Self-Incompatibility. Prospects for a Novel Putative Peptide-Signaling Molecule

Christel R. Schopfer and June B. Nasrallah*

Department of Plant Biology, Cornell University, Ithaca, New York 14853

The recent identification of the long elusive pollen determinant of self-incompatibility (SI) in the Brassicaceae family (Schopfer et al., 1999) marks a major advance in SI research. Not only is it now unambiguously proven that in the genus *Brassica* SI specificity of pollen and stigma is governed by separate genes, but in addition, the availability of the pollen specificity gene opens new avenues for the study of the genetic, biochemical, and physiological aspects of an intriguing cell-cell communication system in plants.

The Brassicaceae comprise many predominantly outbreeding species. These species favor cross-pollination over self-pollination by means of a sophisticated SI system, whereby stigma epidermal cells (the papillar cells) recognize and reject self-related pollen. The SI recognition reaction is controlled genetically by a single locus, the complex and rearranged *S* locus (Nasrallah, 2000, and refs. therein) for which over 50 variants, designated *S* haplotypes, are known. Two sequence-related *S*-locus-encoded proteins, the *S*-locus glycoprotein (*SLG*) and the *S*-locus receptor kinase (*SRK*), are expressed specifically in stigmatic papillar cells. The *SLG* and *SRK* genes display extensive sequence polymorphism, which has been well documented for over 30 alleles (Stein et al., 1991; Kusaba et al., 1997). Genetic and biochemical evidence indicates that both genes are involved in the SI response of stigmas: *SRK* determines SI specificity of the papillar cells, and *SLG* seems to enhance the strength of the SI reaction (Takasaki et al., 2000), possibly by contributing to the stabilization of *SRK* (Dixit et al., 2000). Because *SRK* encodes a plasma membrane-localized receptor-like protein kinase, its biochemical function in SI recognition has been proposed to be the perception of a pollen-borne signal molecule and subsequent activation of a cytoplasmic signal transduction cascade that ultimately causes rejection of self-related pollen. Although the kinase activity of SRK has been confirmed experimentally, its receptor function remains circumstantial in the absence of any evidence that binding of a ligand to its ectodomain is translated into a cytoplasmic signal.

Unlike the molecules responsible for stigma SI specificity, the molecules encoding SI specificity in pollen remained hypothetical until last year, when the *S*-locus Cys-rich (*SCR*) gene, represented by alleles from three *S* haplotypes, was identified and shown in a transgenic approach to confer SI recognition specificity to pollen (Schopfer et al., 1999). Alleles from several other *S* haplotypes subsequently have been identified (Fig. 1A). In agreement with the three prototypic sequences (SCR_6, SCR_8, and SCR_{13}), all SCR alleles encode small polypeptides with a bipartite structure consisting of a conserved, mostly hydrophobic N-terminal part and a variable, mostly hydrophilic C-terminal part (Fig. 1A). The larger portion of the conserved N terminus has the characteristics of a secretion signal, which in the absence of other known localization signals suggests that the SCR polypeptides are processed via the secretory pathway with the apoplast being a likely destination. Unless further processing occurs, the mature gene products of *SCR* can therefore be assumed to be small (<8 kD) and hydrophilic polypeptides.

Genetic control of pollen SI specificity has long been used as a primary criterion in the classification of SI systems. In sporophytically controlled SI systems such as those of the Brassicaceae, the phenotype of a pollen grain reflects the genotype of its diploid parent plant and not its own haploid genotype. Three hypotheses have been put forth to explain sporophytic control in SI. One hypothesis is based on premeiotic expression of the *S*-locus pollen gene within the meiocytes (Pandey, 1970). Another more widely accepted hypothesis proposes that the pollen *S* gene is expressed in the sporophytically derived cells of the tapetum (Heslop-Harrison, 1975), followed by transfer of its product to the pollen coat either by active secretion from the tapetal cells or by release from degenerating tapetal cells. More recently, the observation that a pollen coat protein is expressed gametophytically prompted a hypothesis of pseudosporophytic control of pollen coat proteins, and by inference of the pollen SI determinant (Doughty et al., 1998). This study counters the long-held belief that pollen coat components must be derived from the tapetum and that the thick exine layer presents a barrier to secretion from the pollen grain. In contrast, this new hypothesis assumes the free diffusion and mixing of gametophytically encoded coat proteins between pollen grains.

It was hoped that the expression pattern of the anther-specific *SCR* gene would provide unambiguous support for one of the above hypotheses. However, the *SCR* gene exhibits a complex expression

* Corresponding author; e-mail jbn2@cornell.edu; fax 607–255–5407.

