The Plant Cell, Vol. 19: 3297–3313, October 2007, www.plantcell.org © 2007 American Society of Plant Biologists

Identification and Mutational Analysis of *Arabidopsis* FLS2 Leucine-Rich Repeat Domain Residues That Contribute to Flagellin Perception [W]

F. Mark Dunning,[a,b] Wenxian Sun,[a] Kristin L. Jansen,[a] Laura Helft,[a,b] and Andrew F. Bent[a,b,1]

[a] Department of Plant Pathology, University of Wisconsin, Madison, Wisconsin 53706
[b] Cell and Molecular Biology Program, University of Wisconsin, Madison, Wisconsin 53706

Mutational, phylogenetic, and structural modeling approaches were combined to develop a general method to study leucine-rich repeat (LRR) domains and were used to identify residues within the *Arabidopsis thaliana* FLAGELLIN-SENSING2 (FLS2) LRR that contribute to flagellin perception. FLS2 is a transmembrane receptor kinase that binds bacterial flagellin or a flagellin-based flg22 peptide through a presumed physical interaction within the FLS2 extracellular domain. Double-Ala scanning mutagenesis of solvent-exposed β-strand/β-turn residues across the FLS2 LRR domain identified LRRs 9 to 15 as contributors to flagellin responsiveness. *FLS2* LRR-encoding domains from 15 *Arabidopsis* ecotypes and 20 diverse Brassicaceae accessions were isolated and sequenced. FLS2 is highly conserved across most *Arabidopsis* ecotypes, whereas more diversified functional FLS2 homologs were found in many but not all Brassicaceae accessions. flg22 responsiveness was correlated with conserved LRR regions using Conserved Functional Group software to analyze structural models of the LRR for diverse FLS2 proteins. This identified conserved spatial clusters of residues across the β-strand/β-turn residues of LRRs 12 to 14, the same area identified by the Ala scan, as well as other conserved sites. Site-directed randomizing mutagenesis of solvent-exposed β-strand/β-turn residues across LRRs 9 to 15 identified mutations that disrupt flg22 binding and showed that flagellin perception is dependent on a limited number of tightly constrained residues of LRRs 9 to 15 that make quantitative contributions to the overall phenotypic response.

INTRODUCTION

Proteins with a leucine-rich repeat (LRR) domain play a central role in plant immune systems (Dangl and Jones, 2001; Jones and Takemoto, 2004). LRR proteins are also central in hormone perception, organ formation, and many other growth and development processes in both plants and animals (Kobe and Kajava, 2001; Shiu et al., 2004; Torii, 2004; Pancer and Cooper, 2006). The repeat units of an LRR form a highly regular protein structure resembling a coil or spring. Solvent-exposed residues on the face of this backbone structure determine ligand specificity, and it is becoming apparent that LRR domains are a platform on which binding sites can readily evolve (Binz et al., 2005; Pancer and Cooper, 2006). However, methods to identify the sites within LRRs that determine ligand binding specificity and other functions are not well developed. These sites will be of particular interest for future study and manipulation.

Plant disease resistance depends heavily on two systems, the gene-for-gene and microbe-associated molecular pattern (MAMP; also called PAMP) recognition pathways (Jones and Takemoto, 2004; Zipfel and Felix, 2005; Jones and Dangl, 2006; Bent and Mackey, 2007). These systems rely on intracellular and extracellular receptor proteins (*R* gene products or MAMP receptors) that very often contain a large LRR domain. In a well-developed example, flagellin perception by *Arabidopsis thaliana* is mediated by *FLAGELLIN-SENSING2* (*FLS2*), which encodes a transmembrane receptor–like kinase containing an extracellular LRR (eLRR) domain (Gomez-Gomez and Boller, 2000). In animals, MAMP perception and inflammatory and apoptotic processes can also be mediated by LRR-containing proteins (Ting and Davis, 2005; Trinchieri and Sher, 2007). The distinctions between plant gene-for-gene and MAMP recognition systems may not be universal. For example, the significantly similar proteins Xa21 and FLS2 function as an *R* gene product and a MAMP receptor, respectively, and MAMPs such as flagellin can show the strain-to-strain variation within a pathogen species that is common of many *avr* genes (Mackey and McFall, 2006; Sun et al., 2006).

Proteins containing eLRR domains allow perception of and response to a diverse array of extracellular cues (Shiu and Bleecker, 2003; Torii, 2004; van der Hoorn et al., 2005). The *Arabidopsis* genome encodes >200 apparent eLRR receptor–like kinases and >50 eLRR receptor–like proteins, and the rice (*Oryza sativa*) and poplar (*Populus trichocarpa*) genomes encode even greater numbers of these proteins (Shiu et al., 2004; Fritz-Laylin et al., 2005; Tuskan et al., 2006). LRR domains are known protein–protein interaction and ligand binding surfaces, and direct interaction between plant eLRR-RLKs and their ligands has now been shown in a few cases (Matsubayashi et al., 2002; Scheer et al., 2003; Kinoshita et al., 2005; Chinchilla et al., 2006; Yamaguchi et al., 2006). The precise binding sites for these ligands are not known. Many eLRR proteins associate with other

[1] Address correspondence to afb@plantpath.wisc.edu.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Andrew F. Bent (afb@plantpath.wisc.edu).
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.106.048801

proteins that may be required to form functional binding sites and/or signaling complexes (Jeong et al., 1999; Russinova et al., 2004; Belkhadir and Chory, 2006), but LRR sites that mediate these associations also are not known.

At least seven LRR structural subfamilies have been defined (Kobe and Kajava, 2001). LRR crystal structures reveal a highly ordered tertiary structure with each repeat typically consisting of 22 to 25 amino acids arranged as a β-strand, an α-helix, and additional connecting residues (Kobe and Kajava, 2001; Bell et al., 2003). The β-strands and α-helices of successive repeats align in parallel and form a highly ordered nonglobular structure similar in shape to a metal spring that is curved like a horseshoe and slightly twisted. The solved LRR domain structures show variations from each other, but they all form this curved shape with β-sheets lining the concave side. In plants, each eLRR repeat follows the consensus sequence xxLxxLxxLxxLxxNxLt/sGxIP (with other hydrophobic residues frequently substituted in place of L). The structure for polygalacturonase-inhibiting protein (PGIP), the only plant eLRR structure published to date, is similar to other known LRR structures yet contains structural elements that seem specific to plant eLRR proteins (Di Matteo et al., 2003). The highly similar tertiary structure of LRR proteins in general, and the availability of the plant PGIP structure, allow in silico construction of postulated but relatively valid structures for plant eLRR-containing proteins such as FLS2 (Kajava, 2001; Kajava and Kobe, 2002).

In LRRs, five solvent-exposed residues extend out from the concave face within the so-called β-strand/β-turn region of each repeat (underlined x in LxxLxLxxN). In the majority of the functionally studied examples, the LRR residues implicated in ligand binding or ligand specificity have been at one or more of these five positions in one or more LRR repeats (Kobe and Deisenhofer, 1996; Jones and Jones, 1997; Kobe and Kajava, 2001; Bell et al., 2003). For plant *R* gene products, there is indirect evidence that these solvent-exposed residues along the LRR β-strand/β-turn face may often serve as ligand binding sites. A significant subset of plant *R* genes show evidence of positive evolutionary selection for sequence diversification, and the hypothesized solvent-exposed β-strand/β-turn residues are disproportionately under positive selection in the studied examples (Meyers et al., 1998; Ellis et al., 1999, 2000; Leckie et al., 1999; Mondragon-Palomino et al., 2002; Rose et al., 2004). Furthermore, domain-swapping experiments have shown that specificity determinants of tomato (*Solanum lycopersicum*) Cf-4 and Cf-9 reside in the solvent-exposed residues within the central LRRs (Thomas et al., 1997; van Der Hoorn et al., 2001; Wulff et al., 2001). Similar experiments using site-directed mutagenesis showed that the polygalacturonase specificity of bean (*Phaseolus vulgaris*) PGIP is dependent on residues within the β-sheet/β-turn of the concave inner face (Leckie et al., 1999).

FLS2 was recently shown to directly bind flg22, a peptide based on a conserved N-terminal domain of *Pseudomonas aeruginosa* flagellin (Chinchilla et al., 2006), but the specific site of interaction between FLS2 and flg22 has not been determined. FLS2 diversity within or across taxa has not been extensively investigated. Flagellin or flg22 treatment initiates downstream plant defense responses if a functional FLS2 is present; these responses include mitogen-activated protein kinase signal transduction, the production of ethylene and reactive oxygen species, callose deposition, and pathogenesis-related gene expression (Gomez-Gomez and Boller, 2000; Asai et al., 2002; Navarro et al., 2004; Chinchilla et al., 2006). Pathogenic bacteria such as *Pseudomonas syringae* pv *tomato* produce effector molecules that contribute to bacterial virulence in part by suppressing FLS2-mediated defenses (Asai et al., 2002; Kim et al., 2005; Li et al., 2005; Oh and Collmer, 2005; de Torres et al., 2006), but FLS2 can nevertheless make a significant contribution to the resistance of *Arabidopsis* to virulent *P. syringae* (Zipfel et al., 2004). Some bacteria apparently avoid host flagellin perception, either through the production of unrecognized flagellins or using alternative infection strategies (Felix et al., 1999; Pfund et al., 2004; Sun et al., 2006).

FLS2 mutant alleles that have been studied include the nonfunctional *fls2-24*. The product of *fls2-24* is mutated on the concave face of the tenth LRR and lacks flg22 binding, suggesting that this site is important for FLS2 function and may be part of the flg22 binding site (Bauer et al., 2001; Gomez-Gomez et al., 2001). However, a mutation within the kinase domain of FLS2 also strongly reduces flg22 binding (Gomez-Gomez et al., 2001). It was suggested that this mutation might alter receptor complex stability, and it was recently found that this mutant FLS2 protein fails to accumulate (Robatzek et al., 2007). The kinase domain of FLS2 interacts with the phosphatase KAPP in yeast, and KAPP overexpression in plants causes insensitivity to flg22 (Gomez-Gomez et al., 2001). It is not known whether FLS2, prior to flagellin exposure, is present primarily as a monomer, a homodimer, or in some heteromeric form. After exposure to flg22, FLS2 associates with BAK1 (Chinchilla et al., 2007; Heese et al., 2007) and undergoes endocytosis (Robatzek et al., 2006). The FLS2 protein has not yet been subjected to extensive structure–function analysis.

This project sought to develop efficient methods to identify functional sites within LRR domains. Researchers seeking to functionally manipulate an LRR protein may not have ready access to useful binding assays or even a known ligand but may be able to use molecular genetic methods to focus their work. Given the highly ordered structure of LRRs and the large size of many LRR domains (the inferred LRR of FLS2 contains 28 repeats and consists of >650 amino acids), random mutagenesis is less likely to be productive in the absence of a very-high-throughput assay system. We instead deployed two methods: targeted Ala-scanning mutagenesis and, separately, conserved functional group studies using FLS2 variants from taxonomically related plant species. This identified a β-strand/β-turn region of apparent functional importance. Site-directed randomizing mutagenesis libraries were then generated within this region to provide additional information on the structural and chemical requirements of the FLS2 LRR domain for flg22 perception, and alleles exhibiting a loss of flg22 binding were identified.

