H_2O, 200 μl of 1 M tris-HCl (pH 8.0), and 9 ml of glycerol] was then added to the slide. A coverslip was placed on the slide and sealed with nail varnish. Slides were stored at 4°C until viewed under a fluorescence microscope.
15. The probe was made from the 18*S* ribosomal DNA of *B. phoenicis*. General arthropod primers from C. Simon *et al*. [*Ann. Entomol. Soc. Am*. **87**, 651 (1994)] were used to amplify the 18*S* rDNA from a mixture of DNA from the five isofemale lines of *B. phoenicis*. Once amplified, the product was cloned into a pGEM-T vector (Promega) and sequenced to determine whether it was really mite 18*S* rDNA. The insert was then cut out of the vector using restriction enzymes, cleaned, and labeled with biotin by nick-translation using a BioNick Labeling System (Gibco-BRL, Life Technologies).
16. K. Sahara, F. Marec, W. Traut, *Chromosome Res*. **7**, 449 (1999).
17. Chromosome preparations were digested with RNase A [100 μg/ml in 2× saline sodium citrate (SSC) buffer; 1× SSC is 0.15 M sodium citrate and 0.015 M NaCl (pH 7.0)] for 1 hour at 37°C and washed two times for 15 min in 2× SSC. After FISH, chromosomes were observed in a Zeiss Axioplan 2 microscope. Black and white images were recorded with a cooled charge-coupled device camera MEGA F-View II (Soft Imaging System, Münster, Germany) and then pseudocoloured [blue for 4′,6′-diamidino-2-phenylindole dihydrochloride (DAPI) and red for Cy3] and superimposed with the aid of an image-processing program (Adobe Photoshop Version 5.0).
18. Two hundred immature *B. phoenicis* females were collected from coffee and citrus plantations in Sao Paulo and Minas Gerais, Brazil (two coffee sites and five citrus sites). These were reared in the lab as isofemale lines and screened for genetic variation using the amplified fragment length polymorphism method (AFLP) as mentioned in (*6*). Forty-five genotypically distinct lines were then screened for seven polymorphic microsatellite loci, and two of these lines were screened for two additional loci. Microsatellite loci were isolated using a method developed by A. R. Weeks and J. A. J. Breeuwer. For information regarding this technique, please e-mail the corresponding author.
19. For identifying the bacteria, we used general eubacterial primers to amplify the 16*S* rDNA (fD2 and rP2) from W. G. Weisburg *et al*. [*J. Bacteriol*. **173**, 697 (1991)]. PCR products from each of the five isofemale lines were then cloned into a pGEM-T vector. Owing to the nature of the primers, it is possible that nonendosymbiotic bacteria could be amplified and then wrongly classified as the endosymbiotic bacteria we found through cytology. To control for this, we extracted vectors from five recombinant colonies from each of the five isofemale lines and sequenced all inserts. All 25 inserts contained identical copies of the same sequence (GenBank accession number AF350221).
20. T. J. Kurtti, U. G. Munderloh, T. G. Andreadis, L. A. Magnarelli, *J. Invertebr. Pathol*. **67**, 318 (1996).
21. Two hundred adult female *B. phoenicis* were collected from an egg wave (500 female mites from one of the five isofemale lines were allowed to lay eggs for 2 days, then removed) 2 days after their final molt. One hundred adults were then treated with tetracycline. Leaf discs (1 cm in diameter) of the common bean, *Phaseolus vulgaris*, were floated in a solution of 0.2% tetracycline hydrochloride and distilled water. Five adult mites per leaf disc were left to feed for 72 hours at constant temperature (30°C). At the same time, 60 adult mites were set up in exactly the same way, except that no tetracycline was added to the distilled water, to act as controls. After 72 hours, the tetracycline-treated mites as well as the control mites were transferred to a leaf of *P. vulgaris*, and left to recover for 24 hours at 30°C. Surviving mites from each treatment were then set up as follows: five mites per leaf disc (1 cm in diameter) floated on cotton wool soaked in water at 30°C. Mites were left to lay eggs for 10 days, at which time adults were taken from the leaf disc, leaving the leaf discs and developing eggs at 30°C. Upon reaching the adult stage (2 weeks), the sex of the progeny was scored.
22. R. Stouthamer, J. A. J. Breeuwer, G. D. D. Hurst, *Annu. Rev. Microbiol*. **53**, 71 (1999).
23. Primers were designed based on the sequence in (*19*) and other known bacterial 16*S* rDNA sequences to specifically amplify only this bacterium. The PCR primers were: 5′-CCCATACAATCTTGCTAGCG (forward) and 5′-CTCAAGGTGTCAAAATACATTC (reverse). These primers amplify a fragment 832 bp in length.
24. F. Rousset, D. Bouchon, B. Pintureau, P. Juchault, M. Solignac, *Proc. R. Soc. London Ser. B* **250**, 91 (1992).
25. G. Martin, P. Juchault, O. Sorokine, A. van Dorsselaer, *Gen. Comp. Endocrinol*. **80**, 349 (1990).
26. T. Rigaud, P. Juchault, J. P. Mocquard, *BioEssays* **19**, 409 (1997).
27. S. R. Bordenstein, F. P. O'Hara, J. H. Werren, *Nature* **409**, 707 (2001).
28. E. Zchori-Fein, Y. Gottlieb, M. S. Hunter, personal communication.
29. M. J. Perrot-Minnot, R. A. Norton, *Can. Entomol*. **129**, 691 (1997).
30. T. Wenseleers, *et al.*, *Proc. R. Soc. London Ser. B* **264**, 1447 (1998).
31. T. Wenseleers, J. Billen, *J. Evol. Biol*. **13**, 277 (2000).
32. D. A. Grasso, T. Wenseleers, A. Mori, F. Le Moli, J. Billen, *Ethol. Ecol. Evol*. **12**, 309 (2000).
33. R. L. Blackman, J. M. Spence, *Chromosome Res*. **4**, 314 (1996).
34. E. Suomalainen, A. Saura, J. Lokki, *Cytology and Evolution in Parthenogenesis* (CRC Press, Boca Raton, FL, 1987).
35. Supported by a grant from the Netherlands Foundation for the Advancement of Tropical Research (WOTRO). F.M. was supported by a grant (206/00/0750) from the Grant Agency of the Czech Republic (Prague). We thank M. Sabelis for initiation of this project and insightful discussions and comments; C. Omoto for help in field collections; A. Hoffmann, S. Menken, C. Sgro, T. van Opijnen, and A. Janssen for critical comments; and I. Sauman for methodical suggestions.