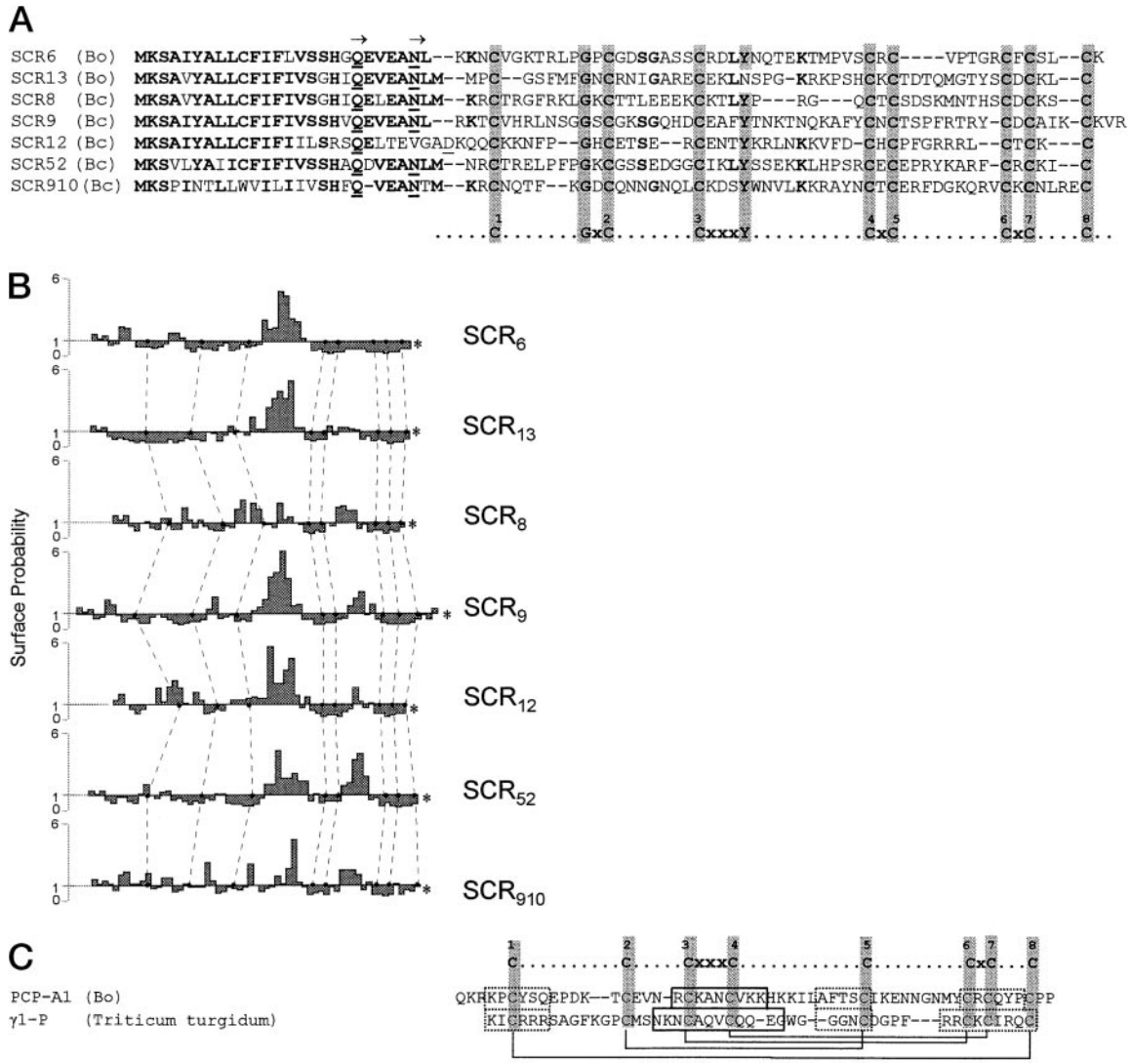


Figure 1. A, SCR sequences are extremely polymorphic except for a putative highly conserved secretion signal. The amino acid sequences of the SCR alleles were deduced from cDNA sequences with the exception of SCR_{910}. Based on homology to SCR_6, SCR_8, and SCR_{13} (Schopfer et al., 1999), we identified exons of *SCR_{910}* in the genomic sequence of the *S_{910}* haplotype (submitted to GenBank by Cui et al., 1999; accession number AJ245479.1). In the absence of an unambiguous splice site in the *SCR_{910}* gene, we tentatively joined the conserved Gln at position 20 in the first exon with Val encoded by the second exon. The SCR allele of the *S_9* haplotype was reported as SP11 without assignment of function (Suzuki et al., 1999). The SCR alleles of the haplotypes *S_{12}* and *S_{52}* were reported as SP11-12 and SP11-52, respectively (Takayama et al., 2000). Gaps were introduced to force the alignment of the conserved Cys residues C_1 to C_8. Residues conserved in at least four of the seven sequences are marked in bold. Classification as secretory proteins and signal peptide prediction was done using the TargetP and SignalP web servers (O. Emanuelsson, H. Nielsen, and G. von Heijne, unpublished data; Nielsen at al., 1997). Residues representing the potential N terminus residues of the mature polypeptide are underlined. The exact signal peptidase cleavage site will have to be determined experimentally. *Bc, Brassica campestris*; *Bo, Brassica oleracea*. B, The predicted secreted SCR polypeptides have a similar surface probability profile. The surface probability of hexapeptide regions of the amino acid sequences was calculated according to Emini et al. (1985) using LASERGENE software (version 4.03, 1999, DNASTAR, Inc., Madison, WI). Values above the default surface decision threshold of 1.0 indicate a likely localization at the surface of the SCR polypeptide. Dotted lines connect the positions of corresponding Cys, which are indicated by black circles. The asterisk marks the C terminus of the SCR polypeptide. C, Cys patterns in other small Cys-rich polypeptides in plants. Pollen-borne polypeptides