

9-2-2003

The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000

C. Robin Buell

Institute for Genomic Research, Rockville, MD

Vinita Joardar

Institute for Genomic Research, Rockville, MD

Magdalen Lindeberg

Cornell University

Jeremy Selengut


Institute for Genomic Research, Rockville, MD

Ian T. Paulsen

Institute for Genomic Research, Rockville, MD

Follow this and additional works at: <https://digitalcommons.unl.edu/plantpathpapers>

See next page for additional authors

 Part of the [Plant Pathology Commons](#)

Buell, C. Robin; Joardar, Vinita; Lindeberg, Magdalen; Selengut, Jeremy; Paulsen, Ian T.; Gwinn, Michelle L.; Dodson, Robert J.; Deboy, Robert T.; Durkin, A. Scott; Kolonay, James F.; Radupu, Ramana; Daugherty, Sean; Brinkac, Lauren; Beanan, Maureen J.; Haft, Daniel H.; Nelson, William C.; Davidsen, Tanja; Zafar, Nikhat; Zhou, Liwei; Liu, Jia; Yuan, Qiaoping; Khouri, Hoda; Fedorova, Nadia; Tran, Bao; Russell, Daniel; Berry, Kristi; Utterback, Teresa; Van Aken, Susan E.; Feldblyum, Tamara V.; D'Ascenzo, Mark; Deng, Wen-Ling; Ramos, Adela R.; Alfano, James R.; Cartinhour, Samuel; Chatterjee, Arn K.; Delaney, Terrence P.; Lazarowitz, Sondra G.; Martin, Gregory B.; Schneider, David J.; Tang, Xiaoyan; Bender, Carol L.; White, Owen; Fraser, Claire M.; and Collmer, Alan, "The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000" (2003). *Papers in Plant Pathology*. 199. <https://digitalcommons.unl.edu/plantpathpapers/199>

This Article is brought to you for free and open access by the Plant Pathology Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Plant Pathology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

C. Robin Buell, Vinita Joardar, Magdalen Lindeberg, Jeremy Selengut, Ian T. Paulsen, Michelle L. Gwinn, Robert J. Dodson, Robert T. Deboy, A. Scott Durkin, James F. Kolonay, Ramana Radupu, Sean Daugherty, Lauren Brinkac, Maureen J. Beanan, Daniel H. Haft, William C. Nelson, Tanja Davidsen, Nikhat Zafar, Liwei Zhou, Jia Liu, Qiaoping Yuan, Hoda Khouri, Nadia Fedorova, Bao Tran, Daniel Russell, Kristi Berry, Teresa Utterback, Susan E. Van Aken, Tamara V. Feldblyum, Mark D'Ascenzo, Wen-Ling Deng, Adela R. Ramos, James R. Alfano, Samuel Cartinhour, Arn K. Chatterjee, Terrence P. Delaney, Sondra G. Lazarowitz, Gregory B. Martin, David J. Schneider, Xiaoyan Tang, Carol L. Bender, Owen White, Claire M. Fraser, and Alan Collmer

The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000

C. Robin Buell^{*†}, Vinita Joardar^{*}, Magdalen Lindeberg[†], Jeremy Selengut^{*}, Ian T. Paulsen^{*}, Michelle L. Gwinn^{*}, Robert J. Dodson^{*}, Robert T. Deboy^{*}, A. Scott Durkin^{*}, James F. Kolonay^{*}, Ramana Madupu^{*}, Sean Daugherty^{*}, Lauren Brinkac^{*}, Maureen J. Beanan^{*}, Daniel H. Haft^{*}, William C. Nelson^{*}, Tanja Davidsen^{*}, Nikhat Zafar^{*}, Liwei Zhou^{*}, Jia Liu^{*}, Qiaoping Yuan^{*}, Hoda Khouri^{*}, Nadia Fedorova^{*}, Bao Tran^{*}, Daniel Russell^{*}, Kristi Berry^{*}, Teresa Utterback^{*}, Susan E. Van Aken^{*}, Tamara V. Feldblyum^{*}, Mark D'Ascenzo[§], Wen-Ling Deng[‡], Adela R. Ramos[‡], James R. Alfano[¶], Samuel Cartinhour[¶], Arun K. Chatterjee^{**}, Terrence P. Delaney[†], Sondra G. Lazarowitz[†], Gregory B. Martin[§], David J. Schneider[¶], Xiaoyan Tang^{††}, Carol L. Bender[‡], Owen White^{*}, Claire M. Fraser^{*}, and Alan Collmer[†]

^{*}The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850; [†]Department of Plant Pathology, Cornell University, Ithaca, NY 14853;

[§]Boyce Thompson Institute for Plant Research, Ithaca, NY 14853-1801; [¶]Plant Science Initiative and Department of Plant Pathology, University of Nebraska, Lincoln, NE 68588-0660; [¶]United States Department of Agriculture Agricultural Research Service, Cornell Theory Center, Ithaca, NY 14853-2901;

^{**}Department of Plant Microbiology and Pathology, University of Missouri, Columbia, MO 65211; ^{††}Department of Plant Pathology, Kansas State University, Manhattan, KS 66502-5502; and [‡]Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078

Edited by Robert Haselkorn, University of Chicago, Chicago, IL, and approved June 23, 2003 (received for review April 4, 2003)

We report the complete genome sequence of the model bacterial pathogen *Pseudomonas syringae* pathovar *tomato* DC3000 (DC3000), which is pathogenic on tomato and *Arabidopsis thaliana*. The DC3000 genome (6.5 megabases) contains a circular chromosome and two plasmids, which collectively encode 5,763 ORFs. We identified 298 established and putative virulence genes, including several clusters of genes encoding 31 confirmed and 19 predicted type III secretion system effector proteins. Many of the virulence genes were members of paralogous families and also were proximal to mobile elements, which collectively comprise 7% of the DC3000 genome. The bacterium possesses a large repertoire of transporters for the acquisition of nutrients, particularly sugars, as well as genes implicated in attachment to plant surfaces. Over 12% of the genes are dedicated to regulation, which may reflect the need for rapid adaptation to the diverse environments encountered during epiphytic growth and pathogenesis. Comparative analyses confirmed a high degree of similarity with two sequenced pseudomonads, *Pseudomonas putida* and *Pseudomonas aeruginosa*, yet revealed 1,159 genes unique to DC3000, of which 811 lack a known function.

The gamma subgroup of the Proteobacteria contains a number of important pathogens, including animal pathogenic *Escherichia*, *Salmonella*, *Shigella*, and *Yersinia* spp. and plant pathogenic *Erwinia*, *Pantoea*, *Xanthomonas*, *Xylella*, and *Pseudomonas* spp. (1). The genus *Pseudomonas* is notable because it contains the clinically important human pathogen *Pseudomonas aeruginosa*, the agriculturally important plant pathogen *Pseudomonas syringae*, and the nonpathogenic bioremediation agent *Pseudomonas putida*. Strains of *P. syringae* are noted for their diverse and host-specific interactions with plants and may be assigned to one of at least 50 pathovars based on host specificity (2). As a pathogen, *P. syringae* typically enters plant leaves through stomata, multiplies in the intercellular space (apoplast), and eventually produces necrotic lesions that are often surrounded by chlorotic halos (2).

P. syringae pv. *tomato* DC3000 (DC3000) causes bacterial speck of tomato (3), a worldwide, economically significant disease that is representative of numerous bacterial plant diseases for which effective controls are still needed (4). DC3000 is also a pathogen of the model plant *Arabidopsis thaliana* (5). Importantly, the pathogenicity of DC3000 resembles that of most animal and plant pathogens in the gamma Proteobacteria, which rely on a type III secretion system (TTSS) to inject virulence effector proteins into host cells (6). Highlighting the role of DC3000 as a model for

plant–pathogen interactions, >35 genes encoding TTSS substrates have been experimentally identified in this strain (7–12), which exceeds the number demonstrated for any other bacterial pathogen to date. Furthermore, the pathogenicity of DC3000 resembles that of many important bacterial and fungal pathogens of plants where host specificity is controlled by “gene-for-gene” interactions in which a dominant allele in the host and a dominant allele in the pathogen condition the outcome of the plant–pathogen interaction (13). In this study, we report the complete genome sequence of DC3000, highlight genes with a likely role in virulence, and compare the virulence gene arsenal of DC3000 with that of other plant and animal pathogens.

Materials and Methods

Sequencing and Annotation Methods. Strain DC3000 was originally isolated in 1960 from tomato in the Channel Islands in Guernsey, United Kingdom (ref. 3; D. Cuppels, personal communication). The culture was confirmed as pathogenic on susceptible host plants before high-throughput sequencing. The complete genome was sequenced as described previously (14) and detailed information is available as *Supporting Sequencing Methods*, which is published as supporting information on the PNAS web site, www.pnas.org. ORFs were identified by using the GLIMMER algorithm (15). Predicted proteins were searched against a nonredundant amino acid database and searched for domains by using HMMER with the Pfam (16) and TIGRfam databases (17). The ORFs were manually curated and assigned to role categories adapted from Riley (18). Paralogous families were constructed as described previously (19). Insertion-sequence (IS) elements were classified according to transposase gene similarity by using BLAST analysis with the ISFinder database (www-is.biotoul.fr/).

Comparative Genomics. The presence of putative orthologs of DC3000 virulence factors in nine completely sequenced genomes of pathogenic proteobacteria was determined by BLASTP analysis by

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CMA, coronamic acid; COR, coronatine; DC3000, *Pseudomonas syringae* pv. *tomato* DC3000; GABA, γ -aminobutyric acid; IS, insertion sequence; Tfp, type IV pili; TTSS, type III secretion system.

Data deposition: The sequence and annotation of *P. syringae* pv. *tomato* DC3000 have been deposited in the GenBank database (accession nos. AE016853–AE016855).

[†]To whom correspondence should be addressed. E-mail: rbuell@tigr.org.