5 March 2001; accepted 9 May 2001

Gene Families from the *Arabidopsis thaliana* Pollen Coat Proteome

Jacob A. Mayfield,[1] Aretha Fiebig,[2] Sarah E. Johnstone,[1] Daphne Preuss[1,3]*

The pollen extracellular matrix contains proteins mediating species specificity and components needed for efficient pollination. We identified all proteins >10 kilodaltons in the *Arabidopsis* pollen coating and showed that most of the corresponding genes reside in two genomic clusters. One cluster encodes six lipases, whereas the other contains six lipid-binding oleosin genes, including *GRP17*, a gene that promotes efficient pollination. Individual oleosins exhibit extensive divergence between ecotypes, but the entire cluster remains intact. Analysis of the syntenic region in *Brassica oleracea* revealed even greater divergence, but a similar clustering of the genes. Such allelic flexibility may promote speciation in plants.

Because self-recognition systems must adapt to the evolution of target molecules, they include some of the most rapidly changing proteins known. Unusual levels of genetic divergence are seen in mate recognition in plants, algae, abalone, and primates (*1–4*); immune responses in animals (*5*); and pathogen defense in plants (*6*). Rapid divergence of molecules controlling mate recognition in flowering plants is essential, considering diversification of most angiosperms occurred only 90 to 130 million years ago. Here, we describe the protein components of the *A. thaliana* pollen coat and show they display remarkable variability.

