New Phytologist

Research

Creating novel specificities in a fungal nonself recognition system by single step homologous recombination events

Jörg Kämper[1,2], Michael W. Friedrich[1,3] and Regine Kahmann[1]

[1]Department Organismic Interactions, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse 10, Marburg 35043, Germany; [2]Department of Genetics, Institute for Applied Biosciences, Karlsruhe Institute of Technology, Fritz Haber Weg 4, Karlsruhe 76131, Germany; [3]Microbial Ecophysiology Group BreMarE, University of Bremen, Leobener Straße 3, Bremen 28359, Germany

Authors for correspondence:
Regine Kahmann
Tel: +49 6421 1788501
Email: kahmann@mpi-marburg.mpg.de

Jörg Kämper
Tel: +49 721 608 45670
Email: kaemper@kit.edu

Received: *9 March 2020*
Accepted: *5 May 2020*

New Phytologist (2020)
doi: 10.1111/nph.16755

Key words: evolution, evolution of new mating specificities, homeodomain protein, mating type, self-incompatibility, self–nonself recognition, *Ustilago maydis*.

Summary

- In many organisms, two component systems have evolved to discriminate self from nonself. While the molecular function of the two components has been elucidated in several systems, the evolutionary events leading to the large number of different specificities for self–nonself recognition found in most systems remain obscure.
- We have investigated the variation within a multiallelic nonself recognition system in the phytopathogenic basidiomycete *Ustilago maydis* by means of sequence analysis and functional studies.
- The multiallelic *b* mating type locus of *U. maydis* ensures outbreeding during sexual development. Nonself recognition is specified by the two homeodomain proteins, bE and bW, encoded by the *b* locus. While bE–bW combinations from the same allele do not dimerize, bE and bW proteins originating from different alleles form a heterodimeric complex that functions as master regulator for sexual and pathogenic development.
- We show that novel specificities of the *b* mating type locus have arisen by single homologous recombination events between distinct *b* alleles that lead to a simultaneous exchange of subdomains involved in dimerization in both bE and bW, altering the specificity of both proteins in a single step.

Introduction

Self–nonself discrimination between organisms is essential for sexual reproduction, defense against pathogen invasion, and the maintenance of individuality. For the majority of self–nonself recognition systems, a large number of different specificities have evolved. It is an accepted view that allelic variance of self-incompatibility (SI) loci is maintained by balancing selection: rare alleles are at a selective advantage, but become disadvantageous when frequent. As a consequence, the loci involved in self–nonself discrimination include the most polymorphic genes known in eukaryotes (Takahata & Nei, 1990; Takahata *et al.*, 1992; Hinata *et al.*, 1995; Wu *et al.*, 1998; Richman, 2000).

In plants, SI systems promote outcrossing by rejecting self-related pollen (Fujii *et al.*, 2016). Within the SI system of *Brassicaceae*, the recognition of self-related pollen is controlled through the activity of the two closely linked genes for the S-locus receptor kinase (SRK) and the S-locus cysteine-rich protein (SCR) localized in the pollen coat. In this system, self-pollen is rejected based on the self-specific molecular interaction between the receptor and the SCR ligand, while a nonself ligand fails to trigger this response (Schopfer *et al.*, 1999; Takasaki *et al.*, 2000; Shiba *et al.*, 2001).

In *Papaveracea*, the female component of the SI system, the small secreted PrsS protein, needs to encounter a male PrpS transmembrane protein originating from a different haplotype; interaction of PrsS and PrpS from the same haplotype results in self-rejection of pollen (Thomas & Franklin-Tong, 2004; Wheeler *et al.*, 2009, 2010; Poulter *et al.*, 2010). In contrast to these self-recognition systems where the two components interact when they originate from the same haplotype, *Solanaceae, Rosaceae* and *Scrophulariaceae* have evolved nonself recognition systems. Here, the female S-determinant is a S-RNase that enters the self-pollen tube and inhibits pollen tube growth via its cytotoxic RNAse activity (Lee *et al.*, 1994; Murfett *et al.*, 1994). The RNAse gene is linked to multiple genes for the male S-determinant, the S-locus F-box (SLF) protein. SLFs are substrate-recognition subunits of the E3 ubiquitin ligase SCF complex. The multiple SLF proteins encoded in each S-locus allow pollen to recognize and detoxify nonself allelic variants of S-RNAse (Sijacic *et al.*, 2004; Kubo *et al.*, 2010; Williams *et al.*, 2014; Fujii *et al.*, 2016).

All of these SI loci are highly polymorphic, and genes for the two components are tightly linked. This assures that they are not separated by recombination, as this would lead to a breakdown of incompatibility. To generate new specificities in self-recognition

systems, both components have to evolve simultaneously while their function needs to be maintained to avoid self-fertilization or sterility. For the nonself recognition system of the *Solanaceae*, stepwise models for the evolution of new specificities have been proposed. First, a mutation modifies the S-RNase to dual specificity, then a second mutation alters one of the SLF genes to acquire the S-specificity for a new haplotype, and finally a third mutation leads to a loss of the original S-RNase specificity (Matton *et al.*, 1999). Alternatively, a two-step model was proposed in which the first mutation alters one of the SLF genes (leading to self-compatibility) which is then restored to SI by a second mutation in the S-RNAse gene (Uyenoyama *et al.*, 2001).

Basidiomycete fungi possess not only one, but two distinct nonself recognition systems to discriminate mating partners and to assure outbreeding (for recent reviews, see Nieuwenhuis *et al.*, 2013; Coelho *et al.*, 2017; Wallen & Perlin, 2018). Among the basidiomycetes, *Ustilago maydis*, a pathogen of maize, harbors one of the least complex systems. It consists of the biallelic *a*-mating type locus, encoding a pheromone/receptor system, and of the multiallelic *b* locus, harboring genes for the HD1- and HD2-type homedomain proteins bE and bW (Gillissen *et al.*, 1992; Raudaskoski & Kothe, 2010). The bE and bW proteins comprise a nonself recognition system controlling sexual as well as pathogenic development. Haploid cells with different *a* loci (*a1* or *a2*) sense the pheromone from a potential mating partner with a different *a* locus and fuse. However, subsequent development resumes only if the resulting dikaryotic cells harbor different alleles of the *b* locus (e.g. *b1* and *b2*). Being heterozygous for *b* allows the formation of two redundant bE/bW heterodimeric transcription factors (bE1/bW2 and bE2/bW1) that both function as key regulators for development. In nature, 18 different specificities of the *b* locus have been described based on pairwise crosses (Puhalla, 1970; for reviews, see Brefort *et al.*, 2009; Raudaskoski & Kothe, 2010; Vollmeister *et al.*, 2012).

