Journal of Cell Science 104, 227-230 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

COMMENTARY

The origin of multiple mating types in mushrooms

Ursula Kües and Lorna A. Casselton*

Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

*Author for correspondence

INTRODUCTION

Having multiple mating types greatly improves the chances of meeting a compatible mating partner, particularly in an organism like the mushroom that has no sexual differentiation and no mechanism for signalling to a likely mate. Having several thousands of mating types, as some mushrooms do, is, however, remarkable - and even more remarkable is the fact that individuals only recognise that they have met a compatible mate after their cells have fused. How are such large numbers of mating types generated and what is the nature of the intracellular interaction that distinguishes self from non-self? Answers to these fascinating questions come from cloning some of the mating type genes of the ink cap mushroom *Coprinus cinereus*.

A successful mating in *Coprinus* triggers a major switch in cell type, the conversion of a sterile mycelium with uninucleate cells (monokaryon) to a fertile mycelium with binucleate cells (dikaryon) which differentiates the characteristic fruit bodies. The mating type genes that regulate this developmental switch map to two multiallelic loci designated *A* and *B* and these must both carry different alleles for full mating compatibility (Casselton, 1978). *A* and *B* independently regulate different steps in the developmental switch, making it possible to study just one component of the system (Swiezynski and Day, 1960) and work in our laboratory has concentrated on understanding the structure and function of the *A* genes. It is estimated that some 160 different *A* mating types exist in nature (Raper, 1966), any two of which can together trigger the *A*-regulated part of sexual development. The first clue to how such large numbers are generated came from classical genetic analysis, which identified two functionally redundant *A* loci, α and β (Day, 1960). Functional redundancy is, indeed, the key to multiple *A* mating types and, as seen in Fig.1, molecular cloning has identified many more genes than was possible by recombination analysis.

GENES OF THE *A* LOCI - THEIR STRUCTURE AND ORGANISATION

Fig. 1A compares the first two *A* gene complexes cloned, *A*42 and *A*43. The basic organisation is similar, the genes that determine *A* specificity (the specificity genes) are separated into two subcomplexes (corresponding to the genetically defined and loci) by a region of non-coding DNA and are flanked by two genes of unknown function *-fg* and *-fg*. The specificity genes appear to encode transcription factors because their proteins have a characteristic DNA binding homeodomain. At first sight it is surprising to find that the number of specificity genes and their direction of transcription are variable. Moreover, there are two classes of genes distinguished by transcript size (2.5 kb and 2.1 kb) and, so far as they have been sequenced, distinctly different amino acid sequences in the homeodomains. We have called these two homeodomain types HD1 and HD2. By transformation into a host cell with a different *A* mating type it can be shown that a single gene of either type is sufficient to trigger the *A*-regulated developmental pathway (Kües et al., 1992; Kües and Casselton, 1992a,b).

The apparently haphazard collection of specificity genes is at first confusing. Fortunately, these genes are sufficiently different in sequence to not cross-hybridise and so we can identify alleles of *A*42 and *A*43 genes present in other *A* gene complexes. What we find leads us to suggest the structure of the archetypal *A* complex illustrated in Fig.1B. This shows four pairs of divergently transcribed HD1 and HD2 genes, which we have called the *a, b, c* and *d gene-pairs* (Kües and Casselton, 1992c), each, perhaps, with no more than four different allelic forms. Significantly, for a particular gene-pair, the same combination of alleles is always found together as an *allele-pair* (i.e. *b1-1* + *b2-1, b1-2* + *b2-2* etc.). The allelic forms of the different gene pairs can, however, be recombined in all possible combinations and this is how the 160 different *A*s are generated. Neither *A*42 nor *A*43 contains all of the eight predicted specificity genes - three are deleted in *A*42 and two in *A*43 Fig. 1A. With so much functional redundancy it is not surprising that specificity genes can be lost.