The complex extracellular pollen coating of many flowering plants uses proteins and lipids to interact selectively with receptive female stigma cells (*7–10*). This coating facilitates communication in plants with dry stigmas, providing a function similar to the lipid-rich exudate on the surface of wet stigmas (*8*). Identification and characterization of the most abundant *Arabidopsis* pollen coat

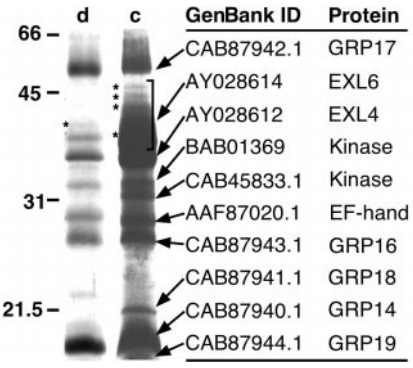


Fig. 1. Purified pollen coat proteins and their identity. Coomassie stained SDS-PAGE with corresponding GenBank ID. Asterisk, EXL6 protein; d, dilute protein sample; c, concentrated protein sample.

[1]Department of Molecular Genetics and Cell Biology, [2]Department of Biochemistry and Molecular Biology, [3]Howard Hughes Medical Institute, The University of Chicago, Chicago, IL 60637, USA.

*To whom correspondence should be addressed. E-mail: dpreuss@midway.uchicago.edu

protein, the lipid-binding oleosin GRP17 (glycine-rich protein), demonstrated its role in initiating pollination (*9*). To discern the role of other *Arabidopsis* pollen coat proteins, we purified and sequenced peptides from each protein >10 kD in size and used the completed genome sequence to progress to the corresponding genes (*11*).

Peptide sequencing revealed the identity of all the detectable coat proteins with mobilities >10 kD; protein identity was verified using mass spectroscopy in some cases (Fig. 1) (*12*). The depletion of these proteins in pollen coat mutants indicated they are extracellular (*13*). Comparisons of peptides to the *Arabidopsis* genomic sequence showed two proteins that corresponded to putative receptor kinases with extracellular domains, one that matched a potential EF-hand Ca^{2+} binding protein, two that fit genomic sequence for lipase proteins (extracellular lipase EXL4 and EXL6), and five that corresponded to oleosins (GRP19, GRP16, GRP17, GRP18, and GRP14). The kinases are similar to each other (BLASTP value $E < 10^{-55}$) and to many other *Arabidopsis* kinases. The EXL proteins share characteristics with the proline/hydroxyproline–rich glycoproteins (HRGP), a loosely defined family found in plant reproductive tissues (*14*), including similarities to the HRGPs APG and CEX (BLASTP values $E < 10^{-70}$), high proline content, and linkage to carbohydrate moieties (*12*). Each EXL also contains a predicted family II lipase domain; these domains can act in

Table 1. Oleosin protein characteristics. 5′ elements are as defined in (*20*). Dash indicates no repetitive motif.

Gene	5′ elements	Amino acid composition		
		Exon 1 (bp)	Exon 2 (bp)	Motifs (copies)
AtGRP20	II, IV	105	48	PAA (6)
AtGRP19	I, II, III, IV	79	27	–
AtGRP16	I, II, III, IV	86	153	GGAS (12)
AtGRP17	I, II, III, IV, V	140	403	EGGM (7), EGMS (5), KSK (11)
AtGRP18	II, III, IV, V	82	146	PAAGAA (5)
AtGRP14	III, IV, V	83	103	LGG (5)
BoGRP1	I, III, IV, V	80	28	–
BoGRP2	I, II, III, IV, V	86	267	PAG (11), DKP (13), AGG (7), KPK(6)
BoGRP3	I, II, III, IV, V	288	73	GSS (6), GGK (9), KHG (10)
BoGRP4	II, III, IV	109	78	AAP (5)
BoGRP5	III, IV, V	86	86	APA (5)