RESULTS

Ala Scanning Reveals LRRs 9 to 15 as a Possible flg22 Perception Domain

A mutagenesis strategy was devised to identify LRR domain functional sites while accommodating the absence of a high-throughput

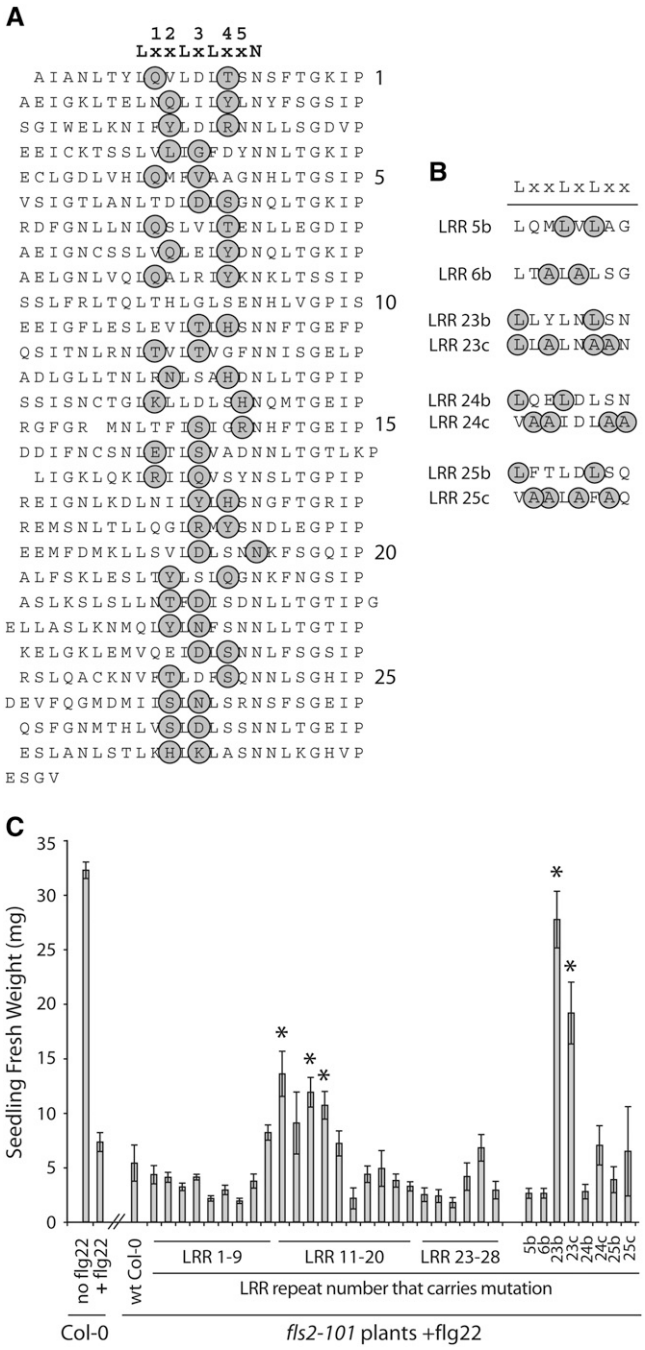

Figure 1. Ala-Scanning Mutagenesis Identifies FLS2 LRRs 9 to 15 and Possibly LRR 23 as Regions Required for Plant Response to the flg22 Flagellin Peptide.

(A) Positions within the FLS2 LRR domain subjected to double-Ala site-directed mutagenesis. Each mutant allele carried only the two mutations indicated on a single line (i.e., carried Ala codons in place of the codons for the two circled amino acids, altering only one repeat of the LRR per allele). This mutagenesis was not successful for the LRR 10 mutant allele. Numbers at right index the LRR repeats; numbers across the top index residues that received further study.

(B) Additional substitutions made for LRRs 5, 6, 23, 24, and 25 targeting L as well as x positions in the LxxLxLxxN consensus motif. Note that letters within circles designate residues encoded by the mutant allele.

(C) FLS2-mediated response of *Arabidopsis* seedlings to flg22 treatment. Mutant *fls2-101* plants were transformed with the *FLS2* alleles shown in **(A)** and **(B)**, controlled by the native *FLS2* promoter and terminator. Response to flg22 was measured using the seedling growth inhibition assay. Each bar reports the average of 10 or fewer independent T1 plants treated with 10 μM flg22 for 12 d. Error bars indicate SE. wt Col-0 transformants received the wild-type Col-0 *FLS2* construct. Significant differences from the response conferred by wt Col-0 are marked with asterisks, based on Fisher's protected LSD ($P = 0.05$). The responses of wild-type Col-0 plants from a separate experiment performed under the same conditions are shown for comparison at far left.

functional screening method and avoiding the potential inefficiencies of random mutagenesis. An LRR-specific method was developed based on structural features that are common among the crystal structures of multiple LRR proteins. The principles that guided our approach were to (1) avoid the mutagenesis of inferred buried hydrophobic residues, (2) avoid the mutagenesis of other LRR positions where the same or chemically similar amino acids recur in most repeats, and (3) among the remaining solvent-exposed residues, focus on the β-strand/β-turn region (the concave face of the LRR) that has frequently been shown or implicated in other LRR proteins to be the key ligand binding surface. Mutation to Ala codons was favored to further minimize the effects on overall LRR structure and increase the chance that function would be disrupted only when significant receptor–ligand or receptor–cofactor interaction sites were mutated. In each mutant *FLS2* allele, only a single repeat of the LRR was mutagenized, but two amino acid residues were changed per allele to increase the likelihood that relevant repeats would be disabled. The five residues corresponding to the variable x positions in the LxxLxLxxN consensus were inspected, and two nonadjacent residues with the highest likelihoods of reactive chemistry were converted to Ala by site-directed mutagenesis (Figure 1A). This strategy allowed for the efficient systematic screening of the entire LRR using a small number of alleles.

To test function in the native context, mutagenized *FLS2* LRR-encoding domains were cloned in place of the wild-type LRR-encoding DNA in a binary vector containing full-length *FLS2* under the control of the native *FLS2* promoter and terminator. The resulting alleles were tested after stable transformation into *Arabidopsis fls2-101* mutant plants, which are homozygous for a T-DNA insertion in the first third of the *FLS2* open reading frame and do not respond to flagellin or flg22 (Pfund et al., 2004). Multiple independent T1 (primary transformant) seedlings for each allele were treated with 10 μM flg22 for 12 d and weighed (Figure 1C) as part of the seedling growth inhibition assay, a convenient assay for FLS2-mediated responses to flagellin that correlates with callose deposition, reactive oxygen production, pathogenesis-related gene expression, and other defense responses (Gomez-Gomez and Boller, 2000; Pfund et al., 2004). The binary vector carrying wild-type ecotype Columbia (Col-0) *FLS2* served as a positive control. For comparison, the responses of untreated and flg22-treated wild-type Col-0 plants of the same age are also shown (Figure 1C). LRR 10 was not

included in this initial screen, as generating the desired mutations in this LRR proved difficult. Mutations within LRRs 21 and 22 were also not available during the initial screening; however, they were later tested and found to be responsive to flg22 (data not shown).

As expected, the plants transformed with a wild-type *FLS2* allele showed strong flg22-induced seedling growth inhibition (Figure 1C). In this and subsequent experiments, no reproducible flg22-independent dwarfing (possible constitutive gain-of-function) mutations were observed. Importantly, most mutant *FLS2* alleles retained a strong flg22 response, validating the hypothesis that Ala-scanning mutation of two putative solvent-exposed β-strand/β-turn residues within a single repeat would leave function unperturbed for most LRR subdomains. Across the full set of primary Ala-scanning alleles in Figure 1A, significantly reduced activity was observed for LRRs 11, 13, and 14 (Figure 1C). Seedlings carrying Ala-scanning alleles for LRRs 9, 12, and 15, while not statistically significant in this initial screen, also tended toward reductions of flg22 sensitivity. Because of the proximity of these LRRs, subsequent studies focused on LRRs 9 to 15 as a candidate flagellin perception domain of FLS2.

At the time of the original mutagenesis, additional *FLS2* mutant alleles were made that carry more extensive changes (Figure 1B). For other LRR proteins, departures from the highly conserved LRR consensus motif have been postulated to suggest ligand binding site proximity (Bell et al., 2003). Very few FLS2 LRRs exhibit substantial departures from the consensus, but we did investigate the clustering of three slightly altered repeats around LRRs 23 to 25. LRR 23 also contains one of only three putative *N*-linked glycosylation sites (NxS/T motifs) along the inferred concave face of the LRR domain, which may affect flagellin perception. For these LRRs, additional mutant alleles were designed that contained Ala substitutions but in which nonhydrophobic residues at some of the LRR consensus motif positions were also changed to Leu (Figure 1B). Additional alleles for LRRs 5 and 6 were also generated (Figure 1B), because LRR 5 contains a substitution in a β-sheet Leu and a very preliminary molecular docking study suggested that the exposed residues within LRR 6 might be able to pair with flg22. When tested in *fls2* mutant plants for their ability to confer a response to flg22, the added LRR 5 and LRR 6 alleles appeared to be fully functional, and the LRR 24 and LRR 25 alleles also conferred a response that was statistically similar to that of wild-type *FLS2* (Figure 1C). The more extreme LRR 23b and LRR 23c alleles did disrupt *FLS2* function; however, the LRR 23 allele carrying only two Ala substitutions conferred full *FLS2* function (Figure 1C).

Variation in the LRR Domains of *Brassica FLS2* Homologs but Not within *Arabidopsis*

As another possible route to identifying residues involved in the FLS2–flagellin interaction, *FLS2* genes from the *Arabidopsis* ecotypes Be-0, Ler-0, No-0, Po-0, Hs-0, Ws-0, Sf-2, Cvi-0, Dra-0, and Su-0 were sequenced and the derived amino acid sequences were compared using ClustalX. We hypothesized that residues within the LRR not involved in the FLS2–flagellin interaction would show greater diversity than those whose function is required for flg22 perception. However, with the exception of stop codons, very little variation in the derived amino acid sequence of the FLS2 LRR domains was present between the tested *Arabidopsis* ecotypes (98.8% identity between the open reading frames of all sequenced ecotypes; see Supplemental Figure 1 and Supplemental Table 1 online). Ws-0 was previously identified as being flg22-insensitive because of a premature stop codon in the *FLS2* kinase domain (Gomez-Gomez et al., 1999). Two additional apparently independent *FLS2* alleles with premature stop codons were found (in Cvi-0, Dra-0, and Po-0; see Supplemental Figure 1 online). Tests of a larger set of 23 *Arabidopsis* ecotypes showed that all were flg22-responsive except Cvi-0 (see Supplemental Figure 2 online; Dra-0 and Po-0 were not included in this test).

In order to expand the diversity of our FLS2 collection, FLS2 homologs from other plant species were sought. DNA gel blot analysis revealed the presence of *FLS2* homologs in *Brassica oleracea* and *Brassica rapa* (data not shown). From a collection of *Brassica* BAC clones carrying possible *FLS2* homologs (a kind gift from the Biotechnology and Biological Science Research Council [BBSRC] *Brassica* Consortium), a full-length FLS2 homolog was amplified from *B. rapa* BAC r97L9 using primers based on the *Arabidopsis* FLS2 sequence. Sequencing of the product revealed close similarity to FLS2 throughout the predicted LRR and kinase domains; however, the sequence predicts a stop codon after the kinase domain but 121 amino acids earlier than the stop position of *Arabidopsis* FLS2 (see Supplemental Figure 3 and Supplemental Table 2 online). The LRR domain is highly similar to FLS2 in both structure and length but contains only 81.1% derived amino acid sequence identity. DNA sequence data for the r97L9-derived gene was used to design new PCR primers and amplify *FLS2* homolog LRR-encoding domains from a collection of *B. oleracea* and *B. rapa* lines. Seven homologous *FLS2* domains were amplified, and after sequencing, five (from *B. rapa* TO1434, TO1441, and TO1450 and *B. oleracea* BI16 and KT18) were predicted to encode full-length LRR domains. Approximately 80% amino acid sequence identity and 90% amino acid similarity were observed between each *Brassica* and *Arabidopsis* FLS2 LRR domain (see Supplemental Figure 4 and Supplemental Table 3 online).

LRR Domains from *Brassica* FLS2 Homologs Can Function in *Arabidopsis*

The LRR-encoding segments of *FLS2* homologs from the *Brassica* accessions TO1434, TO1450, TO1441, and KT18 were cloned in-frame in place of the *Arabidopsis FLS2* LRR-encoding domain within the *Arabidopsis FLS2* gene (including the *Arabidopsis FLS2* promoter and terminator). These constructs were transformed into *Arabidopsis fls2-101* plants, and transgenic seedlings were tested for responsiveness to flg22 using the seedling growth inhibition assay. The *Brassica* LRR domains were able to function in place of the native *Arabidopsis* FLS2 LRR and signal through the *Arabidopsis* FLS2 kinase domain, resulting in flg22-sensitive plants (Figure 2A). Furthermore, the level of seedling growth inhibition seen with the *Brassica* LRR domains was similar in strength to that of the native *Arabidopsis* LRR domain.

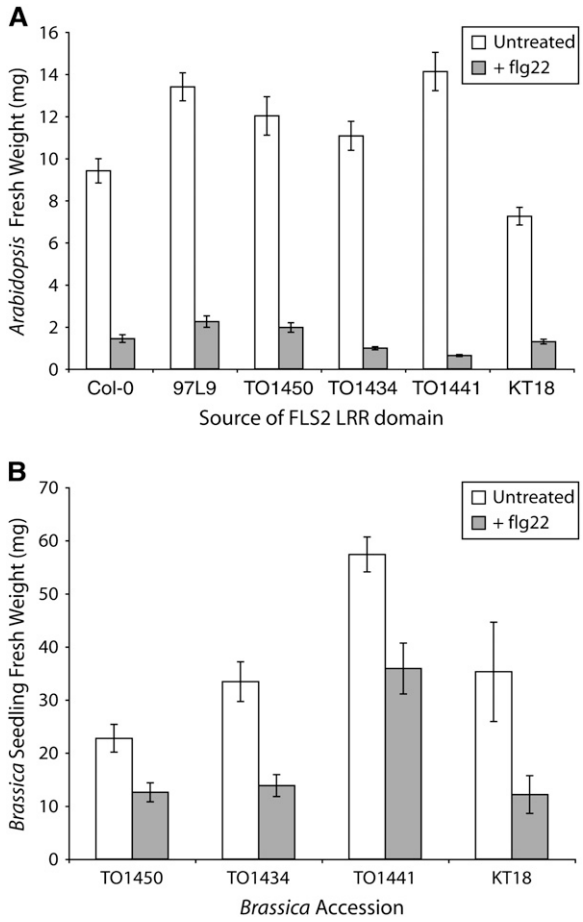

Figure 2. *Brassica* FLS2 LRR Domains Confer flg22 Sensitivity in *Arabidopsis* and Function in *Brassica*.