GENE PAIRS AND THE RULES FOR MATING COMPATIBILITY

To understand the rules of a compatible interaction it is necessary to look only at the *a* gene-pair at the *A* locus, since classical analysis has already shown that genes of the

Key words: basidiomycetes, *Coprinus,* homeodomains, mating type, sexual development

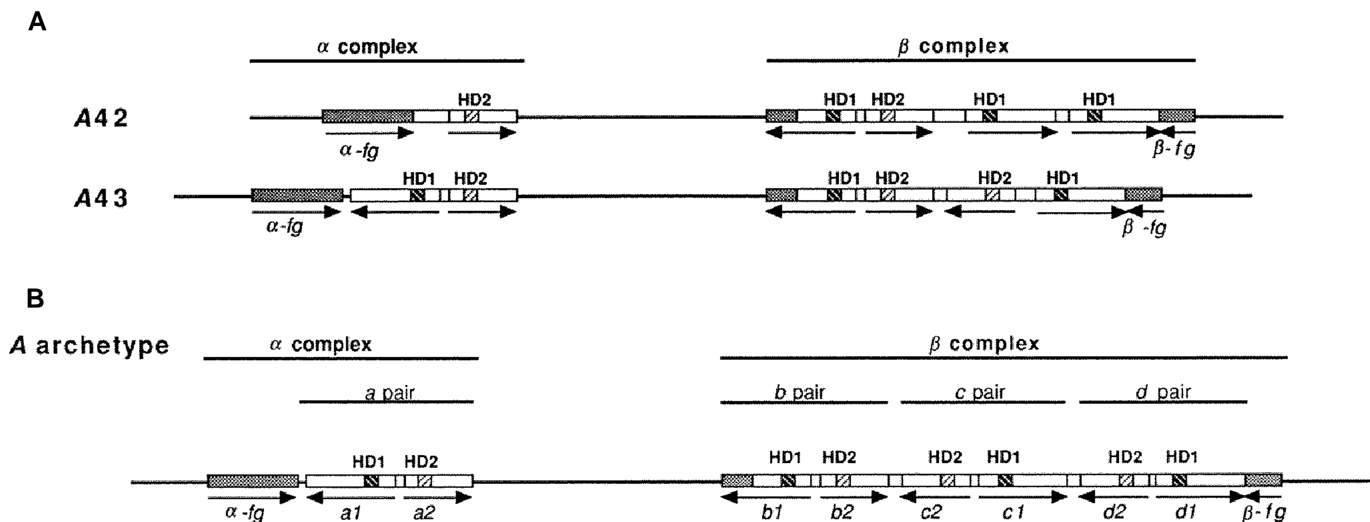


Fig. 1. (A) Organisation of the *A*42 and *A*43 mating type loci of *Coprinus cinereus* cloned by Mutasa et al. (1990) and May et al. (1991), respectively. The specificity genes are shown as open boxes with their direction of transcription indicated by the arrows. The sequences coding for the two classes of homeodomains are indicated by the differently striped boxes; white indicates HD1 and black HD2. Not all *A*43 homeodomains have been confirmed by sequencing but have been deduced from gene transcript sizes of 2.5 kb (HD1) or 2.1 (HD2). Common gene sequences are shown by stippled boxes. (B) Postulated archetypal *A* gene complex showing four divergently transcribed gene-pairs encoding the HD1 and HD2 classes of proteins. The orientation of the gene-pairs was deduced from hybridisation data.

A and *A* locus act quite independently (Day, 1960). *A*42 and *A*43 have different alleles of this gene-pair but, conveniently, the HD1 gene is deleted from *A*42. Introduction of the HD1 gene from *A*43 into an *A*42 host cell is sufficient to switch on *A*-regulated development but introduction of its HD2 partner is not (Kües and Casselton, 1992a). The compatible interaction is a non-allelic one between an HD1 and an HD2 gene from different *a* allele-pairs.

