4.4). The lengths of the predicted proteins are 561, 389 and 427 amino acids for Rcb1, Rcb2 and Rcb3 respectively. An alignment of the three proteins is presented in Figure 4.5. The intron positions are well conserved and are all present in the first 340 amino acids which contain the seven predicted transmembrane helices.

From the alignment of the three *B42* receptor proteins it is evident that they are very dissimilar in length. Comparisons of the amino acid sequences show that there is 53% identity between *rcb1* and *rcb2*, 38% identity between *rcb1* and *rcb3* and 35% identity between *rcb2* and *rcb3*.

Figure 4.4 DNA sequences of the *C. cinereus B42* receptor genes, (**A**) *rcb1*, (**B**) *rcb2*, (C) *rcb3*. The predicted protein sequence is given in one letter code. The introns are indicated in lower case.

A (*rcbl*)

```
AGTCGTGGCCCTTCTTTGTCTTATGATTTTTGAATAGAGCCGTGGGACTGAAGCAAATCGGAATGGAAGCCAAAGGGC  78

TTTGTTTGTACACCTTGATCAATGTTTGGCACCCGCGGCGCTGTGGCTTGTGCCAAACGTTGCCCAGAGACGCGCTTT 156

ACCGGAATGCCTGCCTTCCTCTTGGCTATGCATGAGCAGACATGAGCAGATCCTTCAATAAGGTAAAATACAGGAGAT 234

TTTTACCCAGAAACTGCGAGAAACGGCACTGAAGGGGCCGACCGACGGTGCCTCGCGATCACGATCACGTGACTACAA 312

TGAATCCCGCACTTAGGCCGAAAAAGGAAGGAGCTATTGTCGCGCCAAACAGGTGACGGCATGCCACCTTCCCCTTTG 390

TTCACTTCGCTATTGAGCCAAAGGGCTTTATCACATTATAATGGCTGAGGCAAACTCCTCAAAGGGCACGCATTCCTT 468

GCTTTCTCACTTCCCCACTCCTCGTCATCGACATGAAGTACCCAGCTCTCCCGGTGTTTGCGATCCTCGGCGCTTTAC 546
                                M  K  Y  P  A  L  P  V  F  A  I  L  G  A  L  L

TCGTTCTCATCCCGCTGCCGTGGCACTGGCGGGCTCGTAATGTAGCCACGCTTTCTATCATTGCCTGGCTTTTTGTCA 624
  V  L  I  P  L  P  W  H  W  R  A  R  N  V  A  T  L  S  I  I  A  W  L  F  V  M

TGAATGTTATCTACGCTGTAAACACGATCGTATGGGCTGGTTCACTCCGCGATGTTGCTCCTGTATATTGCGATATCG 702
  N  V  I  Y  A  V  N  T  I  V  W  A  G  S  L  R  D  V  A  P  V  Y  C  D  I

gtatgtctctaggctcgagtattctccagaacgatagctaacgtctcccatccagCGACCAAGCTTATCATTGGTGCA 780
                                                        A  T  K  L  I  I  G  A

TCTCATGCCCTCCCTCTCGCCACTCTCTGCATTTGCAAGCATCTTGAAATGGTCTCTTCTTCACGAACGGTTTCCTAC 858
S  H  A  L  P  L  A  T  L  C  I  C  K  H  L  E  M  V  S  S  S  R  T  V  S  Y

GACGTCTCAGATAAGAAGAGGAGGATGATCTTCGAGGGCGTGATGTGCTTTGTCCTCCCAATGATCTTCATGGCTCTC 936
D  V  S  D  K  K  R  R  M  I  F  E  G  V  M  C  F  V  L  P  M  I  F  M  A  L

CgtaagttattgcttcttatctaatatgtgcgtctgtctgaccatttgtagACTACAACGTTCAGGGCCACAGATACG 1014
                                                    H  Y  N  V  Q  G  H  R  Y  D

ACATCATTCAAGAGTTTGGATGCCAGCCAACCATCTACATTTCTATTCCAGCCATCTTTATCGTCTGGTTCCCCCCTC 1092
  I  I  Q  E  F  G  C  Q  P  T  I  Y  I  S  I  P  A  I  F  I  V  W  F  P  P  L

TTCTCTTCGCAGTTATTTCCTTCATTCTGGCCGgtatgtcctgtctctctattgcttcaactcgtttcgctaacggtc 1170
  L  F  A  V  I  S  F  I  L  A

gccacagCAATGGCCCTTCATCACTTCGTTCGCCGCCGCTTGGTCTTCGCTGCTCACCTGCAAAACGCCAACTCTGCG 1248
     A  M  A  L  H  H  F  V  R  R  R  L  V  F  A  A  H  L  Q  N  A  N  S  A

TTGACCCCTAACCGATACATTCGACTCATTGCTATGGCTCTCACCATCATGACCCGGAATACCAGTCTCACCGCCTTC 1326
L  T  P  N  R  Y  I  R  L  I  A  M  A  L  T  I  M  T  R  N  T  S  L  T  A  F

AACTTGTACAATAACGTTTTCCCCGGTCTCCGCCCCTGGACCAACTGGGCTGACGTGCACTCCAACTTCTCCCGAGTC 1404
N  L  Y  N  N  V  F  P  G  L  R  P  W  T  N  W  A  D  V  H  S  N  F  S  R  V

GACCTCTTCCCGACCGTCTTTCTGCCTGACTACTTCATTCGTGCCATGATGCTCTTCTGGTGGGCGATGCCTGTATCA 1482
D  L  F  P  T  V  F  L  P  D  Y  F  I  R  A  M  M  L  F  W  W  A  M  P  V  S

TCTTTGATCTTCTTCCTGTTCTTCGGGTTCGGTGAAGAGGCAATGAAGGAGTACAGGAAGGTTGGGGCTTGGATCTCC 1560
S  L  I  F  F  L  F  F  G  F  G  E  E  A  M  K  E  Y  R  K  V  G  A  W  I  S

CGAGTGATTCTCAGGAGGAAAACCAACGAGAAGGGCGAATTGCTTGGATCGACCGGAAACAGgttagtcctatttttc 1638
R  V  I  L  R  R  K  T  N  E  K  G  E  L  L  G  S  T  G  N  S

aatcgttggtggaccgtgcgctgacttcaattactattagCCGTCGCCGCCTCCACCTCGTTGACCTCAAGAACAAAA 1716
                                        R  R  R  L  H  L  V  D  L  K  N  K  N

ATGCCTCCTTGGGATCAATGATCAGTAAGCCAATCCCTGGCTCGTTTTGCAAGACTTCCCAGAGTCCTTCGGCGACCG 1794
  A  S  L  G  S  M  I  S  K  P  I  P  G  S  F  C  K  T  S  Q  S  P  S  A  T  V

TTGCTGGTAGCCCTCCTCCGTACAAGAAGTCATTCGACGTCGTTTCTCCAGCCAGCACTTCCAATTGCTCCACCGTCG 1872
  A  G  S  P  P  P  Y  K  K  S  F  D  V  V  S  P  A  S  T  S  N  C  S  T  V  V

TGAACGCCAACAGCCCACTCAAGAAGGCGAATGAGATCGATCTCGAAGCCGGGGACGATTTCTCGTACTACGCCCAAT 1950
  N  A  N  S  P  L  K  K  A  N  E  I  D  L  E  A  G  D  D  F  S  Y  Y  A  Q  S

CGAATGCCTCGACGGCCGCTCAACTCGCTCCACCAGCCATGGTTGAGCGATCGTCATCGACTCGTTCCTCTGGTCCGT 2028
  N  A  S  T  A  A  Q  L  A  P  P  A  M  V  E  R  S  S  S  T  R  S  S  G  P  Y

ACATCTCGGAGCCAATTCCCAGCCCAACGCTCTCCTACTGTTCCATCCTCTCACCTTCAACGGCTCCCTCCACCTCCG 2106
  I  S  E  P  I  P  S  P  T  L  S  Y  C  S  I  L  S  P  S  T  A  P  S  T  S  A

CGTTCTCCCCATCATCATTCCCCGAGTCTGTTGGTCCCGACTCGTTGAACCGTGCTTCCGAGATGGATCTGGTGGAGA 2184
  F  S  P  S  S  F  P  E  S  V  G  P  D  S  L  N  R  A  S  E  M  D  L  V  E  S

GTTACTACGAGATGCAGAGGGACATCCGGGAAGAGGAGACTAGGCGGACACACTTGCGCGGACCTGCCTCCCCCGCCT 2262
  Y  Y  E  M  Q  R  D  I  R  E  E  E  T  R  R  T  H  L  R  G  P  A  S  P  A  Y

ACCACAGGCCGTTCAGCCCAACTCTCTACCCTGTCGCTCTTCCCCATCCCGGCTCAGTACTGCCGCCTGTGAATGGCA 2340
  H  R  P  F  S  P  T  L  Y  P  V  A  L  P  H  P  G  S  V  L  P  P  V  N  G  I

TTCTGGTGACCGTGCACAGACAGGCCTCTGTCGATGAACTGCCTTCGACCCGTTCCTAAGCGCCAATTCACGACTGCT 2418
  L  V  T  V  H  R  Q  A  S  V  D  E  L  P  S  T  R  S  ·

ATTTCCCCCTGATGATCTTTCTTTGTCTCCATTCCACGATATCCCTTCACGAGTACGCACCCATCCAAACGACTCTGA 2496

ACGCCATAACTTATGTCCTACGCCTGCACGACTCTTCATTTCTTTGATACGCTACGGATCTAGATGATCTGGTTCATT 2574

CACGCACCCGTTTTACGCTCATTTAT 2600
```

B (*rcb2*)

```
ATCATGAGGAGGCCTCGAACGAATACATGACGCTACCGCAATTCTCGAAGCACAAAGGGCCAACACAATGGCGAAGTC  78

CCCAGGAGTGGTCAAGCACGACCAAACAGTCGACAGCACATGCTAGCGCAGACACGACGCTGCCATTAGGCGGTCCAC 156

AGTGGGTCAAACTTAGAGAAAGGATACATAACACCAACAATCGCCCCCTTCTCTAAGGTTGATGCTCGTCTAACTTGA 234

GCAGCGAACTGGAGCAGACTTTTGATGCGTACTTTGCCAAATAGGGATGCGCTGGCGGCGGTGCATAGCGTGAGAAGT 312

ATGACATAAACAATGGGTTTTTCTGTTTGGCTTGGCAGAATTACGGCACAAACGACTGTGGCCAAACTCAGTAGGGCT 390

TTGTCCTCAGCCCATTGTCCCTTTCGCCATTGAGCTTGCCGATTCCTACATATACCCTCAATGGACACGATGGAACTC 468

ACGTCTAACGCCCCTTACGACCCTCTTTCACCATGCGTTATGAAATGCCCATTTGCGCTTTCCTCGCTGCCATCCTAG 546
                                  M  R  Y  E  M  P  I  C  A  F  L  A  A  I  L  V

TTCTCATCCCCCTTCCATGGCACTGGCGAGCCAAAAACATTCCTACTCTTTCTCTCAGCATTTGGCTTTTCGTCTCCA 624
  L  I  P  L  P  W  H  W  R  A  K  N  I  P  T  L  S  L  S  I  W  L  F  V  S  N

ATATTATCTACGGAGTGAACTCAATACTCTGGGCGGACAATTACGCTGTCAAGGTTCCAGTTTGGTGCGACATTGgta 702
  I  I  Y  G  V  N  S  I  L  W  A  D  N  Y  A  V  K  V  P  V  W  C  D  I

cgtctttcatgccaccttcatctcaagattcgagacgctgacctccaatccagTCACCAAAATCCAGATTGGGGCCAC 780
                                                        V  T  K  I  Q  I  G  A  T

TCTATCGCTCCCAGCATGCTGTTTATGCCTCTGTATCCGTCTCGAGCGCATCGCCTCAACACGGGCAGCTAGCACCAG 858
 L  S  L  P  A  C  C  L  C  L  C  I  R  L  E  R  I  A  S  T  R  A  A  S  T  S

CGCCACCTCCAAACGGCGAACGGCCATCTTCGACGTTATGATGTGTTGGATTCTCCCCATGATCTACATGGCTGCTCg 936
 A  T  S  K  R  R  T  A  I  F  D  V  M  M  C  W  I  L  P  M  I  Y  M  A  A  H

taagtgtcttcttcatatctataatgctcctttattaatcatacccgccttctcgcgccagACTACATTGTCCAAGG 1014
                                                                  Y  I  V  Q  G

ACATCGCTTCGACATCGTGGAGGACCTCGGCTGCCGACCCTCAATCTACACTTCAATCCCTGCCATCTTTATCATATG 1092
 H  R  F  D  I  V  E  D  L  G  C  R  P  S  I  Y  T  S  I  P  A  I  F  I  I  W

GGTTCCTCCCTTTATCGCCATCTTCCTCACATTCTGCTTTGGAGgtaagtgcgaaatacccgtgctccaccttgtccc 1170
 V  P  P  F  I  A  I  F  L  T  F  C  F  G  G

tgattgcgatccttctgacagGTGTTGCGTTGCACCATTTCTTCCGCCGCCGTCTCACCTTCGCTCGTCACCTCCAAA 1248
                        V  A  L  H  H  F  F  R  R  R  L  T  F  A  R  H  L  Q  N

ACTCCAACTCCGCTATTACCCCAGCGCGCTACTTCCGACTCATGGCAATGGCCATCACCCAGATGCTCTGGGGTGGAA 1326
  S  N  S  A  I  T  P  A  R  Y  F  R  L  M  A  M  A  I  T  Q  M  L  W  G  G  I

TTGTCCTTTCTGTTAACGCGTGGATCACCTTCCGTACAGGCCTTCTGCCGTACACGAGCTGGGAGGATGTTCACTGGG 1404
  V  L  S  V  N  A  W  I  T  F  R  T  G  L  L  P  Y  T  S  W  E  D  V  H  W  G

GTTTCTCTCGTGTTGGGCAATTCCCTCTCATTTTCCTCCCTCGGTCCATGATTCCTTCATTATTCTTCAACTGGTGGT 1482
  F  S  R  V  G  Q  F  P  L  I  F  L  P  R  S  M  I  P  S  L  F  F  N  W  W  S

CGGTGCCTGTCTCAGCACTCATCTTCTTTATCTTCTTCGGCTTTGGCCAGGAGGCCGTCAACGACTACCGCGCCGCAC 1560
  V  P  V  S  A  L  I  F  F  I  F  F  G  F  G  Q  E  A  V  N  D  Y  R  A  A  L

TGCGTCGCGTCAAGTCTCTGTTCTGGAAGCCTCGCCAGGAAACCGAAAAAATgtacgtcgctatccgtatgtttggtc 1638
  R  R  V  K  S  L  F  W  K  P  R  Q  E  T  E  K  M

atggatttcgtctaacctttaatccccagGTCTCTACCTTCATCTGCTTCTAAGCTTTCACCTTTGCCCCCTTGGACT 1716
                              S  L  P  S  S  A  S  K  L  S  P  L  P  P  W  T

CCTCGCACTCCCGACGGGAAGTTCCCTTCAACTGCAACTCTGAACAGGGATACCACTATCTCACTTTCGCTTCCACTT 1794
P  R  T  P  D  G  K  F  P  S  T  A  T  L  N  R  D  T  T  I  S  L  S  L  P  L

GATCTCCATTCCTCCGCTTCTTCAGACGCTACCTGCGTTGATCCGGACAGCCCCTCAAAGCAGACCTTTACGAGCAAA 1872
D  L  H  S  S  A  S  S  D  A  T  C  V  D  P  D  S  P  S  K  Q  T  F  T  S  K

ACGAGTGCATTCACGAGTAACGTCTAGATTTCACGCATATCATTATTCATCCCATGTATCATCAACGACCGTCACAGC 1950
T  S  A  F  T  S  N  V  ·

TAGCTGTACTATTAATCATTCTAATCTTCGTTCACTTTTCATGTACTTTACGCACATTCACGGAAATTCATGCTCTTT 2028

TGCTGGCTCATTTTACATAATTGGTGCGCTTTGGCGGGGTACAGATTCAACTTAATCGTACTTGGGGTGGT 2099
```

C (*rcb3*)

```
GACCCCTCCCGGGGAGGCGCGGGGGCTAGGTCAAAAGGGTTATGTATCTCAGTGTATCTTCGGGTAAATTCTGAGCAT  78

CTACGTGAATCCTCTTTACTATAAGTGTCAATAAGCGACAGGCGCCCTCTTTCCATGAAGGTTTATTAATTATTGGTT  156

ATCAAGCCGGGCATTCAGGAGTAAGGAGGAGAAAAGTCTGTAAGATGAGGGCCTACTGCCCTAACCGATGCGATATCC  234

TGTAAACCCCAGCGATGAGATTTCACGAGATGGCTGCAGGAGTCTGCGCGAGCGCTAAACCCACTGAGAGGAGGGCCC  312

GCCCCCCGGGCCGAGTACAATAGGTCGGAGGTTGATTATATAATCAATGGACAACCTATTGACGCGCTCCTGGCCATT  390

GTTACACCCAAGATGATGCAGAACAAAGGCCTATGGAACCCTCAAAACCTCCAGATGCGAGACGCGAAATCATCGCTC  468

GTCCCCACCCACGATTTCCTTAGACGTCCATCATGTCCTCGTCACGAAAGCTGTCCTTCGACGACCCAACCTACCCCG  546
                              M  S  S  S  R  K  L  S  F  D  D  P  T  Y  P  A

CGTTCCCAGTCCTCTCTTTCATCGGCTTCGTCCTCGTTCTTATTCCTCTACCATGGCATCTTCAAGCATGGAACTCTG  624
  F  P  V  L  S  F  I  G  F  V  L  V  L  I  P  L  P  W  H  L  Q  A  W  N  S  G

GCACATGTCTATACATGATTTGGACAGCACTGGGGTGTCTTAACTTTTTCATCAACTCCATTGTCTGGCATGGGAATG  702
  T  C  L  Y  M  I  W  T  A  L  G  C  L  N  F  F  I  N  S  I  V  W  H  G  N  A

CCATCGACAGGGCACCGCTCTGGTGCGACATCTgtacgtctcccacttttctcgttgctgtttaaccgacttcaagcc  780
  I  D  R  A  P  L  W  C  D  I

gcctgtgccacctcagCGACACGCTTTGTAGTGGGACTCTCTGTGGCCATCCCAGCAGCATCCCTTTGCATCAACCGG  858
                S  T  R  F  V  V  G  L  S  V  A  I  P  A  A  S  L  C  I  N  R

AGACTGTACAAGATTGCTTCGTGTCGGACCGCAAACATTTCGCGCTCTGAAgtgggtagatcctccgtccaaaagtcc  936
R  L  Y  K  I  A  S  C  R  T  A  N  I  S  R  S  E

atggttgttgatcgtctcgtagAAAAGAAAGGCTGTCATGGTTGATCTTGCAATTGGACTCGGAATCCCTCTGGTTCA  1014
                      K  R  K  A  V  M  V  D  L  A  I  G  L  G  I  P  L  V  Q

AATGCCACTGCgtatgttttttcacatcccacgagcgcgcaataacctgacccaccctccagAGTATGTCGTTCAAGG  1092
 M  P  L                                                    Q  Y  V  V  Q  G

CCACCGGTACGATATCTTCCAAGACATCGGCTGCTACGCGACCACTGTCAACACACCGCCTGCGTATCCTCTGGTATT  1170
 H  R  Y  D  I  F  Q  D  I  G  C  Y  A  T  T  V  N  T  P  P  A  Y  P  L  V  F

CCTTTGGCCAACGATCATTGGTCTCGTTTCTGCTGTCTATTGCAgtgagttgcgttcccgatttcaactacgattcct  1248
 L  W  P  T  I  I  G  L  V  S  A  V  Y  C

ctcattgagcttcctttgtccgcagTTCTCACTCTTCGAGAGTTTATGAGACGTCGCGCCCAATTCAGCCAGCTCATC  1326
                         I  L  T  L  R  E  F  M  R  R  R  A  Q  F  S  Q  L  I

TCGTCCAATAGCTCCTCGCTCACCGTCAACCGTTACTTCCGCTTGATGTCTCTGGCCACGCTCGAACTCCTCTTCAAC  1404
S  S  N  S  S  S  L  T  V  N  R  Y  F  R  L  M  S  L  A  T  L  E  L  L  F  N

CTTCCCATTACGACCACGGGACTTTATCTCAGTATCACGTCCCGACCAATCTATCCCTGGACCAACTGGGCCGACATC  1482
L  P  I  T  T  T  G  L  Y  L  S  I  T  S  R  P  I  Y  P  W  T  N  W  A  D  I

CATTACAACTTCAACGTTATCGACACCTACCCCCGCATTCTTTGGGCTCACAAGACGGCAGACGCGATCATTCTCGAG  1560
H  Y  N  F  N  V  I  D  T  Y  P  R  I  L  W  A  H  K  T  A  D  A  I  I  L  E

TTAGGCCGATGGTCAATCGTCTTCTGCGCTCTGTTGTTCTTCGCTTTCTTCGGATTTGCGGAGGAAGCTCGAAAGAAC  1638
L  G  R  W  S  I  V  F  C  A  L  L  F  F  A  F  F  G  F  A  E  E  A  R  K  N

TATCGCAAGGCAATCCTCGCTCTTGCCAAGCGGCTGGGATACAACCCGCGCTCTCCCGACGAGAAGGCGAAAGCCACT  1716
Y  R  K  A  I  L  A  L  A  K  R  L  G  Y  N  P  R  S  P  D  E  K  A  K  A  T

TTTGTgtacgttcttctattatattcattcagcttcccatcctgacaccattgatgctcagTCACGTTGGAAAGCCAG  1794
F  V                                                          H  V  G  K  P  A

CTTCCCTCCCTAAGTTCACGCCTCGTCCTCTTCACCTTGCTTCCACCGCTACCCTTACCGCTACCCAAAGCGACAAGT  1872
  S  L  P  K  F  T  P  R  P  L  H  L  A  S  T  A  T  L  T  A  T  Q  S  D  K  S

CCACCAACTTCACTCGCACGCTCTCGATATCCAGCACCAGCTCAGGCTCCACAAAGGTCGCCTGCAGCCCCCCGGCCT  1950
  T  N  F  T  R  T  L  S  I  S  S  T  S  S  G  S  T  K  V  A  C  S  P  P  A  F

TCCGTCTCGACGACATCAAAGTCAAACTCAGCTCGAGTCCAGGGACTGTGTCTGATTTCCCTTCTTCCCCGACTTCGA  2028
  R  L  D  D  I  K  V  K  L  S  S  S  P  G  T  V  S  D  F  P  S  S  P  T  S  M

TGACAACACGACTCGACGATTTCGACTCTAACGCACACTCGGAAGTCTAAAGACGCCAATCTCCCCGTGGCGCGTATT  2106
  T  T  R  L  D  D  F  D  S  N  A  H  S  E  V  *

CCCGTAGCGCTGATGACCTTTTCTTTCCTTTCCCCCCGTTGTTCTCCTCTGGTTTCTGGCTCACGCATAAAAAACGAA  2184

GCGACCACAGATTCCGACTTGTATTATCTAGCACACACTCACATTCTTTACTTATTATTCACATTCACCGTCATGAAT  2262

TCACGACTATAGGACA  2278
```