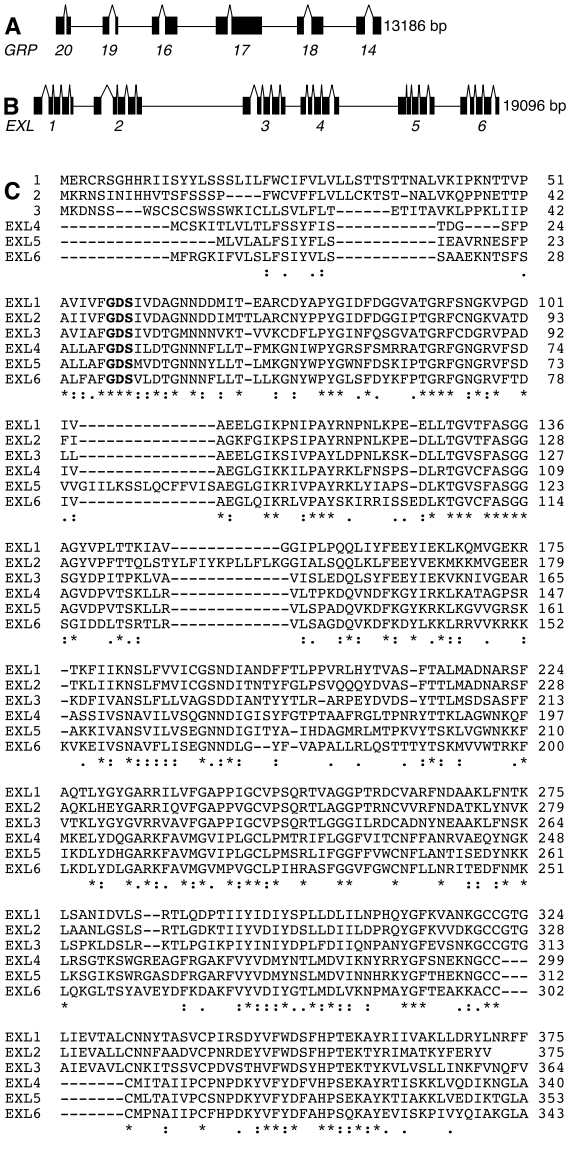


Fig. 2. Gene structure of pollen coat oleosins and lipases. (**A**) Oleosin cluster. (**B**) Lipase cluster. (**C**) Alignment of the EXL proteins by CLUSTALW. Catalytic GDS residues are in bold. (A) and (B) are drawn to scale; genes are transcribed left to right.

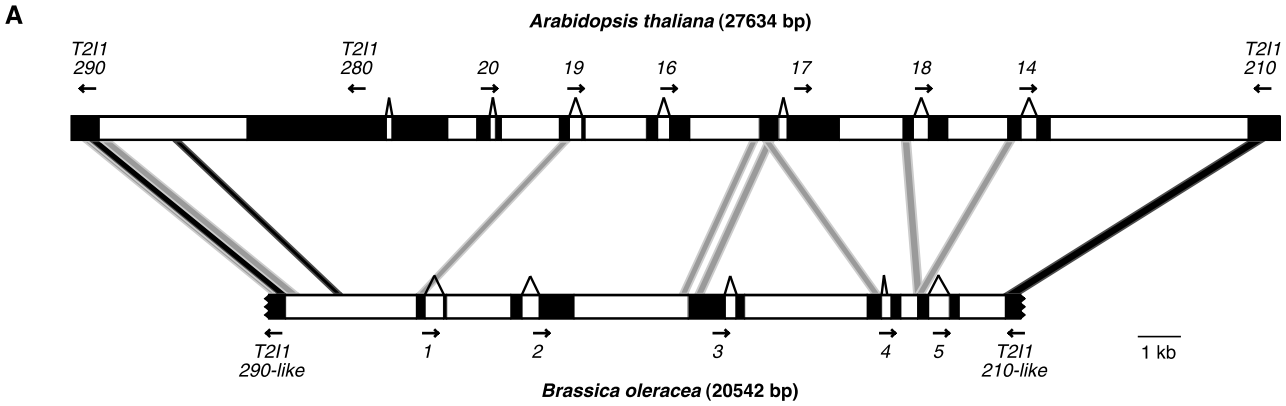
Fig. 3. Sliding window analysis of percent pairwise difference in the *AtGRP* region. Window, 500 bp. Coding regions are shaded. Dashed line, average difference between ecotypes (*22*).

Table 2. Pair-wise analysis of nucleotide variation between *Arabidopsis* ecotypes. Indels give the number of insertions or deletions, with the number that occur in frame given in parenthesis. Π_{ij} is the total number of single nucleotide differences. $\pi = \Pi_{ij}$ /length. N/A, not applicable. Blank cells indicate no change in the compared region.