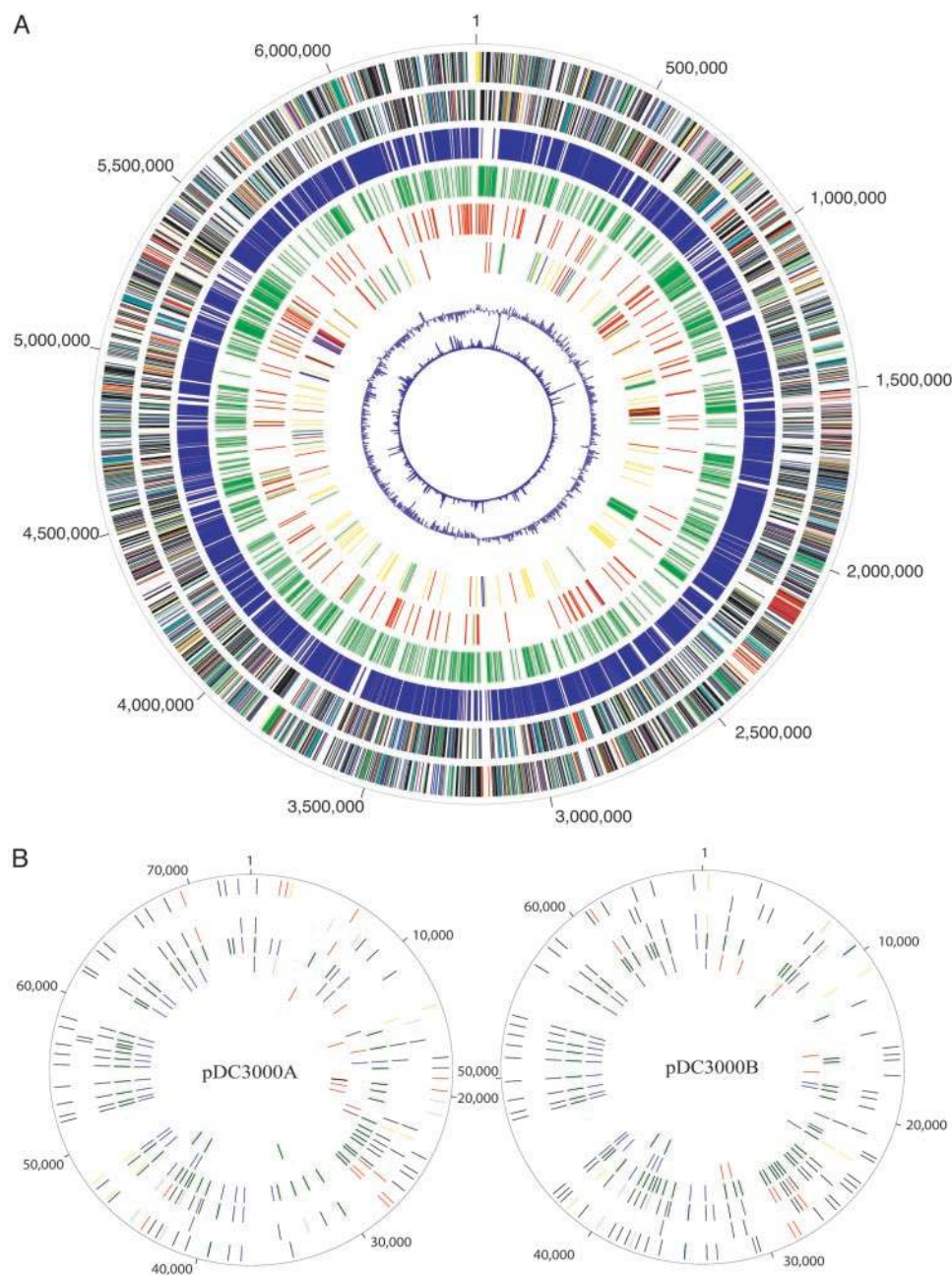


Fig. 1. Features of the DC3000 genome. (A) The chromosome, with the outermost circle depicting ORFs on the positive strand and the second circle depicting ORFs on the negative strand. The ORFs are color-coded based on the major grouping of role categories as follows: amino acid biosynthesis (salmon); biosynthesis of cofactors, prosthetic groups, and carriers (light blue); cell envelope (light green); cellular processes (red); central intermediary metabolism (brown); DNA metabolism (yellow); energy metabolism (green); fatty acid and phospholipid metabolism (purple); protein fate/synthesis (pink); purines, pyrimidines, nucleosides, and nucleotides (orange); regulatory functions (navy blue); transcription (gray); transport and binding proteins (teal); and unknown function, conserved hypothetical, and hypothetical (black). The third circle represents the ORFs with homologs present in all three *Pseudomonas* spp. ($E < 10^{-5}$). The fourth circle depicts the ORFs that are unique to DC3000 compared with *P. aeruginosa* and *P. putida* ($E \geq 10^{-5}$). The fifth circle represents mobile genetic elements, including IS elements (red), phage-related genes (light green), and plasmid-related functions (blue). The sixth circle depicts ORFs associated with virulence, which are color coded as follows: adhesins and other cell surface-associated factors (dark green), miscellaneous virulence factors (yellow), TTSS effectors, candidates, and helpers (red), TTSS regulatory and secretory proteins (brown), and low molecular weight diffusible factors (blue). The seventh circle is the GC skew, and the eighth circle represents atypical nucleotides. (B) The plasmids with the outermost circle representing ORFs on the positive strand and the second circle depicting ORFs on the negative strand, with the same color coding as A. The third circle represents ORFs shared by the two plasmids ($E < 10^{-5}$). The fourth circle depicts the ORFs with homologs present in all three *Pseudomonas* spp. ($E < 10^{-5}$) in navy blue and the ORFs that are unique to DC3000 compared with *P. aeruginosa* and *P. putida* ($E \geq 10^{-5}$) in green. The fifth circle represents mobile genetic elements (color-coded as in A). The sixth circle depicts ORFs associated with virulence (color-coded as in A).

using a cutoff criterion of $E < 10^{-5}$ (20). The reference genomes included the plant pathogens *Ralstonia solanacearum* (21), *Xylella fastidiosa* (22), *Agrobacterium tumefaciens* (23, 24), *Xanthomonas campestris* pv. *campestris* (25), and *Xanthomonas axonopodis* pv. *citri* (25) and the animal pathogens *P. aeruginosa* PAO1 (26), *Salmonella typhimurium* LT2 (27), *Yersinia pestis* CO92 (28), and *Escherichia coli* O157:H7 EDL933 (29).

Results and Discussion

Genome Features. The DC3000 genome is composed of one circular chromosome of 6,397,126 bp and two plasmids of 73,661 bp (pDC3000A) and 67,473 bp (pDC3000B) (Fig. 1 and Table 1). A total of 5,763 ORFs were identified within the DC3000 genome. Among the predicted genes, 3,510 (61%) have been assigned a putative function, with the remaining designated as encoding a protein with unknown function, a conserved hypothetical protein, or a hypothetical protein. Classification of the known proteins into

biological role categories (18) revealed that the largest category was transport and binding, followed by cellular processes (see Fig. 4, which is published as supporting information on the PNAS web site). DC3000 is representative of pathogenic bacteria that have diverse relationships with plants as epiphytes and pathogens and thus likely requires an array of regulatory genes to cope with diverse niches, host responses, and environmental conditions. Indeed, 12% (711) of the DC3000 genes could be assigned to role categories associated with gene regulation, signal transduction, and transcription. A high degree of gene duplication exists in DC3000, and 2,735 genes (48%) were assigned to 687 paralogous families. Although the majority of paralogous families contained two members (see Fig. 5, which is published as supporting information on the PNAS web site), 12 families have >25 members, and many of these large families encode either transporters or regulatory proteins. However, this group also includes families encoding methyl-accepting chemotaxis proteins, oxidoreductases of largely unknown specific-

Feature	Chromosome	pDC3000A	pDC3000B
Size, bp	6,397,126	73,661	67,473
G+C, %	58.4	55.1	56.1
Number of ORFs	5,615	71	77
Percent coding	86.8	81.9	84.7
Average ORF length, nt	988	849	742
No. of ORFs encoding			
Protein with a known function	3,402 (61%)	52 (73%)	46 (60%)
Hypothetical protein	610 (11%)	10 (14%)	16 (21%)
Conserved hypothetical protein	961 (17%)	8 (11%)	12 (16%)
Protein of unknown function	642 (11%)	1 (1%)	3 (4%)

Comparative Genomics with Fluorescent *Pseudomonads*. The complete genome sequences for two other pseudomonads, *P. aeruginosa* PAO1 (26), an opportunistic animal pathogen, and *P. putida* KT2440, a saprophytic soil bacterium (30), are available. Of the 5,608 ORFs in DC3000 that predict a functional protein, 4,449 (79%) have homologs (BLASTP $E < 10^{-5}$) in either *P. aeruginosa* or *P. putida* and 3,797 (68%) are present in all three genomes (Fig. 1). The 1,159 (21% of the total) genes that are unique to DC3000 are heavily enriched in hypothetical proteins (551), conserved hypothetical proteins (203), mobile genetic elements (116), and proteins of unknown function (57). Other genes unique to DC3000 include 96 that are implicated in virulence (see below for more details).

The utilization of γ -aminobutyric acid (GABA) occurs by means of the GABA transaminase-catalyzed deamination of GABA to succinic semialdehyde, which is then oxidized to succinate by succinate semialdehyde dehydrogenase (SSADH) before entering the tricarboxylic acid cycle. In DC3000, there are three copies each of GABA transaminase and SSADH, as well as a GABA permease. Although the composition of the apoplast during a compatible interaction with DC3000 is unknown, GABA is present at millimolar concentration in the apoplast during infection of tomato with the biotrophic fungal pathogen *Cladosporium fulvum* (32). Thus, if GABA is available in the apoplast during infection of tomato, DC3000 appears amply equipped to use this as a carbon and nitrogen source.

The diagram illustrates the interaction between a Bacterial cell and a Plant cell during infection. The Bacterial cell (top) contains regulators of virulence and stress tolerance (GacA/S, HrpS/R/L, CorR/S, and others) which control the Hrp TTSS effectors. These effectors are injected into the Plant cytoplasm (bottom) via the Type IV pilus (TTSS). The plant cell is divided into the Apoplast (top) and the Cytoplasm (bottom). In the Apoplast, the plant has ROS scavengers and repair, ABC exporters, EPS, and FLS2. In the Cytoplasm, the plant has ROS, NO Antimicrobials, Type IV pilus Adhesins, and a Defense suppression pathway. The bacterial cell also has Flagella, ABC importers, and a Defense suppression pathway. The plant cell has a Defense suppression pathway and a Disease pathway. The diagram shows the flow of information and molecules between the bacterial and plant cells, highlighting the role of the Hrp TTSS effectors in plant defense and disease.