This unexpected observation is confirmed by studies with similar genes from two other basidiomycete species, the wood-rotting mushroom *Schizophyllum commune* and the smut fungus *Ustilago maydis*, a pathogen of maize. The mating type genes of *Schizophyllum* are very similar to those of *Coprinus* (Raper, 1966) and recent sequence analysis of the *A* locus genes also confirms that these encode two proteins with similar distinct homeodomains (Stankis et al., 1992). Three different *A* loci have been cloned and one of these, like *A*42, has only an HD2 gene whereas the other two have the complete HD1-HD2 gene-pair like *A*43. Either of the cloned HD1 genes can elicit *A*-regulated development in the host with only the HD2 gene whereas neither of their HD2 gene partners can do so (Specht et al., 1992). The analogous mating type locus in the hemibasidiomycete *Ustilago* is the *b* locus, which regulates development of the pathogenic dikaryotic stage of the life cycle (Banuett, 1992). The *b* locus is simpler than the *A* gene complex of the mushroom fungi with just a single divergently transcribed gene-pair known as *bW* and *bE* (Gillissen et al., 1992) but there are at least 33 different allele-pairs (*bW1-bE1, bW2-bE2,* etc.). By an elegant series of gene disruptions, Gillissen et al. (1992) have shown that it requires just a single *bW* allele from one *b* locus together with a single *bE* allele from another *b* locus (i.e. *bW1-bE2* or *bW2-bE1*) to trigger *b*-regulated development. We are not surprised to discover that the predicted *bE* and *bW* proteins have distinct homeodomains with strong homology to those of the HD1 and HD2 proteins of *Coprinus* and *Schizophyllum* (Kües and Casselton, 1992b).

HOMEODOMAIN PROTEINS AND THE EVOLUTION OF MULTIPLE MATING TYPES

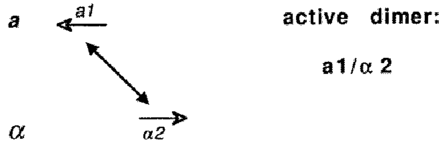
An interaction between two dissimilar homeodomain proteins is reminiscent of one of the best-characterised fungal mating type protein interactions, that between the **a**1 and 2 proteins of the budding yeast *Saccharomyces cerevisiae* (Herskowitz, 1989, 1992; Dolan and Fields, 1991). Like all ascomycete fungi, yeast has only two mating types determined by alternative genes at a single locus. The **a**1 and 2 proteins are coded by genes in different haploid cells but are brought together by mating and dimerise to give a new regulatory protein in diploid cells (Dranginis, 1990). If we compare the yeast homeodomains (Shepherd et al., 1984) with those of the basidiomycete proteins there is sufficient similarity to argue that 2 is analogous to HD1 and **a**1 to HD2 (Kües et al., 1992), and to suggest a common evolutionary origin for the mating type proteins in all these fungi (Kües and Casselton, 1993).

With mating compatibility based on heterodimer formation, multiple mating types could only evolve if both classes of genes resided at the same locus. Every mating partner must be equal and able to contribute both proteins. This leads to functional redundancy on the one hand (see legend to Fig. 2) and a major problem in protein-protein recognition on the other. Why is it that protein pairs encoded by the same gene-pair and present before mating cannot form active dimers whereas new protein pairs generated by mating can? To answer this question would be the key to understanding recognition.

Sequencing of the *Coprinus* genes reveals little sequence similarity but an an overall structural similarity between proteins of the same class. Proteins share three features: a

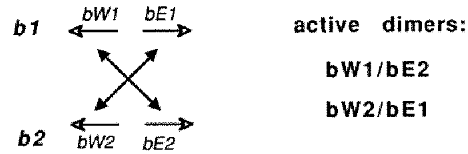
Saccharomyces cerevisiae

2 mating types
2 homeotic genes, 1 allele



Ustilago maydis

33 *b* mating types
2 homeotic genes, 33 alleles



Coprinus cinereus

160 *A* mating types?
4 x 2 homeotic genes, 2-4 alleles of each?