Region	Length (bp)	*Ler* versus Col			Cvi versus Col			*Ler* versus Cvi		
		Indels	Π_{ij}	π ($\times 10^2$)	Indels	Π_{ij}	π (x10^2)	Indels	Π_{ij}	π ($\times 10^2$)
Intragenic	9147	16 (0)	56	0.61	24 (2)	110	1.20	29 (3)	110	1.20
Introns	1511	1 (0)	3	0.20	2 (0)	15	0.99	3 (0)	16	1.06
AtGRP20										
Exon 1	318		1	0.31		5	1.57		6	1.89
Exon 2	144		1	0.69	1 (1)		N/A	1 (1)	1	0.69
AtGRP19										
Exon 1	237			N/A		2	0.84		2	0.84
Exon 2	84			N/A			N/A			N/A
AtGRP16										
Exon 1	257			N/A			N/A			N/A
Exon 2	478	4 (4)	2	0.42	3 (3)	2	0.42	1 (1)		N/A
AtGRP17										
Exon 1	419		3	0.72		2	0.48		3	0.72
Exon 2	1213		3	0.25		15	1.24		13	1.07
AtGRP18										
Exon 1	243			N/A			N/A			N/A
Exon 2	445		5	1.12	1 (1)		N/A	1 (1)	4	0.90
AtGRP14										
Exon 1	249		1	0.40		1	0.40			N/A
Exon 2	334	1 (1)	8	2.40	1 (1)		NA	2 (2)	8	2.40
Total	15079	22 (5)	83	0.55	32 (8)	152	1.01	37 (8)	163	1.08

extracellular environments to perform acyl transfer reactions (*15*). Pollen coat peptides <10 kD have been characterized in other species (*7*), but *Arabidopsis* proteins of this size were not detected by our methods (*12*). Together, the lipases and oleosins make up >90% of the detectable pollen coat protein. The potential for lipid interactions—the oleosins affecting the size and character of lipid droplets and the lipases altering lipid composition—implicate these proteins as mediators of pollen coat behavior.

Genes organized in clusters can promote the generation of adaptive allelic diversity through gene duplication and rearrangement events (*16–18*). The five *GRP* genes are in a chromosome 5 cluster with *GRP20*, a sixth oleosin (Fig. 2A). Likewise, both *EXL* genes are contained within a tandem array of six putative lipases on chromosome 1 (*EXL1–EXL6*) (Fig. 2B). Gene prediction algorithms incorrectly identified the genes in the EXL region; we verified the EXL gene structure by comparison to the expressed sequence tag (EST) database, cDNAs, and reverse transcriptase–polymerase chain reaction (RT-PCR) (*19*). Each EXL gene contains five exons that encode a single lipase domain (Fig. 2, B and C). The EXL proteins share 35 to 69% amino acid identity, with conserved residues throughout (Fig. 2C). Tran-




```
B
   20  EIRRQKPTLKGVMTAFFATEAAICLLLLAGISLTGTAVALFASMPLFLVF 69
   19  -------MF-EIIQAVFSAGVALALLTFAGITLGGSVVACIISTPLFVIF 42
BoGRP1  -------MFFQIIQGVFTGVEALALLAFAGITLGGSAVGLALSTPLFILF 43
AtGRP16 MLSFLIPVV-QFFQVVIAAVASVVFLVFAGITFGASIVGLTIATPLFVIF 49
BoGRP2  MLSSLIQIF-QVFQVTSAVVVTAVLFALAGITLAGSVVGLIVATPLFVIF 49
AtGRP17 MFSFLMPLL-EVIKIIIASVASVIFVGFACVTLAGSAAALVVSTPVFIIF 103
BoGRP3  LFPFLLPMF-EVIKMVVASVASVVYLGFAGVTLSGSAVALAVSTPLFIIF 73
BoGRP4  MFSFIFPLL-NVIKLIIASVTSLVCLAFSCVTLGGSAVALIVSTPLFIIF 71
BoGRP5  MFSFLSPLL-DVIKVVVASVTSVVLFVFAGLTLSGSAVALVVSTPLFLIF 49
AtGRP18 MFSFLIFLL-EAFQVVIATVVSIVFLVFAGLTLVGSATALTITTPLFIIF 49
AtGRP14 MFSFLILQVSEVFQVVIAAVVSIVFLVLAGLTLAGSATALTITTPLFIIF 50
                .   :    :    :  . :: ::: .: ..   : *:*::*

AtGRP20 SPVLVPAGIATTILASGLMAGGTSGVSGLTILMWLY- 105
AtGRP19 SPVLVPATIATTLLASGFTASGSFGATAFTILSWLY- 78
BoGRP1  SPILVPATIATTLLTTGFTTSGGLGMVALRIFWKLF- 79
AtGRP16 SPILVPATIATTFLVGGATAAVALGVTAFALILWLF- 85
BoGRP2  SPVLVPATIASTLLATNLSAGALFGVTAAALIVWLLK 86
AtGRP17 SPVLVPATIATVVLATGFTAGGSFGATALGLIMWLV- 139
BoGRP3  SPILLPAIAATTVLAAGLGSKKVAAAPAASPSLSLLG 110
BoGRP4  SPILVPATIATTLLASGLMAGTTLGLTGIGLITGLV- 107
BoGRP5  SPILVPATIATTLLASGVTAGATLGITAISLIMGLI- 85
AtGRP18 SPILVPATIATAVITTGFTTGGALGAMAVALI----- 81
AtGRP14 SPILVPATIATAVITTGLTAGGALGTMAAASLL---- 83
        **:*:**  *:..:. .  :    .  .
```