Figure 4.5 Alignment of the three predicted *C. cinereus B42* receptors using PILEUP (GCG, version 8.1). Predicted transmembrane helices are indicated by a dashed line. Red, green and blue asterisks signify intron positions for Rcb1, Rcb2 and Rcb3 respectively.

```
          1                                                        50
B42Rcb1   .......... .MKYPALPVF AILGALLVLI PLPWHWRARN VATLSIIAWL
B42Rcb2   .......... .MRY.EMPIC AFLAAILVLI PLPWHWRAKN IPTLSLSIWL
B42Rcb3   MSSSRKLSFD DPTYPAFPVL SFIGFVLVLI PLPWHLQAWN SGTCLYMIWT
                     -----------1----------          --------
          51                                                      100
B42Rcb1   FVMNVIYAVN TIVWAGSLRD VAPVYCDIAT KLIIGASHAL PLATLCICKH
B42Rcb2   FVSNIIYGVN SILWADNYAV KVPVWCDIVT KIQIGATLSL PACCLCLCIR
B42Rcb3   ALGCLNFFIN SIVWHGNAID RAPLWCDIST RFVVGLSVAI PAASLCINRR
          ----2----------              ------------3-----------
                                        *
                                        *
                                       L*
          101                                                     150
B42Rcb1   LEMVSSSRTV SYDVSDKKRR MIFEGVMCFV LPMIFMALHY NVQGHRYDII
B42Rcb2   LERIASTRAA STSATSKRRT AIFDVMMCWI LPMIYMAAHY IVQGHRFDIV
B42Rcb3   LYKIASCRTA NISRSEKRKA VMVDLAIGLG IPLVQMPLQY VVQGHRYDIF
                               ------------4------------
                        *                          *
                                                   *
                                                   *
          151                                                     200
B42Rcb1   QEFGCQPTIY ISIPAIFIVW FPPLLFAVIS FILAAMALHH FVRRRLVFAA
B42Rcb2   EDLGCRPSIY TSIPAIFIIW VPPFIAIFLT FCFGGVALHH FFRRRLTFAR
B42Rcb3   QDIGCYATTV NTPPAYPLVF LWPTIIGLVS AVYCILTLRE FMRRRAQFSQ
                          ------------5-----------
                                                  **
                                                   *
          201                                                     250
B42Rcb1   HLQNANSALT PNRYIRLIAM ALTIMTRNTS LTAFNLYNNV FP.GLRPWTN
B42Rcb2   HLQNSNSAIT PARYFRLMAM AITQMLWGGI VLSVNAWITF RT.GLLPYTS
B42Rcb3   LISSNSSSLT VNRYFRLMSL ATLELLFNLP ITTTGLYLSI TSRPIYPWTN
                               ------------6-----------
          251                                                     300
B42Rcb1   WADVHSNFSR VDLFPTVFLP DYFIRAMM.. LFWWAMPVSS LIFFLFFGFG
B42Rcb2   WEDVHWGFSR VGQFPLIFLP RSMIPSLF.. FNWWSVPVSA LIFFIFFGFG
B42Rcb3   WADIHYNFNV IDTYPRILWA HKTADAIILE LGRWSIVFCA LLFFAFFGFA
                                          -------------7-----------
          301                                                     350
B42Rcb1   EEAMKEYRKV .GAWISRVIL RRK.TNEKGE LLGSTGNSRR RLHLVDLKNK
B42Rcb2   QEAVNDYRAA .LRRVKSLFW KPR.QETEKM SLPSSASKLS PLPPWTPRTP
B42Rcb3   EEARKNYRKA ILALAKRLGY NPRSPDEKAK ATFVHVGKPA SLPKFTPRPL
                                          *     *     *
          351                                                     400
B42Rcb1   NASLGS.... .....MISKP IPGSFCKTSQ SPSATVAGSP PPYKKSFDVV
B42Rcb2   DGKFPSTATL .NRDTTISLS LPLDL.HSSA SSDATCVDPD SPSKQTFTSK
B42Rcb3   H..LASTATL TATQSDKSTN FTRTLSISST SSGSTKVACS PPAFRLDDIK

          401                                                     450
B42Rcb1   SPASTSNCST VVNANSPLKK ANEIDLEAGD DFSYYAQSNA STAAQLAPPA
B42Rcb2   TSAFTSNV.. .......... .......... .......... ..........
B42Rcb3   VKLSSSPGTV SDFPSSPTSM TTRLDDFDSN AHSEV..... ..........
```

4.5 Gene organisation within the *B42* and *B6* loci

The parallel analysis of the *B6* locus of *C. cinereus* (O'Shea et al., 1998) made it possible to compare the gene organisation of *B42* with *B6*. This comparison is given in Figure 4.2. Both loci have the same number of genes and these are arranged in the same order. The genes appear to be organised into three groups, each comprising a receptor gene and two pheromone genes. The three groups have been designated 1, 2 and 3 and the genes within each group are distinguished by these numbers, for example in group 1 the receptor gene is *rcb1* and the two pheromone genes are *phb1.1* and *phb1.2*. The region of shared DNA homology that was shown to exist between *B6* and *B42* (shown in red) covers one complete set of these genes at the right of the map. The region of homology between *B42* and the unsequenced *B3* locus covers the entire middle set of three genes. These sequences are not shared by *B6* and *B42*. The third set of genes is in a region of non-homology between *B6*, *B42* and *B3*. Different colours have been used in Figure 4.2 to distinguish the three groups of genes and where the shared DNA homology exists.

The organisation of the genes into three sets suggests how multiple *B* mating types may be generated in *C. cinereus*. If the three sets of genes are functionally redundant as well as functionally independent, it would require only one set of genes to have different alleles for mating partners to trigger a compatible pheromone-receptor interaction. With four alleles of two sets of genes and five of the other, it would be possible to generate as many as 80 unique allele combinations - more than the 79 predicted *B* mating specificities present in nature (Raper, 1966). Functional redundancy is certainly the basis of multiple *B* mating types in *S. commune*. In contrast to *C. cinereus* there appear to be just two sets of multiallelic genes separated by some 3.5 map units into the discrete α and β loci (Raper, 1966). As predicted for *C. cinereus*, it requires only one

set of genes to have different alleles to promote a compatible pheromone interaction in mating partners. There are nine allelic versions of each locus and a potential 81 unique combinations that generate the 81 *B* mating specifcities known to exist.

4.6 Assay of pheromone and receptor gene function

The six pheromone genes and three receptor genes were separated by cloning smaller restriction fragments to test the function of each gene by transformation. Genes were introduced into *B6* and *B3* host strains and the results of this transformation analysis are summarised in Table 4.2.

In order to interpret the data it is important to recognise which genes lie in DNA sequences that are unique to the *B42* locus and those which lie in a sequence shared with *B6* or with *B3* (Figure 4.2). The three genes that comprise group 1, *rcb1*, *phb1.1* and *phb1.2* lie in the sequence shared by *B6* but not *B3*. All three genes changed the *B* mating specifcity of the *B3* host but not the *B6* host. Similarly, the three genes that comprise group 2 lie in a sequence shared by *B3* but not *B6*. All three genes changed the *B* mating specificity of the B6 host but not the B3 host. These data are consistent with the sequence homology results, which indicate that *B6* and *B42* share the same alleles of the group 1 genes and *B3* and *B42* share the same alleles of the group 2 genes. Where there is no DNA homology, loci have different alleles of the genes and different alleles trigger a compatible *B* mating response in the host strain.

The *B42* genes of group 3 are not in a sequence shared by *B6* or *B3*. It would be expected, therefore, that all three would change the *B* mating specificity of both host strains. This was true for the receptor gene *rcb3*, but not for the pheromone genes. *phb3.1* was only active in the *B3* host whereas *phb3.2* was only active in the *B6* host.

Table 4.2 Functional analysis of the nine genes at the *B42* locus. Genes were introduced into host strains having *B6* (LT2) and *B3* (LCO12) mating type specificities. (+), indicates a change in host *B* mating specificity in 10-80% of the transformants. (-), indicates no change in host *B* mating specificity.

	Host strain specificity	
Gene	*B6*	*B3*
rcb1	-	+
rcb2	+	+
rcb3	+	+
phb1.1	-	+
phb1.2	-	+
phb2.1	+	+
phb2.2	+	+
phb3.1	-	+
phb3.2	+	-

Clearly both genes are functional, and failure to activate a receptor in one of the host strains cannot be attributed to DNA-homology and homoallelism. Interestingly, in *S. commune*, where the *Bα1* and *Bβ1* loci each have three different genes encoding pheromones, it is found that genes may be active in some allelic backgrounds but not in others.

The results of the transformation experiments provided evidence that a single pheromone or single receptor gene can alter the *B* mating specificity of a suitable host strain. The results are also consistent with the prediction that there are three groups of functionally independent genes within the *B* locus and a gene can only trigger *B*-regulated development if the host and donor have different alleles of the gene. In the next chapter, experiments will be described that provide further evidence to support this model for the origin of multiple *B* mating specificities in *C. cinereus*.

Chapter 5

Organisation of the *B* mating type locus

5.1 Introduction

A molecular and functional analysis of the *A* locus by (Pardo et al., 1996) demonstrated that multiple *A* mating types were derived from three sets of functionally redundant genes. Each set contains two genes, whose multiple allelic versions could be combined into different combinations to generate some 160 *A* mating specificities. The organisation of the genes within the *B42* and *B6* loci suggests that multiple *B* mating specificities are also derived from three sets of genes, each set in this case comprising three genes. This chapter describes experiments designed to demonstrate that different *B* mating type loci may share alleles of one or two of the sets of genes identified in the *B6* and *B42* loci.

Transformation experiments with single cloned genes from *B42* showed that a single pheromone or receptor gene was generally sufficient to change the *B* mating specificity of the *B6* or *B3* host, provided that the donor genes lacked sequence homology to corresponding genes in the host genome. in the host strain. It was concluded that where there was no homology, the the host and donor had different allelic versions of the genes and that these were sufficiently different in sequence to be unable to cross-hybridise in Southern blots. Where there was shared DNA-homology, the genes were unable to change the *B* mating specifcity and they were assumed to be homoalleles.

In the analysis of the *A* locus, it was possible to take advantage of naturally deleted loci to generate a partially null background into which genes could be introduced to test for precise gene interactions that promoted development. It would be very difficult to generate similar partially null backgrounds to test interactions between the *B* genes because of the size of the locus, the very low frequency with which gene replacement occurs (> 1 in 500) and the lack of a phenotype with which to recognise such low frequency disruptions.

The analysis of the *B* locus described in this chapter is based on a correlation of hybridisation analysis, in which shared DNA-homology indicates that genes are homoallelic in different loci, and transformation, which detects whether or not the genes are homoallelic or heteroallelic based on their ability to change *B* mating specificity in the host. Strains used contained the following *B* loci, *B1, B3, B5, B6, B40, B41, B42, B43* and *B44*.

A fundamental requirement of peptides and receptor proteins encoded within the same locus is that none of the six pheromones are able to activate the three receptors. With three functionally independent sets of genes, the specificity problem is reduced to the three genes within each set. The three genes within each set must constitute an allele complex analogous to the receptor and pheromone genes found within the *a1* and *a2* loci of *U. maydis*. This implies that the complex cannot be disrupted by genetic recombination and the same alleles of the genes are always found together. The first objective in the analysis of the *B* locus was to establish that where shared alleles could be detected in different *B* loci, known alleles of all three genes remained together.

5.2 Southern blot analyses to detect shared alleles of genes in different *B* loci

Small probes specific to each of the nine *B42* genes and the six different genes present in groups 2 and 3 from *B6* (Table 5.1) were used to probe blots of *Hin*dIII cut genomic DNAs of the nine strains having different *B* mating specificities (Figure 5.1). The results of this analysis are summarised in Figure 5.2.

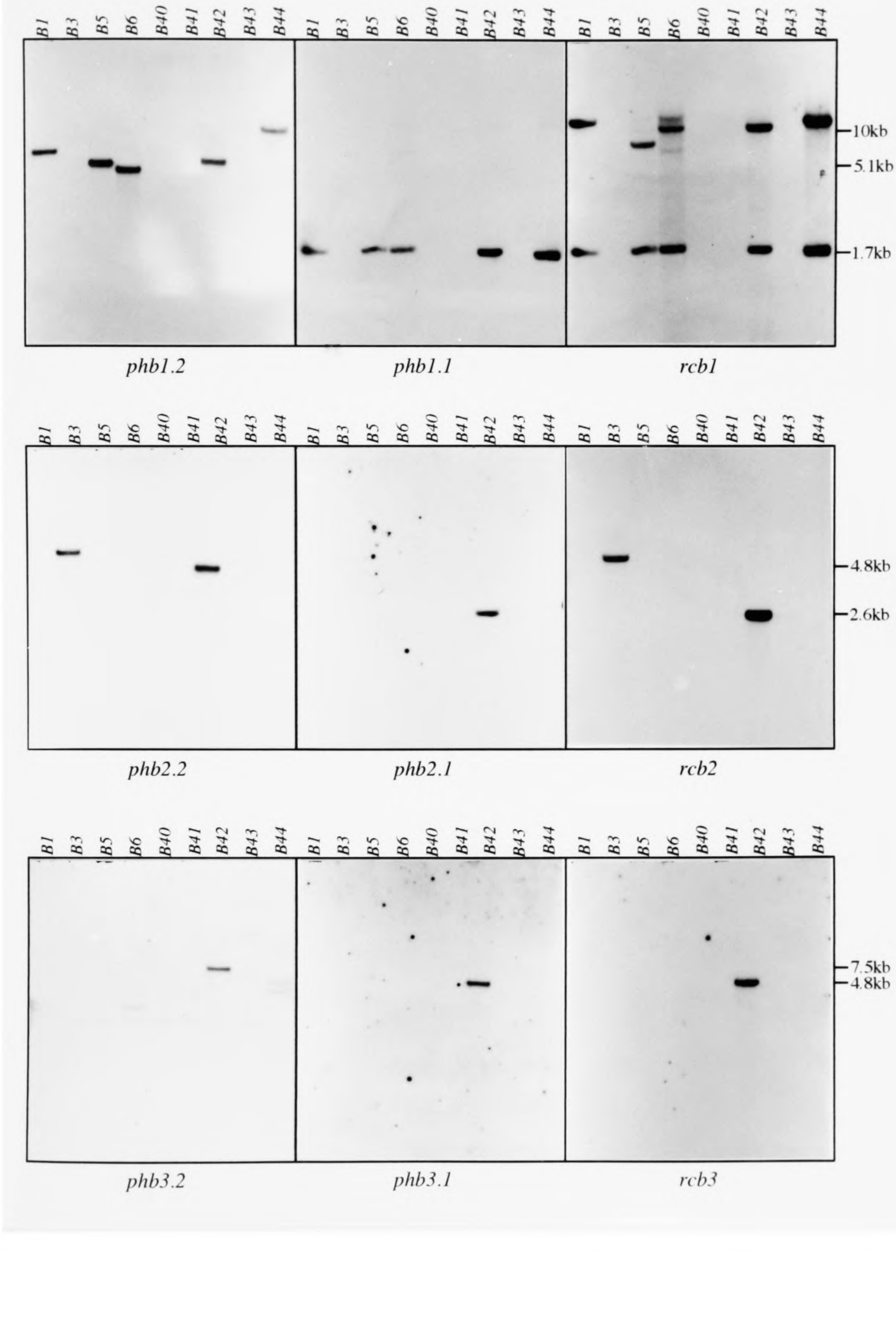
The prediction that the three genes in each group are kept together as an allelic set is strongly supported by the data. The $rcb1^{B42}$, $phb1.1^{B42}$ and $phb1.2^{B42}$ sequences hybridised to genomic DNAs from strains having *B1*, *B5*, *B6*, *B42* and *B44*. None of the three genes hybridised to genomic DNAs from the *B3*, *B40*, *B41* and *B43* strains. This suggests that *B1*, *B5*, *B6*, *B42* and *B44* all share the same allele complex of this set of genes but that *B3*, *B40*, *B41* and *B43* have a different allele complex. Similarly, the three probes used to identify the group 2 genes in *B42*, all hybridised to genomic DNA from the *B3* strain indicating that all three genes are present in the *B3* locus. These three genes were not, however, detected in any of the other DNAs. The relatively weak signal detected on hybridisation of $phb2.1^{B42}$ to *B3* genomic DNA (Figure 5.1A) may be explained by two possible factors. Firstly, the amount of genomic DNA from the *B3* strain on the gel was lower than that of the *B42* strain (photograph not shown), and secondly the exposure time was limited due to a background of non-specific hybridisation. The probes representing the three group 2 genes from *B6* only hybridised to DNA from a *B6* strain, indicating that these alleles were not present in any of the other *B* loci. This was also true for the *B42* alleles of the group 3 genes. The group 3 genes from *B6*, however, all hybridised to corresponding genes in the *B40* and *B41* loci. The data is summarised in Figure 5.2.

Table 5.1 List of the DNA fragments used for the hybridisation analysis

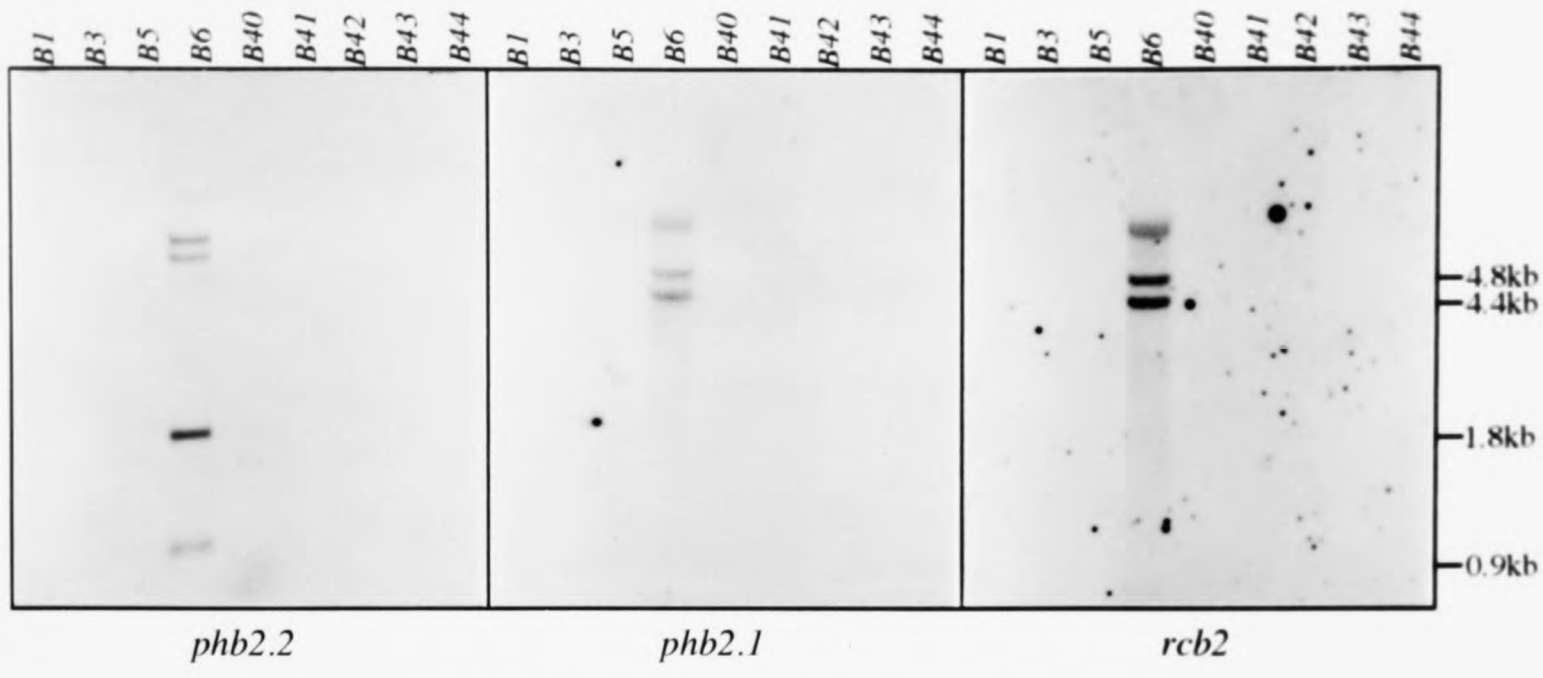
Locus	Gene	Plasmid	Fragment used as probe
B42	*phb1.1*	pJH3	1.7kb *Hin*dIII
B42	*phb1.2*	pJH38	900bp *Eco*RI
B42	*rcb1*	pJH39	2.6kb PCR product
B42	*phb2.1*	pJH31	1.0kb *Hin*dIII-*Sal*I
B42	*phb2.2*	pJH25	1.8kb *Hin*dIII-*Eco*RI
B42	*rcb2*	pJH30	1.6kb *Hin*dIII-*Sal*I
B42	*phb3.1*	pJH26	1.0kb *Hin*dIII-*Pst*I
B42	*phb3.2*	pJH10	1.4kb *Hin*dIII-*Eco*RI
B42	*rcb3*	pJH33	1.6kb *Pst*I
B6	*rcb1*	pSB6-18	1.0kb *Sal*I
B6	*phb2.1*	pSB6-30	900bp PCR product
B6	*phb2.2*	pSB6-34	2.1kb *Eco*RI
B6	*rcb2*	pJB6-29	RT-PCR product
B6	*phb3.1*	pCB6-33	2.1kb *Xho*I-*Bam*HI
B6	*phb3.2*	pSB6-36	1.3kb PCR product
B6	*rcb3*	pSB6-32	400bp *Hin*dIII-*Eco*RI
B3	*B3*L	pJH43	2.6kb *Pst*I-*Sac*II
B3	*B3*R	pJH44	2.3kb *Pst*I-*Sac*II