are represented by PCP-A1 (Doughty et al., 1998) and plant defensins are represented by γ1-P (Colilla et al., 1990). Secondary structure elements of PCP-A1, outlined with dotted boxes (β-strands) and black boxes (α-helix), were predicted based on homology modeling (Doughty et al., 1998) onto the solution structure of γ1-P (Bruix et al., 1993). The brackets indicate disulfide bridges determined for γ1-P. Similar to γ1-P, PCP-A1 is predicted to adopt a triple-stranded β-sheet oriented parallel to the α-helix, with the Cys in the α-helix (motif C_3xxxC_4) connecting to the Cys in the third β-strand (motif C_6xC_7).

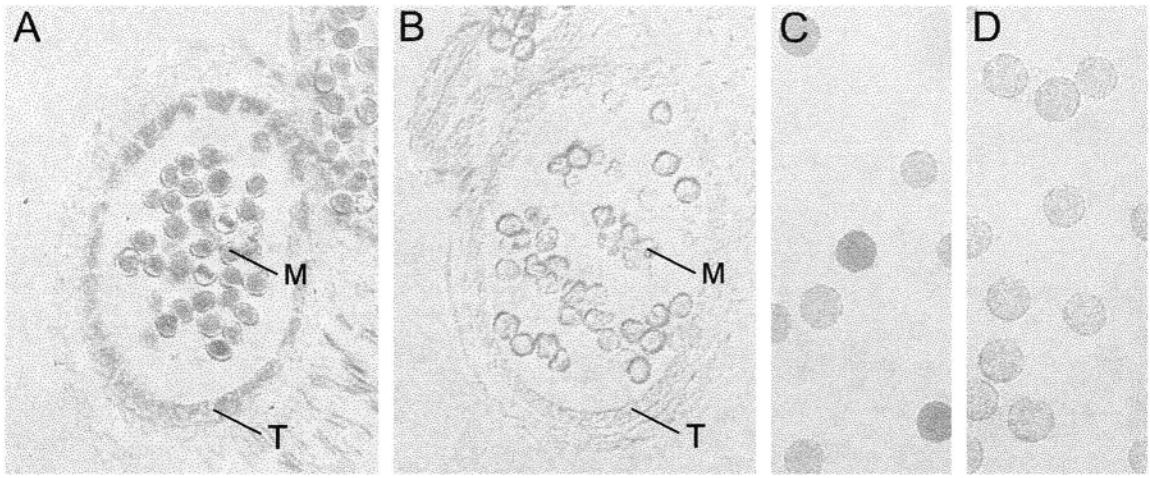


Figure 2. The SCR promoter is active both gametophytically in microspores and sporophytically in the tapetum. Anthers were stained overnight at 37°C (1 mg/mL 5-bromo-4-chloro-3-indolylglucuronide in 0.1 M sodium phosphate buffer [pH 7.0], 0.5% [v/v] Triton X-100, 10% [v/v] methanol, and 0.5 mM ferrocyanide) and embedded in paraffin before sectioning (A and B). Pollen grains isolated from dehisced anthers were stained for 20 min at 37°C (C and D). β-Glucuronidase (GUS) staining of five independent *B. oleracea* transformants carrying the SCR_8:: *uidA* reporter construct showed the patterns of GUS activity illustrated in A and C. In contrast, negative control anthers and microspores showed no GUS staining as shown in B and D. T, Tapetum; M, microspores.

pattern. *SCR* is expressed post-meiotically (Schopfer et al., 1999) in both the microspores and the tapetum as revealed by in situ hybridization (Takayama et al., 2000) and by reporter gene analysis of *SCR* promoter activity (Fig. 2). This dual gametophytic/sporophytic expression pattern precludes an easy resolution of the debate about the basis of sporophytic control in SI. Dissection of the *SCR* promoter and localization of the *SCR* protein during pollen development will be required to determine if tapetal expression is sufficient for sporophytic control, if gametophytic expression is physiologically irrelevant, or if expression in both the tapetum and in microspores is required for robust sporophytic control of the SI phenotype.