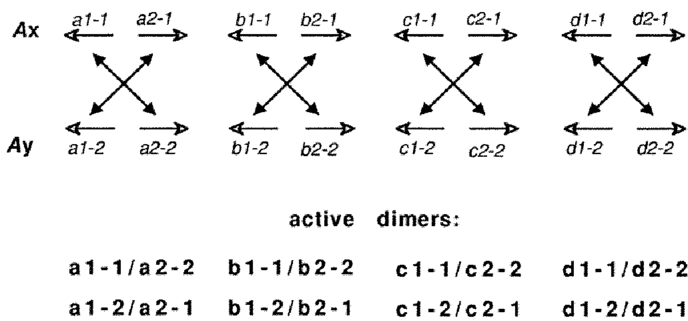


Fig. 2. Heterodimer formation of mating type proteins following a compatible mating in *Saccharomyces cerevisiae*, *Ustilago maydis* and *Coprinus cinereus.* A single heterodimer is sufficient to trigger sexual development in all three fungi. In the simplest case, *S. cerevisiae*, each mate contributes only one component so that only a single heterodimer is formed. In *U. maydis*, each mate contributes one of the alternative protein types and two equivalent heterodimers can be formed. In *C. cinereus*, with a predicted maximum of four gene-pairs, anything from one to eight equivalent heterodimers can be generated depending on whether both members of the gene-pair are present, the actual number of genes present and how many of these have different alleles in the two mates. Note that heterodimer formation is only possible between proteins coded by different alleles of the same gene-pair. (The functions of *a*, *b* and *d* genes have been confirmed by transformation but we still have to confirm that the *c* gene-pairs are active.)

homeodomain, some potentially helical regions that could represent a major dimerisation interface, and a high serine, threonine and proline content in regions that are likely transactivation domains (Fig. 3). The most interesting region is a very unconserved helical domain just N-terminal to the homeodomain designated COP in Fig. 3. The *Ustilago b* proteins have an analogous unconserved region that has been shown to determine their allelic specificity (Dahl et al., and Yee and Kronstad, quoted by Saville and Leong, 1992) and we can show that the N-terminal sequence containing the COP domain is responsible for distinguishing proteins from different *A* gene-pairs in *Coprinus* (unpublished data). Similar regions are present in certain animal homeodomain transcription factors (POU transcription factors) and, by analogy, we predict that this region could participate in DNA target site selection as well as being a dimerization domain with the necessary variability to discriminate compatible from incompatible protein pairings (Tymon et al., 1992; Kües and Casselton, 1992b). Even with such variability demanded of the specificity domains, the conserved homeodomains should recognise the same DNA target site: indeed, it would be very uneconomical to have multiple target sites to accommodate different protein pairs.

DIFFERENT STRATEGIES FOR CREATING MULTIPLE MATING TYPES

Ustilago and *Coprinus* have evolved different strategies for generating multiple mating types. *Ustilago* has evolved some 33 alleles of a single gene-pair (Gillissen et al., 1992), *Coprinus* a few alleles of several gene-pairs. Since only a single HD1-HD2 protein interaction is sufficient to trigger *A*-regulated development, four alleles of each gene-pair would be sufficient to generate 256 different *A* mating types. The *Coprinus* system is more efficient in all respects: firstly, in generating large numbers of mating types; and secondly, in reducing the number of compatible or incompatible interactions that need to be recognised. For *Ustilago* there are1056 compatible interactions (33 × 32; Fincham, 1992) and 33 incompatible ones. For *Coprinus* there are 48

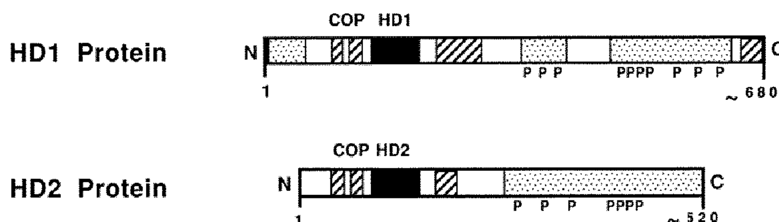


Fig. 3. Schematic presentation of the structures of the HD1 and HD2 *A* mating type proteins of *Coprinus cinereus*. Helical regions are shown by diagonal lines, the DNA binding motifs by black boxes and the stippled regions represent regions rich in serine and threonine. P indicates proline-rich sequences.

compatible interactions if we assume as many as four pairs of alleles at each of *a*, *b*, *c* and *d* ($4 \times 3 \times 4$), and 16 incompatible ones (4×4). In *Coprinus* we assume that there is no need to recognise proteins from different gene-pairs but that there is recognition between proteins from different allele-pairs. Distinguishing active protein combinations from a far fewer number of inactive combinations might be achieved in either of two ways: incompatible proteins are prevented from forming a stable dimer on DNA (Kües and Casselton, 1992b), or they form such a tight dimer in solution that their DNA binding domains are hidden (Gillissen et al., 1992).