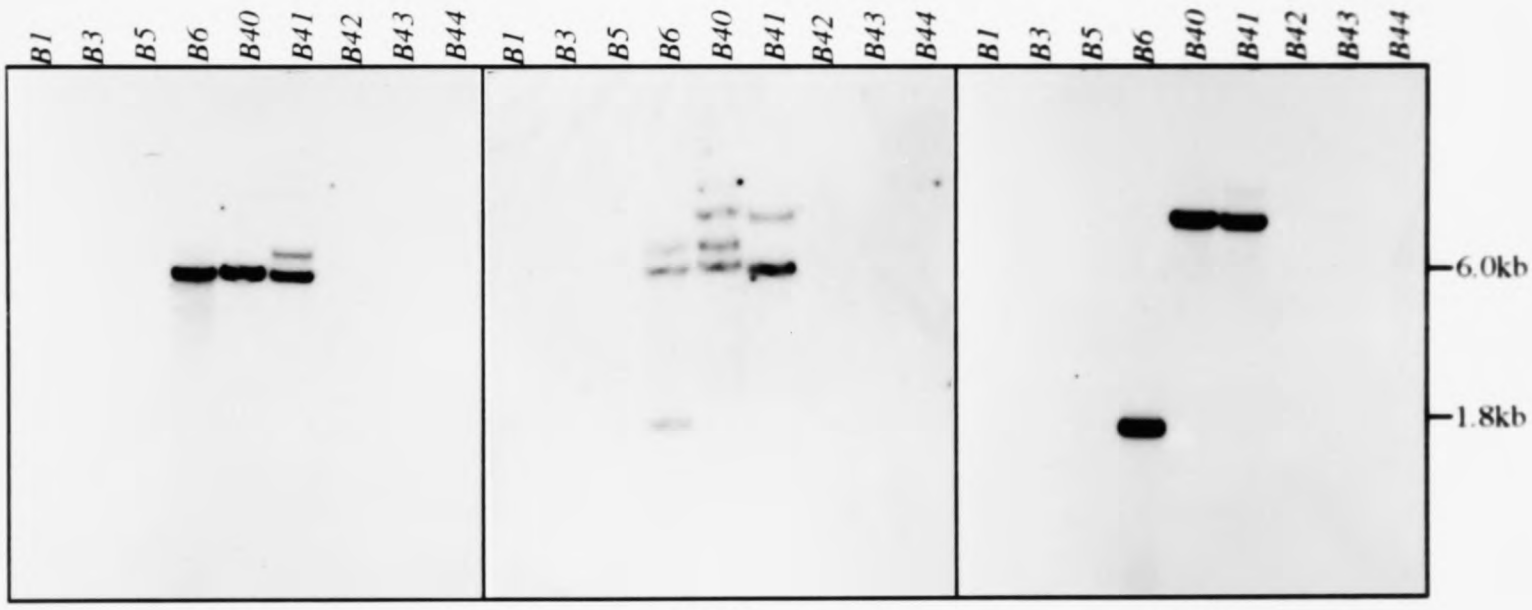
Figure 5.1 Southern blots of *Hin*dIII digested genomic DNA of *C. cinereus* strains 218 (*B1*), LCO12 (*B3*), TC4 (*B5*), LT2 (*B6*), LCO7 (*B40*), LCO5 (*B41*), LN118 (*B42*), AT8 (*B43*) and NL1 (*B44*), probed with cloned DNA sequences containing (**A**) *phb1.1*, *phb1.2*, *rcb1*, *phb2.1*, *phb2.2*, *rcb2*, *phb3.1*, *phb3.2* and *rcb3* pheromone and receptor genes from the *B42* locus, (**B**) *phb2.1*, *phb2.2*, *rcb2*, *phb3.1*, *phb3.2* and *rcb3* from the *B6* locus, (**C**) *B3*L and *B3*R from the *B3* locus.

A

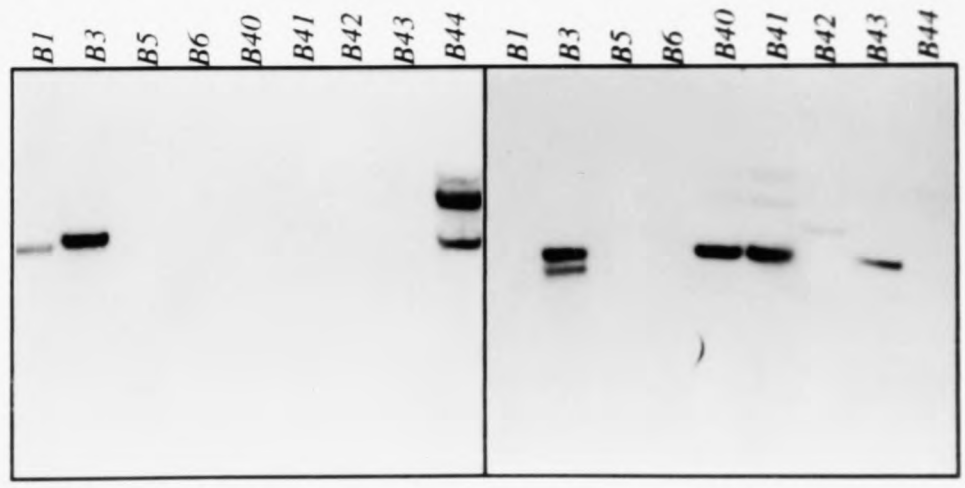

B1 B3 B5 B6 B40 B41 B42 B43 B44
10kb
5.1kb
1.7kb
phb1.2
phb1.1
rcb1
4.8kb
2.6kb
phb2.2
phb2.1
rcb2
7.5kb
4.8kb
phb3.2
phb3.1
rcb3

B


B1 B3 B5 B6 B40 B41 B42 B43 B44
4.8kb
4.4kb
1.8kb
0.9kb
phb2.2
phb2.1
rcb2


B1 B3 B5 B6 B40 B41 B42 B43 B44
6.0kb
1.8kb
phb3.2
phb3.1
rcb3

C


B1 B3 B5 B6 B40 B41 B42 B43 B44
B3L
B3R

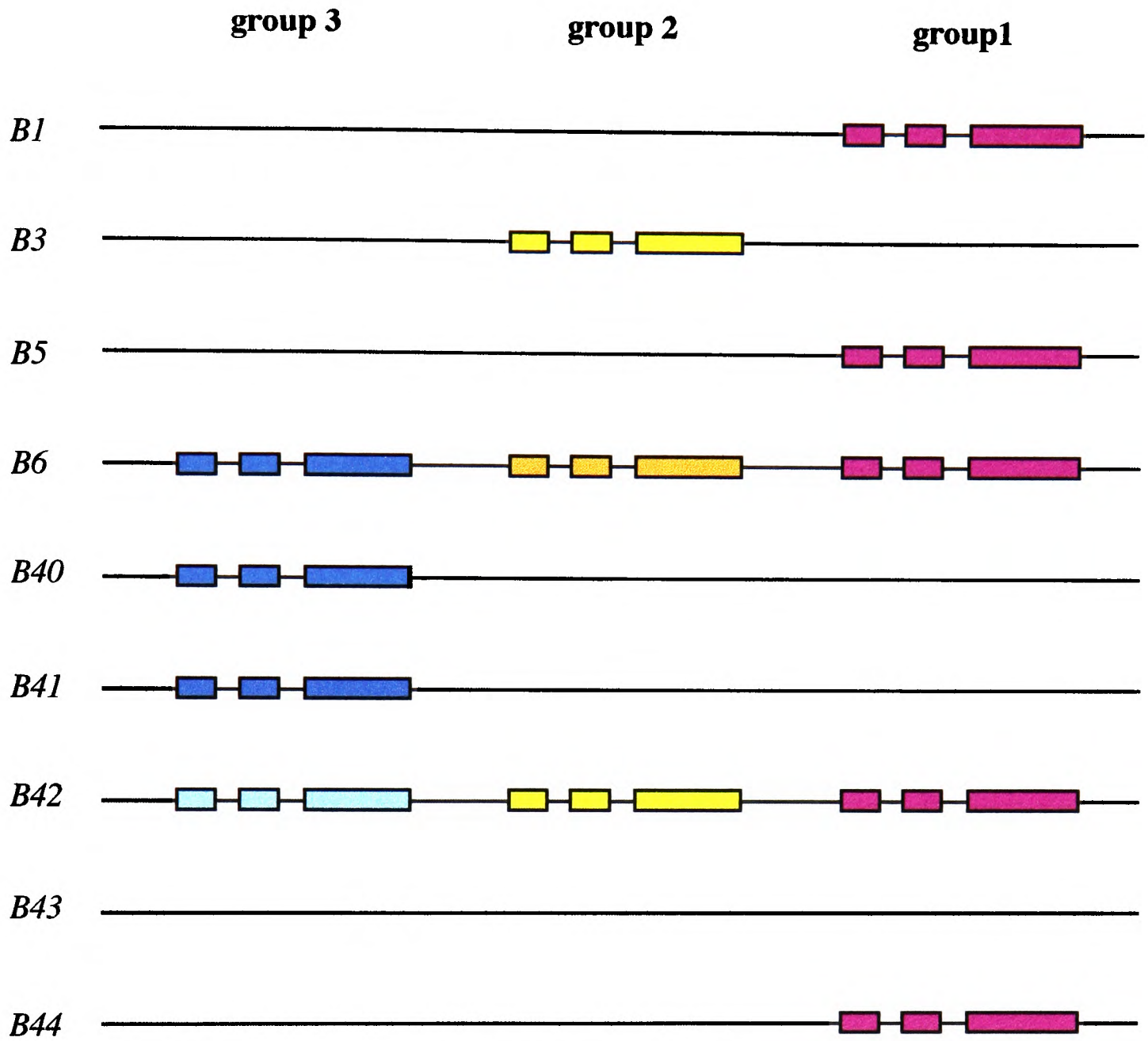


Figure 5.2 Identification of *B42* and *B6* genes in genomic DNAs containing the *B1*, *B3*, *B5*, *B6*, *B40*, *B41*, *B42*, *B43* and *B44* loci by Southern blot analysis. The three groups of genes are assumed to maintain the same relative position in each locus. Colours are used to identify homoalleles.

The results of the hybridisation analysis clearly indicated that the three genes in any one group are kept together as an allele complex. If one of the genes was detected in another *B* locus, the other two genes were also detected.

The data summarised in Figure 5.2, assumes that there are three sets of genes present in all *B* loci and that their relative positions are conserved. The fact that *B42* and *B3* appear to share the middle set of genes provided an excellent opportunity to test this prediction. It was possible to clone the sequences flanking the group 2 genes present in *B3* and to show that these were of the predicted size to contain the group 1 and group 3 genes. It was also possible to demonstrate that they contained genes that conferred *B* mating specificity in transformation experiments, and that these sequences could be detected by hybridisation in some but not all other *B* loci.

The strategy used to clone the group 1 and group 3 genes from the *B3* locus was based on the homology that was shown to exist between the group 2 genes of the *B3* and *B42* loci and the homologous sequences flanking the *B3* locus which had been identified and sequenced by (O'Shea, 1995).

Four PCR primers were designed that could be used to amplify the sequences of interest. The strategy is illustrated in Figure 5.3. In order to amplify the group 3 genes from *B3*, one primer was designed to anneal to the homologous border to the left of the *B* locus, and the other primer was designed to anneal to the *phb2.2* allele which is shared by *B3* and *B42*. One of the primers annealed to the *phb2.2* allele shared in *B3* and *B42* and this together with a primer that annealed to the homologous border to the left of the locus was predicted to amplify a fragment containing the group 3 genes from *B3*. The other two primers annealed to the shared *rcb2* allele in *B3* and *B42* and the homologous border to the right of the *B* locus and were predicted to amplify a fragment containing the group 1

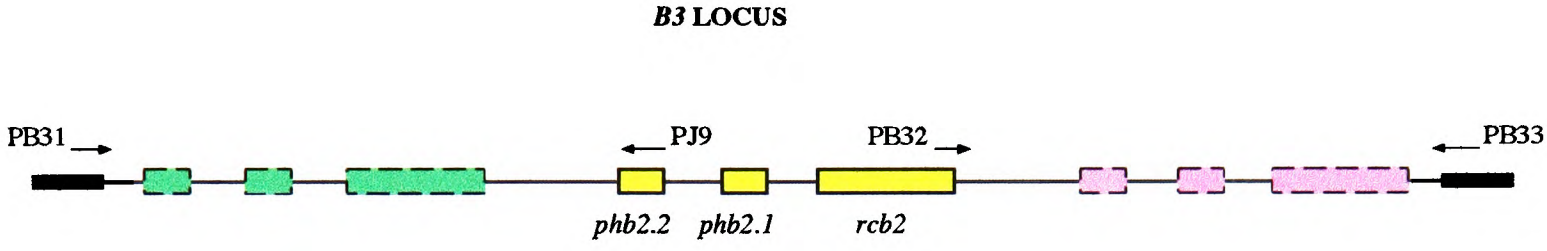


Figure 5.3 Diagram illustrating the PCR strategy used to amplify the left and right sets of genes from the *B3* mating type locus. Homologous flanking regions are represented by thick lines. Primers used are indicated by arrows. Dashed lines are used to represent the genes that are predicted to be present in the *B3* locus.

genes from *B3*. Genomic DNA was isolated from the strain LCO12 (*A2B3*) and used as template for PCR amplification. Products from both amplifications were cloned into pGEM-T Easy (Promega) to give plasmids pJH43 and pJH44 containing the left (*B3*L) and right (*B3*R) regions of the *B3* locus respectively. The ends of the cloned fragments were sequenced using the amplification primers and the data confirmed that the correct products had been amplified and cloned.

Each group of pheromone and receptor genes within the *B6* and *B42* loci extends over some 5-6kb of DNA sequence. The fragments amplified by PCR were 8.5kb (*B3*L) and 9.0kb (*B3*R) respectively, which would be a size expected to contain three genes and the flanking sequences of the adjacent genes of the *B3* locus. To test whether the sequences isolated contained *B*-specific sequences that might be present in other *B* loci, restriction fragments that derived from the internal region of the amplified sequences were used to probe the filters used previously to identify shared *B42* and *B6* genes in other *B* loci.

The *B3*L probes hybridised to homologous sequences in the genomic DNAs from *B1*, *B3* and *B44* strains but failed to hybridise to the other DNAs. The *B3*R probes hybridised to genomic DNAs from *B3*, *B40*, *B41* and *B43* strains. This pattern of hybridisation would be consistent with the *B3* locus sharing the group 3 gene alleles with *B1* and *B44* and the group 1 gene alleles with *B40*, *B41* and *B43*.

The predicted allele constitution of the nine strains represented on the Southern blots is presented in Figure 5.4. It is particularly interesting to note that there appear to be only two allelic versions of the group 1 genes. Five of the *B* loci tested contain the alleles present in *B42* and *B6* whereas the other four loci share the alleles present in *B3*. Three of the *B* loci share the alleles of the group 3 genes found in *B6*, and three share the alleles of these genes present in *B3*. For *B1* and *B44*, only the group 2 genes can have different

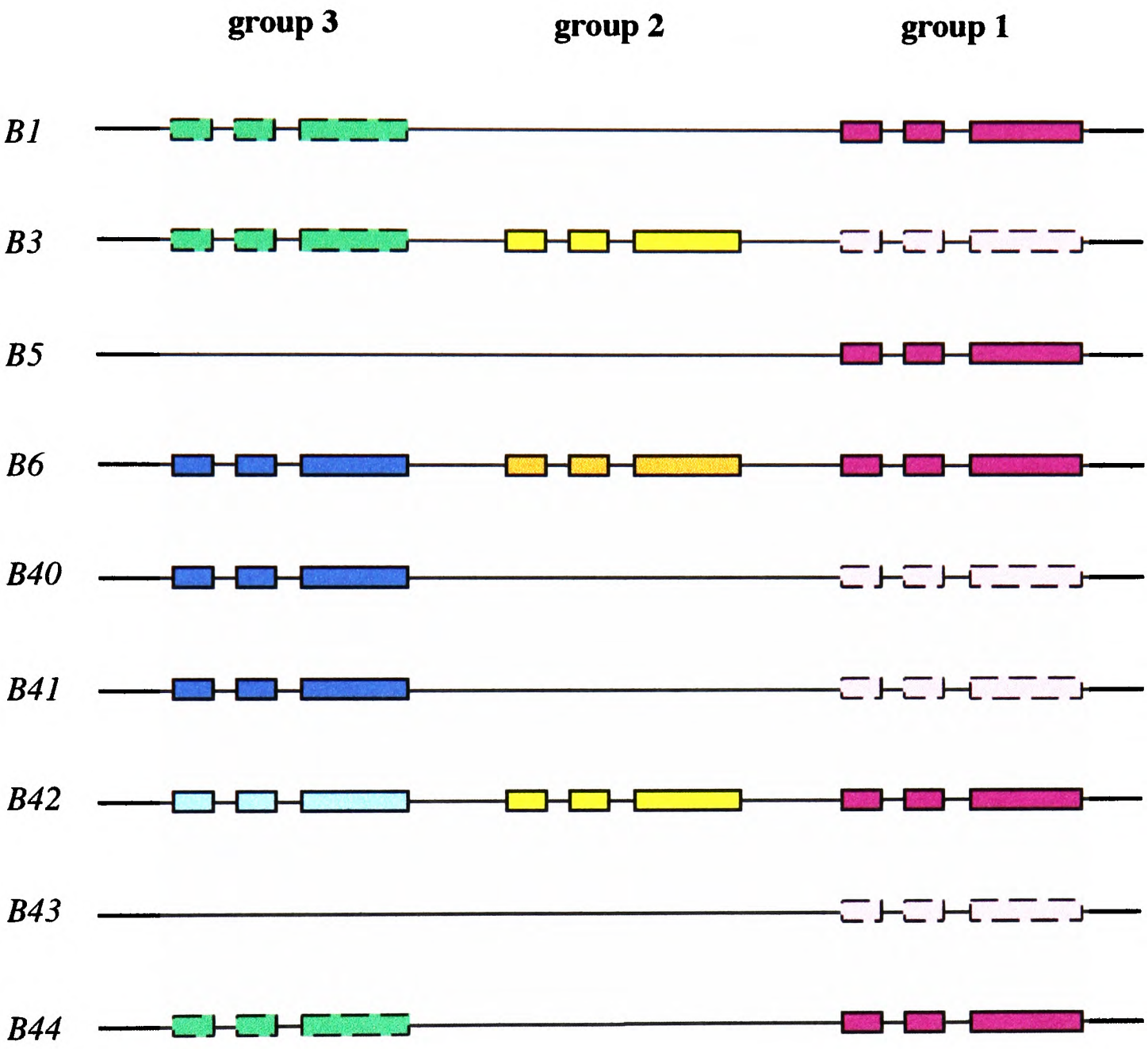


Figure 5.4 Identification of shared alleles in different *B* loci using probes derived from the cloned *B42* and *B6* genes and the *B3*L and *B3*R fragments of the *B3* locus. Dashed boxes indicate genes predicted to be present but not sequenced. Different colours are used to represent DNA homology.

alleles and confer the differences in *B* mating specifcity. This should also be true for *B40* and *B41* which also appear to share the same alleles of the group 1 and group 3 genes.

5.3 Transformation of cloned genes into host strains having five different *B* mating specificities

A correlation between DNA homology and failure to activate *B*-regulated development in transformation was noted in the transformation experiments in which *B42* genes were introduced into *B6* and *B3* host strains. A more extensive analysis was now undertaken to correlate the results of the Southern blot analyses summarised in Figure 5.4 with the ability to change B mating specificity following transformation. All the *B42* and *B6* genes were used in this analysis together with the *B3*L and *B3*R fragments from the *B3* locus. The five host strains used had *B1*, *B3*, *B6*, *B40*, and *B42* mating specificities.Table 5.2 summarises the results of the hybridisation analysis and the transformation analysis for all cloned sequences tested.

It is evident that there is an almost complete correlation between ability of a gene to change the *B* mating specificity of the host strain and lack of DNA homology between host and donor DNAs. For example, the cross-hybridising group 1 genes of *B42* and *B6* also hybridise to corresponding genes in the *B1* locus and none of the genes can activate the *B* pathway in the *B1* host. They can, however, activate the *B* pathway in the *B3* and *B40* hosts which lack homologous gene sequences. In contrast the *B3*R sequence fails to hybridise to *B1*, *B6* and *B42* genes but can activate the *B* pathway in hosts having these specificities. The *B3*R sequence is found in *B40* and, as expected, cannot activate the *B* pathway in a *B40* host. The only exceptions to the predictions made on the basis of DNA non-homology involved two of the pheromone genes from *B42*, $phb3.1^{42}$ and $phb3.2^{42}$. Neither of these genes is present in *B6*, *B3* or *B1*, but $phb3.1^{42}$ was unable to activate the

Table 5.2 Comparative hybridisation and transformation analysis of the cloned *B* genes in *B42*, *B6*, and the uncloned loci *B1*, *B3* and *B44* . H, positive hybridisation and T, fused clamp cell formation after transformation and mating with tester strain. Transformation experiments yet to be completed are left as blank spaces.

		Host strains									
		218 (*B1*)		LCO12 (*B3*)		LT2 (*B6*)		LN118 (*B42*)		NL1 (*B44*)	
Gene	Origin	H	T	H	T	H	T	H	T	H	T
phb1.1	*B42*	+	-	-	+	+	-	+	-	+	-
phb1.2	*B42*	+	-	-	+	+	-	+	-	+	-
rcb1	*B42*	+	-	-	+	+	-	+	-	+	-
phb2.1	*B42*	-	+	-	+	-	+	+	-	-	+
phb2.2	*B42*	-	+	-	+	-	+	+	-	-	
rcb2	*B42*	-	+	-	+	-	+	+	-	-	+
phb3.1	*B42*	-	+	-	+	-	-	+	-	-	+
phb3.2	*B42*	-	-	-	-	-	+	+	-	-	
rcb3	*B42*	-	+	-	+	-	+	+	-	-	+
phb1.1	*B6*	+	-	-	+	+	-	+	-	+	-
phb1.2	*B6*	+	-	-	+	+	-	+	-	+	-
rcb1	*B6*	+	-	-	+	+	-	+	-	+	-
phb2.1	*B6*	-	+	-	+	+	-	-	+	-	+
phb2.2	*B6*	-	+	-	+	+	-	-	+	-	+
rcb2	*B6*	-	+	-	+	+	-	-	+	-	+
phb3.1	*B6*	-	+	-	+	+	-	-	+	-	+
phb3.2	*B6*	-	+	-	+	+	-	-	+	-	+
rcb3	*B6*	-	+	-	+	+	-	-	+	-	+
*B3*L	*B3*	+	-	+	-	-	+	-	+	+	-
*B3*R	*B3*	-	+	+	-	-	+	-	+	-	+

B pathway in a *B6* host and *phb3.2*^{42} was unable to activate the *B* pathway in the *B3* and *B1* hosts which share the same alleles of group 3 genes. This unexpected behaviour of the *phb3.1* and *phb3.2* genes will be discussed in Chapter 7 when the sequences of *B6* and *B42* genes will be considered in more detail.