How might the SCR protein function in SI recognition? Database searches did not yield any meaningful similarity to known proteins. The known SCR sequences are notably devoid of sequence motifs indicative of an enzymatic function. The SCR polypeptides consequently are unlikely to catalyze a conversion of a common precursor molecule into a specific signal molecule by chemical modification, a mode of gene action observed for the rhizobial *nod* genes that specify the host range in symbiotic nodulation (Downie, 1998). Rather the presumably diffusible, apoplastic SCR polypeptides can be expected to function directly as ligand molecules. Strong support for this ligand hypothesis comes from the extensive sequence polymorphism of the *SCR* gene (Schopfer et al., 1999) that is evident from the alignment of the available SCR sequences shown in Figure 1A. Within the predicted mature SCR polypeptide, only the eight Cys residues and one Gly residue are strictly conserved. In addition, one Tyr residue is conserved in six of the seven sequences. This degree of sequence polymorphism exceeds the polymorphism of well-studied peptide signals such as the abalone sperm lysin that determines species specificity in fertilization (Metz et al., 1998; Vacquier, 1998) or the pheromones that determine mating specificity in the unicellular ciliate *Euplotes raikovi* (Luporini et al., 1995). The extreme polymorphism of SCR thus serves to underscore the function of SCR polypeptides as specificity-determining signal molecules. However, the physical interaction of SRK and SCR ultimately must be verified before SCR can be addressed as the pollen-borne ligand for SRK. If such a ligand function can be confirmed, SCR polypeptides would provide novel tools for substantiating the proposed receptor function of SRK, for identifying the primary specificity-determining domains in SCR and SRK, and for understanding how SRK and SCR co-evolve to maintain their functional relationship and thus the SI response itself.

The selective and absolute conservation of eight Cys and one Gly within the predicted secreted region of SCR might reflect an important function for these nine residues in the architecture of the SCR polypeptide. Based on the putative apoplastic localization of SCR, the even-numbered Cys can all be expected to form disulfide bridges and thereby function as anchor residues of the protein fold. The Gly in the motif GxC_2 similarly might indicate a turn of the backbone that might be conserved in SCR polypeptides. Do SCR variants, as a result of the conserved Cys pattern, adopt a common three-dimensional fold, perhaps obscured by the extensive divergence at the primary sequence level?

Conservation of three-dimensional protein structure is in general higher than conservation of the

primary amino acid sequence. For example, the pheromones of *E. raikovi* have been shown to share a common architecture, despite their extensive sequence variability (Luporini et al., 1995). In the case of the SCR sequences, secondary structure prediction tools such as the Profile Network Prediction Heidelberg (Rost and Sander, 1994a, 1994b) or Jnet (J.A. Cuff and G.J. Barton, unpublished data) did not allow reliable prediction of secondary structure elements. Whereas these tools did not yield any hint whether SCR polypeptides have a common tertiary structure, such a structural relationship might be reflected by the similar surface probability profiles of the SCR polypeptides known to date (Fig. 1B). It is noteworthy that these profiles predict a conserved surface-exposed region between C_3 and C_4, which is possibly involved in the interaction with SRK and might be a region of the SCR polypeptide critical for determination of specificity. It is therefore of obvious interest to obtain structural information on SCRs. Among the technologies at hand, NMR spectroscopy seems to be particularly amenable for such small and highly soluble polypeptides (Bruix et al., 1993; Luporini et al., 1995).

Structural data might also provide useful hints on the relationship of the SI recognition reaction to other plant processes. The currently held view is that SI within the Brassicaceae family has a single evolutionary origin with the SI recognition genes having been recruited from preexisting genes. The female recognition genes *SLG* and *SRK* belong to a universal superfamily of genes present in both dicots and monocots (Nasrallah and Nasrallah, 1993). This *S*-gene family notably encompasses genes of unknown function that based on predominantly vegetative gene expression are likely to be involved in functions other than SI recognition (Walker, 1993; Dwyer et al., 1994). Now as a result of the identification of the *SCR* gene, it will be possible to investigate the origin of the pollen recognition gene, and in particular to determine if polypeptides evolutionarily related to SCRs exist that have a function other than SI recognition. Small Cys-rich polypeptides are in fact very common in plants (Garcia-Olmedo et al., 1998). However, their relationship to SCRs is difficult to assess because of the high degree of SCR sequence polymorphism. There is a very real danger that small Cys-rich proteins would be spuriously classified as SCR alleles on the basis of limited and possibly accidental sequence similarity, without the necessary supporting credentials, namely data proving function as SI specificity determinants or at the very least, data demonstrating *S*-locus localization.