C

Oleosin domain pairwise % amino acid identity

	20	19	16	17	18	14	1	2	3	4
BoGRP5	48	49	47	63	56	48	53	55	57	58
BoGRP4	45	46	47	51	45	45	51	47	68	
BoGRP3	50	55	43	49	50	50	58	37		
BoGRP2	40	49	59	54	45	53	45			
BoGRP1	43	61	49	45	46	49				
AtGRP14	39	52	49	46	62					
AtGRP18	43	50	49	49						
AtGRP17	41	48	49							
AtGRP16	36	48								
AtGRP19	50									

T2I1 290	75/81
T2I1 210	76/84

Fig. 4. Synteny between the *A. thaliana* and *B. oleracea* oleosin clusters. (**A**) Comparisons of the oleosin clusters and flanking DNA. Black boxes, exons; arrows, direction of transcription; black connecting bars, regionswith a BLASTN value of $E < 10^{-30}$; grey connecting bars, regions with a value of $E < 10^{-20}$. (**B**) Alignment of oleosin domains by CLUSTALW. (**C**) Pair-wise amino acid identity comparisons for the oleosin domain.

scription from *EXL1* and *EXL3-6* but not *EXL2* was detected in flower buds by RT-PCR (data not shown).

Previously, several *GRP* cDNAs were identified based on their glycine-rich content, and a single genomic clone corresponding to three cDNAs was found (*20*). Our data define a larger array. RT-PCR experiments confirmed expression of all six oleosins in flower buds (data not shown). Despite transcription, GRP20 was not detected in the pollen coating, perhaps due to low abundance (Fig. 1). All six GRP proteins contain a consensus oleosin domain, encoded by exon 1, and share similar 5′ upstream elements (Table 1). Exon 2 varies substantially in both length and character, each protein containing a unique repetitive motif (Table 1). This repetition parallels certain other mating genes, where repeat units within a gene are thought to allow for adaptive changes in interacting molecules. Variation in individual repeat sequences may initiate the evolution of species barriers, whereas unaltered repeat units maintain current interactions (*3*).

Gene duplications occur at approximately 1% per gene per million years, but most duplicated genes subsequently degenerate rapidly to form a pseudogene (*21*). To explore the evolution of *Arabidopsis* oleosins, we compared the sequence of *AtGRP* clusters from five *Arabidopsis* ecotypes, cross-fertile strains of *Arabidopsis* collected from different geographical locations (*22*). Some gene clusters, such as plant defense gene arrays, can contain a mixture of functional and nonfunctional genes (*17*). However, all genes of the *GRP* cluster were functional in the five ecotypes we surveyed. Although insertions or deletions (indels) occurred in both coding and noncoding sequence, coding region indels were constrained to multiples of three nucleotides, maintaining the open reading frames (Table 2). Most amino acid substitutions and indels within the coding sequences occurred in exon 2, as opposed to the exon 1 oleosin domain (Table 2). Indels in exon 2 frequently altered the number or order of repeat units. Nucleotide polymorphisms in exon 1 were primarily silent (16 silent/10 substitution), whereas exon 2 changes usually caused substitutions (20 silent/42 substitution). The average number of polymorphisms between the L*er* and Col-0 ecotypes, counting indels as single events, was 0.70%, higher than the genomic average of 0.57% (*22*). However, pair-wise sequence comparisons revealed regions of higher polymorphism in both coding and noncoding sequences (Fig. 3). The pattern of change differed between ecotypes and could generate mating haplotypes consisting of unique combinations of alleles.