TTSS and Effectors. The TTSS is key to the plant parasitism of *P. syringae* pathovars (Fig. 2). Most of the *hrp* (hypersensitive response and pathogenicity) and *hrc* (*h*rp conserved) genes encoding the TTSS system are essential for pathogenicity, which indicates the collective importance of the effector proteins that are injected into plant cells by the system (6). Hrp effectors are primarily known as Hop (Hrp outer protein) or Avr (avirulence) proteins, based on the

phenotype by which they were discovered (11). Functional analysis of the DC3000 genome has yielded genes for 31 effectors and 7 additional proteins secreted by the Hrp system that likely have a role in translocating effectors across the plant cell wall and plasma membrane (8–12, 36). The associated *hop*, *avr*, and *hol* (Hop-like candidate) genes may provide markers to regions of the genome that are enriched in other virulence genes (Fig. 3). Over half of these TTSS effectors are encoded in one of several regions containing multiple effector genes (Fig. 1). One region with multiple effector genes is located on pDC3000A and contains *avrPpiB2_{Pto}*, *avrPphE_{Pto}*, *hopPtoS1*, and *hopPtoT1*. Interestingly, an identical copy of *avrPpiB2_{Pto}* is located on the chromosome, which could provide a site for integration of the plasmid in a manner similar to that observed with pFKN in *P. syringae* pv. *maculicola* M6 (37). Two other clusters of effector genes are found in the region encoding the phytotoxin coronatine (COR) (Fig. 3). Members of effector gene paralogous families are scattered around the genome and typically have an unusually low G+C content (Table 2), suggesting that such families may have grown by sequential horizontal acquisitions. Interestingly, at least four of the DC3000 effector genes are disrupted by mobile genetic elements (Tables 2 and 3, which are published as supporting information on the PNAS web site), as highlighted by comparison with orthologs in the draft genome of *P. syringae* pv. *syringae* B728a (12). Additional information on DC3000 TTSS effectors and an updated list can be found at <http://pseudomonas-syringae.org>.

Plasmid-Borne Virulence-Implicated Genes. Consistent with previous reports (38), DC3000 contains two plasmids (pDC3000A and pDC3000B) that are similar to the pPT23A family of *P. syringae* plasmids (39). There is substantial redundancy between the two plasmids, with 47 proteins sharing similarity ($E < 10^{-5}$). The majority of the conserved proteins are involved in plasmid replication, conjugation, or transposition-related functions. There also are many genes unique to each of the plasmids, with a clear enrichment of genes implicated in virulence on pDC3000A. However, we found no evidence for a major role of the plasmids in virulence. Using previously described methods (40), we inserted *sacB* in pDC3000A and *repA* in pDC3000B to construct DC3000 derivatives lacking pDC3000A, pDC3000B, or both plasmids. Assays involving dipping of tomato cv. MoneyMaker and *A. thaliana* Col-0 in inoculum (5×10^5 colony-forming units per ml) revealed no major difference in bacterial growth *in planta* or in symptom production (data not shown). This observation is consistent with the finding that most of the known plasmid-borne TTSS effector genes, which are all on pDC3000A, have paralogs on the chromosome.

Siderophores. Siderophores are low molecular weight, high-affinity iron(III) chelators that function as important virulence factors in many bacterial pathogens because of their role in sustaining growth in iron-limited host environments. Synthesis of the siderophore pyoverdinin is well documented among fluorescent pseudomonads (41), and the required genes are present in a single cluster in DC3000. Also, a cluster of genes homologous to those required for biosynthesis of the siderophore pyochelin by *P. aeruginosa* has been found. However, several observations suggest that the encoded proteins actually produce yersiniabactin, a siderophore that shares a common tricyclic core with pyochelin (42). First, DC3000 contains a gene encoding a large polyketide synthase (PKS) in the middle of the pyochelin gene cluster and this has no homolog in *P. aeruginosa* (or *P. putida*). This gene shows extremely high similarity with the PKS gene (*irp1*) found in the yersiniabactin gene cluster of *Y. pestis* (43). Second, there is no homolog of the pyochelin receptor gene of *P. aeruginosa* (*fptA*) in DC3000; however, a receptor gene is present that is highly similar to the pesticin/yersiniabactin receptor gene from *Y. pestis* (*fyuA*). Third, the structure of PSPTO2602, a nonribosomal peptide synthase in this cluster, differs from its *P. aeruginosa* counterpart in that a methylation domain and a terminal

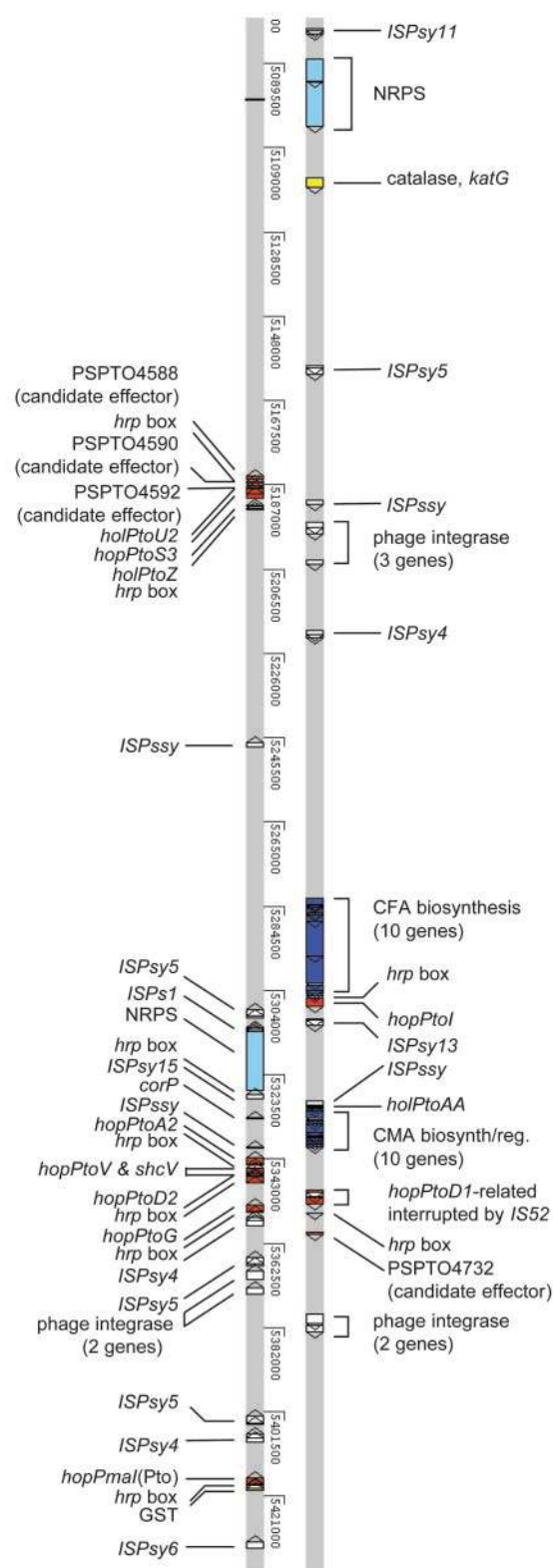


Fig. 3. Detailed view of one region of the DC3000 genome that is enriched in virulence factors and mobile genetic elements. Only virulence-related genes, "hrp box" promoters, and mobile genetic elements are shown. Features are color-coded as in Fig. 1 (circle 6), with the exception that a lighter blue is used for the nonribosomal peptide synthase genes, and the mobile genetic elements are white. The genome display was generated by using ARTEMIS (www.sanger.ac.uk/Software/Artemis/) to view the chromosome (AE016853) with the aid of input files for a variety of virulence-related features that are available at <http://pseudomonas-syringae.org> and <http://monod.cornell.edu>.

thioesterase domain are lacking, consistent with a yersiniabactin synthase.

Phytotoxins and Phytohormones. COR is a chlorosis-inducing phytotoxin produced by several *P. syringae* pathovars, including DC3000 (44). COR is hypothesized to function as a molecular mimic of methyl jasmonate and may promote colonization of host tissue by altering plant defense signaling pathways (45). COR is representative of biologically active small molecules that can diffuse in plant tissues (Fig. 2), and it is responsible for the chlorotic halos surrounding the necrotic lesions caused by DC3000 (44). The structure of COR consists of coronafacic acid and coronamic acid (CMA), joined by an amide linkage (46). Unlike many *P. syringae* pathovars where the COR genes are clustered and plasmid-encoded (46), the COR genes in DC3000 are chromosomally encoded, and the coronafacic acid and CMA structural genes are separated by 26 kb (Fig. 3). In DC3000, the CMA biosynthetic genes are linked with a modified two-component regulatory system, composed of *corS*, *corR*, and *corP* (46). Interestingly, analysis of the CMA region of DC3000 revealed four genes (PSPTO4707, 4708, 4711, and 4713) interspersed among those previously identified with roles in COR synthesis. Consistent with the annotation of the COR region, initial proteomic studies with DC3000 revealed that the HrpL-alternative σ factor, which regulates DC3000 *hrp* genes (8), induces production of a protein encoded by PSPTO4668 in the coronafacic acid region (D.J.S. and S.C., unpublished data; <http://monod.cornell.edu>). There is no evidence in the DC3000 genome for the biosynthesis of any of the known *P. syringae* lipodepsinone-peptide phytotoxins, including syringomycin, syringotoxin, syringostatin, and pseudomycin. Syringopeptin, a lipodepsipeptide produced by *P. syringae* pv. *syringae* (46), is also absent. However, DC3000 does contain five nonribosomal peptide synthetases in three separate gene clusters (PSPTO2829, -2830, -4518, -4519, and -4699) that may direct the biosynthesis of undescribed peptide or lipodepsipeptide products.