The bE and bW proteins of *U. maydis* are unrelated, with the exception of a conserved homeodomain motif. The two genes are closely linked and are divergently transcribed from an intergenic region comprising *c.* 200 bp. Differences between *b* alleles cluster in a region of about 1000 bp covering the N-terminal coding regions of the *bE* and *bW* genes as well as the intergenic spacer (Fig. 1a). As a consequence, the N-terminal 120 amino acids of bE proteins as well as of bW proteins originating from different alleles are highly variable. The variable domains are followed by highly conserved C-terminal regions that include the homeodomain motifs (Kronstad & Leong, 1990; Schulz *et al.*, 1990; Gillissen *et al.*, 1992). Self–nonself discrimination occurs via the N-terminal variable domains and involves discrete regions that allow dimerization of bE and bW when the proteins originate from different alleles (Kämper *et al.*, 1995; Yee & Kronstad, 1998). Substitutions of single amino acids within the variable domain of bE2 (Kämper *et al.*, 1995) can lead to dimerization with bW2, that is, with a partner from the same allele, illustrating that the two proteins are then recognized as nonself. As a consequence, the resulting strains are solopathogenic, that is, they are able to infect the host plant without a mating partner. Such solopathogenic haploid strains can also be obtained by generation of hybrid *b* loci harboring *bE* and *bW* genes from different allelic origins (e.g. *bE1/bW2*) (Bölker *et al.*, 1995; Kämper *et al.*, 2006). In theory, a novel allele could evolve through compensatory mutations that prevent self-dimerization of bE and bW proteins in a solopathogenic strain, similar to the two-step model described for the Solanaceae (Uyenoyama *et al.*, 2001). However, solopathogenic haploid strains have never been isolated from nature. This may reflect the fact that they are at an evolutionary disadvantage, because they are unable to mate (Laity *et al.*, 1995) and produce fewer spores (F. Fukada and R. Kahmann, pers. obs.).

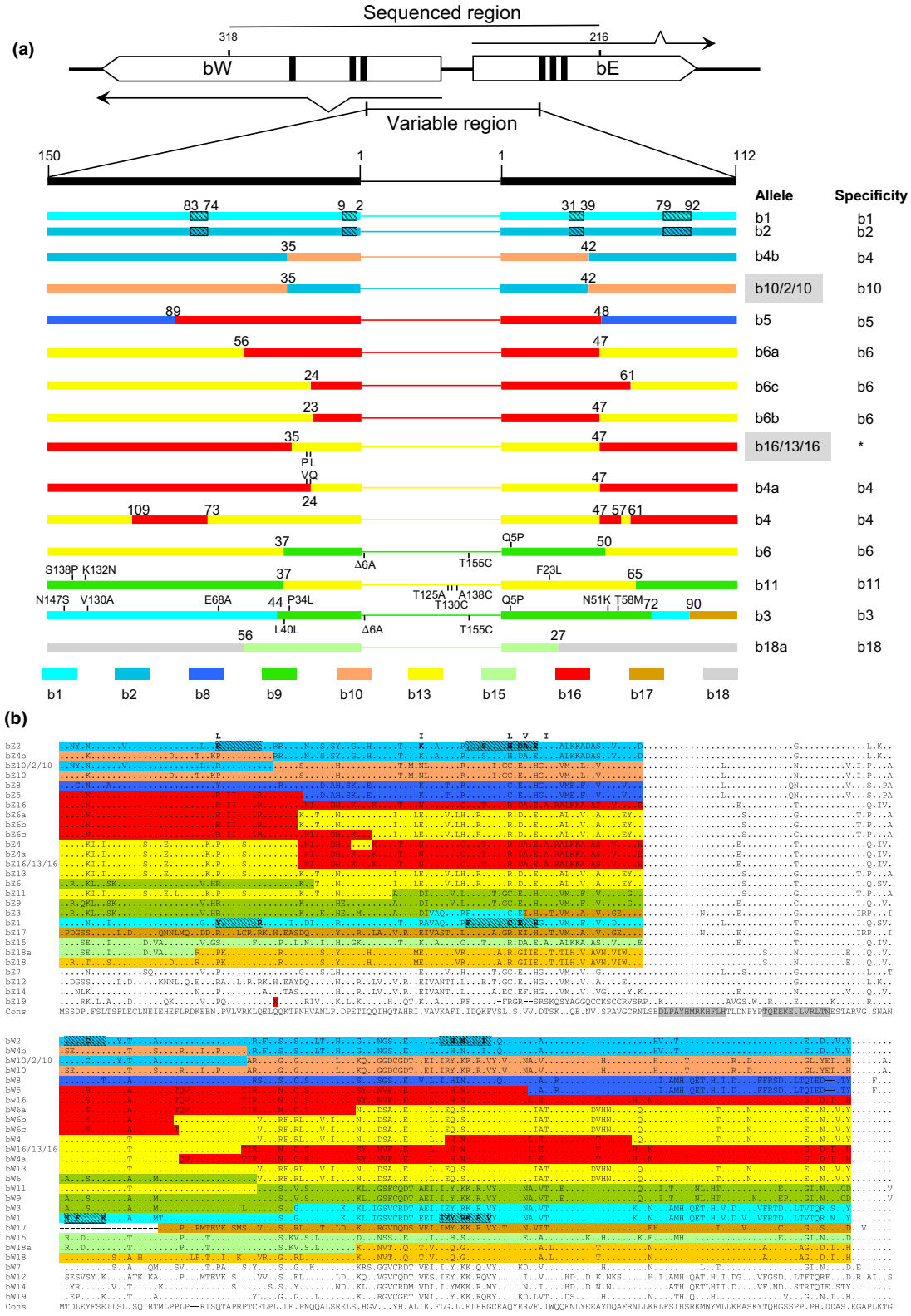
In this study we show that natural isolates of *U. maydis* contain *b* loci that have arisen by single recombination events between distinct *b* alleles. The recombination events observed lead to the simultaneous exchange of subdomains involved in dimerization in both bE and bW proteins, altering the specificity of both proteins in a single step. We also provide evidence for the possible existence of a two-step process leading to a new specificity that involves an intermediate with a nonfunctional bE protein.

Materials and Methods

Strains and growth conditions

Ustilago maydis strains are described in Supporting Information Table S1. Growth conditions, mating and pathogenicity assays, teliospore isolation as well as molecular manipulation of

Fig. 1 Schematic structure of natural *b* mating type genes of *Ustilago maydis* and hybrid genes generated *in vitro*. (a) In the upper panel, organization of the *b* locus with its two divergently transcribed genes, *bE* and *bW*, is depicted. Intron/exon structure and the position of the homeodomain helices (black bars) are given. The colored bars reflect the origin of sequences within the variable domain encoding the N-terminal parts of the bE (112 amino acids) and bW proteins (150 amino acids) and the intergenic spacer. Sequences shown in the same color are identical on both nucleotide and amino acid levels to the parental alleles (lower panel), unless deviations are given (within coding regions, numbers indicate amino acid substitutions at that position, and within the intergenic regions small numbers indicate nucleotide substitutions at that position). Large numbers indicate amino acid positions. Alleles and their specificity are given on the right. The position of the subdomains involved in interaction of b1 and b2 proteins (Yee & Kronstad, 1998) are given as shaded boxes within *b1* and *b2*. Alleles *b10/2/10* and *b16/13/16* (boxed) were generated *in vitro* and tested for functionality in a *U. maydis* strain deleted for the *b* locus. (*) *b16/13/16* strains formed tumors in combination with *b1, b4b* and *b10* strains, while crosses with *b2, b4, b4a, b6, b6a, b6b, b6c, b13, b14* or *b16* strains did not result in tumor induction (Supporting Information Table S2). (b) Alignment of bE and bW proteins. Only deviations from the consensus sequence (lower) are given; amino acids highlighted in grey within the consensus sequence depict the position of the homeodomain helices. In the variable regions, different colors highlight sequences derived from each other by recombination events using the same coloring scheme as in (a). The stop codon in *bE19* is marked with an X highlighted in red. Within the *b1* and *b2* derived sequences, the subdomains involved in interaction of *b1/b2*-derived bE/bW proteins are highlighted by shading. On top of the bE2 sequence, amino acids indicate single point mutations that facilitate interaction of the mutated bE2 protein with bW2. An enlarged version of (b) is available as Fig. S2.