If the oleosin gene cluster constitutes a recognition haplotype, related species should maintain a cluster but allow divergent alleles. We identified a baterial artificial chromosome (BAC) clone containing five putative pollen oleosin genes (*BoGRP1-5*) (Fig. 4A) from *Brassica oleracea*, a species that diverged from *Arabidopsis* between 12 and 19 million years ago (*19*, *22*, *23*); some of these corresponded to previously described cDNAs (*24*). Both the *Arabidopsis* and *B. oleracea* oleosin clusters are flanked by genes sharing ≥75% identity, suggesting synteny (Fig. 4A). The *B. oleracea* oleosins are also transcribed in the same direction, share conserved 5′ upstream elements, have a consensus oleosin domain in exon 1, and have a repetitive exon 2 (Table 1 and Fig. 4A). Oleosin gene homologies reflected position in the cluster, with genes in similar positions showing greater similarity (Fig. 4, A, B, and C). Although the lipid-binding function of the oleosins appears conserved, only 40 to 63% amino acid identity is observed between *B. oleracea* and *Arabidopsis*, and only in exon 1, whereas syntenic regions in these species generally show 85% amino acid identity (*23*). This increased rate of change parallels that observed for alleles of the highly polymorphic self-incompatibility (SI) protein SRK, which shows 50 to 60% amino acid similarity between *Brassica* and *Arabidopsis lyrata*, a closely related, self-incompatible relative of *A. thaliana* (*25*).

The organization of pollen coat genes highlights four trends observed for recognition molecules: higher than average polymorphism, repetition to allow cycles of drift and adaption in interacting molecules, organization in a cluster to facilitate allelic diversity, and generation of unique haplotypes to assist the inheritance of an entire cassette of gene. The oleosin clusters defined here suggest selective pressure to maintain multiple oleosins across species boundaries and functional copies of all six oleosins within *Arabidopsis*. Although wholesale elimination of oleosins was not seen and the observed single amino acid substitutions may barely affect pollination, over extended time accumulated small variations across the entire cluster might lead to speciation. Future experiments involving transfer of mating haplotypes between species will ultimately test species-specificity.