In the future, elucidation of structural characteristics of SCR proteins might provide an additional criterion to distinguish between true homology and accidental similarity. An example of current interest is the relationship between SCRs and the small Cys-rich PCPA1 protein, which is localized in the pollen coat but lacking *S*-locus residency and *S*-haplotype-associated polymorphism cannot have an SI specificity-determining function (Doughty et al., 1998). By homology modeling of PCPA1 onto the defensin γ1-P, a cystine-stabilized α-helix core has been proposed (Fig. 1C). PCPA1 has the same number of Cys as SCRs; however, their positioning is clearly different, as highlighted by a comparison of the C_4 residues in PCPA1 and SCR. In PCPA1, the C_4 residue is predicted to be located in the α-helix and distant from C_5 of the following β-strand, whereas in SCRs, the C_4 residue is part of the C_4xC_5 motif. Whether the different Cys patterns of PCPA1 and SCR reflect different folds might be decided when structural data for SCRs are available. If relatives of SCR can be found among the diversity of small Cys-rich polypeptides in plants, the next question would be: Do they share a similar biochemical function, possibly an interaction with receptor molecules such as orphan receptor-like protein kinases of the *S*-gene family?

Received June 22, 2000; accepted July 31, 2000.

LITERATURE CITED

Bruix M, Jimenez MA, Santoro J, Gonzalez C, Colilla FJ, Mendez E, Rico M (1993) Biochemistry **32:** 715–724

Colilla FJ, Rocher A, Mendez E (1990) FEBS Lett **270:** 191–194

Cui Y, Brugiere N, Jackman L, Bi YM, Rothstein SJ (1999) Plant Cell **11:** 2217–2231

Dixit R, Nasrallah ME, Nasrallah JB (2000) Plant Physiol (in press)

Doughty J, Dixon S, Hiscock SJ, Willis AC, Parkin IAP, Dickinson HG (1998) Plant Cell **10:** 1333–1347

Downie JA (1998) *In* HP Spaink, A Kondorosi, PJJ Hooykas, eds, The Rhizobiaceae, Ed 1. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 387–402

Dwyer KG, Kandasamy MK, Mahosky DI, Acciai J, Kudish BI, Miller JE, Nasrallah ME, Nasrallah JB (1994) Plant Cell **6:** 1829–1843

Emini EA, Hughes J, Perlow D, Boger J (1985) J Virol **55:** 836–839

Garcia-Olmedo F, Molina A, Alamillo JA, Rodriguez-Palenzuela P (1998) Biopolymers **47:** 479–491

Heslop-Harrison J (1975) Annu Rev Plant Physiol **26:** 403–425

Kusaba M, Nishio T, Satta Y, Hinata K, Ockendon D (1997) Proc Natl Acad Sci USA **94:** 7673–7678

Luporini P, Vallesi A, Miceli C, Bradshaw RA (1995) J Eukaryot Microbiol **42:** 208–212

Metz EC, Robles-Sikisaka R, Vacquier VD (1998) Proc Natl Acad Sci USA **95:** 10676–10681

Nasrallah J (2000) Curr Opin Plant Biol (in press)

Nasrallah JB, Nasrallah ME (1993) Plant Cell **5:** 1325–1335
Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Protein Eng **10:** 1–6
Pandey KK (1970) Euphytica **19:** 364–372
Rost B, Sander C (1994a) CABIOS **10:** 53–60
Rost B, Sander C (1994b) Proteins **19:** 55–72
Schopfer CR, Nasrallah ME, Nasrallah JB (1999) Science **286:** 1697–1700
Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB (1991) Proc Natl Acad Sci USA **88:** 8816–8820
Suzuki G, Kai N, Hirose T, Fukui T, Nishio T, Takayama S, Isogai A, Watanabe M, Hinata K (1999) Genetics **153:** 391–400
Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K (2000) Nature **403:** 913–916
Takayama S, Shiba H, Iwano M, Shimisato H, Che F-S, Kai N, Watanabe M, Suzuki G, Hinata K, Isogai A (2000) Proc Natl Acad Sci USA **97:** 1920–1925
Vacquier VD (1998) Science **281:** 1995–1998
Walker JC (1993) Plant J **3:** 451–456

CORRECTIONS

Vol. 124: 935–939, 2000

Schopfer, C.R., and Nasrallah, J.B. Self-Incompatibility. Prospects for a Novel Putative Peptide-Signaling Molecule.