Production of the phytohormone indole-3-acetic acid (IAA) by certain *P. syringae* pathovars has been shown to influence bacterial growth *in planta* and the expression of other virulence factors (47, 48). The two genes required for IAA production, *iaaH* and *iaaM*, are present in DC3000. Interestingly, *iaaL* (IAA-lysine ligase), which can convert IAA to a less active derivative and is widespread in *P. syringae* pathovars (47), is present in the DC3000 genome but lacking in the genomes of all of the other fully sequenced plant pathogens (Table 2). The presence of a functional Hrp promoter upstream of *iaaL* additionally suggests a role in *P. syringae* virulence (8).

Adhesins. Pathogenic bacteria rely on a variety of cell surface-associated virulence factors to facilitate adhesion to the host and provide protection during colonization. Type IV pili (Tfp) represent one of the major adhesins found in pseudomonads (49). Tfp were previously identified in *P. syringae* and appear to play a role in epiphytic fitness (50). Analysis of the 23 Tfp-associated genes suggests that regulation of Tfp production differs significantly between DC3000 and *P. aeruginosa*. Filamentous hemagglutinin (FHA) has also been shown to be an adhesin and a virulence factor in animal pathogenic bacteria and the plant pathogen *Erwinia chrysanthemi* (51). Three genes predicting FHA-like proteins were identified in DC3000 (PSPTO3210, 3214, and 3229).

Extracellular Polysaccharides. Components of the polysaccharide capsule, such as the exopolysaccharide alginate, play a role in both adhesion and protection of the bacterial cells from external stress (52). Among plant pathogenic pseudomonads, alginate production has been linked to epiphytic fitness and the production of water-soaked lesions (53). All of the genes required for alginate biosynthesis in *P. aeruginosa* are present in DC3000, although some of the alginate regulatory genes are absent, indicating that the regulation

of alginate biosynthesis differs between *P. aeruginosa* and *P. syringae*, consistent with previous reports (54). Three genes encoding levansucrases, required for the biosynthesis of the polysaccharide levan, were also identified in DC3000.

Tolerance to Reactive Oxygen Species or Heavy Metals. *P. syringae* and other pathogens can trigger an oxidative burst in plants, which results in the production in the apoplast of reactive oxygen species such as superoxide anion and hydrogen peroxide (H_2O_2). Superoxide dismutases (SODs) convert superoxide to H_2O_2 , which can be detoxified to water by catalase. In addition to its antimicrobial properties, H_2O_2 functions as a signaling molecule within plant cells in processes involving pathogen defense and programmed cell death (55). DC3000 has three catalases and three SODs. Other enzymes in DC3000 that may play a role in scavenging reactive oxygen species include six peroxidases, eight GSTs, and two glutathione reductases. In addition, DC3000 encodes a single peptide methionine sulfoxide reductase, which reduces methionine sulfoxide to methionine, thereby functioning as a repair enzyme for proteins inactivated by oxidation.

One strategy to control bacterial plant pathogens in the field is the use of copper-based bactericides, which are toxic to bacterial cells at high concentrations. For bacterial speck of tomato, this has led to the emergence of copper-resistant strains (56). In *P. syringae* pv. *tomato* strain PT23.2, resistance to copper is mediated by four structural genes, *copA–D*, as well as two regulatory genes, *copR* and *copS*, which encode a two-component regulatory system (57). DC3000 is not resistant to copper (C.L.B., unpublished data), and examination of the genome revealed that although *copA* and *copB* are present, DC3000 lacks the structural proteins CopC and CopD, as well as the regulatory proteins CopR and CopS. Interestingly, homologs of CopR and CopS (PSPTO3603, 1306) are present in DC3000 and may function in regulating responses to other heavy metals.

Other Proteins Relevant to Virulence. Pectic enzymes and cellulases were previously identified in other *Pseudomonas* spp., and pectate lyase has been shown to influence the final symptoms produced by *P. syringae* pv. *lachrymans* (58). Genes encoding cell-wall-degrading enzymes are present in DC3000 and include a pectin lyase, a polygalacturonase, and three enzymes predicted to have cellulolytic activity (PSPTO1029, PSPTO3534, and PSPTO0905). As previously noted, DC3000 has two insecticidal toxin gene clusters (9). Although *P. syringae* is not known to be an insect pathogen, insects may be casual vectors of epiphytic bacteria on wet leaves (2). It is also worth noting that DC3000 lacks a gene encoding the outer-membrane ice nucleation protein, which enables many strains of *P. syringae* to cause frost damage on plants (2).

Mobile Genetic Elements and Their Potential Impact on the Genome. A total of 382 genes (7%) in the DC3000 genome were classified as mobile genetic elements, including genes encoding 270 proteins that function in transposition, 54 proteins with plasmid functions, and 58 proteins with prophage functions. Compared with other published plant pathogen genomes (21–23, 25), this is the highest percentage seen to date, resulting in 24 ORFs interrupted by an IS element or phage protein (Table 3). The transposable element class, which includes the IS elements, accounts for 5% of the DC3000 ORFs (Table 4, which is published as supporting information on the PNAS web site). The IS elements vary widely in their copy number, presumably reflecting a combination of transposition frequency and residence time in the genome. Multiple prophage regions also are present in DC3000, which collectively encompass ~118 kbp of the genome (Table 5, which is published as supporting information on the PNAS web site). The impact of mobile genetic elements on the DC3000 genome is seen by their frequent association with regions that lack homologs in *P. aeruginosa* and/or *P. putida*, that are enriched in atypical nucleotides potentially indicative of hori-

zontal acquisition, and/or that encode either TTSS effectors or proteins of unknown function. One such region, shown in Fig. 3, contains a potpourri of known and candidate virulence factors, including catalase and GST genes, COR biosynthetic genes, a nonribosomal peptide synthase preceded by a Hrp promoter, and two clusters of effector genes, along with multiple mobile genetic elements, one of which has disrupted a *hopPtoD* paralog.

Comparative Genomics of Virulence. We performed a comparative analysis involving 298 putative virulence genes in DC3000 and the complete genome sequences of nine plant and animal pathogenic proteobacteria, which revealed that 86 genes are common to all of these pathogens (Table 2 and Fig. 6, which is published as supporting information on the PNAS web site). The pathogen with the most similar genes was *P. aeruginosa* PAO1 (191 genes; 64% of the virulence factors), which reflects the relatedness of DC3000 and PAO1 and the large number of similar genes involved in flagellum, alginate, and Tfp production. Not surprisingly, among the 65 virulence genes that are unique to DC3000, 33 encode TTSS effectors, helpers, and candidates, which is consistent with the specialized nature of TTSS-based host-pathogen interactions.

Another promising source of virulence candidates is the set of genes unique to DC3000 that were identified in comparative analyses with *P. putida* and *P. aeruginosa*. A search of the 811 genes of unknown function in this category revealed that 635 also were unique to DC3000 in comparisons with all eight of the reference pathogen genomes. However, four of these genes were present in all of the plant pathogenic bacteria, and another set of four were present exclusively in all animal pathogenic bacteria

(Table 6, which is published as supporting information on the PNAS web site).

In considering DC3000 virulence genomics, it is worth noting that pioneering genetic studies with *P. syringae* yielded a paradox that has long frustrated comprehensive exploration of the virulence of this pathogen (59, 60). That is, despite the many indications that *P. syringae* is a highly coevolved plant parasite, screens for mutants with reduced virulence have yielded primarily mutants disabled in a single system, the Hrp TTSS. The complete genome sequence of DC3000 clarifies the basis for this paradox in the redundancy of many of the candidate virulence factors. The sequence will now facilitate the construction of polymutants that can test the role of genes with overlapping function, and it will foster the identification of novel virulence factors through targeted investigation of genes that are unique or duplicated in comparison with *P. aeruginosa* and *P. putida* or in regions showing signatures of horizontal acquisition. In conclusion, *P. syringae* pathobiology has been described as a "highly dynamic, highly variable matrix of interactions of bacteria of widely variable genotypes with a broadly variable group of plants" (61), and the plethora of mobile genetic elements and associated virulence factors that we found provides a genomic basis for such a dynamic association with plants.

We thank the members of The Institute for Genome Research Sequencing Facility, Informatics Department, and Information Technology Support Group who contributed to the work. Funding for the work described in this study was provided by National Science Foundation Grants DBI-0077622 (to A.C., J.R.A., C.R.B., S.C., A.K.C., T.P.D., S.G.L., G.B.M., and X.T.) and IBN-0135286 (to C.L.B.) and National Institutes of Health Grant AI 43311 (to C.L.B.).