References and Notes

1. A. D. Richman, M. K. Uyenoyama, J. R. Kohn, *Science* **273**, 1212 (1996).
2. P. J. Ferris, C. Pavlovic, S. Fabry, U. W. Goodenough, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8634 (1997).
3. V. D. Vacquier, *Science* **281**, 1995 (1998).
4. G. J. Wyckoff, W. Wang, C. I. Wu, *Nature* **403**, 304 (2000).
5. A. L. Hughes, M. Nei, *Nature* **335**, 167 (1988).
6. J. M. McDowell *et al.*, *Plant Cell* **10**, 1861 (1998).
7. H. G. Dickinson, *Trends Genet.* **16**, 373 (2000).
8. M. Wolters-Arts, W. M. Lush, C. Mariani, *Nature* **392**, 818 (1998).
9. J. A. Mayfield, D. Preuss, *Nature Cell Biol.* **2**, 128 (2000).
10. A. Fiebig *et al.*, *Plant Cell* **12**, 2001 (2000).
11. The Arabidopsis Genome Initiative, *Nature* **408**, 796 (2000).
12. Pollen coat proteins from 25,000 L*er* plants were obtained as described (*9*), and purified over C18 resin in cyclohexane. Edman sequencing (Stanford Pan Facility) identified amino acid sequences (*26*) as follows: AtGRP17 (IKKSKAKSKGGLKAXQKKM, GMSGSGGGKHKIGGGKHK); EXL6 (AYEVISKPIVYQIAK, VFTDIVAEGLQIK, GLTSTAVEYDFKDAK); EXL4 (VNGSFPALLAF, LFNSPSDLR, DQVNDFKGYIRK, YVFYDFVHPSEK); BAB01369 (LDDIIQSFSNK, ELYPFXNP); CAB45833.1 (YELYPFLR); AAF87020.1 (AIYDGSLFHQLEK); AtGRP18(DQPGASGDKPGEMSGAGXP); AtGRP16 (WFGPGAADGDASAAGGAPGA, FSXTFGNKPGGGNPFXDISK); AtGRP14 (PGGGGLGGGGVPGGLGGLGGGGLPGGLGGLGGGENPLAK, FGKPGGGGLGGGGVPGGLGGLGGGG); AtGRP19 (TGRDLPKIPGLTPPAPASNPAGXG). Mass spectroscopy confirmed the identities of EXL6, BAB01369, CAB45833.1, and GRP16. An aggregate >95 kD contained the EXL6 peptide AYEVISKPIVYQIAK and several unidentified peaks, and a 32-kD protein contained peptides corresponding to CAB45833.1 and AAF87020.1 (YELYPFLR and AIYDGSLFXQ, respectively). Carbohydrate linkage was determined by concanavalin A blotting.
13. G. M. Zinkl, D. Preuss, *Ann. Bot. (Suppl. A)* **85**, 15 (2000).
14. J. Sommer-Knudsen, A. Clarke, A. Bacic, *Sex. Plant Reprod.* **10**, 253 (1997).
15. C. Upton, J. T. Buckley, *Trends Biochem. Sci.* **20**, 178 (1995).
16. M. Kasahara, J. Nakaya, Y. Satta, N. Takahata, *Trends Genet.* **13**, 90 (1997).
17. R. W. Michelmore, B. C. Meyers, *Genome Res.* **8**, 1113 (1998).
18. X. Wang, A. L. Hughes, T. Tsukamoto, T. Ando, T. H. Kao, *Plant Physiol.* **125**, 1012 (2001).
19. EST sequence data (EXL1, 5 ESTs; EXL3, 31 ESTs; EXL4, 2 ESTs; EXL6, 1 EST), cDNAs (EXL1, 5 cDNAs; EXL4, 18 cDNAs; EXL6, 9 cDNAs), and RT-PCR products using L*er* flower bud mRNA with ubiquitin as a control defined the lipases (AY028609-AY028614). *Brassica oleracea* oleosins were defined using EST (BoGRP1, 2 ESTs; BoGRP3, 3 ESTs; BoGRP4, 5 ESTs; BoGRP5, 1 EST) and cDNA (BoGRP3, 3 cDNAs) sequences.
20. D. E. de Oliveira, *et al.*, *Plant J.* **3**, 495 (1993).
21. M. Lynch, J. S. Conery, *Science* **290**, 1151 (2000).
22. *Arabidopsis* sequence was as published (*11*) or was amplified from L*er*, Ws-2, Kas-1, or Cvi-0 [Arabidopsis Biological Resource Center (ABRC), the Ohio State University; AF362474-AF362487]. Polymorphisms were confirmed by sequencing independent products. Coding regions from Kas-1 and Ws-2 had four and two indels, respectively, as compared to Col. Genomic average of polymorphism between Col-0 and L*er* was from the database of random L*er* sequences [average 552 base pairs (bp); see www.tigr.org/tdb/at/atgenome/Ler.html]. Polymorphisms were confirmed by analysis of both strands. Accumulated sequence (55 kb) revealed π (nucleotide polymorphisms + indels)/length. Seventy-five percent of the sequences had $\pi < 0.57\%$. Cvi averages were estimated (*27*). Subclones and partial digests from an oleosin-containing *B. oleracea* BAC (BAC 37N21, Texas A&M University) were sequenced (AY028608).
23. A. C. Cavell, D. J. Lydiate, I. A. Parkin, C. Dean, M. Trick, *Genome* **41**, 62 (1998).
24. R. K. Ruiter, G. J. Van Eldik, R. M. Van Herpen, J. A. Schrauwen, G. J. Wullems, *Plant Cell* **9**, 1621 (1997).
25. M. Kusaba *et al.*, *Plant Cell* **13**, 627 (2001).
26. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
27. C. Alonso-Blanco *et al.*, *Plant J.* **14**, 259 (1998).

28. We thank S. Hall, M. Johnson, N. Miley, R. Palanivelu, and R. Swanson for critical reading. Supported by an NSF predoctoral fellowship to A.F., an HHMI grant for undergraduate education 52003037 to S.E.J., and an NSF grant 97-35304-4941 to D.P.

22 March 2001; accepted 23 May 2001