Figure 2 was erroneously printed in black and white. Figure 2 has been reprinted in color on p 2204.

Vol. 124: 1007–1017, 2000

Stotz, H.U., Pittendrigh, B.R., Kroymann, J., Weniger, K., Fritsche, J., Bauke, A., and Mitchell-Olds, T. Induced Plant Defense Responses against Chewing Insects. Ethylene Signaling Reduces Resistance of Arabidopsis against Egyptian Cotton Worm But Not Diamondback Moth.

The GenBank accession number of the β-glucosidase gene was not included when this article was first published. The GenBank accession number is AJ251301.

Vol. 124: 1511–1514, 2000

Dennison, K.L., and Spalding, E.P. Glutamate-Gated Calcium Fluxes in Arabidopsis.

Figure 1 was erroneously printed in black and white in the original publication and again in Vol. 125 on p 1151. Figure 1 has been reprinted in color on p 2205.

Vol. 124: 1532–1539, 2000

Gibson, S.I. Plant Sugar-Response Pathways. Part of a Complex Regulatory Web.

In Table I, the line "*sis5* Is allelic to *aba4*" should have appeared as "*sis5* Is allelic to *abi4*." Table I has been reprinted on p 2206.

Vol. 125: 15–19, 2001

Meyerowitz, E.M. Prehistory and History of Arabidopsis Research.

Professor Georges Bernier of the Universite de Liege (Belgium) kindly sent the following corrections for the photographs that appeared as Figures 1 and 2. In Figure 1, the last person on the right of the first row is Silvano Bonotto, not J. Bouharmont; in the third row, between A.R. Kranz and M. Jacobs, the unidentified person is J. Bouharmont. In Figure 2, in the back row, the person identified as Matigne is in fact R. Matagne. We welcome any additional information on the names of those who appear in the photographs.

Vol. 125: 329–338, 2001

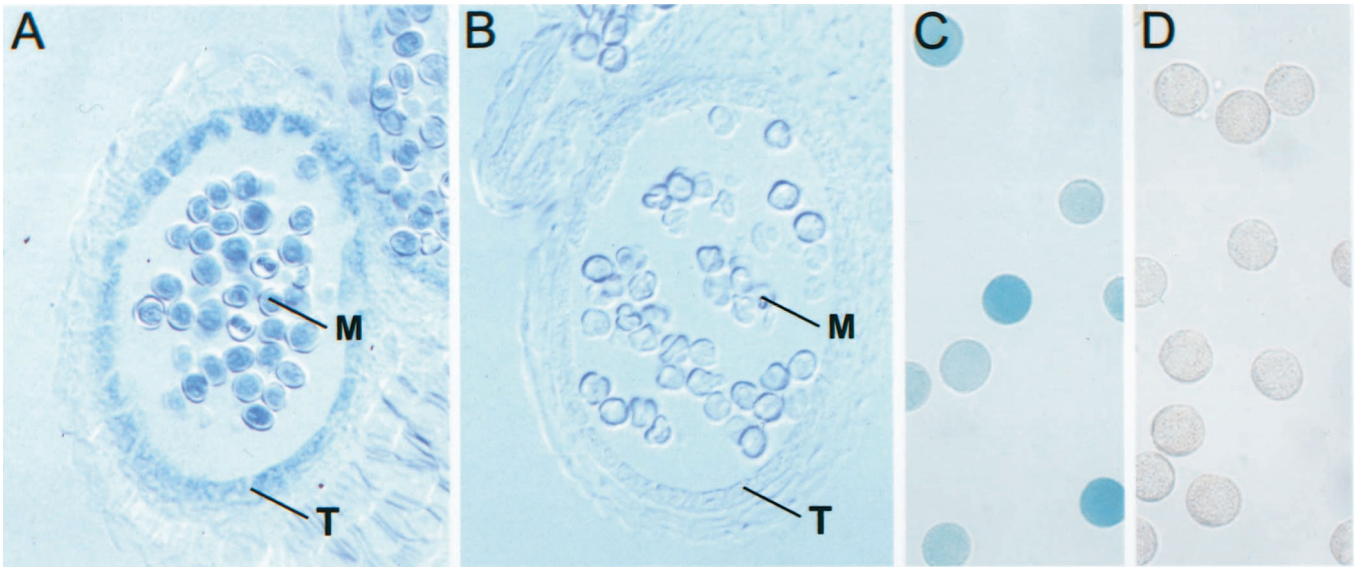
Taylor, A.R., and Assmann, S.M. Apparent Absence of a Redox Requirement for Blue Light Activation of Pump Current in Broad Bean Guard Cells.