- Gupta, R. S. (2000) *FEMS Microbiol. Rev.* **24**, 367–402.
- Hirano, S. S. & Upper, C. D. (2000) *Microbiol. Mol. Biol. Rev.* **64**, 624–653.
- Cuppels, D. A. (1986) *Appl. Environ. Microbiol.* **51**, 323–327.
- Wilson, M., Campbell, H. L., Ji, P., Jones, J. B. & Cuppels, D. A. (2002) *Phytopathology* **92**, 1284–1292.
- Whalen, M. C., Innes, R. W., Bent, A. F. & Staskawicz, B. J. (1991) *Plant Cell* **3**, 49–59.
- Cornelis, G. R. & Van Gijsegem, F. (2000) *Annu. Rev. Microbiol.* **54**, 735–774.
- Boch, J., Joardar, V., Gao, L., Robertson, T. L., Lim, M. & Kunkel, B. N. (2002) *Mol. Microbiol.* **44**, 73–88.
- Fouts, D. E., Abramovitch, R. B., Alfano, J. R., Baldo, A. M., Buell, C. R., Cartinhour, S., Chatterjee, A. K., D'Ascenzo, M., Gwinn, M. L., Lazarowitz, S. G., et al. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 2275–2280.
- Guttman, D. S., Vinatzer, B. A., Sarkar, S. F., Ranall, M. V., Kettler, G. & Greenberg, J. T. (2002) *Science* **295**, 1722–1726.
- Petnicki-Ocwieja, T., Schneider, D. J., Tam, V. C., Chancey, S. T., Shan, L., Jamir, Y., Schechter, L. M., Buell, C. R., Tang, X., Collmer, A., et al. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 7652–7657.
- Collmer, A., Lindeberg, M., Petnicki-Ocwieja, T., Schneider, D. J. & Alfano, J. R. (2002) *Trends Microbiol.* **10**, 462–469.
- Greenberg, J. T. & Vinatzer, B. A. (2003) *Curr. Opin. Microbiol.* **6**, 20–28.
- Keen, N. T. (1990) *Annu. Rev. Genet.* **24**, 447–463.
- Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H., Dodson, R. J., et al. (2001) *Science* **293**, 498–506.
- Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999) *Nucleic Acids Res.* **27**, 4636–4641.
- Bateman, A., Birney, E., Cerruti, L., Durbin, R., Ewiler, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M. & Sonnhammer, E. L. (2002) *Nucleic Acids Res.* **30**, 276–280.
- Haft, D. H., Selengut, J. D. & White, O. (2003) *Nucleic Acids Res.* **31**, 371–373.
- Riley, M. (1993) *Microbiol. Rev.* **57**, 862–952.
- Nierman, W. C., Feldblyum, T. V., Laub, M. T., Paulsen, I. T., Nelson, K. E., Eisen, J. A., Heidelberg, J. F., Alley, M. R., Ohta, N., Maddock, J. R., et al. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 4136–4141.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402.
- Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Manganot, S., Arlat, M., Billault, A., Brottier, P., Camus, J. C., Cattolico, L., et al. (2002) *Nature* **415**, 497–502.
- Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Avarenga, R., Alves, L. M., Araya, J. E., Baia, G. S., Baptista, C. S., et al. (2000) *Nature* **406**, 151–157.
- Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Zhou, Y., Chen, L., Wood, G. E., Almeida, N. F., Jr., et al. (2001) *Science* **294**, 2317–2323.
- Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B. S., Cao, Y., Askenazi, M., Halling, C., et al. (2001) *Science* **294**, 2323–2328.
- Da Silva, A. C., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B., Monteiro-Vitorello, C. B., Slus, M. A., Almeida, N. F., Alves, L. M., et al. (2002) *Nature* **417**, 459–463.
- Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrenner, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., et al. (2000) *Nature* **406**, 959–964.
- McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., et al. (2001) *Nature* **413**, 852–856.
- Parkhill, J., Wren, B. W., Thomson, N. R., Titball, R. W., Holden, M. T., Prentice, M. B., Sebaihia, M., James, K. D., Churcher, C., Mungall, K. L., et al. (2001) *Nature* **413**, 523–527.
- Perna, N. T., Plunkett, G., III, Burland, V., Mau, B., Glasner, J. D., Rose, D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., et al. (2001) *Nature* **409**, 529–533.
- Nelson, K. E. (2002) *Environ. Microbiol.* **4**, 777.
- Palleroni, N. J. (1984) in *Bergey's Manual of Systematic Bacteriology*, eds. Krieg, N. R. & Holt, J. G. (Williams & Wilkins, Baltimore), pp. 141–199.
- Solomon, P. S. & Oliver, R. P. (2001) *Plant* **213**, 241–249.
- Binet, R., Letoffe, S., Ghigo, J. M., Delepelaire, P. & Wandersman, C. (1997) *Gene* **192**, 7–11.
- Sandkvist, M. (2001) *Mol. Microbiol.* **40**, 271–283.
- Preston, G. M. (2000) *Mol. Plant Pathol.* **1**, 263–275.
- Zwiesler-Vollick, J., Plovanih-Jones, A. E., Nomura, K., Brandyopadhyay, S., Joardar, V., Kunkel, B. N. & He, S. Y. (2002) *Mol. Microbiol.* **45**, 1207–1218.
- Rohmer, L., Kjemtrup, S., Marchesini, P. & Dangel, J. L. (2003) *Mol. Microbiol.* **47**, 1545–1562.
- Cuppels, D. A. & Ainsworth, T. (1995) *Appl. Environ. Microbiol.* **61**, 3530–3536.
- Sesma, A., Sundin, G. W. & Murillo, J. (2000) *Microbiology* **146**, 2375–2384.
- Jackson, R. W., Athanassopoulos, E., Tsiamis, G., Mansfield, J. W., Sesma, A., Arnold, D. L., Gibbon, M. J., Murillo, J., Taylor, J. D. & Vivian, A. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 10875–10880.
- Bultreys, A. & Gheysen, I. (2000) *Appl. Environ. Microbiol.* **66**, 325–331.
- Reimann, C., Patel, H. M., Serino, L., Barone, M., Walsh, C. T. & Haas, D. (2001) *J. Bacteriol.* **183**, 813–820.
- Pelludat, C., Rakin, A., Jacobi, C. A., Schubert, S. & Heesemann, J. (1998) *J. Bacteriol.* **180**, 538–546.
- Moore, R. A., Starratt, A. N., Ma, S. W., Morris, V. L. & Cuppels, D. A. (1989) *Can. J. Microbiol.* **35**, 910–917.
- Kloek, A. P., Verbsky, M. L., Sharma, S. B., Schoelz, J. E., Vogel, J., Klessig, D. F. & Kunkel, B. N. (2001) *Plant J.* **26**, 509–522.
- Bender, C. L., Alarcon-Chaidez, F. & Gross, D. C. (1999) *Microbiol. Mol. Biol. Rev.* **63**, 266–292.
- Glickmann, E., Gardan, L. & Dessaux, Y. (1998) *Mol. Plant-Microbe Interact.* **11**, 156–162.
- Mazzola, M. & White, F. F. (1994) *J. Bacteriol.* **176**, 1374–1382.
- Wall, D. & Kaiser, D. (1999) *Mol. Microbiol.* **32**, 1–10.
- Roine, E., Raineri, D. M., Romantschuk, M., Wilson, M. & Nunn, D. N. (1998) *Mol. Plant-Microbe Interact.* **11**, 1048–1056.
- Rojas, C. M., Ham, J. H., Deng, W.-L., Doyle, J. J. & Collmer, A. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 13142–13147.
- Gacesa, P. (1998) *Microbiology* **144**, 1133–1143.
- Yu, J., Peñaloza-Vázquez, A., Chakrabarty, A. M. & Bender, C. L. (1999) *Mol. Microbiol.* **33**, 712–720.
- Peñaloza-Vázquez, A., Kidambi, S. P., Chakrabarty, A. M. & Bender, C. L. (1997) *J. Bacteriol.* **179**, 4464–4472.
- Mittler, R. (2002) *Trends Plant Sci.* **7**, 405–410.
- Bender, C. L. & Cooksey, D. A. (1987) *J. Bacteriol.* **169**, 470–474.
- Silver, S. & Phung, L. T. (1996) *Annu. Rev. Microbiol.* **50**, 753–789.
- Bauer, D. W. & Collmer, A. (1997) *Mol. Plant-Microbe Interact.* **10**, 369–379.
- Niebold, F., Anderson, D. & Mills, D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 406–410.
- Lindgren, P. B., Peet, R. C. & Panopoulos, N. J. (1986) *J. Bacteriol.* **168**, 512–522.
- Hirano, S. S. & Upper, C. D. (1990) *Annu. Rev. Phytopathol.* **28**, 155–177.
- Gomez-Gomez, L. & Boller, T. (2000) *Mol. Cell* **5**, 1003–1011.

Sequencing Methods

Figure 4. Functional classification of annotated ORFs in DC3000.

Figure 5. Distribution of ORFs in paralogous families in DC3000.

Figure 6. Classification of virulence determinants within DC3000 and distribution of these determinants in nine plant and animal bacterial pathogens.

Table 2. Factors implicated in virulence in *P. syringae* pv. *tomato* DC3000 and identification of putative homologs (BLASTPE $<10^{-5}$) in pathogenic bacteria

Table 3. The chromosomal and plasmid genes disrupted by IS elements or phage genes in *P. syringae* pv. *tomato* DC3000

Table 4. IS elements present in *P. syringae* pv. *tomato* DC3000

Table 5. Phage and bacteriocin regions of *P. syringae* pv. *tomato* DC3000

Table 6. Genes with unknown function unique to DC3000 identified from the comparative analysis with the *P. aeruginosa* PAO1 and *P. putida* KT2440 genomes and the identification of putative homologs (BLASTP $E < 10^{-5}$) in pathogenic bacteria

Total *P. syringae* pv. *tomato* DC3000 (DC3000) genomic DNA was nebulized and size-fractionated on agarose gels, and after polishing of the ends with Bal 31 exonuclease and T4 DNA ligase, the DNA was ligated to modified pUC vector as described previously (1). Two libraries, a small insert library (2–3 kbp) and a large insert library (6–14 kbp), were constructed and end sequences were generated using ABI BigDye terminator chemistry using ABI 3700 sequencers (Applied Biosystems). The sequences (73,744 random sequences; average edited length of 697 nt) were assembled by using the TIGR assembler program (2). Sequencing and physical gaps were closed by using a combination of primer walking, transposon-mediated sequencing, PCR, and sequencing of PCR products.

1. Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H., Dodson, R. J., *et al.* (2001) *Science* **293**, 498–506.
2. Sutton, G. G., White, O., Adams, M. D. & Kerlavage, A. R. (1995) *Genome* **1**, 9–19.