(A) *FLS2* homolog LRR-encoding domains from the indicated *B. rapa* and *B. oleracea* sources were cloned in place of the wild-type *Arabidopsis* Col-0 LRR-encoding domain within *Arabidopsis FLS2*. FLS2-mediated responses of independent T1 transformants of *Arabidopsis fls2-101* expressing the indicated *FLS2* transgenes were assayed by measuring seedling growth in the presence or absence of 10 μM flg22 for 12 d.

(B) *B. rapa* and *B. oleracea* seedlings germinated in 10 μM flg22 and grown for 6 d. Each bar represents 10 to 12 plants.

For both **(A)** and **(B)**, values report means ± SE.

Flagellin response assays were also performed on *Brassica* seedlings to test whether the *Brassica* parent plants from which the LRR domains were isolated exhibited flg22-dependent responses. A seedling germination assay detected flg22-dependent growth inhibition (Figure 2B). More traditional seedling growth inhibition assays with pregerminated seedlings did not detect a flg22 response in *Brassica* plants, but in those assays the stem and leaf tissue of the much larger *Brassica* seedlings rapidly grew out of the flg22-containing liquid, whereas *Arabidopsis* seedlings remained in contact with the flg22 for multiple days. Using the germination assay, all of the *Brassica* accessions that contained flg22-responsive LRR domains as tested in *Arabidopsis* were flg22-responsive.

The Brassicaceae Contain Diverse FLS2 LRR Domains

The natural diversity of Brassicaceae FLS2 LRR domains was further investigated by cloning LRR-encoding domains of *FLS2* homologs from highly diverse Brassicaceae family members. FLS2 LRR-encoding domains were amplified and sequenced from species of the Brassicaceae genera *Alliaria*, *Alyssum*, *Aurinia*, *Biscutella*, *Camelina*, *Chorispora*, *Crambe*, *Enarthrocarpus*, *Eruca*, *Erucastrum*, *Erysimum*, *Hesperis*, *Iberis*, *Isatis*, *Lepidium*, *Matthiola*, *Sinapis*, and *Thlaspi* using primers based on the *B. rapa FLS2* LRR domain sequence. Even across such diverse genera, ~90% amino acid similarity was observed between each inferred LRR domain and that of *Arabidopsis* Col-0 FLS2 (see Supplemental Figure 5 and Supplemental Table 4 online). However, taken as a collection of sequences, significant amino acid divergence was observed across the entire LRR domain, leaving universal amino acid identity at only 56%.

The number of individual repeats was highly conserved at 28, but the FLS2 LRR domain homolog from *Iberis amara* contained a complete additional LRR between LRRs 8 and 9 of *Arabidopsis* Col-0 FLS2.

Seedling germination tests were performed to assay the flg22 sensitivity of these diverse Brassicaceae species. *Camelina laxa*, *Chorispora tenella*, *Eruca vesicaria*, *Alyssum alyssoides*, *Lepidium aluaudii*, *Matthiola longipetala*, *Biscutella auriculata*, and *Sinapis alba* all exhibited flg22-dependent seedling growth inhibition (Figure 3). Different species showed different levels of

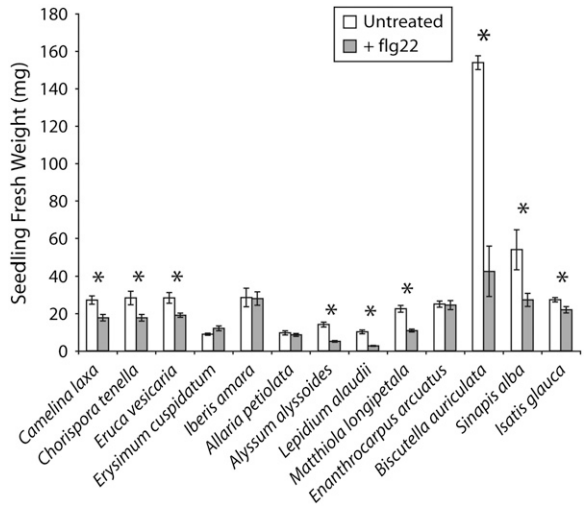

Figure 3. Identification of flg22-Responsive Brassicaceae Accessions.

Seeds of single accessions of the indicated species were germinated overnight and then grown in medium with or without 10 μM flg22 for 5 to 12 d depending on the species. Values report means ± SE for 12 or fewer plants. Accessions exhibiting significant flg22-induced growth inhibition relative to untreated control are marked with asterisks, based on *t* tests ($P < 0.05$).

growth inhibition and overall size, but for these species the observed inhibition was significant and reproducible.

Conserved Functional Group Analysis Reveals a Potential Flagellin Interaction Site

We hypothesized that regions of functional importance in flagellin recognition could be identified among diverse Brassicaceae *FLS2* LRR domains as spatially adjacent clusters of conserved β-strand/β-turn solvent-exposed residues. For this analysis, *FLS2* sequences were only used from species that exhibited an observable response to flg22 (Figures 2 and 3). To account for the prediction that residues from neighboring repeats become adjacent in the three-dimensional folded protein, a sliding-window conservation score was generated for 15- and 9-residue windows, encompassing either all five solvent-exposed β-strand/β-turn residues from three adjacent LRRs or triplet groupings of solvent-exposed β-strand/β-turn residues from three adjacent LRRs. FLS2 exhibited variable conservation across the length of the LRR domain (Figure 4A), with the highest conservation of these residues observed across LRRs 12 to 13 and 23 to 25. Notably, one of the regions where the β-strand/β-turn solvent-exposed residues are most highly conserved (LRRs 9 to 14) matched the result obtained by the separate method of functionally testing Ala-scanning mutagenesis alleles (Figure 1C). For any given set of adjacent LRR repeats, there was usually very little difference in the conservation of the left, right, and middle portions of the concave face (the 9-residue windows) compared with overall conservation in that region for the 15-residue window. However, some left/right divergence was seen across LRRs 23 to 27, and within each conserved region, the left side of the concave LRR face usually showed less conservation.

The method used in Figure 4A represents a simple and general approach for assessing solvent-exposed β-strand/β-turn LRR residue function. However, this approach does not inspect other areas of the LRR domain. It also does not fully exploit the more precise predictions of spatial proximity among residues that can be made using models built upon the solved LRR tertiary structures of other proteins. Therefore, we submitted the FLS2 LRR amino acid sequence to SWISS-MODEL (Schwede et al., 2003), in which one template contained enough homology with FLS2 to allow modeling. This was the crystal structure for bean PGIP, an extracellular LRR protein involved in plant defense. The LRR of PGIP contains many fewer repeats than FLS2; hence, the FLS2 LRR was modeled in three segments. Only the portion of the FLS2 LRR domain from the end of repeat 9 to the early part of repeat 19 (FLS2 residues 298 to 525) contained enough homology for initial model construction. For FLS2 LRRs 1 to 10, a model was successfully generated after divergence between PGIP and FLS2 was artificially reduced by in silico conversion of FLS2 amino acids to those of PGIP in positions where the conserved LxxLxLxxN LRR motif residue differed but was still a hydrophobic Ile, Leu, Val, or Phe residue. After the LRR 1 to 10 model was created, the altered residues were converted back to those of FLS2 within the DeepView program (Schwede et al., 2003). The above approach still did not allow the modeling of FLS2 LRRs 20 to 28, which contain more significant alterations from the conserved LRR motif sequence. The models for the start and

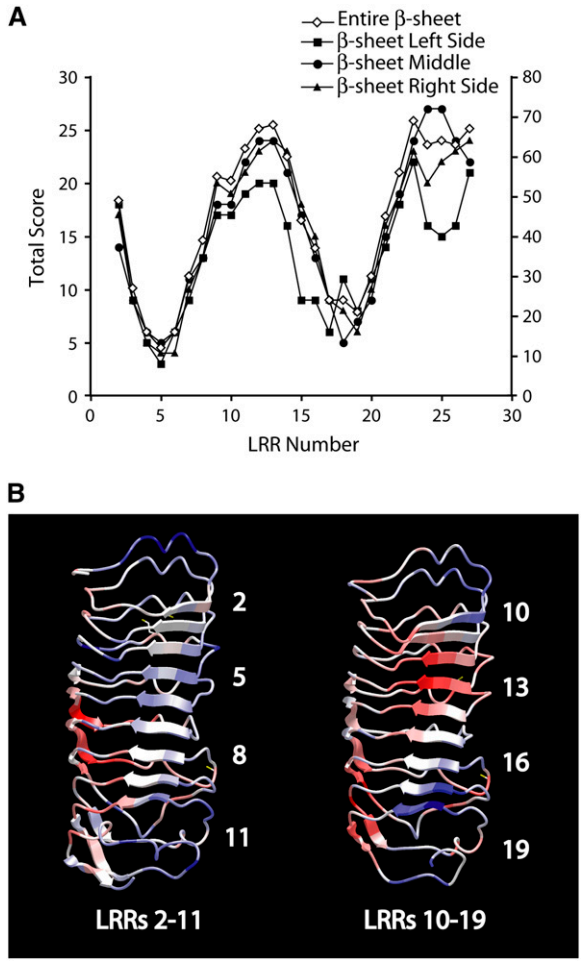

Figure 4. Two Methods to Identify Conserved Regions That Are Predicted Functional Sites in the LRR of FLS2.

(A) Location of conserved blocks of solvent-exposed β-strand/β-turn LRR residues among *FLS2* products from *Arabidopsis*, *Brassica*, and other Brassicaceae species. Composite scores for blocks of adjacent residues were calculated by summing individual residue conservation scores from a ClustalX alignment of FLS2 LRR-derived amino acid sequences from flg22-responsive accessions. Entire β-sheet indicates the conservation score for windows of 15 residues: all five solvent-exposed β-sheet residues of that repeat and the two flanking repeats (plotted against the *y* axis at right). β-sheet left, middle, and right indicate the scores for windows of 9 residues: the first, middle, or last three solvent-exposed β-sheet residues for that LRR and the two neighboring LRRs (plotted against the *y* axis at left).

(B) Location of conserved clusters of spatially proximal FLS2 LRR residues among FLS2 homologs from *Arabidopsis*, *Brassica*, and other Brassicaceae species. Three-dimensional models of repeats 2 to 11 and 10 to 19 of the FLS2 LRR were generated based on the structure of bean PGIP. CFG analysis (Innis et al., 2004) was then performed using these models and the ClustalX alignment of FLS2 LRRs from flg22-responsive accessions. Red regions represent highly conserved regions predicted to form protein functional sites, while white is intermediate and blue represents less conserved regions (note: CFG excludes aliphatic residues such as L, I, and V from the analysis). Index numbers are provided at the right of every third LRR repeat.

Both **(A)** and **(B)** suggest a possible flg22 perception region centered on LRRs 12 to 14.

middle of the FLS2 LRR domain are overall similar but do contain slight variations due to energy minimization after initial threading.

Conserved sites on the FLS2 LRR were then identified using the Conserved Functional Group (CFG) program (Innis et al., 2004) to analyze the aligned FLS2 LRR amino acid sequences from *Arabidopsis* Col-0 and the flg22-responsive Brassicaceae accessions, threaded onto the above structural model. The CFG program tests for small spatial clusters of amino acids that are overall conserved, rather than looking for single conserved amino acids. This exploits the fact that among homologous proteins, spatially clustered amino acids with an important biochemical role (i.e., receptor binding sites or enzyme active sites) tend to display a higher level of functional group conservation (amino acid identity or similarity) than any other area of the protein. For each of the two models (FLS2 LRRs 1 to 10 and LRRs 9 to 19), the CFG program produced an output file containing the hypothetical FLS2 structure with neighborhood conservation scores assigned to each amino acid. Figure 4B shows that a predicted region of high functional group conservation was identified across the β-strand/β-turn solvent-exposed residues at LRRs 12 to 14, again consistent with the functional results obtained using Ala-scan alleles (Figure 1C). Together, these results predict a candidate flagellin binding domain. Additional patches of conservation were found along the convex back side of the LRR domain at LRRs 7 to 8 and 15 to 18 (Figure 4B).