U. maydis have been described (Holliday, 1974; Schulz *et al.*, 1990; Bölker *et al.*, 1992; Gillissen *et al.*, 1992; Brachmann *et al.*, 2001). To determine *b* specificity, *a1* and *a2* strains carrying the *b* allele to be tested were assayed for mating and pathogenicity with all strains harboring compatible *a* loci and specificities *b1* to *b19*. Strains with *b* specificities *b1 to b18* were obtained from the ATCC stock collection. The *b19* strain was isolated from spores collected in Bolivia (see Table S1).

RFLP analysis and functional cloning of *b*-alleles

For the restriction fragment length polymorphism (RFLP) analysis of *b* alleles, we used a modified protocol of the previously published 'barcode' method (Zambino *et al.*, 1997). Fragments (1.4 kb) spanning the variable regions of *bE* and *bW* were amplified by PCR from genomic DNA with primers bW-int-HIII (gagatcatgcactcacccagatag) and bE-HIII (acttcttcagaatatgagaccatc). PCR products were digested with *Hae*III and *Taq*I, respectively, and separated on 2% agarose gels. For sequence analysis, 2.2 kb fragments were amplified with *Pfu*-polymerase (Agilent, Santa Clara, CA, USA) using primers W2Not339 (gcacgcggccgcatgtaatcaaag) and E2Fse235 (gagtggccggccgaggttgtctg). PCR products were directly sequenced using a set of primers located in DNA regions conserved between alleles. Sequences have been submitted to NCBI GenBank, and accession numbers are given in Supporting Information Table S1. Sequences of *b1, b2, b3* and *b4* (Schulz *et al.*, 1990; Gillissen *et al.*, 1992) as well as partial sequences of *bE3, bE5, bE6* and *bE7* (Kronstad & Leong, 1990) have been published previously (Table S1).

For cloning purposes, standard molecular techniques were followed (Sambrook & Russell, 2001; Ausubel *et al.*, 2003), using the *E. coli* K-12 derivative DH5α (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA).

Primers W2Not339 (gcacgcggccgcatgtaatcaaag) and E2Fse235 (gagtggccggccgaggttgtctg), used for PCR amplification of 2.2 kb fragments of the *b* locus, generate silent mutations that introduce a *Fse*I site corresponding to amino acid positions 235/236 in *bE* and a *Not*I site corresponding to amino acid positions 338/339 in *bW*. Digestion of the PCR products with *Not*I and *Fse*I and cloning into respective sites of plasmid pUmbE/bW allowed reconstitution of functional *b* alleles.

For the construction of pUmbE/bW, a *Fse*I-linker was introduced into the *Stu*I site of pSL1180 (GE Healthcare/Amersham Biosciences, Piscataway, NJ, USA). Subsequently, a 1533 bp *Not*I–*Xho*I *bW2* fragment and a 1398 bp *Fse*I/*Eco*RI fragment of *bE2* were integrated into the respective sites of this plasmid. The *Eco*RI site is located 607 bp downstream of the open reading frame (ORF) of *bE2*, the *Xba*I site was introduced 573 bp 3′ of bW2. Recombinant PCR was used to introduce silent mutations to generate the *Fse*I and *Not*I sites at positions corresponding to amino acid position 234/235 in *bE2* and at position corresponding to amino acids 338/339 in *bW2*, respectively, to generate pSL-E2W2. Subsequently, the ip^r allele of the *U. maydis ip* gene (iron sulfur subunit of succinate dehydrogenase; Keon *et al.*, 1991) that confers carboxin resistance was isolated as a 1937 bp *Eco*RV/*Sma*I fragment from pCBX122 (Keon *et al.*, 1991), and inserted into the *Eco*RV site of pSL-E2W2. The resulting plasmid, pUmbE/bW, harbors the constant regions of bE2 and bW2 with unique *Not*I and *Fse*I sites.

Generation of artificial *b*-hybrid alleles

For construction of pTHE10/2/10, carrying a hybrid allele between *b10* and *b2*, recombinant PCR was used to integrate silent mutations generating a *Pst*I site at a position corresponding to amino acid position 38/39 in *bE2* and *bE10*, and an *Avr*II site corresponding to amino acid positions 32/33 in *bW2* and *bW10*, respectively. The sites were used to exchange the central portion of *b10* with the respective fragment of *b2*. The recombinant gene was subsequently amplified by PCR using primers W2Not339 and E2Fse235, and the 2.2 kb fragment was introduced into pUmbE/bW via *Not*I and *Fse*I to yield pTHE10/2/10.

Following the same strategy, a *Pst*I site was generated at a position corresponding to amino acid position 42 in *bE13* and *bE16*, and an *Avr*II site at a position corresponding to amino acids 32/33 in *bW13* and *bW16*, respectively. The sites were used to exchange the central portion of *b16* with the respective fragment of *b13*. Subsequently, the recombinant fragment was introduced into pUmbE/bW to yield pTHE16/13/16, as described for pTHE10/2/10.

Plasmids pUBJ#8 and pUBN#18 harbor 8 kb *Bam*HI fragments of genomic DNA from *b5* and *b8*, respectively, integrated in the *Bam*HI site of pHLN (Schulz *et al.*, 1990).

For construction of pSL-bE5-hyg and pSL-bE8-hyg, the *Hin*dIII site in plasmid pSL1180 (GE Healthcare/Amersham Biosciences) was removed by blunting with T4 polymerase and re-ligation to yield pSL1180ΔH. The *bE5* and *bE8* genes were inserted as 2.3 kb *Bgl*II/*Eco*RI (blunt) fragments from plasmids pUBJ#8 and pUBN#18, respectively, into the *Eco*RV/*Bgl*II digested pSL1180ΔH. Subsequently, a 3 kb *Pvu*II fragment with the hygromycin cassette from plasmid pHL1(Wang *et al.*, 1988) was ligated to *Bam*HI adaptors and inserted into the *Bam*HI site to yield pSL-bE5-hyg and pSL-bE8-hyg.

To construct pMIC-bW5 and pMIC-bW8, a PCR fragment encompassing the 3′ 900 bp of the *bW2* gene and 300 bp of the 3′UTR were inserted into *Not*I/*Bam*HI digested pBluescript KS (+) (Stratagene; Agilent Technologies, Santa Clara, CA, USA). The *Bam*HI site resides in the constant region of the *bW2* gene and is conserved between alleles. The resulting plasmid was digested with *Sma*I and *Bam*HI, and 1.95 kb *Pvu*I (blunt)/*Bam*HI fragments from pUbJ#8 and pUbN#18, harboring the *bW5* and *bW8* genes up to the *Bam*HI site, were inserted, to reconstitute complete *bW* genes. Subsequently, a 3 kb *Pvu*II fragment with the hygromycin cassette from pHL1 (Wang *et al.*, 1988) was inserted into the *Eco*RV sites to yield pMIC-bW5 and pMIC-bW8.