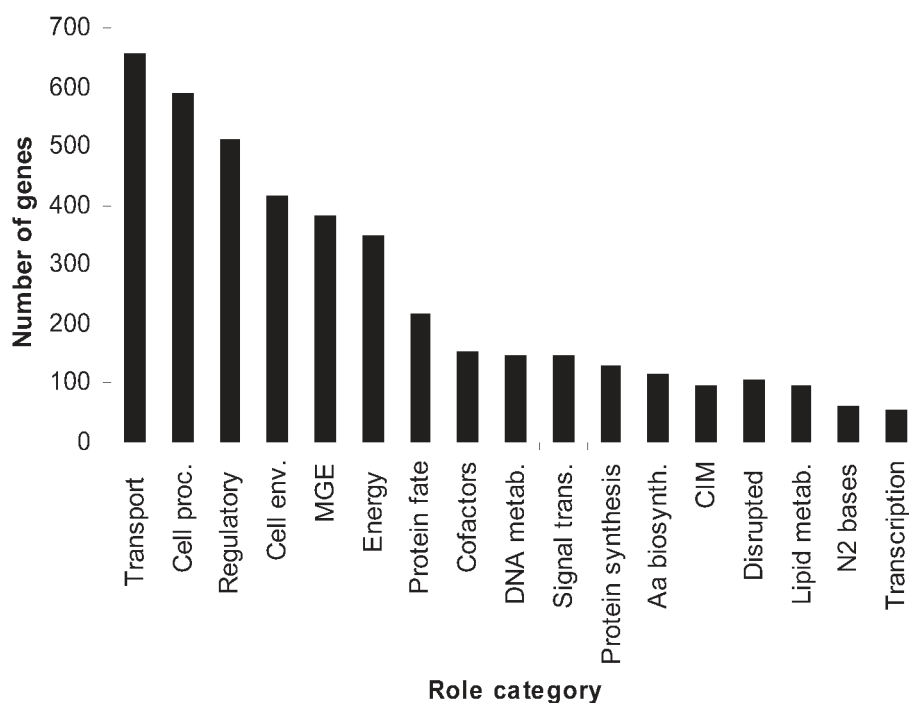


Figure 4. Functional classification of annotated ORFs in DC3000. The ORFs were assigned to role categories adapted from Riley (1) with only the broader categories represented here. Transport, transport and binding function; Cell proc., cellular processes; Regulatory, regulatory functions; Cell env., cellular envelope; MGE, mobile and extrachromosomal element functions; Energy, energy metabolism; Cofactors, biosynthesis of cofactors, prosthetic groups, and carriers; DNA metab., DNA metabolism; Signal trans., signal transduction; Aa biosynth., amino acid biosynthesis; CIM, central intermediary metabolism; Disrupted, disrupted reading frame; Lipid metab., fatty acid and phospholipid metabolism; N2 bases, purines, pyrimidines, nucleosides, and nucleotides.

1. Riley, M. (1993) *Microbiol. Rev.* **57**, 862–952.

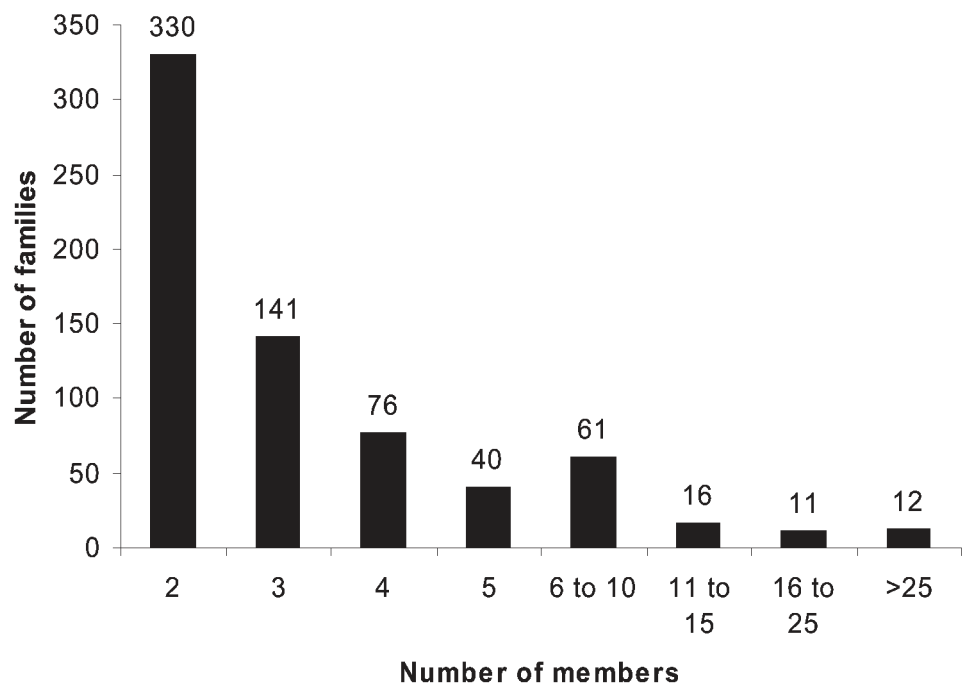


Figure 5. Distribution of ORFs in paralogous families in DC3000. The distribution of ORFs within the 687 paralogous families is shown.

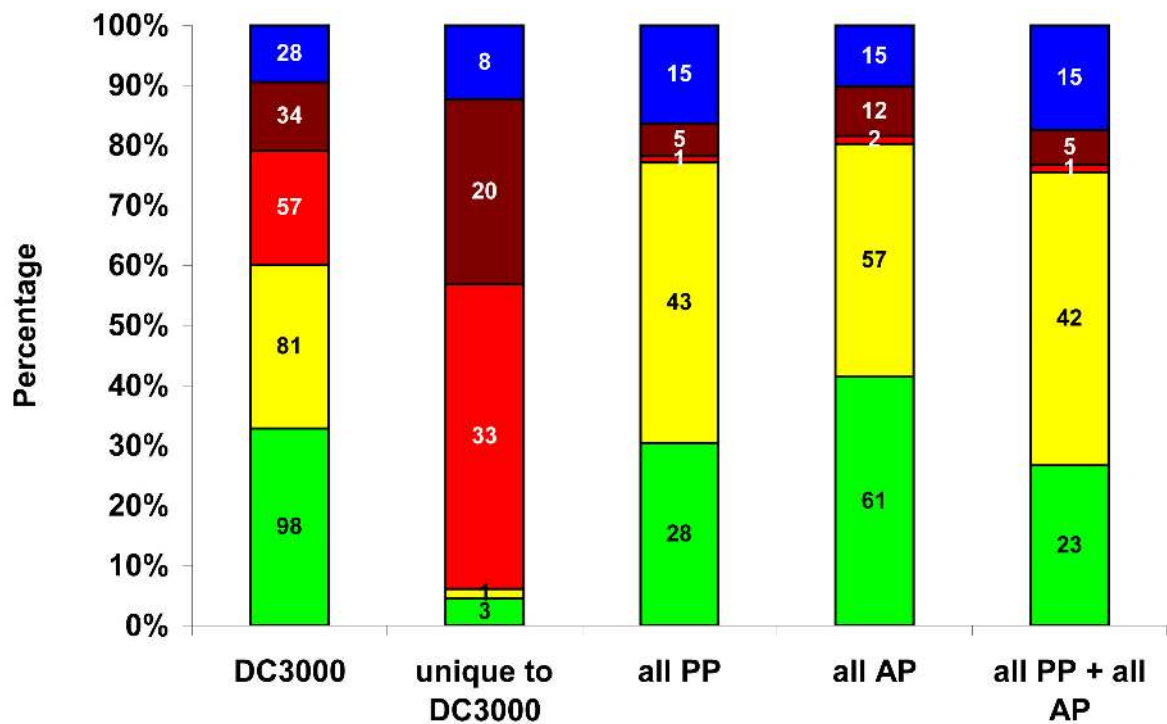


Figure 6. Classification of virulence determinants within DC3000 and distribution of these determinants in nine plant and animal bacterial pathogens. The bars represent all the virulence determinants in DC3000, virulence determinants unique to DC3000, virulence determinants in all plant pathogens (all PP), virulence determinants in all animal pathogens (all AP), and virulence determinants in all plant and animal pathogens (all PP + AP). The DC3000 virulence determinants were categorized into five classes: adhesins and other cell-surface-associated factors (green); miscellaneous virulence factors (yellow); TTSS effectors, candidates, and helpers (red); TTSS regulatory and secretory proteins (brown); and low molecular weight diffusible factors (blue).

Table 2. Factors implicated in virulence in *P. syringae* pv. *tomato* DC3000 and identification of putative homologs (BLASTPE <10⁻⁵) in pathogenic bacteria