Site-Directed Randomizing Mutagenesis Reveals Residues Important for flg22 Perception

Functional sites within the β-strand/β-turn region of *Arabidopsis* FLS2 LRRs 9 to 15 were investigated further by testing multiple amino acids at each solvent-exposed position. Thirty-five separate site-directed randomizing mutagenesis libraries were generated, one for each of the five solvent-exposed β-strand/β-turn residues in LRRs 9 to 15. In each library, the amino acid present at the designated position was randomized. These libraries were numbered with the LRR number and then the position of the solvent-exposed residue in the LxxLxLxxN motif (e.g., LRR 9.3 for the library that randomizes Arg-294, the third x position in the ninth LRR repeat; Figure 1A). Unlike the constructs generated for Ala scanning, these mutant *FLS2* constructs contained a C-terminal 3xHA (hemagglutinin) tag to allow the verification of protein expression. Empty plasmid vector and wild-type *FLS2-HA* control constructs were also generated. The libraries were transformed into the *Arabidopsis fls2-101* line, and transformed T1 plants were tested for flg22 responses using the seedling growth inhibition assay.

Figure 5A presents the raw flg22 response data from a representative experiment. A total of 1535 independent *FLS2* alleles were tested. Seedling flg22 responses were calculated relative to the average weight of positive controls for that experiment date (flg22-treated plants transformed with wild-type *FLS2*), to allow subsequent analysis of data collected across separate assay dates. Data points for 109 seedlings were then omitted because the *FLS2* alleles carried a stop codon or the wild-type residue, or (for the problematic 14.1 library) if correct PCR splicing had not been confirmed. To establish appropriate cutoff values for subsequent data analysis, it was noted that 95% of the positive control plants (plants transformed with wild-type *FLS2* and tested with flg22) gave a flg22 response value of $\leq$1.66, while 95% of the negative control plants gave a flg22 response value of $\geq$2.82 (flg22-treated or nontreated plants transformed with empty vector and nontreated plants transformed with wild-type *FLS2*). Figure 5B was then generated to show for each LRR position the frequency at which amino acid changes disrupted FLS2 function. Ratios were >1.0 when more mutations caused flg22 insensitivity than allowed retention of flg22 sensitivity. Overall, 32% of the tested *FLS2* mutant alleles had lost the capacity to confer a flg22 response, but the distribution across the LRR positions was not even. Mutation of positions 3 and/or 4 in these repeats often disrupted flg22 responsiveness, while outside positions were surprisingly tolerant of random amino acid substitutions. The positions that were most sensitive to amino acid changes were 9.3, 9.4, 10.3, 10.4, 12.4, 13.4, and 14.3. These residues apparently play a key role in determining flagellin responsiveness.

For a subset of the mutant *FLS2* alleles, leaf material was collected and tested for FLS2 protein abundance using the HA epitope tag (Figure 6). When nonresponsive or very weakly flg22-responsive transgenic seedlings from a broad set of libraries (LRR positions) were tested, a strong or weak FLS2 band was detected in 29 of 32 samples. A single band of ~175 kD was detected in each sample, greater than the 132 kD predicted from the FLS2-HA sequence but consistent with the previously observed migration of FLS2 by SDS-PAGE, most likely due to extensive *N*-linked glycosylation (Chinchilla et al., 2006). The three transgenic seedlings from above that did not carry a detectable FLS2-HA band were subsequently determined by DNA sequencing to contain, respectively, no detectable transgenic *FLS2* allele, a transgenic *FLS2* allele with a stop codon, and a frame-shift mutation. This sample of 32 flg22-insensitive seedlings suggests that for the vast majority of the >400 mutant *FLS2* alleles that failed to confer flg22 responsiveness, the transgenic seedlings did carry FLS2 protein. To investigate the possibility of a threshold level below which FLS2 is not effective, we examined multiple *fls2-101* mutant lines transformed with HA-tagged wild-type *FLS2* under the control of the native *FLS2* promoter and terminator (Figure 6B). Plants from multiple transformation and testing dates were sampled. Using the same method that was used with the mutant *FLS2* alleles, all 30 plants that carried a detectable FLS2-HA band were flg22-responsive, while the 1 plant without a detectable FLS2-HA band was flg22-insensitive. Hence, it appeared that if FLS2-HA was detectable by the immunoblot method being used, the transgenic seedlings contained enough FLS2 protein that flg22 responsiveness would be observed if the *FLS2* allele encoded a protein that had not lost function.

Both to verify the successful mutagenesis of the tested constructs and to begin to characterize the chemical properties of LRR residues that are necessary for flg22 responsiveness, DNA sequencing was used to identify the mutations in a large subset of the tested seedlings. The 14.1 library contained a number of alleles carrying a splice error, but all other libraries contained the intended wide range of substitutions at the targeted codon. For further data analysis, the substitutions at each LRR position were

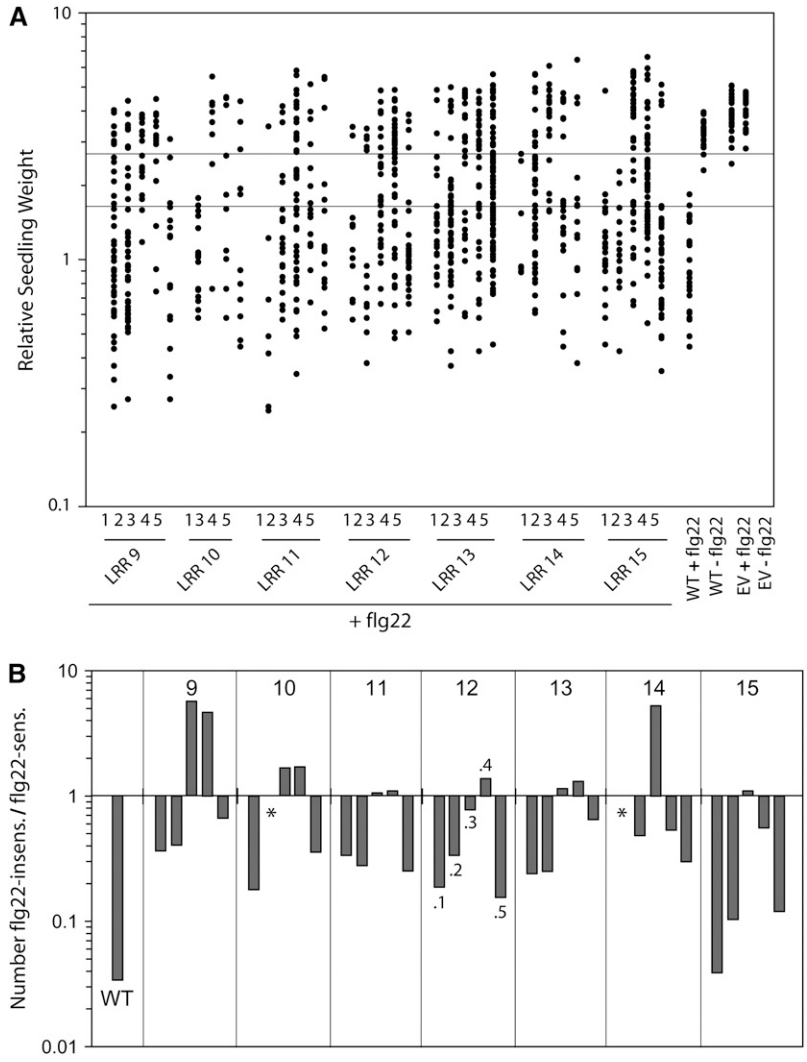

Figure 5. Mutational Analysis of Predicted Solvent-Exposed β-Strand/β-Turn Residues of FLS2 LRRs 9 to 15.

(A) Functional impact of β-strand/β-turn amino acid substitutions at 34 sites within FLS2 LRRs 9 to 15. Each data point reports the flg22 response of an independent T1 seedling of *Arabidopsis fls2-101* transformed with a mutagenized FLS2-HA construct or transformed with wild-type *FLS2-HA* (WT) or with the empty vector (EV) as controls. Columns of data points correspond to single FLS2 libraries mutagenized at the indicated β-strand/β-turn position (the *x* axis uses the numbering system of Figure 1A). The *y* axis reports the FLS2-mediated response as seedling weight after growth in 10 μM flg22, relative to the average growth in flg22 of multiple *fls2-101* T1 seedlings transformed with wild-type *FLS2-HA* and tested on the same date. Controls at far right determined 95% cutoff lines such that relative seedling weight < 1.66 = flg22-responsive and relative seedling weight > 2.82 = flg22-insensitive. Data from one representative experiment are shown.

(B) Proportion of mutants at each β-strand/β-turn residue position that showed flg22 insensitivity, reporting data from all experiments performed as in **(A)**. The *y* axis ratio is the number of T1 seedlings classified as flg22-insensitive divided by the number of T1 seedlings classified as flg22-sensitive. Each bar reports data for a single library that randomized a single β-strand/β-turn solvent-exposed position within FLS2 LRRs 9 to 15. Library index names are shown for the five LRR 12 libraries as an example. Ratios of >1 result when the number of alleles that had lost flg22 responsiveness was greater than the number that retained flg22 responsiveness. WT indicates *fls2-101* plants transformed with the wild-type *FLS2-HA* construct. Sufficient libraries were not available for positions 10.2 and 14.1 (asterisks).

grouped into one of seven categories based on predicted amino acid chemistry. Each sequenced *FLS2* allele had previously been categorized for flg22 sensitivity as in Figure 5A. These data were combined to generate Supplemental Figure 6 online, which starts to highlight the amino acid functional groups that break or maintain flg22 sensitivity for each LRR position. As with any large screening project, some of the observed data points in Figure 5A and Supplemental Figure 6 online may represent false-positives or false-negatives, but the high-throughput method provides insight into LRR structure/function that has otherwise not been available. Figure 5B reports cumulative data and is the key result from this portion of the study, identifying the small

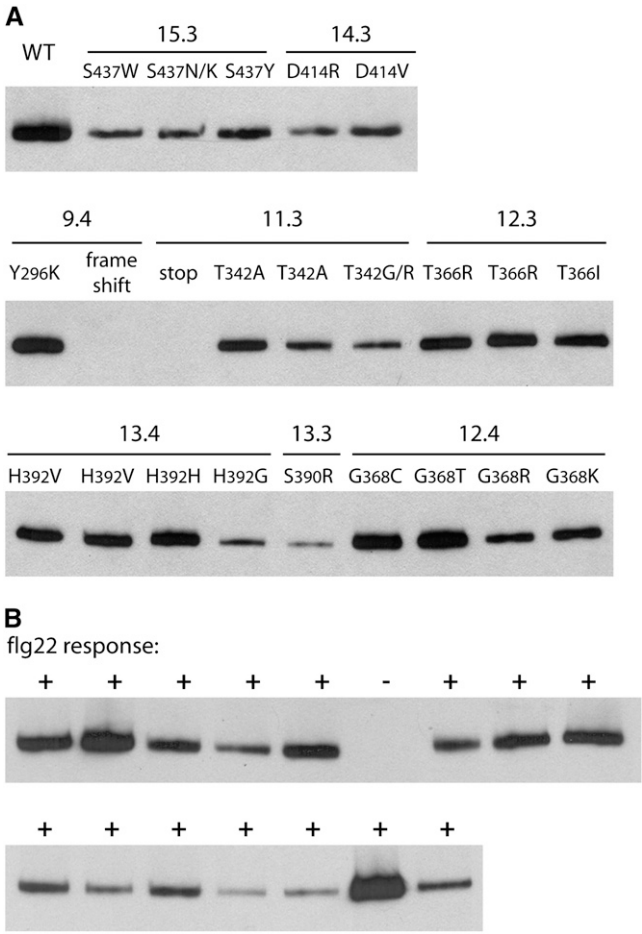

Figure 6. Mutant FLS2-HA Proteins Were Detectably Expressed in Most flg22-Insensitive Transgenic Seedlings, and Transgenic Seedlings That Expressed Wild-Type FLS2-HA Protein at These Levels Were All flg22-Responsive.

(A) Anti-HA protein gel blots of total protein extracts from *Arabidopsis fls2-101* seedlings transformed with mutant *FLS2* alleles. T1 seedlings that were flg22-insensitive and from a broad set of libraries were otherwise chosen at random for testing. Prior to protein extraction, a small portion of tissue was saved for PCR amplification of the *FLS2* transgene and subsequent DNA sequencing. Numbers above the lines indicate the mutant library, and the encoded amino acid change for that allele are listed below the lines. WT indicates the *fls2-101* plant transformed with wild-type Col-0 *FLS2*. Note that for any allele, the level of transgene-derived FLS2-HA will vary between independent transformants, as shown below.

(B) Anti-HA protein gel blots of total protein extracts from independent T1 transformants of *Arabidopsis fls2-101* transformed with wild-type Col-0 *FLS2-HA* flanked by native *FLS2* promoter and terminator. Seedlings were tested for responsiveness to flg22 prior to protein extraction. All plants that carried a detectable FLS2-HA band were flg22-responsive, while the one plant without a detectable FLS2-HA band was flg22-insensitive.

number of solvent-exposed LRR β-strand/β-turn positions at which the majority of amino acid changes cause quantitative losses of FLS2 function.