Generation of recombinant *U. maydis* strains

To avoid positioning effects on *bE/bW* gene expression, pUmbE/bW-derived plasmids were integrated in single copy into the *ip*-locus (Loubradou *et al.*, 2001) of AB2 (*a2Δb*), a strain lacking *bE* and *bW* genes (Romeis *et al.*, 2000). Strains AG1 and AG2 were constructed similar to JB1 and JB2 (Scherer *et al.*, 2006; Wahl *et al.*, 2010). Both are FB1 (*a1b1*) and FB2 (*a2b2*) derivatives, in which the *b*-loci were substituted with a 1.9-kb *Eco*RV/*Sma*I fragment harboring the ip^r carboxin resistance gene from plasmid pCBX122 (Keon *et al.*, 1991), removing the entire bE ORF and the bW ORF to amino acid 624. Integration of recombinant DNA in all strains outlined here were verified by Southern blot analysis.

Plasmids pMIC-bW5, pMIC-bW8, pSL-bE5-hyg and pSL-bE8-hyg were linearized and introduced into strains AG1

(*a1*Δ*b*) and AG2 (*a2*Δ*b*) by ectopic integration, and at least five independent transformants were scored for their mating specificity.

Results and Discussion

b-alleles with novel specificities for self–nonself recognition result from single recombination events of two *b*-alleles with distinct specificities

We have analyzed the *b* loci of 18 strains with different specificities deposited at ATCC (Puhalla, 1970) by RFLP analysis. To maintain consistency with published sequences of *b* alleles (Kronstad & Leong, 1990; Schulz *et al.*, 1990; Gillissen *et al.*, 1992), we changed the letter code of the *b* alleles deposited at ATCC to numbers (Table S1). The unique RFLP pattern (Fig. S1) obtained for each of these 18 *b* alleles was then used to screen > 200 haploid field isolates collected between 1992 and 2002 in Germany, France, Italy, Russia and Bolivia for new *b* alleles. Although seven strains had novel RFLP patterns, only one strain showed a novel specificity in crosses, and this was designated *b19* (Table S2). The total number of alleles identified is consistent with an estimation from field studies that calculated 18–20 alleles with a confidence interval of 95% (Zambino *et al.*, 1997). The remaining six identified strains with novel RFLP patterns had specificities of *b4* (designated *b4a* and *b4b*), *b6* (designated *b6a, b6b, b6c*) and *b18* (designated *b18a*) (Table S2). Subsequently, we sequenced the regions corresponding to the variable domains of bE and bW of strains from the ATCC collection and from the seven new field isolates (Table S1). From pairwise sequence alignments it became apparent that alleles with the specificities *b3*, *b4*, *b5*, *b6* and *b11* can be traced back to two or more parental alleles, suggesting that they have arisen through recombination events (Figs 1a,b, S2). To exemplify this, the sequence of *b4b* is a hybrid of *b2* and *b10* sequences. The central part of the *b4b* sequence, spanning the intergenic spacer and the regions encoding the N-terminal parts of the bE and bW variable domains, shows a 100% match to the sequence of *b10* on the nucleotide level. This central *b10* sequence is flanked on both sides by *b2* sequences. As a result, the N-terminal parts of the variable domains of both *bE4b* and *bW4b* originate from *b10*, while the remaining parts of the variable domain of the two proteins originate from *b2* (Figs 1a,b, S2, S3a). Similarly, the *b5* allele constitutes a hybrid between *b16* and *b8*, with the central part identical to *b16* (Figs 1a,b, S2, S3b). The hybrid genes show a 100% nucleotide sequence match to the two parental genes, in the coding regions as well as the intergenic promoter regions (Fig. S3a,b). This indicates that the recombination events that have led to the formation of *b4b* and *b5* must be of very recent origin.

The three strains with *b6* specificity alleles harbor *b* alleles composed of a central *b16* part and flanking *b13* sequences. Interestingly, the recombination sites, that is, the positions where the sequences shifts from *b16* to *b13*, are different in all three isolates. This demonstrates that these alleles with the same specificity have arisen by independent recombination events between the same parental alleles (Figs 1a,b, S2).

Recombination alters simultaneously both components of the self–nonself recognition system

Remarkably, for all alleles that appear to be generated by recombination events, both recombination breakpoints are within the regions encoding the variable domains of *bE* and *bW* (Figs 1a,b, S2). What are the expected consequences of recombination events that alter both variable domains simultaneously? Dimerization between bE and bW proteins is thought to be mediated by a linear array of subdomains within the variable regions that interfere with or facilitate dimerization. bE and bW proteins from the same allele are postulated to be unable to establish contacts via these subdomains, but to do so in combination with bE and bW proteins from other alleles (Kämper *et al.*, 1995; Yee & Kronstad, 1998, see model in Fig. 2a). By means of randomly generated chimeric bE1/bE2 and bW1/bW2 proteins, Yee & Kronstad (1998) have identified two of these subdomains for the *b1* and *b2* alleles that reside between amino acids 31–39 and 79–92 for bE1 and bE2, and between 2–9 and 74–83 for bW1 and bW2 (Yee & Kronstad, 1998; see Figs 1a,b, S2). Previously, we had identified single amino acid substitutions in bE2 that gained the ability to dimerize with bW2 (Kämper *et al.*, 1995), and consistently, several of these point mutations reside within or close to the domains identified by Yee and Kronstad (Fig. 1b).

As predicted by the model, the exchange of the N-terminal subdomain in bE2 with the N-terminal subdomain of bE1 resulted in a chimeric bE1/2 protein that interacted with both bW1 and bW2 (Yee & Kronstad, 1993), as outlined schematically in Fig. 2(b). In natural occurring *b*-alleles, the exchange of a single subdomain in either bE or bW would lead to self-compatibility, and only the simultaneous exchange of corresponding subdomains in the bE or bW partners would prevent self-compatibility (Fig. 2c). This was shown to be correct for bE1/bE2 chimeric proteins in combination with bW1/bW2 chimeras (Yee & Kronstad, 1998). According to the model in Fig. 2, exchanged subdomains in the hybrid bE and bW proteins are unable to interact with each other because they originate from proteins unable to form contacts (Fig. 2c). In theory, such recombinant alleles could have a novel specificity, that is they could encode products that cannot dimerize with the self-partner but they could maintain the ability to interact with proteins from all other alleles, including the parental bE and bW proteins (Fig. 2d).