Locus	Location ^a	Predicted product	Category	GC%	Rs	Xcc	Xac	At	Xf	Pa	Yp	Ec	St
PSPTO4713	C	alanyl tRNA synthetase-related protein (associated with coronatine)	5	55.3	CAD14690	—	—	AAL42227	—	—	—	AAG57801	—
PSPTO1235	C	alginate biosynthesis protein AlgI	1	60	—	—	—	—	—	AAG06936	—	—	—
PSPTO1241	C	alginate biosynthesis protein Alg44	1	60.5	—	—	—	—	—	AAG06930	—	—	—
PSPTO1242	C	alginate biosynthesis protein Alg8	1	58.8	—	—	—	—	—	AAG06929	—	—	—
PSPTO1232	C	alginate biosynthesis protein AlgA	1	59.8	—	NP_636017	AAM38423	AAL44166	AAF83072	AAG06939	NP_406586	AAG57109	AAL21009
PSPTO1243	C	alginate biosynthesis protein AlgD	1	57.8	CAD14615	NP_636875	AAM36420	AAL44949	AAF84415	AAG06928	NP_405718	AAG57087	AAL20984
PSPTO1239	C	alginate biosynthesis protein AlgE	1	62.1	—	—	—	—	—	AAG06932	—	—	—
PSPTO1233	C	alginate biosynthesis protein AlgF	1	60.4	—	—	—	—	—	AAG06938	—	—	—
PSPTO1238	C	alginate biosynthesis protein AlgG	1	58.1	—	—	—	—	—	AAG06933	—	—	—
PSPTO1234	C	alginate biosynthesis protein AlgJ	1	61.3	—	—	—	—	—	AAG06937	—	—	—
PSPTO1240	C	alginate biosynthesis protein AlgK	1	61.6	—	—	—	—	—	AAG06931	—	—	—
PSPTO1237	C	alginate biosynthesis protein AlgX	1	59.2	—	—	—	—	—	AAG06934	—	—	—
PSPTO1236	C	alginate lyase AlgL	1	57.8	—	—	—	—	—	AAG06935	—	—	—
PSPTO0134	C	alginate regulator AlgQ	1	56.7	—	—	—	—	—	AAG08640	NP_407187	AAG59192	AAL22993
PSPTO0126	C	alginate regulator FimS	1	57.9	CAD15053	NP_638859	AAM35509	—	AAF84434	AAG08647	NP_407384	AAG57194	AAL21063
PSPTO0334	C	alginate regulatory protein AlgB	1	61.8	CAD13860	NP_635584	AAM35118	AAL44574	AAF85342	AAG08868	NP_406415	AAG59201	AAL23002
PSPTO0127	C	alginate regulatory protein AlgR	1	58.9	CAD15054	NP_638858	AAM35510	AAL44596	AAF84435	AAG08646	NP_406755	AAG57193	AAL21062
PSPTO0136	C	alginate regulatory protein AlgR3	1	61.2	CAD16500	NP_638223	AAM38815	AAL42873	—	AAG08638	NP_406242	AAG56007	AAL21483
PSPTO4223	C	alginate regulatory protein MucA	1	59.2	—	NP_636642	—	—	—	AAG04152	NP_406229	AAG57688	AAL21533
PSPTO4222	C	alginate regulatory protein MucB	1	55.3	CAD14759	—	—	—	—	AAG04153	NP_406230	AAG57687	AAL21532
PSPTO2816	C	alkylhydroperoxidase AhpD domain protein	2	61.4	CAD15590	—	—	AAL46114	—	AAG03658	NP_405897	—	—
PSPTO4204	C	amidase family (indoleacetamide hydrolase IaaH)	5	66.1	CAD14119	NP_636310	AAM35198	AAL44059	—	AAG04093	NP_406731	—	—
PSPTO3864	C	autoinducer synthesis protein PsyI	2	55.5	CAD17930	—	—	AAL46278	—	AAG04821	NP_404601	—	—
PSPTO3863	C	autoinducer transcriptional regulator PsyR	2	51.8	CAD17075	NP_638166	AAM37021	AAL43274	—	AAG05287	NP_405991	AAG56931	AAL20862
PSPTO3582	C	catalase KatB	2	58.3	CAD18732	NP_639288	AAM38864	AAL46177	—	AAG08001	NP_404811	AAG56718	AAL20243
PSPTO5263	C	catalase KatE	2	59	CAD18732	NP_636484	AAM36083	AAL46177	—	AAG05535	NP_404811	AAG56718	AAL20243
PSPTO4530	C	catalase/peroxidase KatG	2	61.1	CAD14477	NP_636579	AAM36172	AAL45436	AAF85031	—	NP_406785	AAG59143	AAL22946
PSPTO4683	C	coronafacic acid beta-ketoacyl synthetase component Cfa3	5	67.1	CAD14756	NP_636396	AAM36002	AAL45010	AAF83483	AAG04762	NP_405182	AAG55841	AAL20126
PSPTO4686	C	coronafacic acid polyketide synthase I Cfa6	5	67.6	CAD17792	NP_636396	AAM36002	AAL44484	AAF83483	AAG06356	NP_405471	AAG55841	AAL20126

PSPTO4687	C	coronafacic acid polyketide synthetase II Cfa7	5	68.1	CAD15508	NP_636393	AAM36002	AAL44484	AAF83483	AAG04762	NP_405471	AAG55664	AAL20126
PSPTO4684	C	coronafacic acid synthetase component Cfa4	5	65.3	–	–	–	–	–	–	–	–	–
PSPTO4681	C	coronafacic acid synthetase, acyl carrier protein component Cfa1	5	59	–	–	–	–	–	–	–	–	–
PSPTO4707	C	coronafacic acid synthetase, acyl carrier protein component, putative	5	50.9	–	–	–	–	–	–	–	–	–
PSPTO4682	C	coronafacic acid synthetase, dehydratase component Cfa2	5	62.2	CAD15117	NP_636736	AAM36281	AAL42389	AAF83854	AAG07033	NP_404668	AAG54482	AAL19191
PSPTO4685	C	coronafacic acid synthetase, ligase component Cfa5	5	65	CAD13970	NP_637223	AAM36742	AAL42672	AAF83100	AAG05812	NP_404418	AAG56794	AAL20733
PSPTO4680	C	coronafacic acid synthetase, ligase component Cfl	5	62.5	CAD18416	NP_636198	AAM35769	AAL45616	AAF83100	AAG05943	NP_405626	AAG56688	AAL20275
PSPTO4690	C	coronafacic acid synthetase, thioesterase component Cfa9	5	59.5	CAD15506	–	–	AAL44497	–	AAG05799	NP_405469	–	–
PSPTO4709	C	coronamic acid synthetase, alioisoleucine ligase component CmaA	5	56.9	CAD17793	NP_639207	AAM36952	AAL44495	AAF83100	AAG05790	NP_404403	AAG54921	AAL19539
PSPTO4710	C	coronamic acid synthetase, chlorinating enzyme component, putative	5	55.4	–	–	–	–	–	–	–	–	–
PSPTO4711	C	coronamic acid synthetase, cyclase component, putative	5	57.9	–	–	–	–	–	–	–	–	–
PSPTO4712	C	coronamic acid synthetase, thioesterase component CmaT	5	58.2	CAD15506	–	–	AAL44497	–	AAG07617	NP_404399	–	–
PSPTO4708	C	coronatine biosynthesis protein, putative	5	56.3	–	–	–	–	–	–	–	–	–
PSPTO4706	C	coronatine response regulator CorP	5	57.9	–	–	–	–	–	–	–	–	–
PSPTO4704	C	coronatine response regulator CorR	5	61.3	CAD13820	NP_638695	AAM38323	AAL41503	AAF83782	AAG06992	NP_406531	AAG56081	AAL20682
PSPTO4705 ^b	C	coronatine sensor histidine kinase CorS (authentic frameshift)	5	59.9	CAD16163	NP_638694	AAM38316	AAL41545	AAF83783	AAG05380	NP_403906	AAG57899	AAL20316
PSPTO4689	C	crotonyl-CoA reductase Cfa8 (associated with coronatine)	5	62.3	CAD16839	NP_635424	AAM34923	AAL41642	AAF84543	AAG08812	NP_403969	AAG56765	AAL20485
PSPTO3598	C	dyp-type peroxidase family protein	2	63.5	–	–	–	–	–	–	NP_405420	AAG55637	–
PSPTO1029	C	endo-1,4-beta-D-glucanase	2	65.1	–	–	AAM38359	AAL44120	–	–	–	AAG58673	AAL22477
PSPTO3960	C	endo-polygalacturonase	2	59.4	CAD18031	NP_637621	AAM37226	AAL43945	–	–	–	–	–
PSPTO3210	C	filamentous hemagglutinin family protein	1	55.4	CAD15477	NP_637159	AAM36677	AAL44779	AAF83699	AAG03431	NP_406024	–	–
PSPTO3229	C	filamentous hemagglutinin, intein-containing, putative	1	60.1	CAD17691	NP_637159	AAM36677	AAL44776	AAF83699	AAG05850	NP_404241	AAG55657	AAL21407

PSPTO3214 ^b	C	filamentous hemagglutinin family protein, putative (authentic point mutation)	1	59.1	CAD15477	NP_637159	AAM36677	—	AAF84995	AAG03431	NP_406024	AAG54838	AAL21407
PSPTO1925	C	flagellar anti-transcription factor FlgM	1	54.2	—	NP_637320	AAM36851	—	—	AAG06739	—	—	—
PSPTO1960	C	flagellar assembly protein FlhI, putative	1	58.7	CAD17543	NP_637289	AAM36814	—	—	AAG04492	NP_404348	AAG56955	AAL20883
PSPTO1935	C	flagellar basal-body rod modification protein FlgD	1	57.8	CAD17495	NP_637315	AAM36846	AAL41596	—	AAG04468	NP_405366	AAG55821	AAL20106
PSPTO1933	C	flagellar basal-body rod protein FlgB	1	57.3	CAD17493	NP_637317	AAM36848	—	—	AAG04466	NP_405364	AAG55819	AAL20104
PSPTO1934	C	flagellar basal-body rod protein FlgC	1	56.2	CAD17494	NP_637316	AAM36847	AAL41571	—	AAG04467	NP_404355	AAG55820	AAL20105
PSPTO1939	C	flagellar basal-body rod protein FlgF	1	61.5	CAD17497	NP_637313	AAM36844	AAL41575	—	AAG04470	NP_405368	AAG55823	AAL20108
PSPTO1940	C	flagellar basal-body rod protein FlgG	1	61.2	CAD17498	NP_637312	AAM36843	AAL41569	—	AAG04471	NP_405369	AAG55824	AAL20109
PSPTO1976	C	flagellar biosynthesis protein FlhA	1	58.6	CAD18544	NP_637274	AAM36798	AAL41598	—	AAG04841	NP_405358	AAG56869	AAL20829
PSPTO1975	C	flagellar biosynthetic protein FlhB	1	59.2	CAD18545	NP_637275	AAM36799	AAL41581	—	AAG04838	NP_405357	AAG56870	AAL20830
PSPTO1972	C	flagellar biosynthetic protein FlhP	1	57.1	CAD17526	NP_637281	AAM36806	AAL41563	—	AAG04835	NP_405386	AAG56963	AAL20891
PSPTO1973	C	flagellar biosynthetic protein FlhQ	1	58.8	CAD17525	NP_637280	AAM36804	AAL41597	—	AAG04836	NP_405385	AAG56964	AAL20892
PSPTO1974	C	flagellar biosynthetic protein FlhR	1	59.6	CAD17524	NP_637279	AAM36803	AAL41599	—	AAG04837	NP_405384	AAG56965	AAL20893
PSPTO1936	C	flagellar hook protein FlgE	1	57.5	CAD17496	NP_637314	AAM36845	AAL41591	AAF84783	AAG04469	NP_405367	AAG55822	AAL20107
PSPTO1937	C	flagellar hook protein FlgE	1	60.8	CAD17496	NP_637314	AAM36845	AAL41591	—	AAG04469	NP_405367	AAG55822	AAL20107
PSPTO1944	C	flagellar hook-associated protein FlgK	1	59.8	CAD17502	NP_637308	AAM36839	AAL41592	AAF84783	AAG04475	NP_405373	AAG55828	AAL20113
PSPTO1945	C	flagellar hook-associated protein FlgL	1	59.7	CAD17503	NP_637307	AAM36838	AAL41584	—	AAG04476	NP_405374	AAG55829	AAL20114
PSPTO1951	C	flagellar hook-associated protein FlhD, FRAMESHIFT	1	58.7	CAD17534	NP_637305	AAM36836	AAL44779	AAF84783	AAG04483	NP_405405	AAG56939	AAL20872
PSPTO1957	C	flagellar hook-basal body complex protein FlhE	1	55.4	CAD17540	NP_637292	AAM36817	—	—	AAG04489	NP_405397	AAG56952	AAL20880
PSPTO1966	C	flagellar hook-length control protein FlhK	1	61.7	CAD17546	NP_637286	AAM36811	—	—	AAG04830	NP_405391	AAG56958	AAL20886
PSPTO1941	C	flagellar L-ring protein FlgH	1	61.3	CAD17499	NP_637311	AAM36842	AAL41565	—	AAG04472	NP_404360	AAG55825	AAL20110
PSPTO1959	C	flagellar motor switch protein FlhG	1	58.5	CAD17542	NP_637290	AAM36815	AAL41580	—	AAG04491	NP_405395	AAG56954	AAL20882
PSPTO1969	C	flagellar motor switch protein FlhM	1	58.9	CAD17529	NP_637284	AAM36809	—	—	AAG04832	NP_405389	AAG56960	AAL20888
PSPTO1970	C	flagellar motor switch protein FlhN	1	60.7	CAD17528	NP_637283	AAM36808	AAL41579	—	AAG04833	NP_405388	AAG56961	AAL20889
PSPTO1958	C	flagellar M-ring protein FlhF	1	60.3	CAD17541	NP_637291	AAM36816	AAL41541	—	AAG04490	NP_405396	AAG56953	AAL20881
PSPTO1942	C	flagellar P-ring protein FlgI	1	61.7	CAD17500	NP_637310	AAM36841	AAL41567	—	AAG04473	NP_404361	AAG55826	AAL20111
PSPTO1962	C	flagellar protein FlhJ, putative	1	61.3	CAD17545	NP_637287	AAM36812	—	—	AAG04494	NP_405392	AAG56957	AAL20885