Without a *B*-null host in which to test the compatible gene combinations that activate *B*-regulated development it is not possible to prove conclusively that the three sets of genes within the *B* locus are functionally independent. However, the data summarised in Figure 5.4 and Table 5.2 argue strongly that they are. The ability to activate *B*-regulated development following transformation is dependent on the host not having the same allele of the transforming gene and not on which other genes may be present.

Assuming that the *B* mating type locus derives its many specificities from three sets of paralogous genes in the same way that the *A* mating type locus does, it would require only four alleles of two sets of genes and five of the other in order to generate the predicted 79 *B* mating specificities. The data obtained in this study suggest that there may be two alleles of the group 1 genes, at least four alleles of the group 3 genes but many more alleles of the group 2 genes.

5.4. Regulation of the *B* genes in response to pheromone stimulation

Northern analyses described in Chapter 3 showed that transcription of the majority of the *B* genes was induced when *A*-regulated development was activated. A parallel series of experiments tested the effect of pheromone stimulation on *B* gene transcription. In these experiments RNA was isolated from monokaryons and from a strain that was heterozygous for the *B* mating type genes (*B*-on;*B6/B42*, JH2). The probes were more specific and were designed to identify only the transcripts of the six pheromone genes.

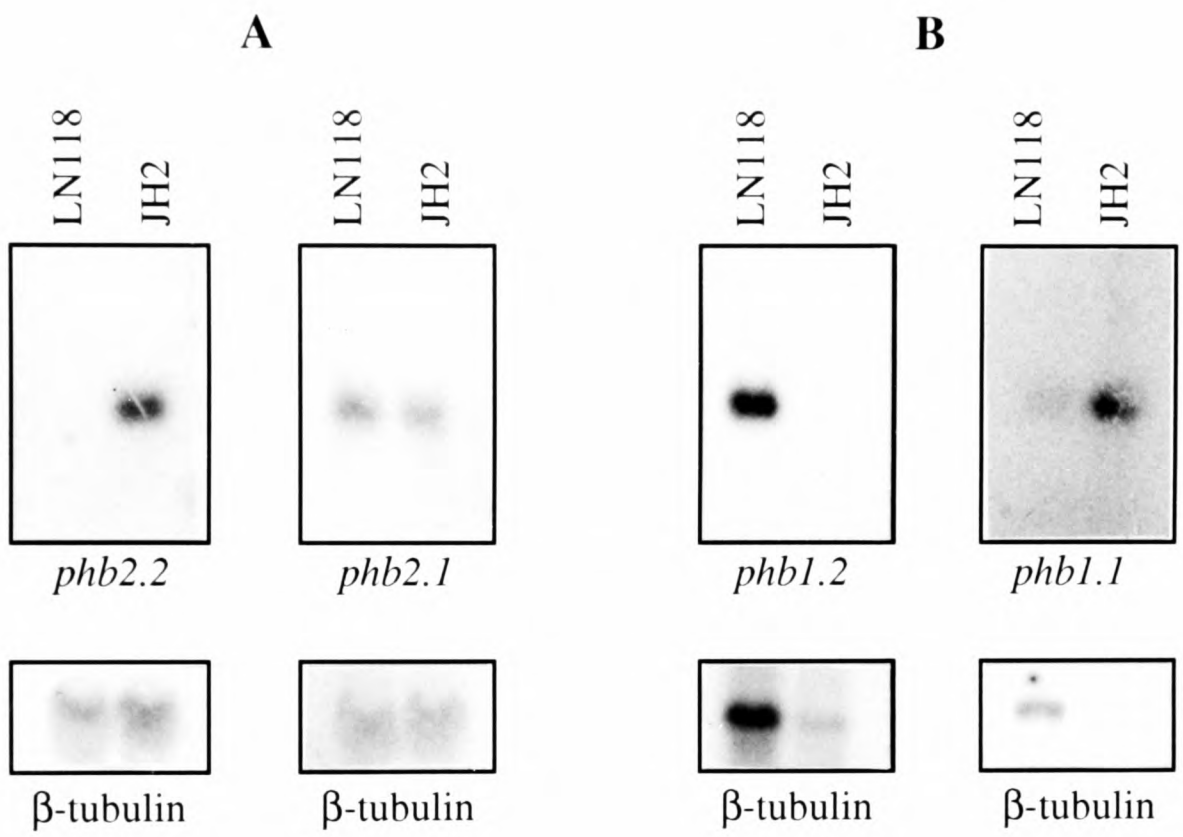


Figure 5.5 Autoradiographs of (**A**) northern blots of total RNA probed with pJH25 (*phb2.2*) and pJH18 (*phb2.1*). The RNA in the left lane was isolated from a monkaryotic strain, LN118 (*A42B42*), and the RNA in the right lane was isolated from a *B*-on strain, JH2 (*A42B6/B42*), in which the *B*-regulated steps of development had been activated by the introduction of a compatible *B* locus; (**B**) northern blots of poly $(A)^+$ RNA isolated from the same LN118 and JH2 strains, probed with pJH38 (*phb1.2*) and pJH3 (*phb1.1*). The same blots probed with the constitutively expressed β-tubulin gene as a loading control are also shown.

Transcripts of four of the genes, *phb3.1, phb3.2, phb2.2* and *phb1.1*, were barely detectable in RNA obtained from the monokaryon but showed strong induction in the *B*-on mycelium. This pattern of regulation would be expected if, as in the yeasts, *S. cerevisiae* and *S. pombe*, pheromone stimulation acts to amplify the signal by increasing the level of transcription of the pheromone and receptor genes. This is also true for *U. maydis* (Hartmann et al., 1996; Urban et al., 1996). Transcription levels of *phb2.1* and *phb1.2* were not as expected. *phb2.1* was equally transcribed in both mycelial types but *phb1.2*, surprisingly, was actually down-regulated in the *B*-on mycelium. Illustrated in Figure 5. 5 are representative autoradiographs to show the transcripts of *phb2.2, phb2.1, phb1.2* and *phb1.1*. The unusual regulation of two of the pheromone genes is not easy to interpret and presumably relates to the position of regulatory squences which have yet to be defined.

Chapter 6

Identification of a membrane spanning transport protein within the *B42* locus

6.1 Introduction

This chapter describes the identification of a tenth gene within the *B42*-specific DNA sequence that flanks the *B42* pheromone and receptor genes. This gene was shown by sequence analysis to be a membrane-spanning transport protein belonging to the major facilitator superfamily (MFS). It is not known whether this gene has a role in mating type gene expression or whether its position within *B42*-specific DNA sequence is fortuitous.

The hybridisation analysis used to identify the borders of the *B42* and *B6* loci revealed that there was a region to the left of the pheromone and receptor genes that was still *B42*- and *B6*-specific, yet did not confer any mating type specificity in transformation experiments. Particularly puzzling was the fact that this region extended for 17kb in the *B42* DNA but only 6.5kb in the *B6* DNA (see Figure 6.1).

6.2 Identification of a tenth transcript derived from *B42*-specific DNA

To determine whether the *B42*-specific sequence contained genes other than those encoding the pheromones and receptors, a search for other transcripts was initiated using the eight fragments detailed in Figure 6.2 as probes onto northern blots. Only three probes identified transcripts. The 3.9kb *Eco*RI fragment present in pJH14 hybridised to

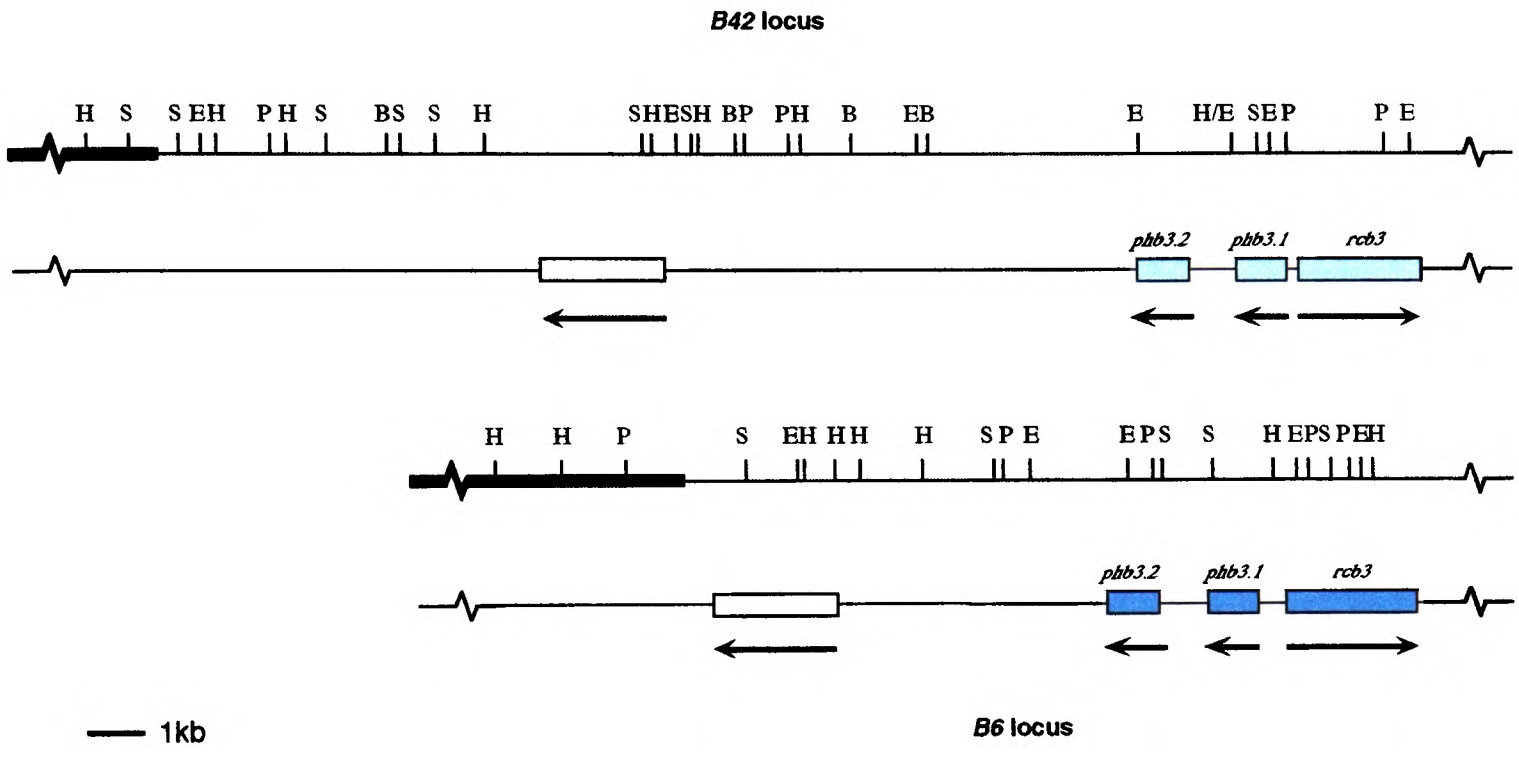


Figure 6.1 Comparison of the region to the left of the *B* specificity genes in the *B42* and *B6* loci. Homologous flanking sequences shown as thick lines. B, *Bam*HI, E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I. Arrows indicate direction of gene transcription. The position of a tenth gene identified by transcript analysis is shown as an open box.

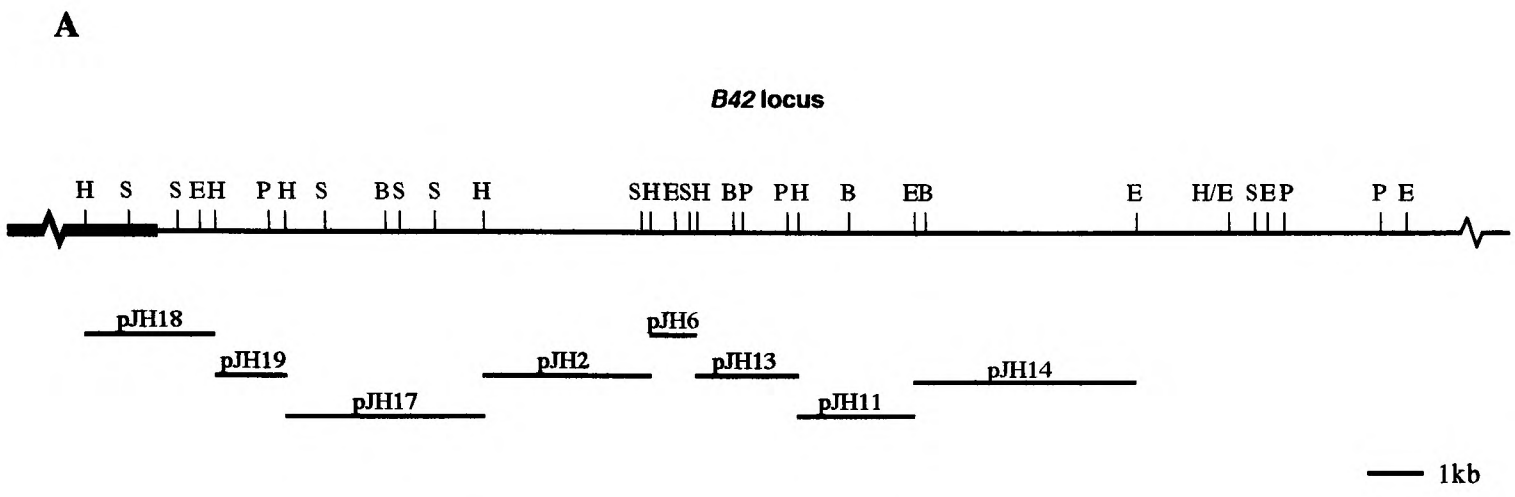


Figure 6.2 Transcript analysis of the unique sequences to the left of the *B42* specificity genes. (**A**) Restriction map of the region. Homologous flanking sequence shown as a thick line. Enzymes are as in Figure 6.1. Cloned sequences used as hybridisation probes are indicated below the map. Autoradiographs of (**B**) northern blots of poly $(A)^+$ RNA isolated from LN118 and JH1 probed with pJH2 and pJH6; (**C**) northern blot of poly $(A)^+$ RNA isolated from LN118 and JH2 probed with pJH2. Blots probed with pCTRL or β-tubulin as loading controls are shown underneath. Strains and loading controls are as described in Figure 3.11 and Figure 5.5.

the *phb3.2* pheromone gene transcript. The two other fragments, a 2.9kb *Hin*dIII fragment in pJH2 and a 0.8kb *Hin*dIII fragment in pJH6, identified a 2.0kb transcript. Although the size of the transcript was similar to that of the receptor genes, there was no clear difference in regulation in response to activation of the *A*- or *B*-regulated pathways of development (Figure 6.2). Moreover, the large 4.2kb *Bam*HI fragment (pJH22) which covers the entire region from which the transcript is derived failed to display any *B42*-specific mating type function when introduced into a *B6* or *B3* host.

6.3 A gene within the *B42* locus encodes a putative drug efflux pump

The 4.2kb *Bam*HI fragment containing the additional *B42*-specific gene was sequenced in order to look for possible open reading frames. The fragment was partially digested with DNAase I, and 500bp-1kb fragments size-selected for cloning into pUC18. Sequence data derived from 24 clones were aligned into a contiguous DNA sequence covering the region of interest using the GELASSEMBLE program (GCG, version 8.1).

Attempts to isolate a cDNA sequence corresponding to the *C. cinereus* gene using the RT-PCR approach were unsuccessful, and for the present the amino acid sequence of the protein remains tentative. The coding sequence of the gene was deduced using the conserved splice borders to predict where the introns were located and BLAST searches to predict protein homologies. The deduced ORF encodes a protein of 546 amino acids and the complete sequence is shown in Figure 6.3.

A BLAST search (Altschul *et al.*, 1997) of the translated sequence revealed significant identity (BLAST E value < $2e^{-30}$) to several bacterial and fungal proteins. The strongest identity was with Yan6, a hypothetical 63.5kDa protein from *Schizosaccharomyces pombe* which has yet to be characterised. Other strong identities were found with the

product of the *TOXA* gene of the plant pathogenic fungus *Cochliobolus carbonum* (Pitkin *et al.*, 1996) which is a proposed HC-toxin efflux pump and Bmr3, a multidrug transporter from *Bacillus subtilis* (Ohki and Murata, 1997). The search also revealed significant similarity to two other characterised proteins: Sge1p, a protein conferring drug resistance in *Saccharomyces cerevisiae* (Ehrenhofer-Murray *et al.*, 1994) and LfrA, a protein which confers low-level fluoroquinolone resistance in *Mycobacterium smegmatis* (Takiff *et al.*, 1996).

6.4 The *B42* gene appears to encode a member of the major facilitator superfamily

The discovery of a gene encoding a putative multidrug transporter was initially exciting. The presence of a CaaX motif at the carboxy-termini of the pheromone precursors suggests that the pheromone will be isoprenylated and by analogy with the isoprenylated **a**-factor of *S. cerevisiae* and M-factor of *S. pombe*, the mature peptide would be expected to be secreted across the the cell membrane. Mature **a**-factor and M-factor are not secreted by the classical secretory pathway but are actively transported across the membrane by an ABC (ATP-binding cassette) transporter. This protein is encoded by the *STE6* gene of *S. cerevisiae* and the *mam1* gene of *S. pombe* (McGrath and Varshavsky, 1989; Christensen *et al.*, 1997). Ste6p and Mam1 belong to a superfamily of proteins that have a wide range of substrates that include small hydrophobic drugs and which use ATP-hydrolysis to pump these molecules out of the cell (Higgins, 1992).

A more detailed analysis of the *B42* gene, now designated *mfs1*, indicates that it belongs to a different family of transporters from the **a**-factor and M-factor transporters. The *C. cinereus* protein lacks certain consensus sequences found in ABC proteins (Decottignies and Goffeau, 1997), and the proteins that were identified as homologous to Mfs1 in