The natural hybrid alleles identified in our work that originate from *b1* or *b2* sequences (*b4b* and *b3*) are consistent with the existence of the two subdomains for *b1* and *b2* identified by Yee & Kronstad (1998). In *b4b*, consisting of *b2* and *b10* (*b2-b10-b2*), the N-terminal subdomain is exchanged for both bE2 and bW2 proteins with sequences originating from *b10* (Figs 1a,b, S2). In *b3*, the N-terminal domains are encoded by *b1* and *b9* sequences (*b1-b9-b3*) (Figs 1a,b, S2), which could combine predicted specificity regions of *b1* and *b9* if they reside in similar positions as shown for *b1* and *b2* (Fig. 1b). This also holds true for most of the other natural alleles shown in Fig. 1(a); that is, recombination events would lead to the exchange of interaction domains if they reside in similar positions as shown for *b1* and

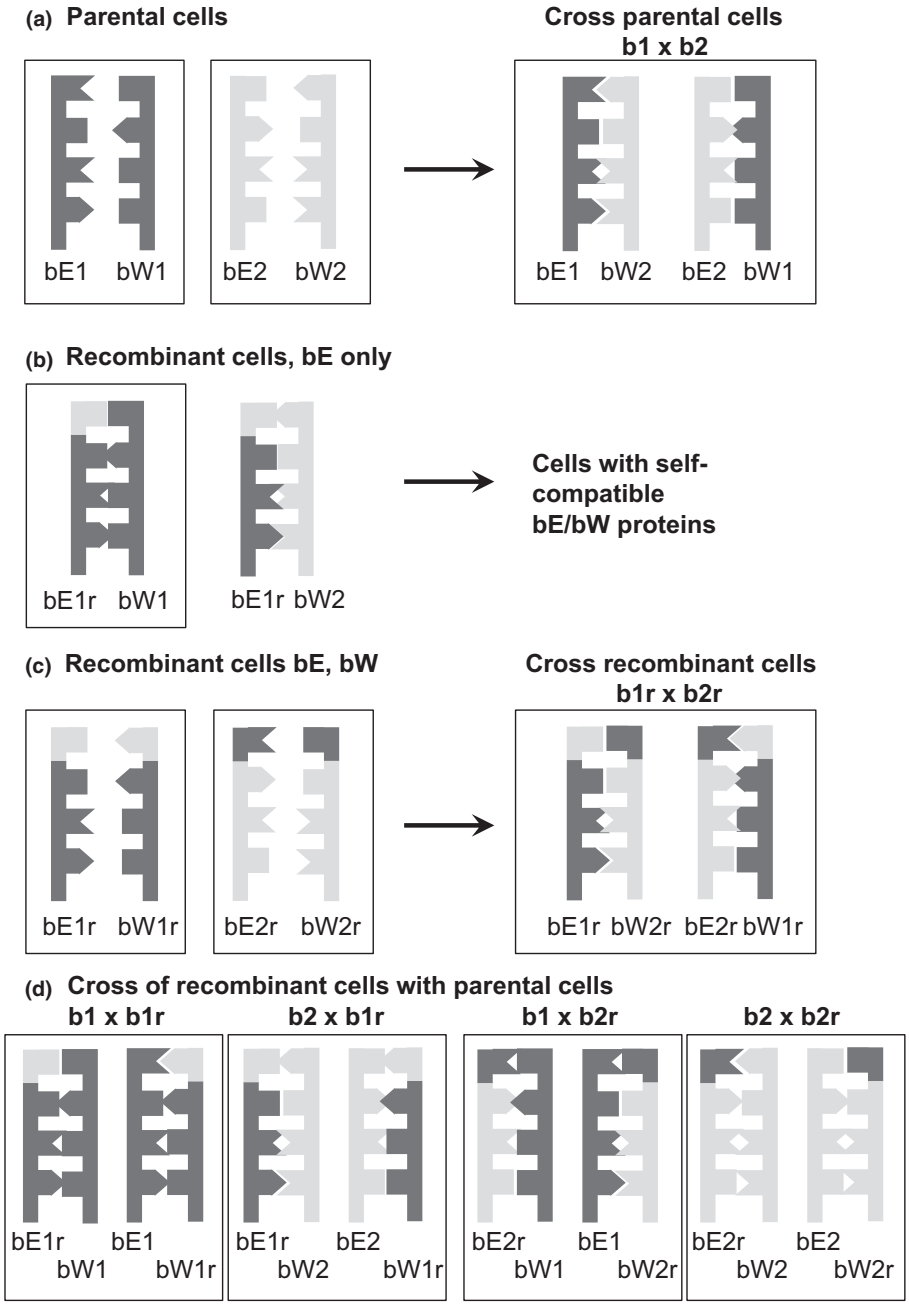


Fig. 2 Model for generation of new specificities of *Ustilago maydis b* mating type proteins by symmetric recombinational exchange of interaction domains. The model depicts schematically how arrays of subdomains within the variable domains of bE and bW lead to dimerization when properly juxtaposed. Dimerization requires a complementary surface in at least one position. (a) Subdomains of bE and bW proteins originating from the same allele (i.e. bE1 and bW1 or bE2 and bW2) do form contacts. By contrast, in combinations bE1/bW2 or bE2/bW2, a complementary array of subdomains facilitates interaction (cross *b1* × *b2*). (b) Recombination events that alter the subdomain of only bE (here bE1r) or bW lead to self-compatible bE/bW protein combinations. (c) Simultaneous exchange of subdomains achieved by recombination in the gene regions encoding the variable domains of *bE* and *bW* genes may lead to two distinct pairs of chimeric proteins (here designated bE1r, bW1r and bE2r, bW2r) that cannot dimerize. (d) Combination of recombinant alleles with each other (*b1r* × *b2r*) or of the recombinant alleles with the parental alleles (*b1r* × *b1*, *b1r* × *b2*, *b2r* × *b1*, *b2r* × *b2*) can lead to activity and sexual and pathogenic development.

b2. Given the multitude of different allelic combinations, we consider it likely that more than two domains are involved in dimerization. This assumption is supported by the finding that single point mutations outside of the b1/b2 interaction domains can also result in dimerization (Fig. 1b; Kämper *et al.*, 1995). Additionally, for bW5, the recombination breakpoint would not exchange a domain when similarly positioned as in bW1/bW2. For a final resolution of the domains one has to await determination of the structure of the bE/bW heterodimeric protein complex that, despite several attempts, has not been resolved so far.

The structural context of interaction domains contributes to nonself recognition

While the results presented so far are fully compatible with the simplistic model that assumes a linear array of contact points in the variable domains of bE and bW presented in Fig. 2, some of the findings outlined in the following can only be explained when considering additional contributions provided by the structural context of the interaction domains. Depending on the context created by recombination, potential contact points may not be accessible or may be incorrectly juxtaposed, and this could prevent interaction. Such an example is the *in vitro*-generated hybrid allele *b10/2/10* that represents a 100% mirror image of *b4b* (Figs 1a,b, S2). This hybrid retains the specificity of *b10* (Table S2), although the N-terminal subdomains of bE2 and bW2 have been transferred, suggesting that the structural context that is altered by the recombination event contributes to the function.

The artificial hybrid allele *b16/13/16* differs from *b4a* only in the recombination breakpoint in *bW*. As a consequence, the bW proteins of *b16/13/16* and *b4a* differ in only two amino acids (Figs 1, 2). A haploid strain carrying *b16/13/16* was not solopathogenic (Table S2). However, the *b16/13/16* strain was not able to mate with the 'parental' *b16* and *b13* strains, or with strains harboring the 'mirror images' *b6a*, *b6b* or *b6c*. Again, we have to assume that the position of the recombination breakpoint impacts the function of the interaction domain by altering the structural context. We can exclude the possibility that the *b16/13/16* allele has lost its function entirely, because combinations with *b1, b4b* and *b10* strains were productive and led to tumor formation (Table S2). The natural *bE4a* allele and the synthetic *b16/13/16* allele encode identical bE proteins, but still *b16/13/16* differs from *b4a* in that it does not function with *b16, b13, b6a, b6b* and *b6c*. Thus, we have to conclude that in the combination of *b4a* with *b16, b13, b6a, b6b* and *b6c*, the interaction occurs exclusively via the bW4a protein, and that in *b16/13/16* the two amino acid exchanges in bW16/13/16 prevent these interactions. This is a first indication that both of the possible bE and bW combinations do not have to be functional in all cases (see later).