As noted above, the CFG analysis identified other regions of substantial conservation along the convex side of the LRR domain at LRRs 7 to 8 and 15 to 18 (Figure 4B), and the conservation plot analysis also revealed high conservation across LRRs 23 to 27 (Figure 4A). Multiple site-directed mutant *FLS2* alleles carrying single- and double-Ala substitutions within these highly conserved regions were tested and found to confer wild-type responses to flg22, as measured using seedling growth inhibition and ethylene production assays (see Supplemental Figure 7 online). To date, outside of the concave face of LRRs 9 to 15, we have not identified any mutations of predicted solvent-exposed LRR residues that affect FLS2 function.

Loss of flg22 Binding and Defense Activation in Identified FLS2 Alleles

A small set of alleles were chosen for more extensive testing to further characterize the flagellin-response phenotype that was initially determined with individual seedlings and to test for effects on flg22 binding. Alleles mutated at positions 12.3, 13.3, 13.4, and 14.3 were chosen for testing because these amino acid positions are at the most conserved area of LRRs 9 to 15 (Figure 4) and are highly sensitive to residue changes (Figure 5B). Transgenic Col-0 *fls2-101* mutants expressing *FLS2* alleles mutated at positions 12.3 (T366K), 13.3 (S390K), 13.4 (H392R), or 14.3 (D414K) all lacked detectable flg22 binding, despite the presence of the transgenic FLS2-HA protein in these lines (Figures 7A and 7B). Plants carrying these alleles also were confirmed in seedling growth inhibition assays to be relatively nonresponsive to flg22 (Figure 7C), lacked ethylene production in response to flg22 (Figure 7D), and failed to produce extra callose in response to flg22 (Figure 7E). A different 12.3 mutant allele (T366Q) also failed to produce flg22 binding but received less study.

Plants expressing *FLS2* alleles with mutations that alter flanking positions of the β-strand/β-turn region of LRRs 12 and 14 were tested and found to bind flg22 (Figure 7F). These were 12.1 (T363D) and 14.1 (K411R). As might be predicted, not all mutations at central positions severely disrupted function. Mutant Col-0 *fls2-101* plants expressing the 13.3 (S390A) or 13.4 (H392N) products were partially responsive to flg22 and could still bind flg22 (Figure 7F; data not shown).

DISCUSSION

LRR Study Methods

In this project, we sought to identify residues within the FLS2 LRR domain that are involved in flagellin perception, and more broadly, to develop useful methods for the study of LRR domain structure and function. While LRR domains are important in many plant proteins, most of our understanding about their structure and function comes from nonplant systems (Kobe and Kajava, 2001). Solved structures of LRR-ligand cocrystals are an ideal

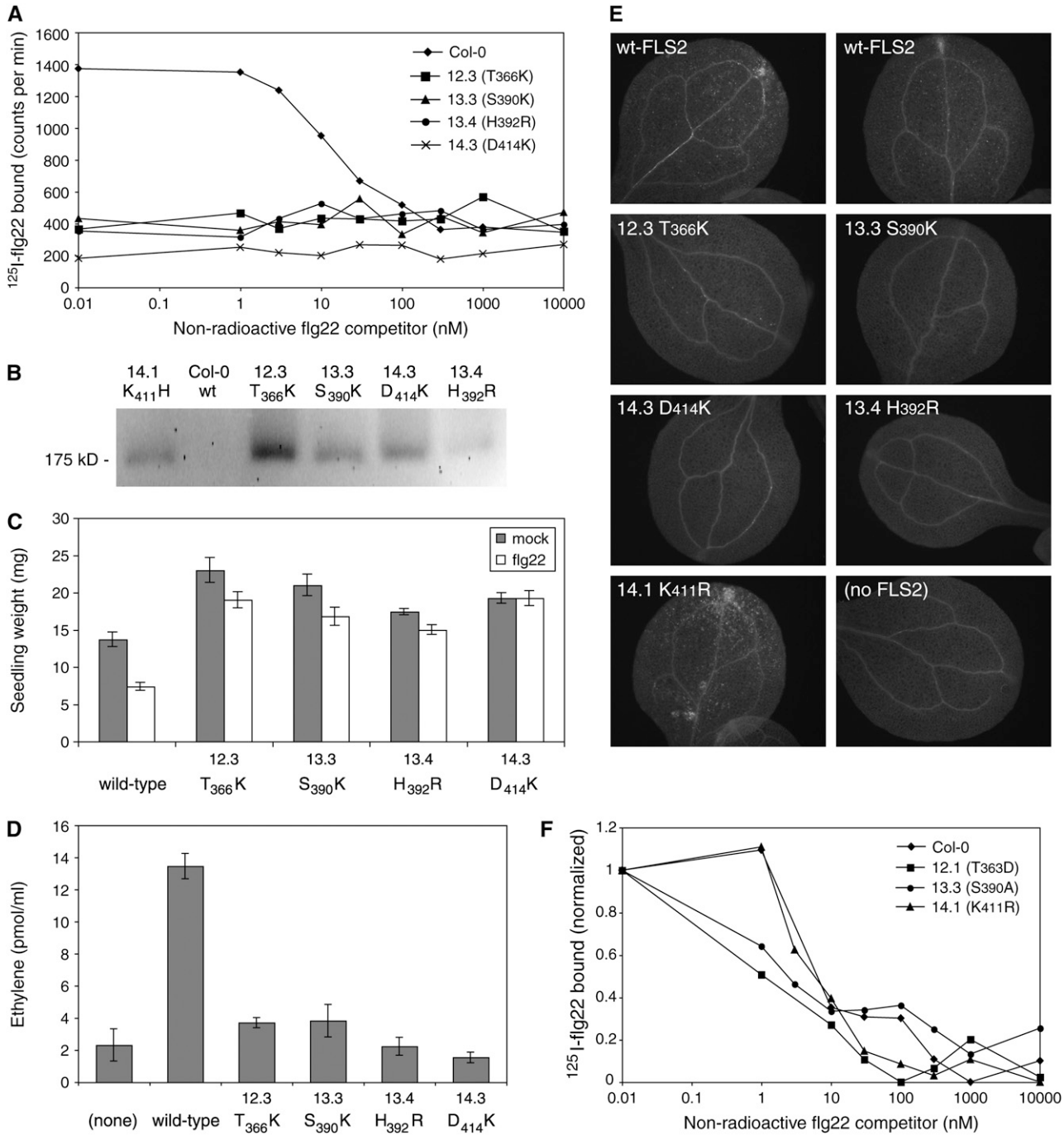

Figure 7. Mutations That Disrupt FLS2-Mediated flg22 Binding and Responsiveness.

(A) Binding of ^{125}I-flg22 in total extracts of *Arabidopsis* Col-0 or Col-0 *fls2-101* transformed with the designated loss-of-function *FLS2-HA* alleles flanked by native *FLS2* promoter and terminator.

(B) Protein gel blot of total protein from plant material of **(A)**, probed with anti-HA antibody. Note that Col-0 is wild-type Col-0 rather than a HA-tagged transgenic line.

(C) FLS2- and flg22-dependent seedling growth inhibition in transgenic *Arabidopsis* Col-0 *fls2-101* transformed with wild-type or mutant *FLS2-HA* alleles flanked by native *FLS2* promoter and terminator. Seedlings were grown in 10 μM flg22.

(D) FLS2-dependent ethylene production after 5 h of exposure to 1.0 μM flg22. Leaf strips were from six separate 5-week-old *Arabidopsis* plants (described in **[C]**).

(E) Callose deposition visualized as fluorescent spots on leaves of *Arabidopsis* seedlings (described in **[C]**) at 24 h after exposure to 3 μM flg22. Two replicate wild-type Col-0 leaves are shown for comparison.

(F) Binding of ^{125}I-flg22 in total extracts of transgenic *Arabidopsis* Col-0 *fls2-101* transformed with wild-type or mutant *FLS2-HA* alleles that still respond to flg22. Data were normalized by dividing counts per minute for each data point by the counts per minute for the maximally binding sample (nonradioactive competitor [flg22] = 0.01 nM) for that plant extract to allow comparison of experiments from separate dates.

resource, but such structures often are not readily available. Studies comparing tomato Cf-4 and Cf-9 have demonstrated the power of domain swapping and shuffling to elucidate LRR domain function (Wulff et al., 2001). Other studies of plant *R* genes have also used deletions, site-directed mutations, and domain swapping to understand function, but with less of a focus on LRR dissection (Rairdan and Moffett, 2006). Approaches have not been demonstrated that efficiently identify LRR functional sites when only a single ligand specificity is available (as opposed to swapping among related genes encoding different known ligand specificities). The crystal structures of LRR proteins, and of PGIP in particular, offer a useful framework within which plant eLRR proteins can be studied (Wulff et al., 2001; Di Matteo et al., 2003; van der Hoorn et al., 2005). Using site-directed Ala-scan mutagenesis of the LRR and a separate analysis of the conservation of spatially clustered FLS2 LRR residues across diverse Brassicaceae species, we identified a key subdomain within the FLS2 LRR. We then initiated a study of the specific amino acid positions at this site that contribute to function.

The rationale for the Ala-scan mutational survey across the FLS2 LRR was described in Results; briefly, the screen targeted solvent-exposed residues along the concave face, mutated two residues in a single repeat in each allele, and tested allele function in a native setting in stably transformed *Arabidopsis*. This strategy allows for relatively efficient screening across very large LRR domains, such as in FLS2. A small set of 28 alleles was sufficient to systematically survey all of the large FLS2 LRRs. Most LRR proteins would require a smaller set of alleles. As long as the appropriate biological stimulus is known, the above type of screen can be done without access to a known purified ligand or biochemical binding assays.

The screen to dissect evolutionary conservation exploited an underutilized resource: derived amino acid sequences of orthologous proteins that retain similar function despite sequence divergence. Genes with sufficient divergence were required, pushing us beyond *Arabidopsis* and *Brassica* species and into broader taxa. Very recently, a comparison of *Arabidopsis* FLS2 with the newly reported FLS2 homologs from tomato and poplar pointed to the same β-strand/β-turn region, centered on LRRs 10 and 11, as the most conserved cluster of solvent-exposed LRR residues (A. Bent and P. Bittel, unpublished data). For this study, it was necessary to PCR-amplify and study only the LRR-encoding segments of the relevant genes. The CFG program was then used to identify the most highly conserved spatially clustered sets of residues. A large proof-of-concept study using known proteins had previously found that the CFG program fully or partially predicted correct functional site locations (receptor binding sites or enzyme active sites) in 96% of 470 test cases (Innis et al., 2004). One key advantage of CFG over other programs that identify conserved residue clusters is that CFG omits hydrophobic residues from the analysis. This largely removes from study those buried residue clusters that may be conserved because of contributions to overall protein shape and instead focuses the analysis on the solvent-accessible residues that tend to be sites of interaction with ligands and partner proteins. Omission of conserved, buried hydrophobic residues is particularly appropriate for the study of LRRs.

We confirmed the function of *Brassica* FLS2 LRRs by splicing them in place of the original LRR-encoding domain in *Arabidopsis FLS2* and testing in *Arabidopsis*, and for other Brassicaceae species we confirmed that the FLS2-homologous LRR under study was from a plant line that was flg22-responsive. This precaution may or may not have been necessary. Computer programs such as CFG do not require functional identity among the tested sequences for successful identification of enzyme active sites: the data set only needs to be drawn from structurally similar enzymes or receptors (Innis et al., 2004). However, it may be important to seek functionally identical genes when studying a more generic protein fold such as an LRR, as conserved regions of diversified LRRs that recognize different ligands might only identify LRR sites for receptor complex formation and not for ligand recognition. Conversely, CFG still might identify LRR ligand binding sites among functionally distinct proteins if the divergent ligands are structurally related (as is now predicted for *Arabidopsis RPP13* and for the *L*/avr flax-flax rust system used by H. Flor to first formulate the gene-for-gene hypothesis [Rose et al., 2004; Dodds et al., 2006]). These remain open questions for future investigation.

Flagellin Binding Site?

It was recently shown that FLS2 directly binds flg22 (Chinchilla et al., 2006), but the flg22 binding site within FLS2 has not otherwise been determined. The extracellular portion of FLS2 is 811 amino acids in length, consisting of a 675–amino acid eLRR domain as well as flanking segments on either side of the LRR, any of which may form the flagellin binding site (Gomez-Gomez and Boller, 2000; van der Hoorn et al., 2005). Our results here reveal a solvent-exposed site in FLS2 that is extensively conserved across genera and identify mutations within that site for which the FLS2 protein remains present yet flg22 binding is disrupted. Double-Ala substitutions across this site also affected function, while double-Ala mutations at analogous positions in other repeats of the LRR domain did not detectably disrupt function. Based on the above, this site is the likely flg22 binding site.