(*mfs1*)

```
GTGGGCACATAAATGCCATCAATGGCTTTGGGCAACAACGCCAAACTGTCCCAGTTCCGGTATCGCTACTTGATTCTT  78

CCATCCTTCAACGTACGTCTCGCGCAGCTGCTAGAAGCATTTTTTTTTTATCGCCATGGTCCGCAACAAAAGAATTCGA 156

TGCTCCACACACAGCTGTGCACACAAACGTTCTTCCTGGACTCGAACGCACCCACACACCTTAAACTAGTCGACAGTT 234

TCGCCGCTTGAACAGCGACGACCACAACCCTGGCACTACTTTCCGAACTCGTCGCATTGGCCGATCCACCTAGATACG 312

AACCGAAACAGCTTCGATGATACATCTTTATTGAGACGTAATCCACACCGCCTACTCTCGTTCTGACGTTCTTTGGTG 390

TGTCCCCACAAATARGGCCACCTGCATGGTCAAACACAGCATCGATCTCGGAGAAGAAGCTTATAGGAGAGAAGGAAT 468

CTCTCACTCGTCCACCAAACCCTGMCTTCGCCATGACCACCTATGCCTCTGAATCCCGAGAGCAACGCCTAGGACCTC 546
                                M  T  T  Y  A  S  E  S  R  E  Q  R  L  G  P  P

CAAGTCCTAGTGGGAAAAAGGGGGCCAGGTTTTGGCTCATATTTGCGTCTCTCAGTGTCACCCTGTTCCTCAGCGCTC 624
  S  P  S  G  K  K  G  A  R  F  W  L  I  F  A  S  L  S  V  T  L  F  L  S  A  L

TTGACTACgtaggtcgttcctccaatgatcacgttccgcgctaatacctttgttagACCGCTCTCGCCTCCGCGCTTC 702
  D  Y                                                    T  A  L  A  S  A  L  P

CTGTCATTGTCCATGATTTGAAGGGCTCGGACTTCTCGTGGGTTGGAACTGCATATGCGgtaagtgttccagcgcara 780
  V  I  V  H  D  L  K  G  S  D  F  S  W  V  G  T  A  Y  A

ttaccccgtccactttcaagttctaatcagactagCTCTCGTCGACGGCGTTCCTGCCCATGAGTGGCGGACTTGCTC 858
                                    L  S  S  T  A  F  L  P  M  S  G  G  L  A  Q

AGATTGTTGGAAGGCGAATCGCGTTTATCATCGCGCTGGGCTTCTTTGCACTGGGAAGTGCGCTATGTGGAGCGGCCA 936
  I  V  G  R  R  I  A  F  I  I  A  L  G  F  F  A  L  G  S  A  L  C  G  A  A  R

GATCGATGAATTGGCTGATTGGTGCTCGGGgcacgtcgttcccattcgaaaggccggtttgaagactgacgtttctca
  S  M  N  W  L  I  G  A  R

gCGATCCAGGGTGTGGGCGGGGGCGCCATCATCTCATTGGCAAGCATCGTTATTAGCGACTTGgtaagcctttcccga 1094
A  I  Q  G  V  G  G  G  A  I  I  S  L  A  S  I  V  I  S  D  L

ctgactgaggaatttcaactcatcatttcgttcagGTCACTTTGCAGGAGCGAGGAACCTACAACGGTTTGCTTGGAA 1170
                                    V  T  L  Q  E  R  G  T  Y  N  G  L  L  G  M

TgtacgctttggcatcccttcaaactggttgtttgtctaacagtgactctactagGACCTGGGCTGTGGCGGTTGCTC 1248
                                                        T  W  A  V  A  V  A  L

TTGGCCCCATTGTGGGAGGCGCTCTAGCAGAGAACGGGGCATGGCGTTGGCTTTTTTgtaggacaaattacatccatc 1326
  G  P  I  V  G  G  A  L  A  E  N  G  A  W  R  W  L  F

caaacatattgactgacgcgaatggatagATCTCAACCTCCCGATATGTGCGGTCTCGTTCATTCTCGTCGTCGTCT 1404
                               Y  L  N  L  P  I  C  A  V  S  F  I  L  V  V  V  F

TCATGTCACTTCCAACGCCTGCGGGAACGTTCAAAGAAAAGTTTTTTGCAATGGACTGGATgtaagcatggttatttc 1482
  M  S  L  P  T  P  A  G  T  F  K  E  K  F  F  A  M  D  W  I

agattcgccctcccacactgaactgtctgttagTGGTAATTTCCTTGTCATTGGCAGCTCGAGTGCACTTGTCATCGC 1560
                                    G  N  F  L  V  I  G  S  S  S  A  L  V  I  A

TCTTACTTGGGGAGGTGGTCGGTATGGATGGGGCGAAGCCCCAGTTCTGGTTCCGTTAATTCTTGGAATCGTTGGCTT 1638
 L  T  W  G  G  G  R  Y  G  W  G  E  A  P  V  L  V  P  L  I  L  G  I  V  G  L

GGTAGTTTGGCTATTTTACGAGGTCAAATATTCTTCACATCCGACCgtaagttgataagtcgattggcgttctccgct 1716
 V  V  W  L  F  Y  E  V  K  Y  S  S  H  P  T

tcaacccagctgattttcgttcagGTTGCCCCCGAATTACTCGGAAACCGAACGAGCGTGTCCGGgtatgctcgcttc 1794
                          V  A  P  E  L  L  G  N  R  T  S  V  S  G

acatcttctattggtctacataatctcactttgtctaagTTATCTACAGGTTTTCTTCAACTCAGTCATCGTTGTCGC 1872
                                        Y  L  Q  V  F  F  N  S  V  I  V  V  A

TTGGGTCTgtgagtgaaaaagtgacctttcataagcatgttgttgactatccattcacgcagACTACATGCCCGTGT 1950
 W  V                                                    Y  M  P  V  Y

ACTTCCAGGCCTGCAAGGATGCTGGCCCTATCCAGTCCGGAATCGATATGTTTGGAGGAACTTTGTCTATCGGgtacg 2028
  F  Q  A  C  K  D  A  G  P  I  Q  S  G  I  D  M  F  G  G  T  L  S  I  G

tagattgctgactcctttcgatgcttaacgagcataacatcttgatttcaagTCCGTTCATCGTCCTCGCTGGGGTTA 2106
                                                    P  F  I  V  L  A  G  V  S

GCGTGACAGTTTTGAAGATTTACCGTCCACAACTATGGATTGGCTGGGTCCTTCTGACTGCCGGAATGGGAATTATGG 2184
  V  T  V  L  K  I  Y  R  P  Q  L  W  I  G  W  V  L  L  T  A  G  M  G  I  M  G

GCCTATTGAGAGCAAATACACCGCTCGCCCAAGGAATCGGGTTCTCAGTAATCATCGCATCCGGCGCCGGCATGATAT 2262
  L  L  R  A  N  T  P  L  A  Q  G  I  G  F  S  V  I  I  A  S  G  A  G  M  I  F

TCGgtaggacatttctttatcacctggttggaatcgacgttcatcatccagtttagCGGTTACTTACTTCCCTGTCCT 2340
                                                        A  V  T  Y  F  P  V  L

ATCCCCGCTTCCGGTTACATTGAACGCCCAAGCATTGGCGCTGTTCGCGTTTACGCGTGCGCTTGCCCAGgtacgctt 2418
 S  P  L  P  V  T  L  N  A  Q  A  L  A  L  F  A  F  T  R  A  L  A  Q

ctccttgctcatcgcgtttcgcatctcattttgggtgccattagATTTGGGGTATTGCCTTCGGAAGCGCGATTCTAA 2498
                                        I  W  G  I  A  F  G  S  A  I  L  T

CGAATGAACTGAAGAAGTACCTTCCCGATCAATATGTTCAGTTCGCCGCCGGGGGCATTGAGAAGGTGTATGCTCTCA 2576
  N  E  L  K  K  Y  L  P  D  Q  Y  V  Q  F  A  A  G  G  I  E  K  V  Y  A  L  I

TTCCTGAGATCCCCGCCCTACCTGCCCCTTTGAAGCAAGAAGTTCAGTACGCCTTTGGCGAGAGTATTCGCCCAATTT 2652
  P  E  I  P  A  L  P  A  P  L  K  Q  E  V  Q  Y  A  F  G  E  S  I  R  P  I  W

GGTACACTCTTGCCGGGATCGGTTTCCTTGGCCTGGTATCGTCGGCGCTGATGAGGGATATCCCTTTGCACAATTACG 2730
  Y  T  L  A  G  I  G  F  L  G  L  V  S  S  A  L  M  R  D  I  P  L  H  N  Y  V

TGGACGAGCAATGGAAGTTGGAAGAGAAAAATTCCAACGGAAGTAACGTGGTCTTGAAAAACCAGACTGATGTCGATA 2808
  D  E  Q  W  K  L  E  E  K  N  S  N  G  S  N  V  V  L  K  N  Q  T  D  V  D  I

TCGAAGAAGTCCGTCATTAGGATGGACGTAGGACCGCTCGTTTCACTTTCTACCTGATCACTCCATGTTCTTCATCTC 2886
  E  E  V  R  H  ·

GAGGTCGCGCACTCCCATTTGTATTTAACATCCCTCATTACTACAATGATCACCATGCCTACCCCCTAATAGCATAAA 2964

TGTAACCTTTTGGCATCATAGGCATCGACGATCTGCCACATTCACGGGGTCTGTGCGGGTCTTG  3028
```

Figure 6.3 DNA sequence of the *C. cinereus mfs1* gene. The predicted protein sequence is given in one letter code and the introns are indicated in lower case.

BLAST searches do not belong to the ABC superfamily. The *mfs1* gene product clearly belongs to a different class of transporters known as the major facilitator superfamily (MFS). This superfamily includes the drug resistance proteins, sugar facilitators, facilitators for Krebs cycle intermediates, phosphate ester-phosphate antiporters and oligosaccharide-H^+ symporters (Marger and Saier, 1993; Lewis, 1994).

The MFS transporters contain 12 or 14 transmembrane helices. Unlike the ABC transporters which utilise ATP hydrolysis, members of the major facilitator superfamily couple substrate transport to a proton electrochemical gradient (proton motive force). The *C. cinereus* Mfs1 protein is predicted to have 14 transmembrane domains. It also contains two consensus sequences present in multidrug transporters, the drug translocase consensus sequence $GXXXD^{R}{}_{K}XG^{R}{}_{K}$ (Rouch *et al.*, 1990) and a drug consensus sequence GXhyhyGPXXGG (Yamaguchi *et al.*, 1992) (X is any amino acid, hy is a hydrophobic amino acid) (Figure 6.4). They also contain a WRW motif just in front of the sixth transmembrane domain which is conserved in the members of this superfamily found in *S. cerevisiae* (Goffeau *et al.*, 1997). The alignment shown in Figure 6.4 reveals one further conserved motif, FXXNLPIGA. The data presented in this chapter provides good evidence for the presence of a gene encoding a membrane-spanning transport protein within a DNA sequence that forms part of the *B42* locus of *C. cinereus*. It would require gene disruption to ascertain whether this gene has a role to play in pheromone function or whether its location within the *B42* locus is incidental.

A similar gene (*mfs2*) has been identified within the corresponding 6kb of *B6*-specific sequence to the left of the *B6* pheromone and receptor genes (Halsall and Chaure, unpublished data). The presence of this transporter is thus a general property of *C. cinereus* cells, and not unique to cells with a *B42* mating type locus. Hybridisation analysis using the *B6* gene as a probe to genomic DNA from all nine available strains

Figure 6.4 Alignment produced using PILEUP (GCG, version 8.1) of the deduced amino acid sequences of the *msf1* gene product from the *B42* mating type locus of *C. cinereus*, Yan6 hypothetical protein from *S. pombe*, Bmr3 from *B. subtilis*, the *TOXA* gene product from *C. carbonum* and Sge1 from *S. cerevisiae*. The positions containing identical amino acids in all five sequences are indicated by a **c**. The positions where two or more sequences share the same amino acid as Mfs1 are indicated by an asterisk. The translocase consensus sequence is shown in red, the drug consensus sequence is shown in green, and the WRW and FXXNLPIGA motifs are shown in blue.

```
      1                                                            50
Mfs1  .......... .......... .......... .......... ..........
Yan6  MNPSSSPNRS LNPTRGLNLY STGNEVTWFS STVDEVNELK PAKSKAFSIS
Bmr3  .......... .......... .......... .......... ..........
ToxA  .......... .......... .......... .......... .MDEQIVSAS
Sge1  .......... .......... .......... .......... ..........

      51                                                           100
Mfs1  ..MTTYASES REQRLGPPSP SGKKGA.... .......RFW LIFASLSVTL
Yan6  AQDQIYGTSS LKSRFSTHEA SSEKDDSSDF TLVLPSNRMY IVIPGLMLSI
Bmr3  .......... .......MDT TTAKQASTK. .......... FVVLGLLLGI
ToxA  SNVKDGVEKQ PVKDREDVDA NVVPPHSTPS LPKISLISLF SIVMSLGAAA
Sge1  .......... .......... .......... ..MKSTLSLT LCVISLLLTL
                                    *                      c
      101                                                          150
Mfs1  FLSALDYTAL ASALPVIVHD LK.GSDFSWV GTAYALSSTA FLPMSGGLAQ
Yan6  FLSALDQTVI TTAIPTIVAN LDGGSSYSWI GTAYTLAQTS ILPFCGIMSE
Bmr3  LMSAMDNTIV ATAMGNIVAD LGSFDKFAWV TASYMVAVMA GMPIYGKLSD
ToxA  FLGALDATVV AVLTPTLAQE FHSVDAVAWY GAIYLLMSGT TQPLFGKLYN
Sge1  FLAALDIVIV VTLYDTIGIK FHDFGNIGWL VTGYALSNAV FMLLWGRLAE
      ***c*c *   * * * **   *       c  ** c *       *  c *
      151                                                          200
Mfs1  IVGRRIAFII ALGFFALGSA LCGAARSMNW LIGARAIQGV GGGAIISLAS
Yan6  VVGRKIVLYT SIVLFLFGSA MCGAAQNMLW LVLCRAVQGI GGGGITSLVT
Bmr3  MYGRKRFFLF GLIFFLIGSA LCGIAQTMNQ LIIFRAIQGI GGGALLPIAF
ToxA  EFSPKWLFIT CLIVLQLGSL VCALARNSPT FIVGRAVAGI GAGGILSGAL
Sge1  ILGTKECLMI SVIVFEIGSL ISALSNSMAT LISGRVVAGF GGSGIESLAF
        **   *    * * cc*  ** *  *   **  c* *c* c** * ***
      201                                                          250
Mfs1  IVISDLVTLQ ERGTYNGLLG MTWAVAVALG PIVGGALAEN GAWRWLFYLN
Yan6  IVIADITPLQ TRPYYTGCMG VTWGLASVMG PLIGGAISQK ASWRWIFFIN
Bmr3  TIIFDLFPPE KRGKMSGMFG AVFGLSSVLG PLLGAIITDS ISWHWVFYIN
ToxA  NIVALIVPLH HRAAFTGMIG ALECVALIIG PIIGGAIADN IGWRWCFWIN
Sge1  VVGTSIVREN HRGIMITALA ISYVIAEGVG PFIGGAFNEH LSWRWCFYIN
       ** *** *   c*   *  *     *    c c  c**       c*c c* c
      251                                                          300
Mfs1  LPICAVSFIL VVV..... ....FM.... ......SL.. ....PTPAGT
Yan6  LPTGGLSLAL LIF..... ....FL.... ......NL.. ...VPKPTVS
Bmr3  VPIGA..... LSL..... ....FF.... .......I.. ...IRYYKES
ToxA  LPIGAAVCAI LLF..... ....FFHPPR STYSASGV.. ...PRSYSEI
Sge1  LPIGAFAFII LAFCNTSGEP HQKMWLPSKI KKIMNYDYGE LLKASFWKNT
      *c* *                     *
      301                                                          350
Mfs1  FKEKFFAMDW IGNFLVIGSS SALVIALTWG GGRYGWGEAP VLVPLILGIV
Yan6  FRVFLRDFDF VGIITITTGV VLFLVGLNIG STTGHWAHAN VLCYLIFGIL
```

```
      Bmr3  LEHRKQKIDW GGAITLVVSI VCLMFALELG GKTYDWNSIQ IIGLFIVFAV
      ToxA  LGN....LDY IGAGMIISSL VCLSLALQWG GTKYKWGDGR VVALLVVFGV
      Sge1  FEVLVFKLDM VGIILSSAGF TLLMLGLSFG GNNFPWNSGI IICFFTVGPI
            *       c   c       *    *  *c  c *  * c      *   ** * *
      351                                                          400
      Mfs1  GLVVWLFYEV K......... ..YSSHPTVA PELLGNRTSV SGYLQVFFNS
      Yan6  CIAGFVVNEF Y......... ..TTRTRIIT PSAFKTLSLT SVMVTSFLHY
      Bmr3  FFIAFFIVE. R......... ..KAEEPIIS FWMFKNRLFA TAQILAFLYG
      ToxA  LFLSASGHQY W......... ..KGEKALFP TRLLRQRGFL LSLFNGLCFG
      Sge1  LLLLFCAYDF HFLSLSGLHY DNKRIKPLLT WNIASNCGIF TSSITGFLSC
                  *                    *          *          *
      401                                                          450
      Mfs1  VIVVAWV.YM PVYFQACKDA GPIQSGIDMF G.GTLSIGFI VLAGVSVTVL
      Yan6  YIVSTITYYI PVYFQNIKGD GPLMSGVHTL SLAVVSSLLS AVSGMGIGKL
      Bmr3  GTFIILAVFI PIFVQAVYGS SATSAGFILT PMMIGSVIGS MIGGI.FQTK
      ToxA  GVQYAALYYL PTWFQAIKGE TRVGAGIQML PIVGAIIGVN IVAGITISFT
      Sge1  FAYELQSAYL VQLYQLVFKK KPTLASIHLW ELSIPAMIAT MAIAYLNSKY
                  *   *  *c* *    *   **          *          *
      451                                                          500
      Mfs1  KIYRPQLWIG WVLLTAGMGI MGLLRANTPL AQGIGFSVII ASGAGMIFAV
      Yan6  KNYRYPMIGG WIVLLAGTGS MIAIYYNTDI SRTMGFLALT AVGTGNLFQP
      Bmr3  ASFRNLMLIS VIAFFIGMLL LSNMTPDTAR VWLTVFMMIS GFGVGFNFSL
      ToxA  GRLAPFIVIA TVLASVGSGL LYTFTPTKSQ ARIIGYQLIY GAGSGAGVQQ
      Sge1  GIIKPAIVFG VLCGIVGSGL FTLI..NGEL SQSIGYSILP GIAFGSIFQA
            **   **       c *         *       ***  *    * c  *
      501                                                          550
      Mfs1  TYFPVLSPLP VTLNA..... .QALALFAFT RALAQIWGIA FGSAILTNEL
      Yan6  NLIAVQASVP PALIA..... .TSCSAFMLL RNMGASVGIS MGAVIYDQQL
      Bmr3  LPAASMNDLE PRFRG..... .TANSTNSFL RSFGMTLGVT IFGTVQTNVF
      ToxA  AFIGAQAALD PADVT..... .YASASVLLM NSMS...GVI TLCVCQ.NLF
      Sge1  TLLSSQVQIT SDDPDFQNKF IEVTAFNSFA KSLGFAFGGN MGAMIFTASL
                  *                 * *   *  *      c    *  * ** *
      551                                                          600
      Mfs1  KKYLPDQYVQ FAAGGIEKVY ALIPEI.... ........PA LPAPLKQEVQ
      Yan6  TTLLKG..TK YSAGLSYSQI ASIPNV.... ........SE .....KNFVF
      Bmr3  TNKLNDAFSG MKGSAGSGAA QNIGDPQEIF QAG...TRSQ IPDAILNRII
      ToxA  TNRINALTEV LPGVT..... ......KETL QSGFAFLRST LTPAEFGVAI
      Sge1  KNQMRSSQLN IPQFTSVETL LAYS...... ........TE HYDGPQSSLS
              *                     *
      601                                                          650
      Mfs1  YAFGESIRPI WYTLAGIGFL GLVSSALMRD IPLHNYVDEQ WKLEEKNSNG
      Yan6  NAYANAIRMI WIVNCPVAGV GMLLSFFTKQ EKLSQSVTE. YKEKDKGFKD
      Bmr3  DAMSSSITYV FL....LALI PIVLAAVTIL FMGKARVKTT AEMTKKAN..
      ToxA  QTFNSAIQDA FL....VAIV LSCASVLGWP FLSWASVKGQ KKMNK.....
      Sge1  KFINTAIHDV FYCALGCYAL SFFFGIFTSS KKTTISAKKQ Q.........
             *    c                      *           *  *  *   *
```

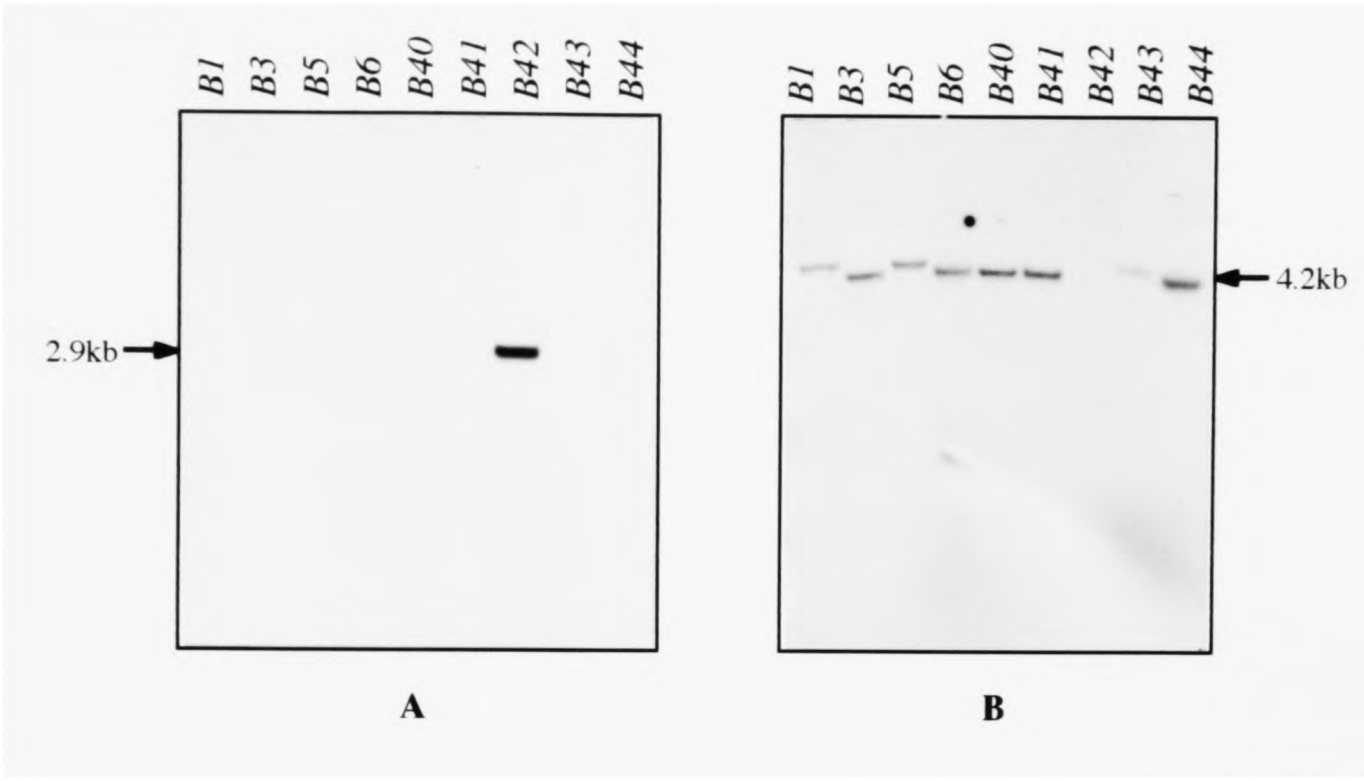


Figure 6.5 Southern blots of *Hin*dIII digested genomic DNA of *C. cinereus* strains 218 (*B1*), LCO12 (*B3*), TC4 (*B5*), LT2 (*B6*), LCO7 (*B40*), LCO5 (*B41*), LN118 (*B42*), AT8 (*B43*) and NL1 (*B44*) probed with cloned DNA sequences containing (**A**) *mfs1* and (**B**) *mfs2*.

having different *B* mating specificities, showed that this gene was present in all but the *B42* strain (Figure 6.5). It seems likely therefore, that the gene generally lies within the homologous flanking sequence of the *B* locus rather than within the locus itself. Perhaps, as a result of a DNA rearrangement, the gene at some time became translocated from homologous flanking region into the unique DNA sequence of the *B42* mating type locus. Flanked by *B42*-specific sequences its chance of being recombined into other genomic backgrounds would be reduced and there would then be no selection against the accumulation of nucleotide substitutions that have little effect on gene function. Over the course of time, a sufficient loss of DNA sequence homology would occur with the result that the *B42* gene would no longer cross-hybridise to other alleles of the gene and would remain trapped in its current position.

Chapter 7

Discussion

The *B* mating type genes of the homobasidiomycetes play an essential role in the recognition of a compatible mating partner. The fusion of compatible monokaryotic cells leads to a reciprocal exchange of nuclei and the migration of donor nuclei through the established monokaryotic cells of the recipient. A compatible *B* gene interaction is required to promote this nuclear migration. Migration is rapid and necessitates the enzymic disruption of the complex dolipore septa that normally prevent movement of organelles between cells. Once the dikaryon has been established, the *B* genes continue to play an essential role in maintaining the characteristic growth of this mycelium. Division of each apical cell of the mycelium involves the accurate distribution of the two nuclei that constitute the dikaryotic pair. In *C. cinereus* and many other species including *Schizophyllum commune*, this involves the formation of the clamp connection through which one of the daughter nuclei must pass to join its partner in the newly formed subapical cell. Compatible *B* genes are essential for fusion of the clamp cell to the subapical cell.

This study has helped to establish that the *B* mating type genes encode components of a signalling pathway, namely peptide mating pheromones and their transmembrane receptors. Molecular cloning of the separate *Bα1* (Wendland *et al.*, 1995) and *Bβ1*(Vaillancourt *et al.*, 1997) loci of *Schizophyllum commune*, and the *B6* (O'Shea 1995; O'Shea *et al.*, 1998) and *B42* (this study) mating type loci of *Coprinus cinereus* has provided the essential sequence to identify gene functions. It has, moreover, revealed

that these loci contain large gene complexes with a considerable degee of functional redundancy.