The artificial *b16/13/16* allele illustrates that artificial exchange of subdomains can lead to strains with a severely restricted mating spectrum. In field populations of *U. maydis*, such events would confer a selective disadvantage, and this may be the reason why such strains have never been isolated from nature. All natural alleles have a specificity that allows function with all other 18 alleles known.

Also for naturally occurring alleles we have to assume that the recombination breakpoint *per se* impacts dimerization. This must be the case for *b4a (b16-b13-b16)* and *b4b* (*b2-b10-b2*) which have the same specificity, although they have clearly been generated by recombination between different parental alleles (Fig. 1a). The subdomains that normally facilitate the interaction within the C-terminal parts of the bE and bW proteins of *b16* and *b2* or between the N-terminal parts of *b13* and *b10* must have been altered by the recombination event in a way that precludes interaction between *b4a* and *b4b*. By assuming a loss of contact points as a result of the recombination event, it can also be explained that alleles with *b6* specificity have been generated in *b6* (*b13-b9-b13*) as well as in alleles *b6a, b6b* and *b6c* (all *b13-b16-b13* with different recombination breakpoints) (Fig. 1a).

Alleles with 'mirror images' are probably the result of independent recombination events

In principle, one reciprocal recombination event by a double crossover could lead to two alleles with novel specificity. Indeed, the alleles *b4a* and *b6b* are almost mirror images (Figs 1a,b, S2), supporting the idea that two new specificities can be generated from the same two parental alleles. Consistent with the model in Fig. 2, b4a and b6b are intercompatible (Table S2). However, as the two recombination breakpoints in *b4a* and *b6b* are > 350 bp apart, it is unlikely that the two alleles have been generated by reciprocal recombination events because such events should be prevented by crossover interference. For *Saccharomyces cerevisiae* it has been demonstrated that the heteroduplex region leading to gene conversion upon double-strand breaks was in the range 7.2–32.3 kb in 20 independent experiments (Hum & Jinks-Robertson, 2017). Although comparable data are not available for *U. maydis*, we consider it more likely that the alleles have arisen independently by gene conversion events. The mirror images of two *b*-alleles would then have to be explained by restraints for 'breaks' between interaction domains and selection for function of the individual interaction domains.

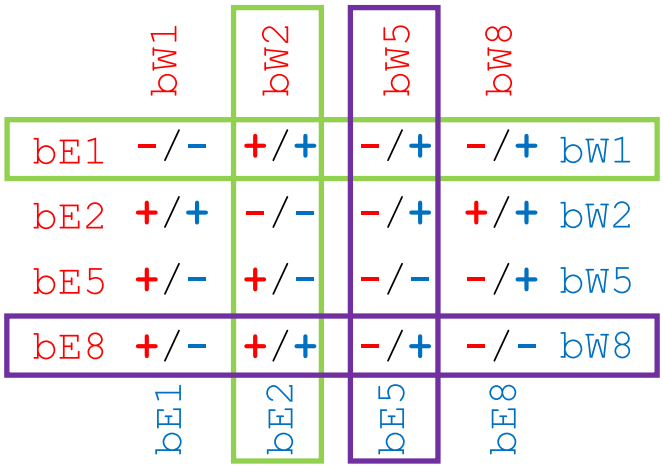


Fig. 3 Interaction between pairs of *Ustilago maydis* bE and bW proteins from different alleles is not always redundant. Strains expressing a single bE or bW protein from *b1*, *b2*, *b5* or *b8* alleles were crossed with strains harboring a compatible *a* locus and *b1*, *b2*, *b5* or *b8* loci. Function of bE and bW was assessed by the formation of the filamentous dikaryon on PD-plates containing charcoal (Fuz-reaction). The colored (+) and (−) indicate heterodimer formation or no heterodimer formation between proteins given in the same color. In the *b1* × *b2* combination (highlighted by green boxes), bE1 interacts with bW2 (indicated by a red +) and bW1 interacts with bE2 (indicated by a blue +). For *b8* × *b5* (highlighted by purple boxes), only bE5 and bW8 interact (blue +), while the reciprocal combination bE8 × bW5 does not (red −). The original data for this compilation are listed in Supporting Information Table S3.

In various compatible interactions, only one of the two potential bE–bW protein combinations is functional

After fusion of two compatible *U. maydis* strains, two different bE/bW heterodimers can be produced, and for bE1/bW2 and bE2/bW1 both were shown to be active (Gillissen *et al.*, 1992). To assess whether formation of two redundant heterodimeric complexes is a general principle, we monitored the activity of single bE or bW proteins. To this end we tested engineered strains harboring either single *bE1, bE2, bE5* and *bE8* genes or single *bW1, bW2, bW5* and bW8 genes for mating with strains with a compatible *a*-mating type harboring *b1, b2, b5* or *b8* loci (Fig. 3; Table S3). bW1 and bW2 showed activity with all nonself-allelic bE combinations tested. However, although bE1 functioned with bW2, it showed no activity with bW1, bW5 or bW8. Thus, the productive mating of *b1* strains with *b5* or *b8* strains relies on the bW1/bE5 and bW1/bE8 heterodimers, respectively, while heterodimers of bE1 and bW5 or bE1 and bW8 do not form. The same holds true for other combinations of proteins: bW5 functions with bE8, but does not function with bE1, bE2 or bE5. Mating of *b5* with *b1* or *b2* strains depends on bW1/bE5 and bW2/bE5 heterodimers, respectively. Successful mating of *b5* and *b8* strains relies on bW5/bE8, while the reciprocal combination of bE5 and bW8 proteins is nonfunctional.

These results show that not all bE and bW proteins are active in all combinations, but it is assured that at least one active pair is formed. It is likely that the redundant function of the two protein pairs involved in nonself recognition allows a higher degree of freedom for the evolution of new alleles.

New b-alleles can arise from a two-step mechanism that avoids self-compatible intermediates

Of the 25 *b* allele sequences analyzed, 14 cannot be traced back to recent recombination events. One explanation could be that we only sequenced a limited number of natural *b* alleles based on differences in RFLP pattern. This might have prevented the detection of alleles having undergone multiple recombination events. Alternatively, new alleles might also have arisen from a two-step mechanism. In laboratory experiments it has been demonstrated that single amino acid substitutions in the variable domain of bE2 can lead to dimerization with bW2. As a result, the respective haploid strain became solopathogenic (Kämper *et al.*, 1995), that is it did not need a mating partner to infect plants. To convert such a self-activating allele to a novel allele, compensatory mutations would have to occur in bW2 to prevent dimerization with the mutant bE2 protein. However, such a sequential mechanism is unlikely to operate in nature because the obligatory solopathogenic intermediate strains are unable to mate (Laity *et al.*, 1995), and would be excluded from the mating population. For SI systems in plants, the evolution through self-compatible intermediates or intermediates with dual specificity is also controversially discussed because of the disadvantages that such intermediates face in populations (Charlesworth, 2000; Uyenoyama & Newbigin, 2000). Only in a population with strong inbreeding depression and low rates of self-pollination, mutations that lead to self-compatible pollen that is not rejected by any other haplotype may be stabilized along with self-incompatible haplotypes (Uyenoyama *et al.*, 2001).