Firm confirmation that this is the flg22 binding site would require ligand-receptor cross-linking/protein fragmentation/mass spectrometry identification or a cocrystal structure of FLS2 and flg22, but additional molecular genetic data can support the hypothesis. For example, mutation of other conserved sites has not yet been observed to disrupt function. Alleles carrying Ala substitutions at conserved regions along the convex side of LRRs 7 to 8 or 15 to 18 (Figure 4B) or the concave face of LRRs 23 to 27 (Figure 4A) all retained flg22 responsiveness similar to that of wild-type FLS2 (Figure 1C; see Supplemental Figure 7 online). This suggests that these other sites have less of an impact on flg22 binding. Furthermore, the broader set of mutational data reported in this study suggests that mutations in solvent-exposed residues within the β-strand/β-turn and other conserved regions of FLS2 do not generically disrupt FLS2 function. It remains possible that the LRR 12 to 14 alleles described in Figure 7 alter flg22 binding only indirectly, due to gross perturbation of FLS2 structure or to changes in FLS2 positioning relative to associated proteins in a postulated receptor complex. However, an equally likely hypothesis is that the β-strand/β-turn

region centered on LRRs 12 to 14 forms the flagellin binding site of FLS2, especially given that β-strand/β-turn subdomains are known to contribute to ligand binding and/or receptor specificity in many other LRR proteins. The Ala-scan and conserved site methods that we used in this study may be of widespread utility in identifying ligand binding sites and other functional domains in a diverse range of LRR proteins.

Adjacent Mutations Suggest a Larger Site, and Conservation Reveals Multiple Sites

Residues across the entire 28 repeats of the LRR domain could contribute directly to flagellin binding. However, the absence of any obvious non-LRR hinge region within the FLS2 LRRs, and the fact that the 22–amino acid flg22 peptide carries full elicitation activity, suggest that a smaller set of nearby LRR residues determine flagellin binding. Computer docking programs using homology-based models of FLS2 and flg22 suggest that flg22 could contact three to seven adjacent repeats, depending on how it lies across the concave inner face of the LRR domain (data not shown).

Consistent with this, our data suggest that contributions to flg22 binding arise across a limited set of multiple adjacent repeats of the LRR. Most double-Ala LRR substitutions produced FLS2 receptors that confer a flg22 response similar to that of the wild type. Even the double-Ala alleles for LRRs 9 to 15 that affected the response caused only partial reduction in FLS2 function, and partial responsiveness was common with the randomizing mutagenesis alleles as well (Figures 1C and 5). A very small amount of seedling growth inhibition was observed with three of the nonbinding mutants in Figure 7 after stimulation with a very high dose of flg22 (Figure 7C). This suggests that they may retain slight residual or low-affinity binding of flg22 that is below the detection of the binding assay, again supporting the idea that multiple contact points contribute to binding site formation. LRRs are highly evolvable protein interaction surfaces that can be altered to recognize a wide array of ligands (Pancer et al., 2004), but to the extent that the LRR consensus is maintained, they form these diverse binding sites on a relatively generic β-sheet platform. More so than for many typical enzyme or receptor active sites, LRR/ligand specificity may rely on additive interactions at many amino acids that each contribute a relatively small amount to overall specificity. The partial loss of activity that we observed with many alleles demonstrates in a different way an observation already made by Felix et al. (1999), that a range of slightly different flagellins can be recognized by any single *FLS2* allele, a situation of adaptive benefit to the plant. It also suggests that shifts in ligand specificity across natural or synthetic *FLS2* alleles in many cases might be quantitative and gradual rather than qualitative.

As noted above, the diversity survey of Brassicaceae FLS2 homolog LRR domains identified two conserved β-strand/β-turn regions, centered at LRRs 12 to 13 and 23 to 24, yet the double-Ala mutant alleles revealed a depletion of in planta function only for the first region. We hypothesize that some functional contribution has driven the evolutionary conservation of residues along the concave face of FLS2 around LRRs 23 to 24, but any such function remains undocumented at present.

This study concentrated on putative solvent-exposed β-strand/β-turn residues along the convex LRR face because of their predicted role in ligand specificity. However, the CFG analysis also highlighted regions of conserved functional groups on the outer convex face of the FLS2 LRR domain. While testing of site-directed mutant alleles for many of these conserved residues has not yet revealed effects on FLS2 function, the conservation at these other sites does suggest a functional role, perhaps as flagellin interaction sites or as residues that drive intramolecular FLS2 structure or intermolecular receptor complex formation. For instance, some of these other conserved regions are located in or around putative *N*-linked glycosylation motifs along the outer face of the LRR domain (Leconte et al., 1994; van der Hoorn et al., 2005).

Residues and Chemistries That Determine flg22 Perception

After identifying specific LRRs involved in flg22 perception, we used site-directed randomizing mutagenesis on the exposed β-strand/β-turn residues of LRRs 9 to 15 to assay their contribution to flg22 perception and investigate the range of chemistries that could be present without breaking flg22 perception. It was not surprising that severe mutations, for example to Pro, tended to break function at any position. However, the mutagenesis revealed tight constraints on only a minority of the studied positions: approximately two-thirds of the 35 adjacent positions tolerated most mutations with little or no functional consequence. The residues on the outer edges of each β-sheet (x.1, x.2, or x.5) were as a trend less functionally constrained than those on the two inner positions of the convex LRR face (Figure 5B). This result may hint at the orientation and position of flg22 during docking, or it may generically reflect the more restricted spatial constraints along the tightest curve of the LRR exterior, and/or it may be attributable to other unknown causes. If positions x.3 and x.4 form the core of the flg22 binding site, then in future work, complementary chemistries should be found at the corresponding sites on the bound ligand.

Variation in MAMP Perception and Other LRR-Based Perception Systems

The FLS2 LRR domain is highly conserved throughout *Arabidopsis* ecotypes, with the only major variation seen in alleles that contain premature stop codons. This high conservation suggests an important role for flagellin-induced basal immunity in overall *Arabidopsis* defense responses. Many of the *B. oleracea*, *B. rapa*, and diverse Brassicaceae accessions that were tested also respond to flg22 and have a functional FLS2 LRR domain. The presence of an *FLS2* ortholog in tomato (Felix et al., 1999; Robatzek et al., 2007) and the presence of FLS2 homologs in the genomes of poplar and rice imply that flagellin perception is adaptive in these species. The presence of variable (defense-eliciting or noneliciting) flagellins in plant-associated bacteria, including variation among different strains of the Brassicaceae pathogen *Xanthomonas campestris* pv *campestris*, further suggests that there is selective pressure for bacteria to evade flagellin detection (Felix et al., 1999; Pfund et al., 2004; Sun et al., 2006). However, it is interesting that *Arabidopsis* ecotypes such

as Ws-0, Dra-0, and Po-0 contain nonfunctional *FLS2* alleles yet appear reasonably immunocompetent. Other Brassicaceae accessions were also identified that were unresponsive to flg22. These results argue that flagellin perception is important, but not required, for effective plant defense responses. As research proceeds on the mechanisms of flagellin perception, basal defense activation, and bacterial suppression of basal defenses, it will be of increasing interest to study other MAMP perception systems, such as the Brassicaceae-specific systems that detect bacterial EF-Tu (Zipfel and Felix, 2005; Zipfel et al., 2006). It will be relevant to assess the varying roles that the detection of flagellin and the detection of other MAMPs play in overall plant resistance/bacterial virulence.

This study identified exposed β-strand/β-turn residues of FLS2 that have functional roles in flagellin perception. This project also demonstrated ways in which structural modeling, site-directed mutagenesis, and evolutionary conservation studies can be combined to dissect structure/function relationships within LRR proteins. This type of work may highlight LRR residues to target for in vitro evolution, with the goal of developing modified LRR proteins with new ligand specificities. The approaches taken in this study should be applicable to research on plant disease resistance, on the numerous other LRR-RLK receptors that occur in plants, and on many other types of LRR-containing proteins.

METHODS

Construction of pHD3300 and pHD3300HA

The vector pHD3300 is based on pCAMBIA3300 (http://www.cambia.org) and contains 1 kb of native *Arabidopsis thaliana* Col-0 *FLS2* promoter, the N-terminal coding region of *FLS2* from the start codon to codon 60, an engineered linker flanked by *Asc*I and *Pac*I with the sequence 5′-GGCGCGCCTTCTGCAGTTAATTAA-3′, the C-terminal coding region of *FLS2* from codon 771 to the stop codon, and 505 bp of native *FLS2* terminator. To engineer the *Asc*I and *Pac*I sites, the Leu at amino acid 58 of FLS2 was changed to an Ala and the Asn at amino acid 781 was changed to a Leu. The vector pHD3300HA is identical to pHD3300 except that the *FLS2* stop codon was replaced with an in-frame 3xHA tag and stop codon. The 3xHA tag was amplified from pGWB14 (courtesy T. Nakagawa, Shimane University, Matsue, Japan) and inserted using PCR splice overlap extension. Site-directed mutagenesis was used to destroy the *Asc*I and *Pac*I sites located within the pGWB14 3xHA tag (GGCGCGCC converted to GGCGAGCC and TTAATTAA converted to TTAATCAA). Unless noted otherwise, standard molecular biology protocols were used here and throughout (Ausubel et al., 1997). DNA sequencing reactions were performed using Big Dye Terminator cycle sequencing (ABI) and analyzed by capillary electrophoresis at the University of Wisconsin–Madison Biotechnology Center (http://www.biotech.wisc.edu).

Mutagenesis of the FLS2 LRR

The wild-type *FLS2* LRR-encoding domain was amplified by PCR using Pfu turbo (Stratagene) with the primers 5′-GGCGCGCCACTGTAATTGGACCGGAATCA-3′ and 5′-TTAATTAATTTGAACACCCCGGATTCAGG-3′ containing engineered *Asc*I and *Pac*I sites. Gel-purified DNA was then cloned into pCR-Blunt II-TOPO using the Zero Blunt TOPO PCR cloning kit (Invitrogen) to yield pTOPO:FLS2LRR.

For the Ala-scanning mutagenesis, inverse PCR was performed with Pfu turbo using 21- to 25-mer primers designed such that they abutted each other at the 5′ end but on opposite strands and primed DNA synthesis outward from this site. Ala replacements were inserted by changing wild-type codons to AGC in the primer sequence. Primers were positioned such that each contained a single Ala mutation with the altered sequence positioned close to the 5′ end; this method was successful for all alleles except LRR 10. Mutant LRRs were generated from pTOPO:FLS2LRR template using Pfu turbo with the following program: 95°C for 5 min; 40 cycles of 93°C for 30 s, 56°C for 1 min, and 72°C for 10 min; 72°C for 10 min; hold at 10°C. The major vector-sized product was purified by agarose gel electrophoresis, excised and extracted using the QIAquick gel extraction kit (Qiagen), treated with T4 polynucleotide kinase (New England Biolabs), ligated together, and transformed into TOP10 chemically competent cells (Invitrogen). Purified plasmid DNA was extracted using the Wizard Plus SV Miniprep kit (Promega), sequenced to verify the mutations, and digested with *Asc*I and *Pac*I (New England Biolabs) to excise the mutant DNA. Mutant LRR-encoding fragments were gel-purified, cloned into pHD3300, and transformed into *Agrobacterium tumefaciens* strain GV3101 (pMP90) for plant transformation.

For site-directed randomizing mutagenesis, 35-mer oligonucleotide primers were designed such that they pointed in opposite directions with the 5′ ends complementary across 18 nucleotides. In each forward primer, the codon immediately after the overlapped region was aligned with the desired mutation site and changed to the degenerate NNB codon (where N = A, T, C, or G; B = A, T, or C) (Neylon, 2004). Mutant LRRs were generated by PCR from a pTOPO:FLS2LRR template using Pfu turbo with the same temperature cycling as above. Ten units of *Dpn*I (New England Biolabs) were then added, and after 1 h at 37°C, amplified products were run on an agarose gel and the major vector-sized product was excised and extracted. Extracted DNA was denatured, annealed, and extended in the following conditions: 21.5 μL of DNA, 2.5 μL of 10× Pfu buffer, 0.5 μL of 10 mM deoxynucleotide triphosphate, and 0.5 μL of Pfu turbo with the following program: 94°C for 2 min; 490 intervals of 4 s initially at 94°C, then decreasing by 0.1°C at each interval; 270 intervals of 2 s initially at 45°C, then increasing by 0.1°C at each cycle; 72°C for 20 min; hold at 10°C. An aliquot of this reaction was then transformed into electrocompetent TOP10 cells (Invitrogen). Transformation selection plates typically carried several thousand colonies, after which all clones for a single LRR were pooled in 5 mL of Luria-Bertani medium. Multiple attempts at mutagenesis of the 10.2 position were not successful. For the other 34 libraries, purified plasmid DNA was isolated from 3 mL of the bacterial suspension and digested with *Asc*I and *Pac*I to excise the saturated mutagenesis LRR clones, gel-purified, bulk-cloned into pHD3300HA, and transformed into TOP10 cells. Transformation selection plates typically carried several thousand colonies, after which all clones for a single LRR were pooled in 5 mL of Luria-Bertani medium. Purified plasmid DNA was isolated from 3 mL of the bacterial suspension, and an aliquot was transformed into *Agrobacterium* GV3101 (pMP90) for plant transformation.