Pheromone signalling in mating plays an essential role in other fungal species. In the unicellular yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, mating cells secrete peptide pheromones as chemoattractants. Binding of the pheromone to receptors on a compatible cell of different mating type elicits the major changes in gene expression that permit cell fusion. Finally, in the hemibasidiomycete *Ustilago maydis* which, like the homobasidiomycetes, sequesters the genes encoding the pheromones and receptors into a mating type locus, also secretes pheromones that play an essential role in preparing cells for fusion. Like the yeast species, there only two versions of the pheromones and receptors in *U. maydis*.

The homobasidiomycetes are unique in that the pheromone signalling pathway appears to be activated only after cells have fused. The components of the pathway are also unusual with several muliallelic genes encoding an array of *B* mating type specificities. Providing mating partners have different alleles of the *B* genes, the pheromone response is activated and *B*-regulated development takes place. A considerable degree of specificity on the part of the pheromones and receptors is required to discriminate incompatible products from compatible ones, a level of complexity that is found in higher eukaryotes (Dulac and Axel, 1995) but not in other fungal species.

7.1 Organisation of the *B* locus

At the outset of this project little was known about the molecular structure of the *B* locus of *C. cinereus*. Unlike *S. commune*, there was no genetic evidence to suggest that the *B* genes might be separated into two discrete loci (Haylock *et al.*, 1980). Attempts to

demonstrate recombination had proved unsuccessful. It seemed likely, therefore, that the genes would map to a single complex locus.

The *B6* locus was isolated by O'Shea (1995) using a genomic subtraction technique. The prediction was made that the multiple alleles of the *B* genes, like those of the *A* genes (Pardo *et al.*, 1996), would share little DNA sequence homology and thus fail to cross-hybridise. Excess sheared genomic DNA from a *B3* strain was used to subtract out all common sequences from enzyme digested genomic DNA derived from a *B6* strain. Sequences unique to the *B6* strain were then recovered by cloning into a suitable vector followed by transformation into *Escherichia coli.* A fragment of *B6*-specific DNA recovered identified cosmid clones in which the region of DNA non-homology between *B6* and *B3* was shown to extend over 17kb. In this study, the *B42* locus was isolated using a homologous sequence from the border of the *B6* locus to identify the first of a number of overlapping clones covering the entire *B42* locus.

DNA sequence analysis was used to identify nine genes that confer *B42* mating specificity. Six of these encode pheromone precursors and three encode transmembrane receptors. The parallel analysis of the *B6* locus identified the same number of genes, six encoding pheromone precursors and three encoding receptors. A BLAST search showed that the three receptor proteins encoded in each locus had significant homology to the Ste3p receptor of *S. cerevisiae* and the Map1 receptor of *S. pombe*. In the yeast species there are two receptor proteins corresponding to the two different mating types. Whilst they have a common tertiary structure, the DNA sequences are quite dissimilar. Ste3p and Map1 are activated by the isoprenylated **a**-factor and M-factor, respectively whereas the Ste2p and Mam2 receptors bind the non-modified α-factor and P-factor, respectively. It is perhaps not surprising that all six pheromone precursor peptides encoded in the *C.*

cinereus B loci contain the CaaX motif at the carboxy-terminus which indicates that the mature pheromones will be isoprenylated like **a**-factor and M-factor.

The pheromone receptors encoded by the *a* mating type loci of *U. maydis* and the *Bα1* and *Bβ1* loci of *S. commune* also belong to the same sub-family of proteins that bind isoprenylated pheromones (Bölker *et al.*, 1992; Wendland *et al.*, 1995; Vaillancourt 1997). In both these basidiomycete fungi only genes encoding pheromone precursors that are predicted to be isoprenylated have been identified. This would seem to be a characteristic of pheromones produced by basidiomycete species.

Figure 7.1 shows an alignment of several pheromone precursors from the *B6* and *B42* loci of *C. cinereus* and the *Bα 1* and *Bβ 1* loci of *S. commune*. The CaaX motifs are highlighted in red. A conserved ER/EH/DR motif is highlighted in green. This motif represents a possible processing recognition site. Proteolytic cleavage adjacent to this motif would generate mature peptide pheromones of the expected size, 11-13 residues. The mature *U. maydis* pheromones have 9 and 13 amino acids and each has an amino terminal NR or GR. It is interesting to note that when the *S. commune* pheromone sequences are aligned with those of *C. cinereus* this EH/ER/DR motif can be seen.

Clearly, it will be of interest to ask whether the maturation pathway of the *C. cinereus* pheromones is similar to that described for *S. cerevisiae* and *S. pombe*. The processing pathway for the CaaX modified **a**-factor of *S. cerevisiae* is presented in Figure 1.4 and described in the introduction. It involves five steps, of which, four are membrane associated. The final step in the pathway is the secretion mediated by the Ste6p transporter. The pathway in *S. cerevisiae* was determined using mutants defective in mating function and with a knowledge of the sequence and structure of the mature peptide.

```
           1                                                  50
B42Phb1.1  .MDSFDSLDSLNLSVEETTLQTLLESMDTTDA..ASE.............
B42Phb1.2  ....MDSFQQLNLFVEETIQQSLPEAIPSSDSTDTGA.............
B6Phb3.1   ........................MDSFAAIDFAEFGETPVELPSSAPSKD
B42Phb3.2  MSDAFTTLDTVDLFIEENEQE..VVEVPSCPPPRRPSFSSADAESIFLTV
Bap1(2)    ........................MRSRASAEGIAVLGLRRRGESPVCRRR
Bap1(3)    .MDDFAEFFPTLV.............LDEPEVARRPA...........RD
Bbp1(1)    .MDAFTAMFPELFPIEEGLEDALVGSLSDTSAASASATHTSPASTDTFDD
Bbp1(3)    ......MASSVLARPGPSTVLPAMTRPPPPMAHRAAATPSFARSAQPQTD

           51                              78
B42Phb1.1  SELDAILVNSER...DPGFTSKGFCVIA
B42Phb1.2  SERDATPVNTER...HLGFTTKGFCVVS
B6Phb3.1   DDIMQRLVDSDRRLPDSYYGGA..CVIA
B42Phb3.2  EEVNDLPVDYERRTQGGGGLTW.FCVIA
Bap1(2)    NVVVCFWG..DR.SCVER...EG.CVRG
Bap1(3)    AEVLAILADAERPG.GSNCTA..WCVVA
Bbp1(1)    ADILAILADAEHWR.GGNTTAHGWCVVA
Bbp1(3)    DAVLALLANAEHTEAGEETTARGWCVVA
```

Figure 7.1 Alignment of four *C. cinereus* and four *S. commune* pheromone precursor sequences using PILEUP (GCG, version 8.1). The CaaX motif is displayed in red, the ER/EH/DR motif is displayed in green.

A similar approach to identifying the pathway in *C. cinereus* presents many problems. Firstly, with the observed levels of functional redundancy it would be difficult to select for gene mutations that disrupt pheromone function and secondly, how can one purify a single mature peptide pheromone when each monokaryon produces several?

There appears to be no requirement for large amounts of pheromones to be secreted into the surrounding environment. Mating cells fuse irrespective of mating type and the pheromone response is initiated intracellularly. Pheromone secretion is expected to play a role in clamp cell fusion but this would not require diffusion of pheromone beyond the cell wall. If pheromones are localised within the mycelium, the actual concentration may be too low to permit the purification procedures used for the yeast pheromones. One of the challenges for future work on this system is to identfy where the pheromones and receptors interact in *C. cinereus*.

A novel approach to understanding how the *C. cinereus* pheromones are processed is currently being undertaken in our laboratory by asking whether a *C. cinereus* pheromone gene product can be processed in *S. cerevisiae* (N. Olesnicky, unpublished). If these experiments are successful, it will imply that the pathway is highly conserved and the genes involved may be identified by homology to those in *S. cerevisiae*.

By analogy with *S. cerevisiae* and *S. pombe*, it would be expected that the pheromones would be transported across a cell membrane. This would require a protein homologous in function to Ste6p and Mam1. The fortuitous discovery of a gene encoding a transport protein within the *B42* locus was initially of great interest as a possible candidate for a Ste6p/Mam1 homologue. Unfortunately, the sequence analysis of the predicted protein identified it as a different class of transporter and unlikely to have a role in pheromone export. One way of eliminating this possibility would be to knock out the gene by homologous recombination and then look for mating defects.

7.2 The origin of multiple *B* mating specificities

Each of the nine genes identified in the *B42* locus (this study) and the *B6* locus (O'Shea *et al.*, 1998) was shown by transformation to be able to change the *B* mating specificity of a suitable host strain. A single pheromone or single receptor is thus sufficient to activate the pheromone response when expressed in a host cell. A comparison of the *B6* and *B42* loci showed that gene order was conserved (Figure 4.2). Moreover, these results suggested that the genes might be organised into three groups each comprising a receptor gene and two pheromone genes. There is clearly considerable redundancy in function. In the *A* locus, a similar redundancy in function is the basis for the large numbers of *A* mating specificities (Pardo *et al.*, 1996). If this is also true in the *B* locus, then many unique allele combinations could be generated from three functionally independent groups of genes. In support of this idea, Southern blot analyses established that a region containing one of the proposed groups of *B42* genes (group 1) shared DNA-homology with the same region of the *B6* locus. None of the three genes in this region of shared homology could activate the pheromone response in a *B6* host suggesting that the genes were homoallelic. Similarly, a region containing the middle group of three genes in *B42* (group 2) shared DNA homology with a corresponding region in the uncloned *B3* locus and these three genes were unable to activate the pheromone response in a *B3* host.

Since all the genes identified are able to activate the pheromone response in a suitable host, it is evident that mating partners require only one group of genes to have different alleles in order to have a compatible *B* gene interaction. A few different alleles of each group of genes would be sufficient to generate a large number of unique allele combinations and hence *B* mating specificities. Four alleles of two of the groups of genes and five of the other would be sufficient to generate the 79 predicted *B* mating specificities.

DNA sequence analysis confirmed the prediction that *B6* and *B42* share the same alleles of the three genes comprising group 1. Figure 7.2 shows a direct comparison of the Phb1.1 and Phb2.1 pheromone precursor sequences and the Rcb1 protein sequences. In the case of the pheromone precusors, the carboxy-terminal residues predicted to represent the mature peptide pheromones are identical. In the case of the two receptor proteins, the only significant differences in sequence are a single non-conservative amino acid substitution at position 130 and an insertion at position 336 in the *B6* protein.

Although the *B3* locus has not been cloned, a single putative pheromone precursor gene was amplified by PCR. *B42* and *B3* are predicted to share alleles of the middle group of genes. Comparison of the Phb2.2 peptides from *B42* and *B3* clearly identifies them as functionally identical (Figure 7.2B). The organisation of the rest of the *B3* locus can only be predicted but there is evidence (Chapter 5) to suggest that there are three groups of genes arranged in the same order as those in *B42* and *B6*. Hybridisation analysis identified all three genes that constitute the middle group of genes (group 2). By using PCR to amplify the sequences flanking the middle group of genes, it was possible to confirm that these regions contained genes that conferred *B* mating specificity in certain host strains but not others. Where the genes failed to activate the pheromone response, there was shared DNA homology between them and sequences in the genomic DNAs of the hosts.

Only nine different *B* mating specificities are represented in the strains of *C. cinereus* available in the Oxford collection. The combination of Southern blot analyses and transformation with cloned genes from *B42* and *B6* and fragments from *B3* identified at least one group of genes in each locus and often two (Figure 5.4). Four of the loci, *B1/B44* and *B40/B41* share the group 1 and group 3 gene alleles and their different specificities can only reside in the middle group of genes. To demonstrate conclusively

A

```
B42Phb1.1 MDSFDSLDSLNLSVEETTLQTLLESMDTTDAASESELDAILVNSERDPGFTSKGFCVIA 59
          ||||||||||||||||||||||||:||||||||||||| ||||:||||||||||||||||
B6Phb1.1  MDSFDSLDSLNLSVEETTLQTLIESMDTTDAASESERDAILINSERDPGFTSKGFCVIA 59

B42Phb1.2 MDSFQQLNLFVEETIQQSLPEAIPSSDSTDTGASERDATPVNTERHLGFTTKGFCVVS 58
          ||||||||||||||:.||||||||||||||||||.||||||||||||||||||:|
B6Phb1.2  MDSFQQLNLFVEETIHRSLPEAIPSSDSTDTGASERDTTPVNTERHLGFTTKGFCVIS 58
```

B

```
B42Phb2.2 MDNFTVDLATLFEEFPELQEIQATASEHCSQDQYGSCEGPPINQERPGSGVNRAFCVIA 59
          |||| |||||||:||||| |||:..||.|: ||.|||:|:||||||||||||||||||
B3Phb2.2  MDNFLVDLATLLEEFPEHQEIEVSASAHSFQDGYGSSEAPPINQERPGSGVNRAFCVIA 59
```

C

```
B6Rcb1  MKYPALPVFAILGALLVLIPLPWHWRARNVATLSIIAWLFVMNVIYAVNT 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
B42Rcb1 MKYPALPVFAILGALLVLIPLPWHWRARNVATLSIIAWLFVMNVIYAVNT 50

B6Rcb1  IVWAGSIRDVAPVYCDIATKLIIGASHALPLATLCICKHLEMVSSSRTVS 100
        ||||||:||||||||||||||||||||||||||||||||||||||||||||
B42Rcb1 IVWAGSLRDVAPVYCDIATKLIIGASHALPLATLCICKHLEMVSSSRTVS 100

B6Rcb1  YDVSDKKRRMIFEGVMCFVLPMIFMALHYIVQGHRYDIIQEFGCQPTIYI 150
        |||||||||||||||||||||||||||| ||||||||||||||||||||
B42Rcb1 YDVSDKKRRMIFEGVMCFVLPMIFMALHYNVQGHRYDIIQEFGCQPTIYI 150

B6Rcb1  SIPAIFIVWFPPLLFAVISFILAAMALHHFVRRRLVLAAHLQNANSALTP 200
        ||||||||||||||||||||||||||||||||||||:|||||||||||||
B42Rcb1 SIPAIFIVWFPPLLFAVISFILAAMALHHFVRRRLVFAAHLQNANSALTP 200

B6Rcb1  RYIRLIAMALTIMTWNTSLTAFNLYNNVFPGLRPWTNWADVHSNFSRVD 250
        ||||||||||||||:|||||||||||||||||||||||||||||||||||
B42Rcb1 NRYIRLIAMALTIMTRNTSLTAFNLYNNVFPGLRPWTNWADVHSNFSRVD 250

B6Rcb1  LFPTVFLPDYFIRAMMLFWWAMPVSSLIFFLFFGFGEEAMKEYRKVGAWI 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
B42Rcb1 LFPTVFLPDYFIRAMMLFWWAMPVSSLIFFLFFGFGEEAMKEYRKVGAWI 300

B6Rcb1  SRVILRRKTNEKGELLGSTGNSRRRLHLVDLKNKNA...SMISKPIPGSF 347
        |||||||||||||||||||||||||||||||||||   ||||||||||||
B42Rcb1 SRVILRRKTNEKGELLGSTGNSRRRLHLVDLKNKNASLGSMISKPIPGSF 350

B6Rcb1  CKTSQSPSATVAGSPPPYKKSFDVVSPTSTSNCSTVVNANSPLKKANEID 397
        |||||||||||||||||||||||||||.|||||||||||||||||||||||
B42Rcb1 CKTSQSPSATVAGSPPPYKKSFDVVSPASTSNCSTVVNANSPLKKANEID 400

B6Rcb1  LEAGDDFSYYAQSNASTATQLAPPAMVERSSSTRSSGPYISEPIPSPTLS 447
        ||||||||||||||||||.|||||||||||||||||||||||||||||||
B42Rcb1 LEAGDDFSYYAQSNASTAAQLAPPAMVERSSSTRSSGPYISEPIPSPTLS 450

B6Rcb1  YCSILSPSTAPSTSAFSPSSFPESVGPDSLNRASEMDLVESYYEMQRDIR 497
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
B42Rcb1 YCSILSPSTAPSTSAFSPSSFPESVGPDSLNRASEMDLVESYYEMQRDIR 500

B6Rcb1  EEETRRTHLRGPASPAYHRPFSPTLYPVALPHPGSALPPVNGILVTVHRQ 547
        |||||||||||||||||||||||||||||||||||.|||||||||||||||
B42Rcb1 EEETRRTHLRGPASPAYHRPFSPTLYPVALPHPGSVLPPVNGILVTVHRQ 550

B6Rcb1  ASVDELPSTHS 558
        |||||||||:|
B42Rcb1 ASVDELPSTRS 561
```

Figure 7.2 Alignment of *C. cinereus B* proteins encoded by homoalleles, using GAP (GCG, version 8.1). (**A**) Phb1.1^{B42} with Phb1.1^{B6} and Phb1.2^{B42} with Phb1.2^{B6}. (**B**) Phb2.2^{B42} with Phb2.2^{B3}. (**C**) Rcb1^{B42} with Rcb1^{B6}. | indicates identical amino acids and : or . indicate conserved changes.

that the three groups of genes belong to three functionally independent sub-families would require the generation of a *B*-null host in which to test compatible gene combinations. However, the data are consistent with this interpretation.

The lack of DNA homology between different alleles of the *B* genes and the fact that the three genes in each group are embedded in allele-specific flanking sequences, means that they cannot be separated by recombination. This is essential for maintaining self-incompatibility. If the three groups of genes are functionally independent, recombining whole groups might generate a new *B* mating specificity in a population and add to genetic diversity. The only possibility for homologous recombination within the *B* locus is where a group of genes is shared. This will only be effective in generating a new mating specificity if the shared genes are the middle group and recombination effects new allele combinations of genes in the flanking groups. Alternatively, new alleles could arise through random point mutations. Although the frequency of these events is expected to be very low, redundancy within the *B* locus may allow non-functional mutations to be tolerated and new specificity to arise through nucleotide substitutions and genetic drift (Ohta, 1996).

Each group of three genes extends over some 5-6kb, sufficient to permit a low frequency of homologous recombination but one that would be difficult to detect in a standard genetic cross. It should, however, be possible to detect recombinant products in the spore progeny using PCR. Of the nine *B* loci available, only *B42* and *B3* share the same alleles of the middle group of genes and have different alleles of the flanking genes. This presents a perfect opportunity to look for recombination within the *B* locus. A PCR strategy that could be used to detect recombinant DNA molecules is illustrated in Figure 7.3.

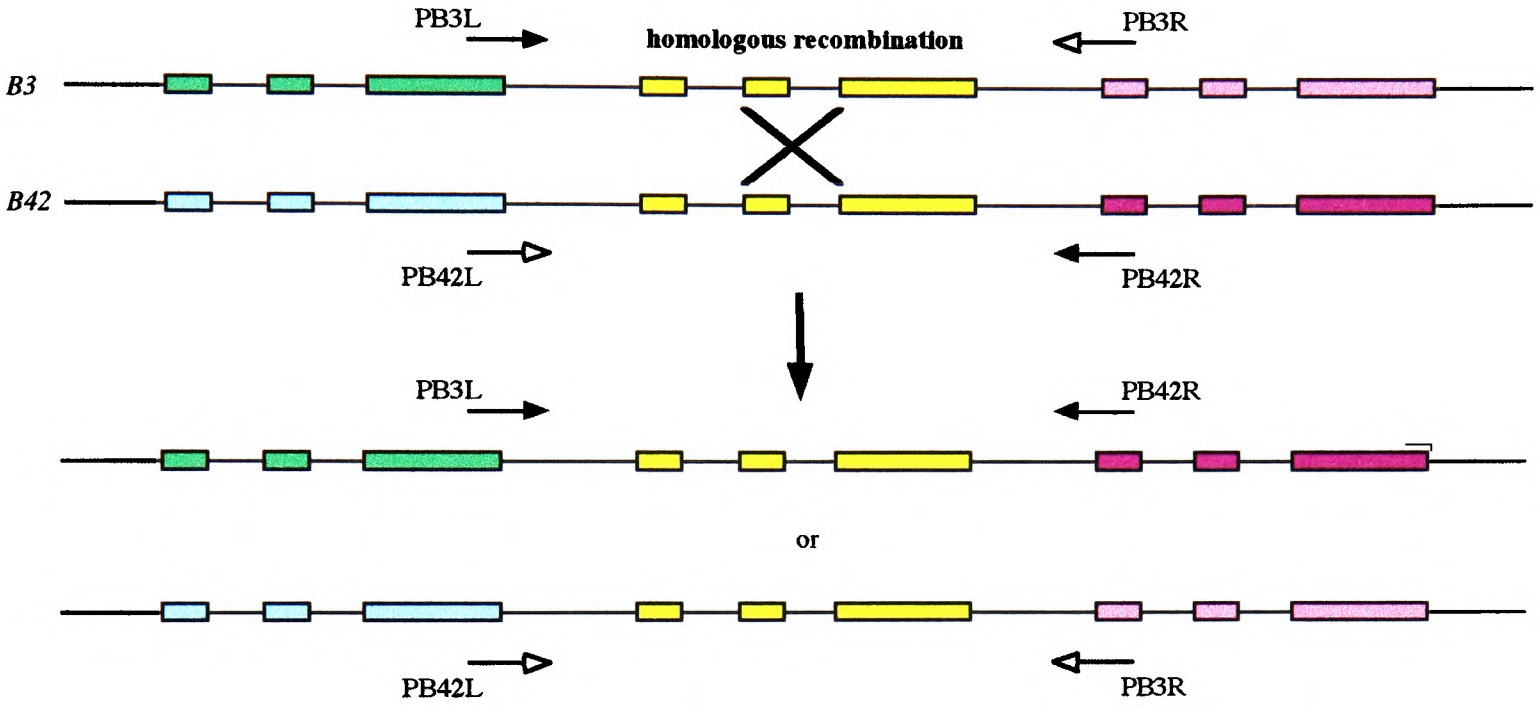


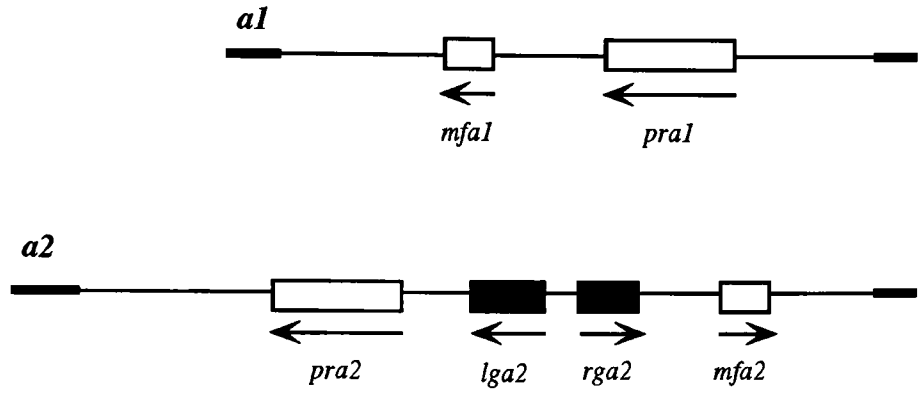
Figure 7.3 Diagram illustrating a PCR strategy that could be used to detect recombination within a set of genes shared by the *B42* and *B3* loci. Primers are indicated by arrows. PB42L and PB42R are *B42* specific primers, PB3L and PB3R are *B3* specific primers.