Interestingly, the sequence of *b19* suggests an alternative route to a new allele that bypasses the solopathogenic stage. In *b19*, the *bE* gene is nonfunctional as a result of a stop codon at amino acid position 42 (Figs 1b, S4; Table S4). The functional bW19 protein in this strain is sufficient to interact with all bE proteins with different specificity and thus allows mating and pathogenic development (Tables S2, S4). However, because of the nonfunctional *bE19*, *bW19* would be liberated from the restraints of coevolution and could accumulate mutations without creating a situation where a constitutive allele could arise through mutations facilitating interaction with bE19. At the same time, the inactive *bE19* gene could accumulate mutations. Reversion of the stop codon (or recombination) could eventually return bE19 to function, creating a new allele without going through a constitutively active intermediate.

It is generally believed that the accumulation of point mutations is the major driving force for allelic diversity in self–nonself recognition systems (May & Matzke, 1995; Bergstrom *et al.*, 1998; Awadalla & Charlesworth, 1999; Schierup *et al.*, 2001; Wang *et al.*, 2001). Only in few systems, recombination has been proposed to contribute. One such example is the *A* mating type locus of *Coprinopsis cinereus*, which comprises several homeobox gene pairs arranged like the *U. maydis bE* and *bW* genes but functioning independently of each other. In this system new specificities have been shown to arise through recombination in intergenic regions that separate the individual homeobox gene pairs (May & Matzke, 1995; Pardo *et al.*, 1996). It is important to stress that in such recombination events the specificity of the individual gene pairs remains unaltered, which makes it mechanistically distinct from what we have described here for the *b* locus of *U. maydis*.

The *b* mating type alleles in *U. maydis* represent a convincing example where new specificities for nonself recognition are generated by recombination events that alter the two components of the system simultaneously. The collection of natural *b* mating type alleles not only allows the visualization of the evolutionary events that have led to specificity diversification, but also supports the mechanistic model proposed for nonself recognition of bE and bW proteins (Kämper *et al.*, 1995; Yee & Kronstad, 1998). It will be revealing to support this genetic model and the consequences of the observed recombination events with structural data.

Acknowledgements

We would like to thank Angelika Waldschmidt, Volker Vincon and Petra Happel for excellent technical assistance, and all students of the molecular mycology classes who with great enthusiasm searched for new *b* specificities. This work was supported by generous funds from the Max Planck Society.

Author contributions

JK and RK designed and supervised the project. JK conducted the experiments, and JK and MF analyzed data. JK and RK wrote

the manuscript. All authors edited and approved the final version of this report.

ORCID

Michael W. Friedrich https://orcid.org/0000-0002-8055-3232
Regine Kahmann https://orcid.org/0000-0001-7779-7837
Jörg Kämper https://orcid.org/0000-0002-8161-5431

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl E. 2003. *Current protocols in molecular biology*. New York, NY, USA: John Wiley & Sons.

Awadalla P, Charlesworth D. 1999. Recombination and selection at *Brassica* self-incompatibility loci. *Genetics* **152**: 413–425.

Bergstrom TF, Josefsson A, Erlich HA, Gyllensten U. 1998. Recent origin of HLA-DRB1 alleles and implications for human evolution. *Nature Genetics* **18**: 237–242.

Bölker M, Genin S, Lehmler C, Kahmann R. 1995. Genetic regulation of mating, and dimorphism in *Ustilago maydis*. *Canadian Journal of Botany* **73**: 320–325.

Bölker M, Urban M, Kahmann R. 1992. The *a* mating type locus of *U. maydis* specifies cell signaling components. *Cell* **68**: 441–450.

Brachmann A, Weinzierl G, Kämper J, Kahmann R. 2001. Identification of genes in the bW/bE regulatory cascade in *Ustilago maydis*. *Molecular Microbiology* **42**: 1047–1063.

Brefort T, Doehlemann G, Mendoza-Mendoza A, Reissmann S, Djamei A, Kahmann R. 2009. *Ustilago maydis* as a pathogen. *Annual Review of Phytopathology* **47**: 423–445.

Charlesworth D. 2000. How can two-gene models of self-incompatibility generate new specificities? *Plant Cell* **12**: 309–310.

Coelho MA, Bakkeren G, Sun S, Hood ME, Giraud T. 2017. Fungal sex: the Basidiomycota. In: Heitman J, Howlett BJ, Crous PW, Stukenbrock EH, James TY, Gow GAR, eds. *Microbiology spectrum*. Sterling, VA, USA: ASM Press, 147–176.

Fujii S, Kubo K, Takayama S. 2016. Non-self- and self-recognition models in plant self-incompatibility. *Nature Plants* **2**: 16130.

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

Hinata K, Watanabe M, Yamakawa S, Satta Y, Isogai A. 1995. Evolutionary aspects of the S-related genes of the *Brassica* self-incompatibility system: synonymous and nonsynonymous base substitutions. *Genetics* **140**: 1099–1104.

Holliday R. 1974. *Ustilago maydis*. In: King RC, ed. *Handbook of genetics*. New York, NY, USA: Plenum Press, 575–595.

Hum YF, Jinks-Robertson S. 2017. Mitotic gene conversion tracts associated with repair of a defined double-strand break in *Saccharomyces cerevisiae*. *Genetics* **207**: 115–128.

Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O *et al.* 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* **444**: 97–101.

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

Keon JP, White GA, Hargreaves JA. 1991. Isolation, characterization and sequence of a gene conferring resistance to the systemic fungicide carboxin from the maize smut pathogen, *Ustilago maydis*. *Current Genetics* **19**: 475–481.

Kronstad JW, Leong SA. 1990. The *b* mating-type locus of *Ustilago maydis* contains variable and constant regions. *Genes & Development* **4**: 1384–1395.

Kubo K, Entani T, Takara A, Wang N, Fields AM, Hua Z, Toyoda M, Kawashima S, Ando T, Isogai A *et al.* 2010. Collaborative non-self recognition system in S-RNase-based self-incompatibility. *Science* **330**: 796–799.

Laity C, Giasson L, Campbell R, Kronstad J. 1995. Heterozygosity at the *b* mating-type locus attenuates fusion in *Ustilago maydis*. *Current Genetics* **27**: 451–459.

Lee HS, Huang S, Kao T. 1994. S proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature* **367**: 560–563.

Loubradou G, Brachmann A, Feldbrügge M, Kahmann R. 2001. A homologue of the transcriptional repressor Ssn6p antagonizes cAMP signalling in *Ustilago maydis*. *Molecular Microbiology* **40**: 719–730.

Matton DP, Luu DT, Xike Q, Laublin G, O'Brien M, Maes O, Morse D, Cappadocia M. 1999. Production of an S RNase with dual specificity suggests a novel hypothesis for the generation of new S alleles. *Plant Cell* **11**: 2087–2097.