Figures 2, 3, and 4 were not printed in the correct order. The correctly numbered figures with legends are reprinted on pp 2207–2209.

Acknowledgment

Vol. 125: No. 1, ii, 2001

We would like to acknowledge Jan Zeevaart, who supplied the photograph of the morning glory flower that appears on the cover of the January 2001 75th Anniversary Special Issue.


A
M
T
B
M
T
C
D

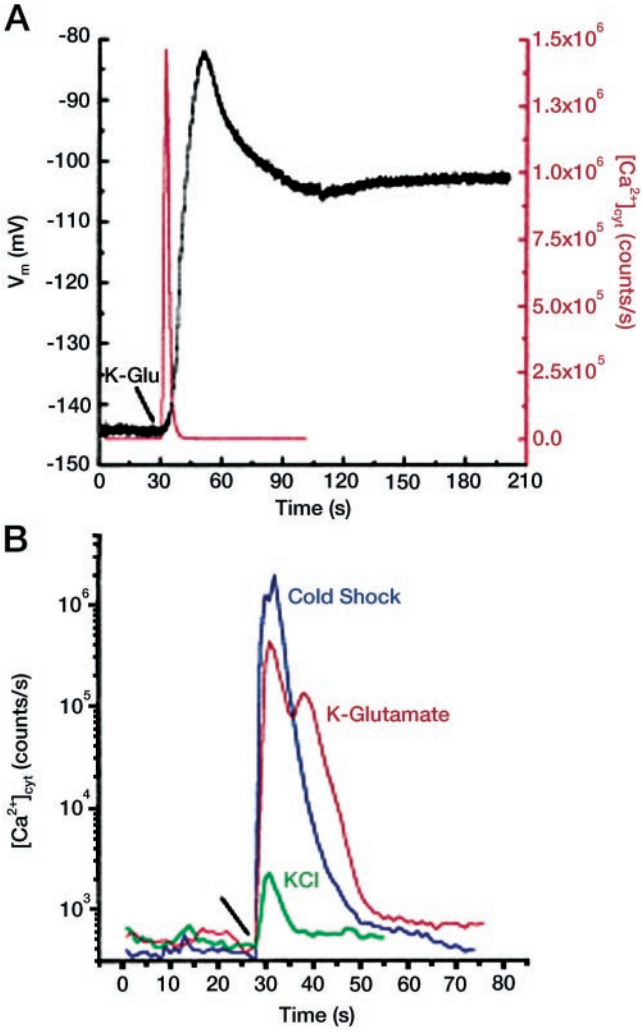

A
-80
-90
-100
-110
-120
-130
-140
-150
V_m (mV)
K-Glu
0
30
60
90
120
150
180
210
Time (s)
1.5x10^6
1.3x10^6
1.0x10^6
7.5x10^5
5.0x10^5
2.5x10^5
0.0
$[Ca^{2+}]_{cyt}$ (counts/s)
B
10^6
10^5
10^4
10^3
$[Ca^{2+}]_{cyt}$ (counts/s)
Cold Shock
K-Glutamate
KCl
0
10
20
30
40
50
60
70
80
Time (s)

Table I. *Sugar-response mutants and corresponding loci*

Mutants	Originals Selection	Loci	References
rsr	Reduced sensitivity to Suc induction of patatin expression		Martin et al., 1997
lba	Reduced sensitivity to Suc induction of β-amylase expression		Mita et al., 1997b
hba	Increased sensitivity to Suc induction of β-amylase expression		Mita et al., 1997a
sun	Reduced sensitivity to Suc repression of plastocyanin expression	*sun6* ls allelic to *abi4*	Dijkwel et al., 1997; Huijser ete al., 2000
sis	Reduced sensitivity to Glc or Suc-mediated inhibition of early seedling development	*sis1* ls allelic to *ctr1* *sis4* ls allelic to *aba2* *sis5* ls allelic to *abi4*	Laby et al., 2000; S. Gibson, R. Laby, and D. Kim, unpublished data
gin	Reduced sensitivity to Glc-mediated inhibition of early seedling development	*gin1* ls allelic to *aba2* *gin6* ls allelic to *abi4*	Zhou et al., 1998; Arenas-Huertero et al., 2000; J. Sheen, personal communication
prl	Increased sensitivity to sugar-mediated inhibition of early seedling development	*PRL1* Encodes a WD-40 protein	Németh et al., 1998; Bhalerao et al., 1999