It is essential to the evolution of multiple mating types in both *Ustilago* and *Coprinus* that the two genes of a pair (e.g. *bW-bE* of *Ustilago*, *a1-a2, b1-b2* etc. of *Coprinus*) must be inseparable, otherwise recombination would bring together active allele combinations and replace the need to mate! Homologous recombination between *bW* and *bE* is prevented by the fact that they are divergently transcribed, they are separated by 260 bp of variable DNA sequence and both genes have alleles with highly variable 5 sequences (Gillissen et al., 1992; Schulz et al., 1990; Kronstad and Leong, 1990). In *Coprinus* a much greater degree of sequence variablity exists, the alleles of the genes have little homology as indicated by their failure to cross-hybridise and the flanking sequences are equally variable (Kües et al., 1992). The restraints are far greater in *Coprinus* where not only must there be no recombination between the genes of a pair, but there must be no irregular recombination involving genes from different pairs.

Classical genetic studies suggested that the multiple mating types of the basidiomycetes arose by gene duplication and mutation (Raper, 1966). Molecular data show that the genes that exist today have diverged very widely and it is not surprising that many of the traditional studies aimed at generating new alleles by mutation met with failure (Casselton, 1978). Considerable evolutionary effort has been expended to evolve genes with as little as 50% DNA homology between alleles (*Schizophyllum A* alleles; Stankis et al., 1990) and even less between genes of the same *A* complex (*Coprinus*, unpublished data) yet still coding for functionally equivalent proteins.

We thank our colleagues Rachel Asante-Owusu, Berthold Göttgens, Suzanne O'Shea, Eneida Pardo and Wendy Richardson for allowing us to discuss their unpublished data. Work in our laboratory is funded by AFRC, SERC and The Gatsby Charitable Foundation.

REFERENCES

Banuett, F. (1992). *Ustilago maydis*, the delightful blight. *Trends. Genet.* **8**, 174-180.

Casselton, L.A. (1978). Dikaryon formation in higher basidiomycetes. In *The Filamentous Fungi,* vol. 3 (ed. J.E. Smith and D.R. Berry), pp. 275-297. Edward Arnold: London.

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

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

Dranginis, A.M. (1990). Binding of yeast **a**1 and 2 as a heterodimer to the operator DNA of a haploid-specific gene. *Nature*, **347**, 682-685.

Fincham, J.R.S. (1992). Distinguishing self from non-self. *Current Biol.* **2**, 377-378.

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

Herskowitz, I. (1989). A regulatory hierarchy for cell specialization in yeast. *Nature* **342**, 749-757.

Herskowitz, I. (1992). Yeast branches out. *Nature* **357**, 190-191.

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

Kües, U. and Casselton, L.A. (1992a). Molecular and functional analysis of the *A* mating type genes of *Coprinus cinereus.* In *Genetic Engineering, Principles and Methods*, vol. 14 (ed. J.K. Setlow), pp. 251-268, Plenum Press: New York.

Kües, U. and Casselton, L. A. (1992b). Homeodomains and regulation of sexual development in basidiomycetes. *Trends Genet.* **8**, 154-155.

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

Kües, U. and Casselton, L.A. (1993). Regulation of fungal development by mating type genes. In *The Eukaryotic Microbial Genome* (ed. P.M. Broda, S.G. Oliver and P. Sims). *SGM Symposium,* Cambridge University Press (in press).

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

May, G., LeChevanton, L. and Pukkila, P.J. (1991). Molecular analysis of the *Coprinus cinereus* mating type *A* factor demonstrated a unexpectedly complex structure. *Genetics* **128**, 529-538.

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.

Raper, J.R. (1966). *Genetics of Sexuality in Higher Fungi*. The Ronald Press: New York.

Saville, B.J. and Leong, S.A. (1992). Molecular biology of pathogenesis in *Ustilago maydis*. In *Genetic Engineering, Principles and Methods*, vol. 14 (ed. J.K. Setlow), pp. 139-162. Plenum Press: New York.

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

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

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*α locus of *Schizophyllum commune. Proc. Nat. Acad. Sci. USA* **89**, 7174-7178.

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

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

Tymon, A.M., Kües, U., Richardson, W.V.J. and Casselton, L.A. (1992). A fungal mating type protein that regulates asexual and sexual development contains a POU-related domain. *EMBO J.* **11**, 1805-1813.