May G, Matzke E. 1995. Recombination and variation at the *A* mating type of *Coprinus cinereus*. *Molecular Biology and Evolution* **12**: 794–802.

Murfett J, Atherton TL, Mou B, Gasser CS, McClure BA. 1994. S-RNase expressed in transgenic *Nicotiana* causes S-allele-specific pollen rejection. *Nature* **367**: 563–566.

Nieuwenhuis BP, Billiard S, Vuilleumier S, Petit E, Hood ME, Giraud T. 2013. Evolution of uni- and bifactorial sexual compatibility systems in fungi. *Heredity* **111**: 445–455.

Pardo EH, O'Shea SF, Casselton LA. 1996. Multiple versions of the A mating type locus of *Coprinus cinereus* are generated by three paralogous pairs of multiallelic homeobox genes. *Genetics* **144**: 87–94.

Poulter NS, Wheeler MJ, Bosch M, Franklin-Tong VE. 2010. Self-incompatibility in *Papaver*: identification of the pollen S-determinant PrpS. *Biochemical Society Transactions* **38**: 588–592.

Puhalla JE. 1970. Genetic studies on the *b* incompatibility locus of *Ustilago maydis*. *Genetics Research* **16**: 229–232.

Raudaskoski M, Kothe E. 2010. Basidiomycete mating type genes and pheromone signaling. *Eukaryotic Cell* **9**: 847–859.

Richman A. 2000. Evolution of balanced genetic polymorphism. *Molecular Ecology* **9**: 1953–1963.

Romeis T, Brachmann A, Kahmann R, Kämper J. 2000. Identification of a target gene for the bE/bW homeodomain protein complex in *Ustilago maydis*. *Molecular Microbiology* **37**: 54–66.

Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*. Cold Spring Harbour, NY, USA: Cold Spring Harbour Laboratory Press.

Scherer M, Heimel K, Starke V, Kämper J. 2006. The Clp1 protein is required for clamp formation and pathogenic development of *Ustilago maydis*. *Plant Cell* **18**: 2388–2401.

Schierup MH, Mable BK, Awadalla P, Charlesworth D. 2001. Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*. *Genetics* **158**: 387–399.

Schopfer CR, Nasrallah ME, Nasrallah JB. 1999. The male determinant of self-incompatibility in *Brassica*. *Science* **286**: 1697–1700.

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

Shiba H, Takayama S, Iwano M, Shimosato H, Funato M, Nakagawa T, Che FS, Suzuki G, Watanabe M, Hinata K *et al.* 2001. A pollen coat protein, SP11/SCR, determines the pollen S-specificity in the self-incompatibility of Brassica species. *Plant Physiology* **125**: 2095–2103.

Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao TH. 2004. Identification of the pollen determinant of S-RNase-mediated self-incompatibility. *Nature* **429**: 302–305.

Takahata N, Nei M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**: 967–978.

Takahata N, Satta Y, Klein J. 1992. Polymorphism and balancing selection at major histocompatibility complex loci. *Genetics* **130**: 925–938.

Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K. 2000. The S receptor kinase determines self-incompatibility in *Brassica stigma*. *Nature* **403**: 913–916.

Thomas SG, Franklin-Tong VE. 2004. Self-incompatibility triggers programmed cell death in *Papaver* pollen. *Nature* **429**: 305–309.

Uyenoyama MK, Newbigin E. 2000. Evolutionary dynamics of dual-specificity self-incompatibility alleles. *Plant Cell* **12**: 310–312.

Uyenoyama MK, Zhang Y, Newbigin E. 2001. On the origin of self-incompatibility haplotypes: transition through self-compatible intermediates. *Genetics* **157**: 1805–1817.

Vollmeister E, Schipper K, Baumann S, Haag C, Pohlmann T, Stock J, Feldbrügge M. 2012. Fungal development of the plant pathogen *Ustilago maydis*. *FEMS Microbiology Reviews* **36**: 59–77.

Wahl R, Zahiri A, Kämper J. 2010. The *Ustilago maydis b* mating type locus controls hyphal proliferation and expression of secreted virulence factors *in planta*. *Molecular Microbiology* **75**: 208–220.

Wallen RM, Perlin MH. 2018. An overview of the function and maintenance of sexual reproduction in dikaryotic fungi. *Frontiers in Microbiology* **9**: 503.

Wang J, Holden DW, Leong SA. 1988. Gene transfer system for the phytopathogenic fungus *Ustilago maydis*. *Proceedings of the National Academy of Sciences, USA* **85**: 865–869.

Wang X, Hughes AL, Tsukamoto T, Ando T, Kao T. 2001. Evidence that intragenic recombination contributes to allelic diversity of the S-RNase gene at the self-incompatibility (S) locus in *Petunia inflata*. *Plant Physiology* **125**: 1012–1022.

Wheeler MJ, de Graaf BH, Hadjiosif N, Perry RM, Poulter NS, Osman K, Vatovec S, Harper A, Franklin FC, Franklin-Tong VE. 2009. Identification of the pollen self-incompatibility determinant in *Papaver rhoeas*. *Nature* **459**: 992–995.

Wheeler MJ, Vatovec S, Franklin-Tong VE. 2010. The pollen S-determinant in *Papaver*: comparisons with known plant receptors and protein ligand partners. *Journal of Experimental Botany* **61**: 2015–2025.

Williams JS, Natale CA, Wang N, Li S, Brubaker TR, Sun P, Kao TH. 2014. Four previously identified *Petunia inflata* S-locus F-box genes are involved in pollen specificity in self-incompatibility. *Molecular Plant* **7**: 567–569.

Wu J, Saupe SJ, Glass NL. 1998. Evidence for balancing selection operating at the het-c heterokaryon incompatibility locus in a group of filamentous fungi. *Proceedings of the National Academy of Sciences, USA* **95**: 12398–12403.

Yee AR, Kronstad JW. 1993. Construction of chimeric alleles with altered specificity at the b incompatibility locus of *Ustilago maydis*. *Proceedings of the National Academy of Sciences, USA* **90**: 664–668.

Yee AR, Kronstad JW. 1998. Dual sets of chimeric alleles identify specificity sequences for the bE and bW mating and pathogenicity genes of *Ustilago maydis*. *Molecular and Cellular Biology* **18**: 221–232.

Zambino P, Groth JV, Lukens L, Garton JR, May G. 1997. Variation at the *b* mating type locus of *Ustilago maydis*. *Phytopathology* **87**: 1233–1239.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 RFLP analysis of *U. maydis b*-alleles.

Fig. S2 Alignment of bE and bW proteins of natural *b* mating type genes and hybrid genes generated *in vitro* (enhanced version of Fig. 1b).

Fig. S3 Alignment of protein and DNA sequences of the alleles *b10, b4b* and *b2*, and of *b16, b5* and *b8*.

Fig. S4 The open reading frame of *bE19* harbors a nonsense mutation.

Table S1 Sequenced *b*-alleles and *Ustilago maydis* strains used in the study.

Table S2 Disease symptoms in maize plants after infection with mixtures of indicated strains.

Table S3 Genetic interactions of bE and bW proteins from different alleles are not redundant.

Table S4 Disease symptoms in maize plants after infection of *b19* strain with strains harboring single *bE* or *bW* genes.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.