Identification and Cloning of *Brassica* FLS2 LRR Homologs

The *Arabidopsis FLS2* LRR-encoding domain was nominated to the BBSRC *Brassica* Consortium (http://www.brassica.bbsrc.ac.uk/IGF/) as a BAC probe; the Consortium chose to probe its *Brassica rapa* and *Brassica oleracea* BAC genomic libraries with a segment of the *FLS2* kinase-encoding domain and identified 23 and 32 strongly hybridizing clones from the respective libraries, likely to represent multiple separate *FLS2* homologs as well as genes for other LRR-RLK proteins. Primers designed to *Arabidopsis* Col-0 *FLS2* were used for PCR with BACs o33m6, r97L9, r46D20, r41K3, r17I6, r79B6, o9P24, o3J23, o79F22, and r18O2. The primers corresponded to the first 20 and last 19 bases of FLS2, 5′-AAAAGTCGACATGAAGTTACTCTCAAAGAC-3′ and 5′-AAGGATCCCTAAACTTCTCGATCCTCG-3′, and also contained engineered *Sal*I and *Bam*HI restriction sites. Three putative *FLS2* homologs were

amplified from BACs o33m6, r97L9, and r46D20 and cloned into pCR-BluntII-TOPO using the Zero Blunt TOPO PCR cloning kit (Invitrogen). The product from r97L9, named bFLS2, was fully sequenced using primer walking and aligned to *Arabidopsis* FLS2 using ClustalX. Primers to amplify the region of *Brassica FLS2* corresponding to the *Arabidopsis FLS2* LRR-encoding domain with engineered *Asc*I and *Pac*I sites for insertion into pHD3300 were designed with the sequences 5′-GGC-GCGCCACTGCAGTTGGACCGGAATC-3′ and 5′-TTAATTAATTTGAA-CACACCGGACTCAGG-3′. These primers were then used to amplify bFLS2 LRR-encoding domains from genomic DNA templates from *B. oleracea* accessions BI106, KT18, and IMB218DH and *B. rapa* accessions TO434, TO450, TO486, TO441, and TO980 (provided courtesy of Thomas Osborn). Products were cloned into pCR-BluntII-TOPO and fully sequenced, and the derived amino acid sequences were aligned with the *Arabidopsis* FLS2 LRR domain. The extra *Iberis amara* LRR was deleted from the alignment prior to amino acid identity scoring. bFLS2 LRR-encoding domains for KT18, TO1441, TO1434, and TO1450 were excised with *Asc*I and *Pac*I, cloned into pHD3300, and transferred into *Agrobacterium* GV3101.

Transformation and Testing of Mutant *FLS2* Constructs in *Arabidopsis*

For most *FLS2* constructs, individual *Agrobacterium* clones were transformed into *Arabidopsis* Col-0 *fls2-101/fls2-101* (Pfund et al., 2004) by floral dip transformation (Clough and Bent, 1998). For the randomizing mutagenesis libraries, freshly transformed *Agrobacterium* (several thousand colonies) or bacteria from −80°C stocks made directly from initial *Agrobacterium* transformation plates with no further growth were resuspended and used for transformation. To select transformant plants, T1 seeds for the Ala-scanning mutants was surface-sterilized and plated on 0.8% agar plates carrying 0.5× Murashige and Skoog (MS) basal salts, 2% sucrose, 1 mL/L 1000× MS vitamin solution (Sigma-Aldrich), and 200 mg/L cefotaxime. Plates were kept at 4°C for 2 d and then grown in 9 h of light per day at 23°C and, after 5 d of initial growth, sprayed every other day for 7 d with sterile 200 mg/L BASTA (phosphinothricin; Liberty; AgrEvo). For the randomizing mutagenesis pools, a simpler selection was used in which surface-sterilized T1 seeds were plated on 0.8% agar/0.5× MS plates with 1 mL/L vitamins, 200 mg/L cefotaxime, and 10 mg/L BASTA. Plates were kept at 4°C for 2 d and then grown in 16 h of light per day at 23°C for 1 week. In each case, healthy green seedlings were then used for flg22 response testing in the seedling growth inhibition assay. DNA sequences of randomizing mutagenesis *FLS2* alleles were obtained by direct sequencing of PCR products amplified from T1 seedling tissue.

Seedling growth inhibition flg22 response assays were performed as described (Gomez-Gomez and Boller, 2000; Pfund et al., 2004) with BASTA-resistant seedlings grown in 400 μL of liquid 0.5× MS with 10 μM flg22 or sterile dH_2O for 12 to 14 d, briefly blotted dry, and then weighed. Typically, 10 or more independent T1 seedlings per genotype were tested except in the randomizing mutagenesis, for which each T1 seedling represents a unique genotype.

Ethylene production in response to flg22 (Felix et al., 1999) was monitored using a Shimadzu GC-14A gas chromatograph, a C-R4A Chromatopac, and an aluminum oxide column. Multiple transverse leaf strips (~1 mm wide) were prepared from one leaf from each of six to eight 4- to 6-week-old plants of a given genotype (selected for phosphinothricin resistance as seedlings), and leaf strips were pooled together by genotype and floated on a water surface for 14 h. For each replicate, six strips chosen at random were then floated together in 0.5 mL of water in a 13- × 100-mm tube and flg22 was added immediately prior to gas-tight capping of tubes. The tubes were rocked gently for 5.0 h, and then the ethylene concentration in the air space was determined.

Callose deposition in response to flg22 (Felix et al., 1999) was monitored in approximately six *Arabidopsis* seedlings per treatment, grown on 0.5× MS, 2% (w/v) sucrose, 1× Gamborg's vitamins, 25 mg/mL phosphinothricin, and 0.6% agar medium for 7 d, and then transferred to 24-well plates (one seedling per well) containing 400 μL of liquid medium (no agar) with flg22. At 24 h after treatment, seedlings were fixed overnight in 1% (v/v) glutaraldehyde, 5 mM citric acid, and 90 mM Na_2HPO_4, pH 7.4, and then cleared and dehydrated with 100% ethanol. Callose was visualized using UV epifluorescence microscopy as described (Gomez-Gomez et al., 1999).

Brassica flg22 Response Assays

The above seedling growth inhibition assay was poorly reproducible with most Brassicaceae accessions, apparently because pregerminated seedlings rapidly grew out of the flg22 solution. In an alternative but also less successful method, young seedlings were transferred inverted into 18- × 150-mm test tubes containing 5 mL of 0.5× MS medium with or without 10 μM flg22 (cotyledons immersed), vacuum-infiltrated, and then seedlings were placed upright, grown for 7 d, and weighed. Instead, reproducible data were obtained when seeds were added, immediately after surface sterilization, to 18- × 150-mm test tubes containing 3 mL of 0.5× MS and 0.6% agar with or without 10 μM flg22. The seeds were grown on the surface of the agar at 23°C, their ability to germinate was scored, and their growth was measured.

Alignment, Conservation Score, and CFG Analysis of Brassicaceae FLS2 LRR Homologs

Derived amino acid sequences of FLS2 LRRs from flg22-responsive accessions were aligned using ClustalX (Thompson et al., 1997). To calculate the conservation score in Figure 4A, a subscore for the three triplet groupings (left, right, and middle) and the total for all five solvent-exposed β-sheet residues of each LRR were computed by adding scores for the amino acid similarity for these positions, derived from the ClustalX alignment (* = 3; : = 2; . = 1; blank = 0 points). The total score for an LRR is the sum of the subscores for that LRR plus the two adjacent LRRs.

Homology-based structural modeling of the FLS2 LRR domain was performed using the SWISS-MODEL (http://www.expasy.org/spdbv/) (Schwede et al., 2003) first-approach method based on the *Phaseolus vulgaris* PGIP template, 1OGQA, as template (Di Matteo et al., 2003). This model was viewed and further refined using the DeepView program. Only the portion of the FLS2 LRR domain from residues 298 to 525 showed sufficient homology for initial modeling, corresponding to the region from the end of repeat 9 to the beginning of repeat 19. Modeling of LRRs 1 to 10 is described in Results. Coordinate data were submitted to the CFG (Innis et al., 2004) program with an inclusion cutoff of 0.9. Structural models and functional group prediction were analyzed using DeepView (Schwede et al., 2003).

Protein Extraction and Protein Gel Blotting

Protein was extracted from *Arabidopsis* seedlings as described by Kim et al. (2005). Protein concentrations were determined using the BCA protein assay (Pierce), and 100 μg of total protein in each sample was resolved on a 7.5% SDS-PAGE gel and transferred to Hybond-P (GE Healthcare). Protein gel blotting was performed using standard procedures. Anti-HA antibody (Covance) was used at a 1:1000 dilution.

^{125}I-flg22 Binding

Transgenic T2 seedlings were selected on 0.5× MS salts (Sigma-Aldrich product No. M-5519), 1% sucrose, 0.5 g/L MES, pH 5.8, 0.8% agar, and

25 mg/mL phosphinothricin, and then 16 to 20 seedlings were grown in liquid medium (no agar or phosphinothricin) at 120 rpm for 10 d and harvested by brief blotting, immediate freezing in liquid N_2, and storage at –80°C. Binding of ^{125}I-flg22 was monitored at 4°C using ground total tissue extracts as described (Bauer et al., 2001). Aliquots of the same extracts were stored at –20°C and then used for protein gel blotting, loading equal total protein per SDS-PAGE lane.

Accession Numbers

The Arabidopsis Genome Initiative number for FLS2 is At5g46330.1. DNA sequence data for the Brassicaceae *FLS2* genes can be found in the GenBank data library under accession numbers EU189456 to EU189483.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. ClustalX Alignment of *FLS2* Alleles from 11 *Arabidopsis* Ecotypes.

Supplemental Figure 2. Responsiveness of 23 *Arabidopsis* Ecotypes to flg22, Tested Using the Seedling Growth Response Assay.

Supplemental Figure 3. ClustalX Alignment of Full-Length *Arabidopsis* FLS2 with the *Brassica* FLS2 Homolog Isolated from BAC r97L9 Reveals 80.1% Sequence Identity.

Supplemental Figure 4. ClustalX Alignment of FLS2 LRR Domains from *Arabidopsis* and Five *B. rapa* and *B. oleracea* Accessions.

Supplemental Figure 5. ClustalX Alignment of FLS2 LRR Domains from Diverse Brassicaceae Genera Revealing ~90% Amino Acid Identity between the *Arabidopsis* FLS2 LRR and Any Single Other Sequence but ~55% Amino Acid Identity within the Entire Set.

Supplemental Figure 6. Amino Acid Changes That Do or Do Not Disrupt FLS2 Function, Determined by Sequencing of a Subset of the *FLS2* Alleles That Had Been Tested for the Ability to Confer flg22 Response after Transformation into *Arabidopsis fls2-101*.

Supplemental Figure 7. Functional Tests of *FLS2* Alleles Carrying Ala Mutations at Other Conserved Sites of the FLS2 LRR Domain.

Supplemental Table 1. Alignment Data from Supplemental Figure 1.

Supplemental Table 2. Alignment Data from Supplemental Figure 3.

Supplemental Table 3. Alignment Data from Supplemental Figure 4.

Supplemental Table 4. Alignment Data from Supplemental Figure 5.

ACKNOWLEDGMENTS

We thank Christine Pfund for her important role in the early formation of this project; Adedayo Fashoyin, Marie-Stella Essilfie, and James Murphy for their experimental contributions; the BBSRC *Brassica* Consortium for BAC clones; the USDA-Agricultural Research Service North Central Regional Plant Introduction Center for diverse Brassicaceae accessions; the ABRC for *Arabidopsis* accessions; Rick Amasino, Patrick Masson, Patrick Krysan, Thomas Boller, and Delphine Chinchilla for suggestions during the course of the work; and Thomas Boller for hosting A. Bent for sabbatical research. This research was supported by U.S. Department of Energy Energy Biosciences.

Received November 4, 2006; revised September 13, 2007; accepted September 19, 2007; published October 12, 2007.

REFERENCES

Asai, T., Tena, G., Plotnikova, J., Willmann, M.R., Chiu, W.L., Gomez-Gomez, L., Boller, T., Ausubel, F.M., and Sheen, J. (2002). MAP kinase signalling cascade in Arabidopsis innate immunity. Nature **415:** 977–983.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1997). Current Protocols in Molecular Biology. (New York: John Wiley & Sons).