7.3 Comparative analysis of the basidiomycete *B* and *a* mating type loci

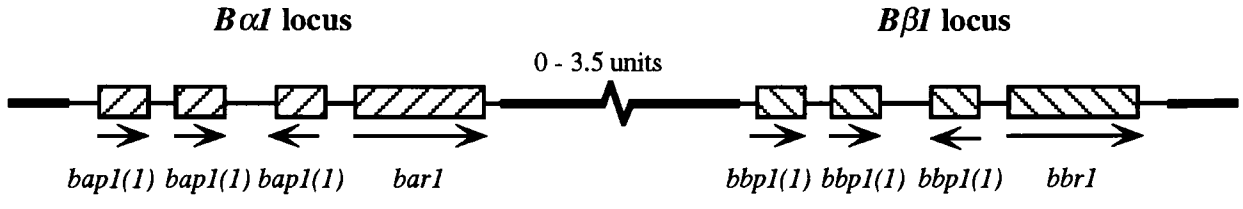
Figure 7.4 compares the organisation of the multiallelic *B* loci of *C. cinereus* and *S. commune* with the much simpler biallelic *a* locus of *U. maydis*. The *B* loci of *S. commune* each contain a receptor gene and three pheromone precursor genes. It is clearly evident that these two sets of genes are functionally independent and that the nine different alleles of each locus are sufficient to generate the 81 *B* mating specificities. Like the *C. cinereus* genes, different alleles are very dissimilar in sequence and an allele complex is kept together by a general level of DNA non-homology across the entire 8kb which constitutes each locus. The *a* loci of *U. maydis* each contain a single receptor and pheromone precursor gene and the overall sequence of each locus, including the genes, is very dissimilar. As shown in *C. cinereus*, the pheromone and receptor genes cannot be separated by recombination. The two genes designated *lga2* and *rga2* in the *a2* locus are of unknown function. The fact that they are sequestered into a mating type locus suggests that they may play some role in mating and interestingly, they are both up-regulated in mating cells (Urban *et al.*, 1996).

The complex multiallelic *B* loci that are present in *C. cinereus* and *S. commune* are assumed to have evolved by gene duplication and amplification from the simple biallelic organisation seen in *U. maydis*. In *S. commune*, there are 9 $B\alpha$ and 9 $B\beta$ alleles - equal numbers of genes in the two groups. In *C. cinereus*, three groups of genes have evolved. From the small sample of *B* loci that it was possible to screen in this analysis it is clear that the middle group, group 2, has many more alleles than group 1 and group 3 and genes of this group probably gave rise to the others. Since the genes are embedded in allele specific sequence, recombination and rearrangement would not be expected to occur. The fact that in some groups, the pheromone genes may be in different orientation

(A) *Ustilago maydis*



(B) *Schizophyllum commune*



(C) *Coprinus cinereus*

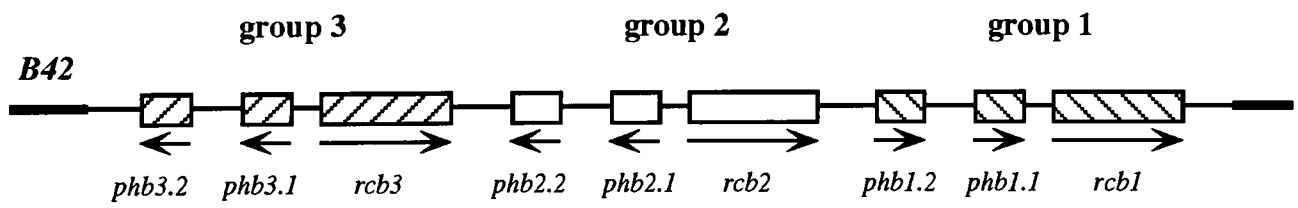


Figure 7.4 Organisation of the *a* mating type locus of *Ustilago maydis* and the *B* loci of *Schizophyllum commune* and *Coprinus cinereus*. Arrows indicate the direction of gene transcription. Thick lines indicate homologous DNA sequences.

with respect to the receptor, may indicate that some rearrangement occurred early in the evolution of different alleles.

7.4 Allelic specificity

With three functionally redundant groups of genes within the *C. cinereus B* locus there must be a considerable degree of specificity exhibited by the individual pheromones and receptors. A comparison of the three receptors sequences encoded by the *rcb1*, *rcb2* and *rcb3* genes in the *B42* locus revealed variable levels of homology. The comparisons are summarised in Table 7.1 together with sequence comparisons for the *B6* receptor proteins and three receptors from cloned *S. commune* loci, two versions of the Bar1 protein encoded by the *Bα1* and *Bα2* loci and the receptor encoded by the *Bβ1* locus.

The level of homology displayed by the three receptors encoded by the *B42* or *B6* genes varies from 35% identity to 53%. Since it is concluded that these represent members of three functionally independent sub-families, a higher degree of identity would not be expected. More striking is the comparison of the sequences encoded by the two alleles of *rcb3*. These are illustrated in Figure 7.5A. Here it can be seen just how conserved the two sequences are with 78% identity and 92% similarity. Such similar sequences must pose problems for pheromone specificity. The *rcb3* receptors bind the Phb3.1 and Phb3.2 peptides. Interestingly, if one compares the two *B42* pheromone precursors and the corresponding *B6* precursors, they show very little sequence similarity and it is surprising that they can both activate the same receptor. In fact, the Phb3.1 and Phb3.2 pheromones did not always show the same spectrum of activity. For example, Phb3.1 activated the *B3* and *B1* receptor but not the *B6*. Phb3.2 activated the *B6* receptor but not the *B1* or *B3*. However, when one compares the sequences of the Phb3.1 and Phb3.2 alleles present in *B6* and *B42*, the conservation in sequence is striking with only four

Table 7.1 Comparison of the predicted *C. cinereus* receptor proteins. The first 300 amino acids of the predicted proteins were aligned with each other using GAP (GCG, version 8.1), then with the *S. commune* receptors Bar1, Bar2 and Bbr1. Percentage identity and similarity scores are displayed as identity/similarity. Where the similarity is greater than 70% the values are shown in red.

	Percent Identity / Percent Similarity							
Receptor	*B42*Rcb2	*B42*Rcb3	*B6*Rcb1	*B6*Rcb2	*B6*Rcb3	Bar1	Bar2	Bbr1
*B42*Rcb1	53%/72%	38%/64%	99%/99%	37%/61%	38%/64%	35%/63%	23%/52%	49%/68%
*B42*Rcb2	-	35%/64%	53%/73%	37%/59%	35%/61%	37%/65%	26%/51%	46%/67%
*B42*Rcb3			38%/65%	52%/73%	78%/92%	49%/70%	35%/58%	32%/58%
*B6*Rcb1				36%/62%	38%/65%			
*B6*Rcb2				-	52%/72%	52%/72%	32%/53%	34%/57%

Figure 7.5 Alignment of *C. cinereus* predicted receptor proteins produced by the GAP program (GCG, version 8.1). (**A**) Rcb3^{B42} with Rcb3^{B6}, (**B**) Rcb2^{B42} with Rcb2^{B6}. | indicates identical amino acids and **:** or **.** indicate conserved changes. Predicted transmembrane domains are shown as dashed lines.

A

```
B42Rcb3 MSSSRKLSFDDPTYPAFPVLSFIGFVLVLIPLPWHLQAWNSGTCLYMIWT 50
B6Rcb3  ....MKLSFDDPSYPAFPIFAFIGFILVLIPLPWHLQAWNSGTCLFMIWT 46
                    ----------1---------       --------

B42Rcb3 ALGCLNFFINSIVWHGNAIDRAPLWCDISTRFVVGLSVAIPAASLCINRR 100
B6Rcb3  AIGCLNLFINSIIWHGNAIDWAPIWCDISTRLIVGISVAIPAASLCINRR 96
        ----2---------           -----------3----------

B42Rcb3 LYKIASCRTANISRSEKRKAVMVDLAIGLGIPLVQMPLQYVVQGHRYDIF 150
B6Rcb3  LYKIASCQTATISRAQKRRAVMIDLAIGLGIPFLQMPLQFIVQGHRYDIF 146
                          ----------4----------
B42Rcb3 QDIGCYATTVNTPPAYPLVFLWPTIIGLVSAVYCILTLREFMRRRAQFSQ 200
B6Rcb3  QDVGCYPTTVNTPPAYPLVFFWPPIICLVSAVYCILTLRAFMKRRAQFGQ 196
                     ----------5----------

B42Rcb3 LISSNSSSLTVNRYFRLMSLATLELLFNLPITTTGLYLSITSRPIYPWTN 250
B6Rcb3  LISSN.NSLTVHRYFRLMSLATLELVLHLPITTYGLYLCFTSRPIYPWTS 245
                     ----------6----------

B42Rcb3 WADIHYNFNVIDTYPRILWAHKTADAIILELGRWSIVFCALLFFAFFGFA 300
B6Rcb3  WSDIHHNWYIIDTFPRVLWGHDYVQVVTLELARWALVFCAFLFFGFFGFA 295
                                    ----------7----------

B42Rcb3 EEARKNYRKAILALAKRLGYNPRSPDEKAKATFVHVGKPASLPKFTPRPL 350
B6Rcb3  QEARKQYRKVFLFIAKPFGYNPPSSNEK.KNTFVHVGKPPSIPTFTPRPL 344

B42Rcb3 HLASTATLTATQSDKSTNFTRTLSISSTSSGSTKVACSPPAFRLDDIKVK 400
B6Rcb3  HLASTPTL....NDKSTSFTRTLSISSNSSGSTKVACSPPAFRLDDIKVK 390

B42Rcb3 LSSSPGTVSDFPSSPTSMTTRLDDFDSNAHSEV 433
B6Rcb3  LSSSPGTVSDFPSSPTSMTTRFDDFDSNAHSKV 423
```

B

```
B6Rcb2  MAPSHDPTYPLFPIFSFLGFILSAIPFSWHLQAWNSGTCAFMIWTGLACL 50
B42Rcb2 ......MRYEM.PICAFLAAILVLIPLPWHWRAKNIPTLSLSIWLFVSNI 43

B6Rcb2  NGFINSLVWKGNSENPAPVWCDISSKLIIGVSIGIPAATLCISRRLYALT 100
B42Rcb2 IYGVNSILWADNYAVKVPVWCDIVTKIQIGATLSLPACCLCLCIRLERIA 93

B6Rcb2  SVSTVSVTRADKRRMVIVDLCISFGIPAIIMVLHYVVQGHRFDILEDVGC 150
B42Rcb2 STRAASTSATSKRRTAIFDVMMCWILPMIYMAAHYIVQGHRFDIVEDLGC 143

B6Rcb2  YPVVYNTLLAYFLYIAWPVILGIFSFVFSCLTLRSFWIRRAQFS.QLVTS 199
B42Rcb2 RPSIYTSIPAIFIIWVPPFIAIFLTFCFGGVALHHFFRRRLTFARHLQNS 193

B6Rcb2  NSSMSMSRYLRLMILALVDIVCTVPLSAYQIYLGTHGVPLHPWVSWEETH 249
B42Rcb2 NSAITPARYFRLMAMAITQMLWGGIVLSVNAWI.TFRTGLLPYTSWEDVH 242

B6Rcb2  YGFSRVDRIPALFWRSSTTYVVGVELTRWLPVLCAFLFFALFGFASEAQK 299
B42Rcb2 WGFSRVGQFPLIFLPRSM..IPSLFFNWWSVPVSALIFFIFFGFGQEAVN 290

B6Rcb2  QYKRAYWAIVKPFGIEPHSAKQPIKVALPSWVKSANLDSTLSPKPHMRQH 349
B42Rcb2 DYRAALR.RVKSLFWKPR..QETEKMSLPSS......ASKLSPLPPWTPR 331

B6Rcb2  KQHVSLASTANASVDIASTVVGSDLEKSACRSSFSSIRSAPPTYSPSVPH 399
B42Rcb2 TPDGKFPSTATLNRD.........................TTISLSLP. 354

B6Rcb2  HVALDVDIDQRSLFPSSPLSSDSTVFCPESPREKLPEPYTYSLPKSQFTR 449
B42Rcb2 ...LDLH.........SSASSDATCVDPDSPSKQ.....TFTSKTSAFTS 387

B6Rcb2  TSFERSNPSSPLSPSNHPGFHRPFSPPITYPPTTAHPSAAVAVDVIRATL 499
B42Rcb2 NV.......................................... 389
```

conservative changes in the residues that constitute the predicted pheromone sequence (Figure 7.6A and B).

By comparison with *U. maydis, S. cerevisiae* and *S. pombe*, it is surprising to find that each group of genes within the *B* locus has two rather than one pheromone gene. In *S. commune*, even more functional redundancy is observed. There are two groups of genes in this species, separated into the $B\alpha$ or $B\beta$ loci and each is comprised of one receptor and three pheromone genes. As with the *B42* group 3 pheromones, those encoded in any one allelic version of $B\alpha$ or $B\beta$ show different spectra of activity with respect to which other $B\alpha$ or $B\beta$ receptors they can activate. Pheromone redundancy may be a necessary consequence of the high level of specificity that is demanded of pheromone-receptor recognition in these species. During the evolution of this large family, pheromone silencing may have occurred to prevent very similar peptides activating the wrong receptor. Without an active pheromone gene, mating would be impaired but this would be overcome by having a second (or third) gene encoding a pheromone with a complementary activity. This is exactly the situation one sees with the two *B42* group3 pheromones.

When the sequences of the two allelic versions of the Rcb2 proteins are compared they show much less conservation (Figure 7.5B). There is 37% identity and 59% similarity. The pairs of pheromones that activate these receptors, Phb2.1 and Phb2.2, are very similar in sequence unlike the group 3 peptides. Not surprisingly, since the receptor proteins are so different, the different alleles of the *phb2.1* and *phb2.2* genes are not very similar in sequence. It can be noted also that the Phb1.1 and Phb1.2 sequences are conserved (Figure 7.6C). It is not known whether different versions of the Rcb1 receptor proteins are similar, because *B42* and *B6* have the same alleles of the gene.

A

```
B42Phb3.1 .MDSGNTVDLAELCDMDPNIGFTPDSSAP.................TEDNVAKQLVDSDQRLPGGYYGGQ...CIIA 56
B42Phb3.2 MSDAFTTLDTVDLFIEENEQEVVEVPSCPPPRRPSFSSADAESIFLTVEEVNDLPVDYERRTQGG..GGLTWFCVIA 75

B6Phb3.1 .......MDSFAAIDFAEFGETPVELPSSAPSKDDDIMQRL....VDSDRRLPDSY....YGGA.....CVIA 53
B6Phb3.2 MSDSFISFDSVVGPAHSEASET.IAIVDSQSSQLSAIDPRLSSTSLDELNDLPVEFERRTHGGNGLTFWCVIA 72
```

B

```
B42Phb3.1 MDSGNTVDLAELCDMDPNIGFTPDSSAPTEDNVAKQLVDSDQRLPGGYYGGQCIIA 56
B6Phb3.1  MDSFAAIDFAEFGETPVEL...PSSAPSKDDDIMQRLVDSDRRLPDSYYGGACVIA 53

B42Phb3.2 MSDAFTTLDTVDLFIEENEQEVVEVPSCPPPRRPSFSSADAESIFLTVEEVNDLPVDYERRTQGGGGLTWFCVIA 75
B6Phb3.2  MSDSFISFDSV...VGPAHSEASETIAIVDSQSSQLSAIDPRLSSTSLDELNDLPVEFERRTHGGNGLTFWCVIA 72
```

C

```
B42Phb1.1 MDSFDSLDSLNLSVEETTLQTLLESMDTTDA..ASESELDAILVNSERDPGFTSKGFCVIA 59
B42Phb1.2 ...MDSFQQLNLFVEETIQQSLPEAIPSSDSTDTGASERDATPVNTERHLGFTTKGFCVVS 58

B42Phb2.1 MDSFSTLSLPATLGNEEAQQTTVIATVAQ...PQESPSSSGTPVDSERPGAGKVRAFCIIA 58
B42Phb2.2 MDNF.TVDL.ATLFEEFPELQEIQATASEHCSQDQYGSCEGPPINQERPGSGVNRAFCVIA 59
```

Figure 7.6 Alignments of *C. cinereus* predicted pheromone precursor sequences produced by the GAP program (GCG, version 8.1). (**A**) Phb3.1 with Phb3.2. (**B**) *B42* group 3 pheromones with the *B6* group 3 pheromones. (**C**) *B42* group 1 and group 2 pheromone pairs. | indicates identical amino acids and : or . indicate conserved changes.

7.5 Regulation

Studies on the regulation of the pheromone and receptor genes are still somewhat preliminary. However, there is good evidence that for most genes in this large complex there is independent regulation by the *A* and *B* genes. The presence of pheromone and receptor gene transcripts were initially examined in two mycelial types, the monokaryon and a strain heterozygous for *A5/A6* (*A*-on). As seen in Figure 3.11, in the monokaryon most pheromone and receptor gene transcripts were undetectable. This reinforces the conclusion that the pheromone response plays no role before cell fusion. Following A gene activation (*A*-on), transcripts for all three receptors and most of the pheromone genes were highly expressed. In an *A*-on mycelium, transcription of these genes cannot be influenced by the pheromone response pathway and must relate to the fact that the *A* protein heterodimer is present in the cells and *A*-regulated development is activated. Similar experiments were carried out to detect levels of *B* gene transcripts when the pheromone response pathway had been activated by the presence of two different *B* loci (*B42/B6*). In most cases it was clear that the transcript levels were up-regulated, although there was the exception for two of the pheromone genes which were expressed at higher levels in the monokaryon (Figure 5.5).

Why would the transcription of the pheromone and receptor genes be regulated independently by the *A* and the *B* genes? This question still has to be answered. In *U. maydis*, it was shown that transcription of the pheromone and receptor genes was up-regulated some 10-40 fold by activation of the pheromone response pathway (Urban *et al.*, 1996). As in the yeasts, once pheromone binding has activated the pheromone response pathway, transcription of the receptor and pheromone genes themselves is up-regulated to amplify the signal. In *U. maydis* Prf1 plays a pivotal role in co-ordinating the pheromone response and activating transcription of the *b* mating type genes. It is

suggested that continued pheromone signalling after dikaryon formation is necessary only to maintain the expression of the *b* proteins (Hartmann *et al.*, 1995). These form the heterodimeric transcription factor whose activity is essential for filamentous growth and pathogenicity of the dikaryon. This model is not applicable to *C. cinereus* since pheromone signalling is not necessary for cell fusion and the *A* proteins that generate the important heterodimeric transcription factor are expressed constitutively in both unmated and mated cells (Richardson *et al.*, 1993). Activation of the pheromone response pathway is essential to maintain the growth of the homobasidiomycete dikaryon and this would imply that its role is far more fundamental than just regulating transcription levels of the homeobox mating type genes.

7.6 Future perspectives

An important aspect to be addressed in the near future is the cellular localisation of the pheromones and receptors. This may not be an easy task, and has proved difficult in other species. An attempt to use antibodies raised to one of the outside loops of the receptor to localise the receptors of *S. commune* has highlighted the problems of non-specific antibody binding (E. Kothe, pers. com.). An alternative approach would be to make in-frame translation fusions between the receptors and commercially available peptide tags such as His or HA sequences for which the antibodies are available. Localisation of the pheromones is more difficult since tagging is likely to impair the conformation and processing of these molecules.

The identification of the pheromone response factor and its downstream targets is another important area for future research. The pheromone response factor of *U. maydis* was shown to be a possible homologue of the *S. pombe* Ste11 protein (Hartmann *et al*, 1996), as outlined in the introduction. It would clearly be of interest to look for the

corresponding gene in *C. cinereus*. This would open up the possibility of examining the downstream targets involved in maintaining growth of the dikaryon, something that is in itself an interesting developmental pathway, and one that cannot be studied in *U. maydis* because it occurs within the plant host. It is of considerable interest to determine how the activities of the *A* and *B* genes are co-ordinated during dikaryotic growth and to understand why *A*-regulated development leads to induction of the *B* genes independently of the pheromone response.

The genes of the *B* mating type locus encode a large family of 7-transmembrane receptors and peptide pheromone precursors. The organisation of the *C. cinereus* locus indicates that there are three sub-families of these signalling molecules and that only members of the same sub-family can activate the pheromone response pathway. The specificity of the pheromone-receptor interaction is truly remarkable. Receptors can distinguish between pheromones to the extent that they reject those encoded by their own alleles and by those encoded by genes in other sub-families, yet be stimulated by a large number of different pheromones from within their own sub-family. The basidiomycetes are not the only organisms to have evolved large families of receptors and their ligands. Indeed, using this fungal system to analyse such remarkably specific interactions may be extremely useful in understanding their complexity, as it offers all the advantages of genetic and molecular manipulations.

References

Altschul, 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.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. (1997). Gapped BLAST and PSI-BLAST: a new gneration of protein database search programs. Nucl. Acids Res. **25**, 3389-3402.

Aono, T., Yanai, H., Miki, F., Davey, J. and Shimoda, C. (1994). Mating pheromone-induced expression of the *mat1-Pm* gene of *Schizosaccharomyces pombe*: identification of signalling components and characterisation of upstream controlling elements. Yeast **10**, 757-770.

Asante-Owusu, R.N., Banham, A.H., Bohnert, H.U., Mellor, E.J.C. and Casselton, L.A. (1996). Heterodimerization between two classes of homeodomain proteins in the mushroom *Coprinus cinereus* brings together potential DNA-binding and activation domains. Gene **172**, 25-31.

Aviv, H. and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**, 1408-1412.

Bakkeren, G. and Kronstad, J.W. (1994). Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. Proc. Natl. Acad. Sci. USA. **91**, 7085-7089.

Baldwin, J.M. (1994). Structure and function of receptors coupled to G proteins. Curr. Opin. Cell Biol. **6**, 180-190.