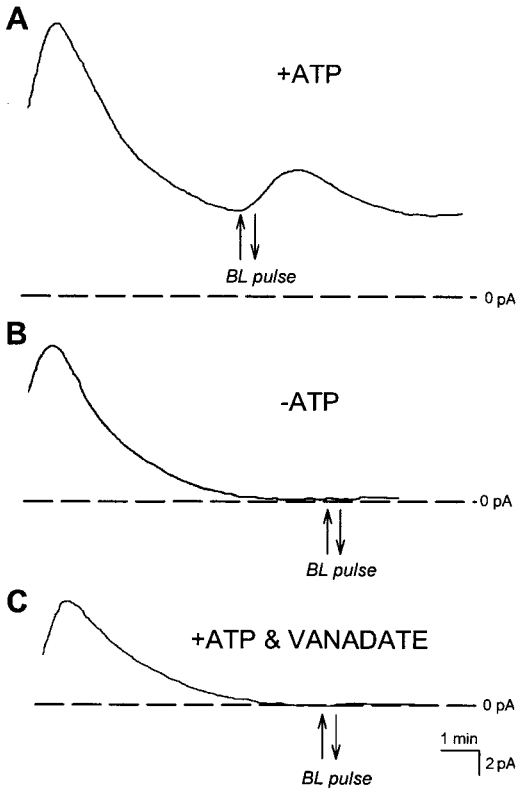


Figure 2. Steady-state- and BL-stimulated pump currents require ATP and are inhibited by vanadate. A, A typical recording with 5 mM ATP in the pipette under saturating RL. The cell responded to a 30-s pulse of BL with a typical transient increase in pump current. B, When ATP is absent from the pipette, cell currents quickly decay to 0 pA under saturating RL and are unresponsive to a pulse of BL. C, Inclusion of ATP and 20 μM vanadate in the pipette causes inhibition of pump current. All cells where pump current was inhibited by vanadate were unresponsive to BL pulses.

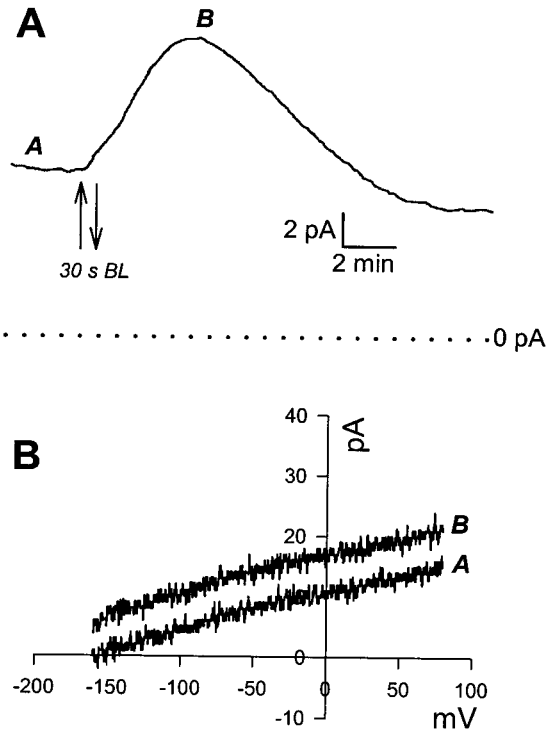


Figure 3. H^+-ATPase activation by a pulse of BL. Saturating RL background illumination was switched on before the beginning of the trace. A, Once stable baseline current is achieved a pulse of BL causes a transient increase in pump current. B, I/V ramps conducted before (*A*) and at the peak (*B*) of the response in A show the parallel shunt in pump current.

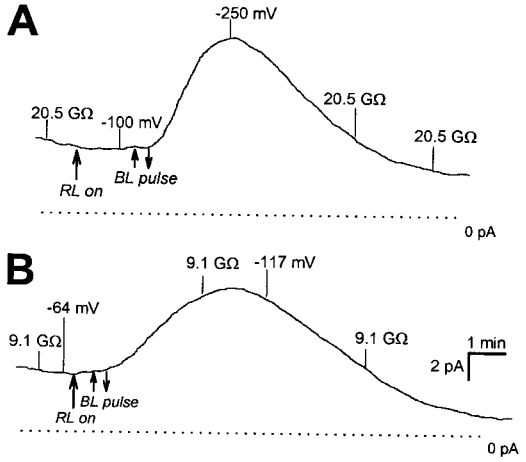


Figure 4. The effect of plasma membrane H^+-ATPase currents on membrane potential. The two traces show the pump current measured with ATP in the pipette. Membrane potential and input resistance are indicated on the traces at steady state and during BL-activated stimulation of pump current. Note the insensitivity to saturating RL illumination.