Bauer, Z., Gomez-Gomez, L., Boller, T., and Felix, G. (2001). Sensitivity of different ecotypes and mutants of *Arabidopsis thaliana* toward the bacterial elicitor flagellin correlates with the presence of receptor-binding sites. J. Biol. Chem. **276:** 45669–45676.

Belkhadir, Y., and Chory, J. (2006). Brassinosteroid signaling: A paradigm for steroid hormone signaling from the cell surface. Science **314:** 1410–1411.

Bell, J.K., Mullen, G.E., Leifer, C.A., Mazzoni, A., Davies, D.R., and Segal, D.M. (2003). Leucine-rich repeats and pathogen recognition in Toll-like receptors. Trends Immunol. **24:** 528–533.

Bent, A.F., and Mackey, D. (2007). Elicitors, effectors, and R genes: The new paradigm and a lifetime supply of questions. Annu. Rev. Phytopathol. **45:** 399–436.

Binz, H.K., Amstutz, P., and Pluckthun, A. (2005). Engineering novel binding proteins from nonimmunoglobulin domains. Nat. Biotechnol. **23:** 1257–1268.

Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The Arabidopsis receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell **18:** 465–476.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nurnberger, T., Jones, J.D., Felix, G., and Boller, T. (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. Nature **448:** 497–500.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

Dangl, J.L., and Jones, J.D. (2001). Plant pathogens and integrated defense responses to infection. Nature **411:** 826–833.

de Torres, M., Mansfield, J.W., Grabov, N., Brown, I.R., Ammouneh, H., Tsiamis, G., Forsyth, A., Robatzek, S., Grant, M., and Boch, J. (2006). *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in Arabidopsis. Plant J. **47:** 368–382.

Di Matteo, A., Federici, L., Mattei, B., Salvi, G., Johnson, K.A., Savino, C., De Lorenzo, G., Tsernoglou, D., and Cervone, F. (2003). The crystal structure of polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. Proc. Natl. Acad. Sci. USA **100:** 10124–10128.

Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Teh, T., Wang, C.I., Ayliffe, M.A., Kobe, B., and Ellis, J.G. (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proc. Natl. Acad. Sci. USA **103:** 8888–8893.

Ellis, J., Dodds, P., and Pryor, T. (2000). Structure, function and evolution of plant disease resistance genes. Curr. Opin. Plant Biol. **3:** 278–284.

Ellis, J.G., Lawrence, G.J., Luck, J.E., and Dodds, P.N. (1999). Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. Plant Cell **11:** 495–506.

Felix, G., Duran, J.D., Volko, S., and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J. **18:** 265–276.

Fritz-Laylin, L.K., Krishnamurthy, N., Tor, M., Sjolander, K.V., and Jones, J.D. (2005). Phylogenomic analysis of the receptor-like proteins of rice and Arabidopsis. Plant Physiol. **138:** 611–623.

Gomez-Gomez, L., Bauer, Z., and Boller, T. (2001). Both the extracellular leucine-rich repeat domain and the kinase activity of FSL2 are required for flagellin binding and signaling in Arabidopsis. Plant Cell **13:** 1155–1163.

Gomez-Gomez, L., and Boller, T. (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. Mol. Cell **5:** 1003–1011.

Gomez-Gomez, L., Felix, G., and Boller, T. (1999). A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. Plant J. **18:** 277–284.

Heese, A., Hann, D.R., Gimenez-Ibanez, S., Jones, A.M., He, K., Li, J., Schroeder, J.I., Peck, S.C., and Rathjen, J.P. (2007). The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. Proc. Natl. Acad. Sci. USA **104:** 12217–12222.

Innis, C.A., Anand, A.P., and Sowdhamini, R. (2004). Prediction of functional sites in proteins using conserved functional group analysis. J. Mol. Biol. **337:** 1053–1068.

Jeong, S., Trotochaud, A.E., and Clark, S.E. (1999). The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. Plant Cell **11:** 1925–1934.

Jones, D.A., and Jones, J.D.G. (1997). The role of leucine-rich repeat proteins in plant defences. In Advances in Botanical Research Incorporating Advances in Plant Pathology, Vol. 24. (London: Academic Press), pp. 89–167.

Jones, D.A., and Takemoto, D. (2004). Plant innate immunity—Direct and indirect recognition of general and specific pathogen-associated molecules. Curr. Opin. Immunol. **16:** 48–62.

Jones, J.D., and Dangl, J.L. (2006). The plant immune system. Nature **444:** 323–329.

Kajava, A.V. (2001). Review. Proteins with repeated sequence–structural prediction and modeling. J. Struct. Biol. **134:** 132–144.

Kajava, A.V., and Kobe, B. (2002). Assessment of the ability to model proteins with leucine-rich repeats in light of the latest structural information. Protein Sci. **11:** 1082–1090.

Kim, M.G., da Cunha, L., McFall, A.J., Belkhadir, Y., DebRoy, S., Dangl, J.L., and Mackey, D. (2005). Two *Pseudomonas syringae* type III effectors inhibit RIN4-regulated basal defense in Arabidopsis. Cell **121:** 749–759.

Kinoshita, T., Cano-Delgado, A., Seto, H., Hiranuma, S., Fujioka, S., Yoshida, S., and Chory, J. (2005). Binding of brassinosteroids to the extracellular domain of plant receptor kinase BRI1. Nature **433:** 167–171.

Kobe, B., and Deisenhofer, J. (1996). Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A. J. Mol. Biol. **264:** 1028–1043.

Kobe, B., and Kajava, A.V. (2001). The leucine-rich repeat as a protein recognition motif. Curr. Opin. Struct. Biol. **11:** 725–732.

Leckie, F., Mattei, B., Capodicasa, C., Hemmings, A., Nuss, L., Aracri, B., De Lorenzo, G., and Cervone, F. (1999). The specificity of polygalacturonase-inhibiting protein (PGIP): A single amino acid substitution in the solvent-exposed beta-strand/beta-turn region of the leucine-rich repeats (LRRs) confers a new recognition capability. EMBO J. **18:** 2352–2363.

Leconte, I., Carpentier, J.L., and Clauser, E. (1994). The functions of the human insulin receptor are affected in different ways by mutation of each of the four N-glycosylation sites in the beta subunit. J. Biol. Chem. **269:** 18062–18071.

Li, X., Lin, H., Zhang, W., Zou, Y., Zhang, J., Tang, X., and Zhou, J.M. (2005). Flagellin induces innate immunity in nonhost interactions that is suppressed by *Pseudomonas syringae* effectors. Proc. Natl. Acad. Sci. USA **102:** 12990–12995.

Mackey, D., and McFall, A.J. (2006). MAMPs and MIMPs: Proposed classifications for inducers of innate immunity. Mol. Microbiol. **61:** 1365–1371.

Matsubayashi, Y., Ogawa, M., Morita, A., and Sakagami, Y. (2002). An LRR receptor kinase involved in perception of a peptide plant hormone, phytosulfokine. Science **296:** 1470–1472.

Meyers, B.C., Shen, K.A., Rohani, P., Gaut, B.S., and Michelmore, R.W. (1998). Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell **10:** 1833–1846.

Mondragon-Palomino, M., Meyers, B.C., Michelmore, R.W., and Gaut, B.S. (2002). Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. Genome Res. **12:** 1305–1315.

Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T., and Jones, J.D. (2004). The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. Plant Physiol. **135:** 1113–1128.

Neylon, C. (2004). Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: Library construction methods for directed evolution. Nucleic Acids Res. **32:** 1448–1459.

Oh, H.S., and Collmer, A. (2005). Basal resistance against bacteria in *Nicotiana benthamiana* leaves is accompanied by reduced vascular staining and suppressed by multiple *Pseudomonas syringae* type III secretion system effector proteins. Plant J. **44:** 348–359.

Pancer, Z., Amemiya, C.T., Ehrhardt, G.R., Ceitlin, J., Gartland, G.L., and Cooper, M.D. (2004). Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature **430:** 174–180.

Pancer, Z., and Cooper, M.D. (2006). The evolution of adaptive immunity. Annu. Rev. Immunol. **24:** 497–518.

Pfund, C., Tans-Kersten, J., Dunning, F.M., Alonso, J.M., Ecker, J.R., Allen, C., and Bent, A.F. (2004). Flagellin is not a major defense elicitor in *Ralstonia solanacearum* cells or extracts applied to *Arabidopsis thaliana*. Mol. Plant Microbe Interact. **17:** 696–706.

Rairdan, G.J., and Moffett, P. (2006). Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. Plant Cell **18:** 2082–2093.

Robatzek, S., Bittel, P., Chinchilla, D., Kochner, P., Felix, G., Shiu, S.H., and Boller, T. (2007). Molecular identification and characterization of the tomato flagellin receptor LeFLS2, an orthologue of Arabidopsis FLS2 exhibiting characteristically different perception specificities. Plant Mol. Biol. **64:** 539–547.

Robatzek, S., Chinchilla, D., and Boller, T. (2006). Ligand-induced endocytosis of the pattern recognition receptor FLS2 in Arabidopsis. Genes Dev. **20:** 537–542.

Rose, L.E., Bittner-Eddy, P.D., Langley, C.H., Holub, E.B., Michelmore, R.W., and Beynon, J.L. (2004). The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in *Arabidopsis thaliana*. Genetics **166:** 1517–1527.

Russinova, E., Borst, J.W., Kwaaitaal, M., Cano-Delgado, A., Yin, Y., Chory, J., and de Vries, S.C. (2004). Heterodimerization and endocytosis of Arabidopsis brassinosteroid receptors BRI1 and AtSERK3 (BAK1). Plant Cell **16:** 3216–3229.

Scheer, J.M., Pearce, G., and Ryan, C.A. (2003). Generation of systemin signaling in tobacco by transformation with the tomato systemin receptor kinase gene. Proc. Natl. Acad. Sci. USA **100:** 10114–10117.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. **31:** 3381–3385.

Shiu, S.H., and Bleecker, A.B. (2003). Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Plant Physiol. **132:** 530–543.

Shiu, S.H., Karlowski, W.M., Pan, R., Tzeng, Y.H., Mayer, K.F., and Li, W.H. (2004). Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. Plant Cell **16:** 1220–1234.

Sun, W., Dunning, F.M., Pfund, C., Weingarten, R., and Bent, A.F. (2006). Within-species flagellin polymorphism in *Xanthomonas campestris* pv *campestris* and its impact on elicitation of Arabidopsis FLAGELLIN SENSING2-dependent defenses. Plant Cell **18:** 764–779.

Thomas, C.M., Jones, D.A., Parniske, M., Harrison, K., Balint-Kurti, P.J., Hatzixanthis, K., and Jones, J.D. (1997). Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. Plant Cell **9:** 2209–2224.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:** 4876–4882.

Ting, J.P., and Davis, B.K. (2005). CATERPILLER: A novel gene family important in immunity, cell death, and diseases. Annu. Rev. Immunol. **23:** 387–414.

Torii, K.U. (2004). Leucine-rich repeat receptor kinases in plants: Structure, function, and signal transduction pathways. Int. Rev. Cytol. **234:** 1–46.

Trinchieri, G., and Sher, A. (2007). Cooperation of Toll-like receptor signals in innate immune defence. Nat. Rev. **7:** 179–190.

Tuskan, G.A., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science **313:** 1596–1604.

van Der Hoorn, R.A., Roth, R., and De Wit, P.J. (2001). Identification of distinct specificity determinants in resistance protein Cf-4 allows construction of a Cf-9 mutant that confers recognition of avirulence protein Avr4. Plant Cell **13:** 273–285.

van der Hoorn, R.A., Wulff, B.B., Rivas, S., Durrant, M.C., van der Ploeg, A., de Wit, P.J., and Jones, J.D. (2005). Structure-function analysis of cf-9, a receptor-like protein with extracytoplasmic leucine-rich repeats. Plant Cell **17:** 1000–1015.

Wulff, B.B., Thomas, C.M., Smoker, M., Grant, M., and Jones, J.D. (2001). Domain swapping and gene shuffling identify sequences required for induction of an Avr-dependent hypersensitive response by the tomato Cf-4 and Cf-9 proteins. Plant Cell **13:** 255–272.

Yamaguchi, Y., Pearce, G., and Ryan, C.A. (2006). The cell surface leucine-rich repeat receptor for AtPep1, an endogenous peptide elicitor in Arabidopsis, is functional in transgenic tobacco cells. Proc. Natl. Acad. Sci. USA **103:** 10104–10109.

Zipfel, C., and Felix, G. (2005). Plants and animals: A different taste for microbes? Curr. Opin. Plant Biol. **8:** 353–360.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T., and Felix, G. (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. Cell **125:** 749–760.

Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E.J., Jones, J.D., Felix, G., and Boller, T. (2004). Bacterial disease resistance in Arabidopsis through flagellin perception. Nature **428:** 764–767.