Ballance, J. (1990). Transformation systems for filamentous fungi and an overview of fungal gene structure. In Molecular Industrial Mycology - Systems and applications for filamentous fungi., S.A. Leong and R.M. Berka (Marcel Dekker Inc.), 1-29.

Banham, A.H., Asante-Owusu, R.N., Göttgens, B., Thompson, S.A.J., Kingsnorth, C.S., Mellor, E.J.C. and Casselton, L.A. (1995). An N-terminal dimerization domain permits homeodomain proteins to choose compatible partners and intiate sexual development in the mushroom *Coprinus cinereus*. The Plant Cell **7**, 773-783.

Banuett, F. and Herskowitz, I. (1994a). Identification of Fuz7, an *Ustilago maydis* MEK/MAPKK homolog required for *a*-locus-dependent and -independent steps in the fungal life cycle. Genes Devel. **8**, 1367-1378.

Banuett, F. and Herskowitz, I. (1994b). Morphological transitions in the life cycle of *Ustilago maydis* under genetic control by the *a* and *b* loci. Exp. Mycol. **18**, 247-266.

Bender, A. and Sprague, G.F., Jr. (1986). Yeast peptide pheromones, **a**-factor and α-factor, activate a common response mechanism in their target cells. Cell **47**, 929-937.

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

Birnbaumer, L. (1992). Receptor-to-effector signalling through G proteins: roles for βγ dimers as well as α subunits. Cell **71**, 1069-1072.

Bölker, M., Urban, K. and Kahmann, R. (1992). The *a* mating-type locus of *Ustilago maydis* specifies cell signalling components. Cell **68**, 441-450.

Brake, A.J., Brenner, C., Najarian, R., Laybourne, P. and Merryweather, J. (1985). Structure of genes encoding precursors of the yeast peptide mating pheromone **a**-factor. In Protein Transport and Secretion, M. Gething (Cold Spring Harbour Laboratory Press), 103-108.

Büküsoglu, G. and Jenness, D.D. (1996). Agonist-specific conformational changes in the yeast **a**-factor pheromone receptor. Mol. Cell. Biol. **16**, 4818-4823.

Buller, A.H.R. (1931) Researches on fungi. (Longmans Green).

Bullock, W.O., Fernandez, J.M. and Short, J.M. (1987). XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with β-galactosidase selection. Biotechniques **5**, 376-379.

Burrows, D.M. (1991). Transformation studies in the basidiomycete fungi *Coprinus cinereus* and *Coprinus bilanatus*. PhD, University of London (Queen Mary and Westfield College).

Caldwell, G.A., Naider, F. and Becker, J.M. (1995). Fungal lipopeptide mating pheromones: a model system for the study of protein prenylation. Microbiol. Rev. **59**, 406-422.

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

Chen, P., Sapperstein, S.K., Choi, J.D. and Michaelis, S. (1997). Biogenesis of the Saccharomyces cerevisiae mating pheromone **a**-factor. J. Cell Biol. **136**, 251-269.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochem. **18**, 5294-5299.

Christensen, P.U., Davey, J. and Nielsen, O. (1997). The *Schizosaccharomyces pombe mam1* gene encodes an ABC transporter mediating secretion of M-factor. Mol. Gen. Genet. **255**, 226-236.

Coughlin, S.R. (1994). Expanding horizons for receptors coupled to G proteins: diversity and disease. Curr. Opin. Cell Biol. **6**, 191-197.

Dagert, M. and Ehrlich, S.D. (1979). Prolonged incubation in calcium improves the competence of *Escherichia coli* cells. Gene **6**, 23-28.

Davey, J. (1992). Mating pheromones of the fission yeast *Schizosaccharomyces pombe*: purification and structural characterisation of M-factor and isolation and structural characterisation of two genes encoding the pheromone. EMBO J. **11**, 951-960.

Day, P.R. (1960). The structure of the *A* mating type locus in *Coprinus lagopus*. Genet. **45**, 645-650.

Day, P.R. (1963). Mutations of the *A* mating locus in *Coprinus lagopus*. Genet. Res. Camb. **4**, 55-64.

Decottignies, A. and Goffeau, A. (1997). Complete inventory of the yeast ABC proteins. Nature genetics **15**, 137-145.

Dohlman, H.G., Thorner, J., Caron, M.G. and Lefkowitz, R.J. (1991). Model systems for the study of seven-transmembrane segment receptors. Ann. Rev. Biochem. **60**, 653-688.

Dulac, C. and Axel, R. (1995). A novel family of genes encoding putative pheromone receptors in mammals. Cell **83**, 195-206.

Edmonds, M., Vaughan, J.M.H. and Nakazato, H. (1971). Polyadeylic acid sequences in the heterogenous nuclear RNA and rapidly-labelled polysomal RNA of HeLa cells: possible evidence for a precursor relationship. Proc. Natl. Acad. Sci. USA **68**, 1336-1340.

Ehrenhofer-Murray, A.E., Würgler, F.E. and Sengstag, C. (1994). The *Saccharomyces cervisiae SGE1* gene product: a novel drug-resistance protein within the major facilitator superfamily. Mol. Gen. Genet. **244**, 287-294.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to a high specificity. Anal. Biochem. **132**, 6-13.

Fujimura-Kamada, K., Nouvet, F.J. and Michaelis, S. (1997). A novel membrane-associated metalloprotease, Ste24p, is required for the first step of NH_2-terminal processing of the yeast **a**-factor precursor. J. Cell Biol. **136**, 271-285.

Giesy, R.M. and Day, P.R. (1965). The septal pores of *Coprinus lagopus* in relation to nuclear migration. Amer. J. Bot. **52**, 287-293.

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

Glisin, V., Crkvenjakov, R. and Byus, C. (1974). Ribonucleic acid isolated by caesium chloride centrifugation. Biochem. **13**, 2633-2637.

Goffeau, A., Park, J., Paulsen, I.T., Jonniaux, J.-L., Dinh, T., Mordant, P. and Saier, M.H., Jr. (1997). Multidrug-resistance transport proteins in yeast: complete inventory and phylogenetic characterisation of yeast open reading frames within the major facilitator superfamily. Yeast **13**, 43-54.

Golz, J.F., Clarke, A.E. and Newbigin, E. (1995). Self-incompatibility in flowering plants. Curr. Opin. Genet. Dev. **5**, 640-645.

Gottesman, M.M. and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. Ann. Rev. Biochem. **62**, 385-427.

Grosschedl, R., Giese, K. and Pagel, J. (1994). HMG domain proteins: Architectural elements in the assembly of nucleoprotein structures. TIG **10**, 94-100.

Gurr, S.J., Unkles, S.E. and Kinghorn, J.R. (1987). The structure and organisation of nuclear genes of filamentous fungi. In Gene Structure in Eukaryotic Microbes., J.R. Kinghorn (IRL Press), 93-139.

Hagen, D.C., McCaffrey, G. and Sprague, G.F., Jr. (1986). Evidence the yeast *STE3* gene encodes a receptor for the peptide pheromone **a**-factor. Gene sequence and implications for structure of the presumed receptor. Proc. Natl. Acad. Sci. USA **83**, 1418-1422.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**, 557-580.

Hartmann, H.A., Kahmann, R. and Boelker, M. (1996). The pheromone response factor coordinates filamentous growth and pathogenicity in *Ustilago maydis*. EMBO J. **15**, 1632-1641.

Hartwell, L. (1980). Mutants of *Saccharomyces cerevisiae* unresponsive to cell division control by polypeptide mating hormone. J. Cell Biol. **85**, 811-822.

Haylock, R.W., Economou, A. and Casselton, L.A. (1980). Dikaryon formation in *Coprinus cinereus*: selection and identification of *B* factor mutants. J. Gen. Microbiol. **121**, 17-26.

Herskowitz, I. (1988). Life cycle of the budding yeast *Saccharomyces cerevisiae*. Microbiol. Rev. **52**, 536-553.

Herskowitz, I. (1995). MAP kinase pathways in yeast: for mating and more. Cell **80**, 187-197.

Higgins, C.F. (1992). ABC transporters: from microorganisms to man. Ann. Rev. Cell Biol. **8**, 67-113.

Ho, C., 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.

Hoge, J.H.C., Springer, J., Zantinge, B. and Wessels, J.G.H. (1982). Absence of differences in polysomal RNAs from vegetative monokaryotic cells of the fungus *Schizophyllum commune*. Exp. Mycol. **6**, 225-232.

Imai, Y. and Yamamoto, M. (1994). The fission yeast mating pheromone P-factor: its molecular structure, gene structure and physiological activities to induce gene expression and G1 arrest in the mating partner. Genes. Dev. **8**, 328-338.

Jeansonne, N.E. (1994). Yeast as a model system for mammalian seven-transmembrane segment receptors. Proc. Soc. Exp. Biol. & Med. **206**, 35-44.

Johnson, A.D. (1995). Molecular mechanisms of cell-type determination in budding yeast. Curr. Opin. Genet. Dev. **5**, 552-558.

Julius, D., Schekman, R. and Thorner, J. (1984) Glycosylation and processing of prepro-α-factor through the yeast secretory pathway. Cell **36**, 309-318.

Kämper, J., Reichmann, M., Romeis, T., Bölker, M. and Kahman, R. (1995). Multiallelic recognition: nonself-dependant dimerization of the bE and bW homeodomain proteins in *Ustilago maydis*. Cell **81**, 73-83.

Kelly, M., Burke, J., Smith, M., Klar, A. and Beach, D. (1988). Four mating-type genes control sexual differentiation in the fission yeast. EMBO J. **7**, 1537-1547.

Kingsnorth, C. S. (1996). Identification of genes regulated by the *A* mating type of *Coprinus cinereus*. DPhil, University of Oxford.

Kitamura, K. and Shimoda, C. (1991). The *Schizosaccharomyces pombe mam2* gene encodes a putative pheromone receptor which has significant homology with the *Saccharomyces cerevisiae* STE2 protein. EMBO J. **10**, 3743-3751.

Kjærulff, S., Davey, J. and Nielsen, O. (1994). Analysis of the structural genes encoding the M-factor in the fission yeast *Schizosaccharomyces pombe*: identification of a third gene *mfm3*. Mol. Cell. Biol. **14**, 3895-3905.

Koltin, Y. (1978). In Genetics and morphogenesis in the basidiomycetes, M.N. Schwalb and P.G. Miles (Academic Press), 31-54.

Koltin, Y., Stamberg, J. and Lemke, P.A. (1972). Genetic structure and evolution of the incompatibility factors in higher fungi. Bacteriol. Rev. **36**, 156-171.

Kuchler, K., Sterne, R.E. and Thorner, J. (1989). *Saccharomyces cerevisiae STE6* gene product: a novel pathway for protein export in eukaryotic cells. EMBO J. **8**, 3973-3984.

Kües, U. and Casselton, L.A. (1992). Fungal mating type genes-regulators of sexual development. Mycol. Res. **96**, 993-1006.

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

Kües, U., Tymon, A.M., Richardson, W.V.J., May, G., Gieser, P.T. and Casselton, L.A. (1994). *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.

Kurjan, J. (1993). The pheromone response pathway in *Saccharomyces cerevisiae*. Annu. Rev. Genet. **27**, 147-179.

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**, 105-132.

Laskey, R.A. and Mills, D. (1977). Enhanced autoradiographic detection of ^{32}P and ^{125}I using intensifying screens and hypersensitized film. FEBS Lett. **82**, 314-316.

Leberer, E., Thomas, D.Y. and Whiteway, M. (1997). Pheromone signalling and polarized morphogenesis in yeast. Curr. Opin. Genet. Dev. **7**, 59-66.

Leupold, U., Nielsen, O. and Egel, R. (1989). Pheromone-induced meiosis in P-specific mutants of fission yeast. Curr. Genet. **15**, 403-405.

Lewis, K. (1994). Multidrug resistance pumps in bacteria: variations on a theme. TIBS **19**, 119-123.

Li, X., Paech, N., Nield, J., Hayman, D. and Langridge, P. (1997). Self-incompatibility in the grasses: evolutionary relationship of the *S* gene from *Phalaris coerulescens* to homologous sequences in other grasses. Plant Mol. Biol. **34**, 223-232.

Little, P.F.R. (1987). Choice and use of cosmid vectors. In DNA cloning: Volume III, a practical approach, D.M. Glover (IRL Press), 19-42.

MacKay, V.L. and Manney, T.R. (1974). Mutations affecting sexual conjugation and related processes in *Saccharomyces cerevisiae*. II. Genetic analysis of non mating mutants. Genet. **76**, 273-288.

Magae, Y., Novotny, C. and Ullrich, R. (1995). Interaction of the *Aα Y* mating-type and *Z* mating-type homeodomain proteins of *Schizophyllum commune* detected by the two-hybrid system. Biochem. Biophys. Res. Comm. **211**, 1071-1076.

Marger, M.D. and Saier, M.H., Jr. (1993). A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. TIBS **18**, 13-20.

McGrath, J.P. and Varshavsky, A. (1989). The yeast *STE6* gene encodes a homologue of the mammalian multidrug resistance P-glycoprotein. Nature **340**, 400-404.

Mellon, F.M., Little, P.F.R. and Casselton, L.A. (1988). Gene cloning and transformation in the basidiomycete *Coprinus cinereus*: isolation and expression of the isocitrate lyase gene (*acu-7*). Mol. Gen. Genet. **210**, 352-357.

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.

Nasrallah, J.B., Stein, J.C., Kandasamy, M.K. and Nasrallah, M.E. (1994). Signalling the arrest of pollen tube development in self-incompatible plants. Science **266**, 1505-1508.

Newbigin, E., Anderson, M.A. and Clarke, A.E. (1993). Gametophytic self-incompatibility systems. Plant Cell **5**, 1315-1324.

Nielsen, O. and Davey, J. (1995). Pheromone communication in the fission yeast *Schizosaccharomyces pombe*. Semin. Cell. Biol. **6**, 95-104.

Nielsen, O., Davey, J. and Egel, R. (1992). The *ras1* function of *Schizosaccharomyces pombe* mediates pheromone-induced transcription. EMBO J. **11**, 1391-1395.

Ohta, T. (1996) The current significance and standing of neutral and nearly neutral theories. BioEssays **18**, 673-684.

Ohki, R. and Murata, M. (1997). *bmr3*, a third multidrug transporter gene of *Bacillus subtilis*. J. Bacteriol. **179**, 1423-1427.

Omer, C.A. and Gibbs, J.B. (1994). Protein prenylation in eukaryotic microorganisms: genetics, biology and biochemistry. Mol. Microbiol. **11**, 219-225.

O'Shea, S. F. (1995). Isolation and characterisation of the *B* mating type locus of *Coprinus cinereus*. DPhil, University of Oxford.

O'Shea, S.F., Chaure, P.T., Halsall, J.R., Olesnicky, N.S., Leibbrandt, A., Connerton, I.F. and Casselton, L.A. (1998). A large pheromone and pheromone receptor gene complex determines *B* mating specificity in *Coprinus cinereus*. Genet. **in press**.

Papazian, H.P. (1951). The incompatibility factors and a related gene in *Schizophyllum commune*. Genet. **36**, 441-459.

Pardo, E.H., O'Shea, S.F. and Casselton, L.A. (1996). Multiple versions of the *A* mating type locus of *Coprinus cinereus* are generated by three paralogous pairs of multiallelic homeobox genes. Genet. **144**, 87-94.

Petersen, J., Weilguny, D., Egel, R. and Nielsen, O. (1995). Characterization of *fus1* of *Schizosaccharomyces pombe*: a developmentally controlled function needed for conjugation. Mol. Cell. Biol. **15**, 3697-3707.

Pitkin, J.W., Panaccione, D.G. and Walton, J.D. (1996). A putative cyclic peptide efflux pump encoded by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum*. Microbiol. **142**, 1557-1565.

Puhalla, J.E. (1970). Genetic studies of the *b* incompatibility locus of *Ustilago maydis*. Genet. Res. **16**, 229-232.

Quintanilha, A. and Pinto-Lopes, J. (1950). Aperçu sur l'état actuel de nos connaissances concernant la "conduite sexuelle" des espèces d'Hymenomycetes. Bol. Soc. Broteriana **24**, 115-290.

Rao, P.S. and Niederpruem, D.J. (1969). Carbohydrate metabolism during morphogenesis of *Coprinus lagopus* (*sensu* Buller). J. Bacteriol. **100**, 1222-1228.

Raper, J.R. (1966) Genetics of sexuality in higher fungi. (The Ronald Press).

Raper, J.R., Baxter, M.G. and Ellingboe, A.H. (1960). The genetic structure of the incompatibility loci in *Schizophyllum commune*: the *A* factor. Proc. Natl. Acad. Sci. USA **46**, 833-842.

Raper, J.R., Baxter, M.G. and Middleton, R.B. (1958). The genetic structure of the incompatibility loci in *Schizophyllum*. Proc. Natl. Acad. Sci. USA **44**, 889-900.

Raymond, M., Gros, P., Whiteway, M. and Thomas, D.Y. (1992). Functional complementation of yeast *ste6* by a mammalian multidrug resistance *mdr* gene. Science **256**, 232-234.

Regenfelder, E., Spellig, T., Hartmann, A., Lauenstein, S., Bölker, M. and Kahman, R. (1997). G proteins in *Ustilago maydis*: transmission of multiple signals? EMBO J. **16**, 1934-1942.

Richardson, W.V.J., Kües, U. and Casselton, L.A. (1993). The *A* mating-type genes of the mushroom *Coprinus cinereus* are not differentially transcribed in monokaryons and dikaryons. Mol. Gen. Genet. **238**, 304-307.

Rosen, K.M. and Villa-Komaroff, L. (1993). An alternative method for the visualisation of RNA in formaldehyde agarose gels. Focus **12**, 23-24.

Ross, I. K. (1976) Nuclear migration rates in *Coprinus congregatus*: a new record? Mycologia **68**, 418-422.

Roth, A.F. and Davis, N.G. (1996). Ubiquitination of the yeast α-factor receptor. J. Cell Biol. **134**, 661-674.

Rouch, D.A., Cram, D.S., DiBerardino, D., Littlejohn, T.G. and Skurray, R.A. (1990). Efflux-mediated antiseptic resistance gene *qacA* from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar-transport proteins. Mol. Microbiol. **4**, 2051-2062.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning. A laboratory manual. (Cold Spring Harbour Laboratory Press).

Scott, M.P., Tamkum, J.W. and Hartzell, G.W., III (1989). The structure and function of the homeodomain. Biochim. Biophys. Acta **989**, 25-48.

Shahriari, H. and Casselton, L.A. (1974). Suppression of methionine mutants in *Coprinus* I. Complementation and allele specificity as criteria of suppressor gene action. Mol. Gen. Genet. **134**, 85-92.

Snider, P.J. and Raper, J.R. (1958). Nuclear migration in the basidiomycete *Schizophyllum commune*. Amer. J. Bot. **45**, 538-546.

Specht, C.A. (1995). Isolation of the $B\alpha$ and $B\beta$ mating-type loci of *Scizophyllum commune*. Curr. Genet. **28**, 374-379.

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 $A\alpha$ locus of *Schizophyllum commune*. Proc. Acad. Natl. Sci. USA **89**, 7174-7178.

Specht, C.A., Stankis, M.M., Novotny, C.P. and Ullrich, R.C. (1994). Mapping the heterogeneous DNA region that determines the 9 $A\alpha$ mating-type specificities of *Schizophyllum commune*. Genet. **137**, 709-714.

Spellig, T., Bölker, M., Lottspeich, F., Frank, R.W. and Kahmann, R. (1994). Pheromones trigger filamentous growth in *Ustilago maydis*. EMBO J. **13**, 1620-1627.

Sugimoto, A., Lino, Y., Maeda, T., Watanabe, Y. and Yamamoto, M. (1991). *Schizosaccharomyces pombe ste*11^{+} encodes a transcription factor with an HMG motif that is a critical regulator of sexual development. Genes & Dev. **5**, 1990-1999.

Swiezynski, K.M. and Day, P.R. (1960a). Heterokaryon formation in *Coprinus lagopus*. Genet. Res. **1**, 114-128.

Swiezynski, K.M. and Day, P.R. (1960b). Migration of nuclei in *Coprinus lagopus*. Genet. Res. **1**, 129-139.

Takiff, H.E., Cimino, M., Musso, M.C., Weisbrod, T., Martinez, R., Delgardo, M.B., Salazar, L., Bloom, B.R. and Jacobs, W.R., Jr. (1996). Efflux pump of the proton antiporter family confers low-level fluoroquinolone resistance in *Mycobacterium smegmatis*. Proc. Natl. Acad. Sci. USA **93**, 362-366.

Tanaka, K., Davey, J., Imai, Y. and Yamamoto, M. (1993). *Schizosaccharomyces pombe map3*$^{=}$ encodes the putative M-factor receptor. Mol. Cell. Biol. **13**, 80-88.

Troemel, E.R., Chou, J.H., Dwyer, N.D., Colbert, H.A. and Bargmann, C.I. (1995). Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. Cell **83**, 207-218.

Urban, M., Kahmann, R. and Boelker, M. (1996). Identification of the pheromone response element in *Ustilago maydis*. Molecular & General Genetics **251**, 31-37.

Vaillancourt, L.J., Raudaskoski, M., Specht, C. and Raper, C.A. (1997). Multiple genes encoding pheromones and a pheromone receptor define the *Bβ1* mating-type specificity in *Schizophyllum commune*. Genet. **146**, 541-551.

Wang, Z. and Rossman, T.G. (1994). Isolation of DNA fragments from agarose gel by centrifugation. Nucl. Acids Res. **22**, 2862-2863.

Wendland, J., Vaillancourt, L.J., Hegner, J., Lengeler, K.B., Laddison, K.J., Specht, C.A., Raper, C.A. and Kothe, E. (1995). The mating-type locus *Bα1* of *Schizophyllum commune* contains a pheromone receptor gene and putative pheromone genes. EMBO J. **14**, 5271-5278.

Whitehouse, H.L.K. (1949). Multiple-allelomorph heterothallism in the fungi. New Phytol. **48**, 212-244.

Yamaguchi, A., Someya, Y. and Sawai, T. (1992). Metal-tetracyclin/H^+ antiporter of *Escherichia coli* encoded by transposon *Tn10*. J. Biol. Chem. **267**, 19155-19162.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**, 103-119.

Yuan, Y.O., Stroke, I.L. and Fields, S. (1993). Coupling of cell identity to signal response in yeast: interaction between the α1 and STE12 proteins. Genes Dev. **7**, 1584-1597.

Zolan, M.E. and Pukkila, P.J. (1986). Inheritance of DNA methylation in *Coprinus cinereus*. Mol. Cell. Biol. **6**, 195-200.