PSPTO1968	C	flagellar protein FliL, putative	1	58.3	CAD17530	NP_637285	AAM36810	—	—	AAG04831	NP_405390	AAG56959	AAL20887
PSPTO5230	C	flagellar protein FliL, putative	1	56	—	NP_637285	AAM36810	—	—	AAG08618	—	—	—
PSPTO1971	C	flagellar protein FliO	1	63.2	CAD17527	NP_637282	AAM36807	—	—	AAG04834	NP_405387	AAG56962	AAL20890
PSPTO1952	C	flagellar protein FliS	1	58.7	CAD17535	NP_637304	AAM36835	—	—	AAG04484	NP_405404	AAG56940	AAL20873
PSPTO1926	C	flagellar protein, putative	1	58.6	CAD17492	NP_637319	AAM36850	AAL41568	—	AAG06738	NP_404353	—	AAL20103
PSPTO1954	C	flagellar regulator FleQ	1	59.3	CAD17379	NP_637298	AAM36829	AAL42452	AAF85342	AAG04486	NP_403688	AAG59201	AAL23002
PSPTO1956	C	flagellar response regulator FleR	1	61.6	CAD13860	NP_635584	AAM38082	AAL42452	AAF84654	AAG04488	NP_404344	AAG59201	AAL23002
PSPTO1955	C	flagellar sensor histidine kinase FleS	1	63	CAD13859	NP_638780	AAM35574	AAL45099	AAF84264	AAG04487	NP_404513	AAG59200	AAL23001
PSPTO1978	C	flagellar synthesis regulator FleN	1	60.6	CAD16862	NP_637272	AAM36796	AAL44064	AAF84130	AAG04843	NP_405629	AAG56026	AAL20730
PSPTO1950	C	flagellin FlaG, putative	1	51	—	—	—	—	—	AAG04482	—	—	—
PSPTO1949	C	flagellin FliC	1	55.9	CAD17533	NP_637306	AAM36837	AAL41559	AAF84325	AAG04481	NP_405406	AAG56938	AAL20871
PSPTO1961	C	flagellum-specific ATP synthase FliI	1	64.1	CAD17544	NP_637288	AAM36813	AAL41574	AAF83953	AAG04493	NP_405393	AAG56956	AAL20884
PSPTO3637	C	FlhB domain protein	1	63.9	CAD18545	NP_637275	AAM36799	AAL41581	—	AAG04862	NP_405357	AAG56870	AAL20830
PSPTO3307	C	general secretion pathway protein GspD	2	60.9	CAD16680	NP_636062	AAM38377	AAL41244	AAF84336	AAG04074	NP_404440	AAG58491	AAL22350
PSPTO3317	C	general secretion pathway protein GspE	2	62.5	CAD16824	NP_636052	AAM38387	AAL46277	AAF84326	AAG06065	NP_404439	AAG54411	AAL19107
PSPTO3316	C	general secretion pathway protein GspF	2	62.8	CAD16825	NP_636053	AAM38386	—	AAF84327	AAG04076	NP_406889	AAG54410	AAL19106
PSPTO3315	C	general secretion pathway protein GspG	2	59.2	CAD16017	NP_636054	AAM38385	—	AAF84328	AAG06489	NP_404438	—	—
PSPTO3314	C	general secretion pathway protein GspH	2	62.5	—	NP_636055	AAM38384	—	AAF84329	—	—	—	—
PSPTO3313	C	general secretion pathway protein GspI, putative	2	62.7	—	NP_636056	AAM38383	—	—	AAG06061	—	—	—
PSPTO3312	C	general secretion pathway protein GspJ, putative	2	63.8	—	NP_636057	AAM38382	—	AAF84331	—	—	—	—
PSPTO3311	C	general secretion pathway protein GspK, putative	2	60.9	—	NP_636058	AAM38381	—	AAF84332	—	—	—	—
PSPTO3310	C	general secretion pathway protein GspL, putative	2	63.1	—	NP_636059	AAM38380	—	AAF84333	—	—	—	—
PSPTO3309	C	general secretion pathway protein GspM, putative	2	64.5	—	NP_636060	AAM38379	—	AAF84334	—	—	—	—
PSPTO3308	C	general secretion pathway protein GspN, putative	2	66.1	—	—	—	—	AAF84335	—	—	—	—
PSPTO1172	C	glutathione peroxidase family protein	2	57.4	CAD16381	NP_636786	AAM36327	—	AAF84696	AAG04227	NP_405957	AAG56697	AAL20266
PSPTO1719	C	glutathione peroxidase family protein	2	57.4	CAD13835	NP_636873	AAM36418	—	AAF84413	AAG04676	NP_405957	AAG56697	AAL20266

PSPTO1781	C	glutathione peroxidase family protein	2	56.1	CAD16381	NP_636786	AAM36327	—	AAF84696	AAG06214	NP_405957	AAG56697	AAL20266
PSPTO3112	C	glutathione reductase	2	58.4	CAD13542	NP_637927	AAM37595	AAL42613	AAF83678	AAG05413	NP_407417	AAG58632	AAL22457
PSPTO4254	C	glutathione reductase	2	54.7	CAD13542	NP_637927	AAM37595	AAL42613	AAF83678	AAG05413	NP_407417	AAG58632	AAL22457
PSPTO3343	C	glutathione S-transferase domain protein	2	56.5	—	NP_638907	AAM35474	AAL43250	—	AAG05998	—	AAG58235	AAL22106
PSPTO0929	C	glutathione S-transferase family	2	56.8	CAD14519	NP_636315	AAM37083	AAL45840	AAF84020	AAG04855	NP_405903	AAG56624	AAL20373
PSPTO2032	C	glutathione S-transferase family	2	56.7	CAD15723	NP_636804	AAM36344	AAL41850	AAF84020	AAG05044	NP_407094	AAG57431	AAL22014
PSPTO3856	C	glutathione S-transferase family	2	62	CAD14686	NP_639111	AAM38661	AAL44512	AAF84020	AAG06201	NP_404909	AAG55214	AAL19798
PSPTO4398	C	glutathione S-transferase family	2	62.1	CAD13614	NP_636213	AAM35782	AAL42400	—	AAG07789	NP_405903	AAG57430	AAL21249
PSPTO0703	C	glutathione S-transferase family protein	2	58.9	CAD15723	NP_639301	AAM38882	AAL44397	—	AAG04422	NP_407094	AAG58126	AAL22014
PSPTO1365	C	glutathione S-transferase family protein	2	58	CAD14137	NP_637641	AAM37246	AAL41850	—	AAG06201	NP_404909	AAG55214	AAL23091
PSPTO4777	C	glutathione S-transferase family protein	2	58.5	CAD14686	NP_639111	AAM39182	AAL41850	AAF84020	AAG05279	NP_404909	AAG55214	AAL19798
PSPTO0448	C	glutathione S-transferase, putative	2	61.4	CAD16428	—	—	AAL42145	—	AAG03856	—	AAG58736	AAL22543
PSPTO3534	C	glycosyl hydrolase, family 5 (cellulolytic activity)	2	62.7	—	—	—	—	—	AAG05625	—	—	—
PSPTO3230	C	hemolysin activator protein, HlyB family	2	60.8	CAD14590	NP_637158	AAM36676	—	AAF85347	AAG05851	NP_407171	AAG55658	—
PSPTO1309	C	HlyD family secretion protein	2	61.4	CAD18348	NP_639132	AAM38690	AAL43469	AAF85183	AAG04824	NP_406355	AAG57137	AAL21030
PSPTO2753	C	HlyD family secretion protein	2	58.6	CAD18348	NP_639132	AAM38690	AAL43469	AAF85183	AAG05916	NP_406355	AAG57137	AAL21030
PSPTO3620	C	HlyD family secretion protein	2	63.9	CAD16433	NP_639503	AAM39135	AAL46225	AAF84892	AAG06692	NP_407139	AAG56633	AAL22234
PSPTO3703	C	HlyD family secretion protein	2	62.8	CAD16433	NP_638835	AAM35530	AAL42100	—	AAG04626	NP_406737	AAG57794	AAL21699
PSPTO0371	C	indoleacetate-lysine ligase IaaL	5	53.3	—	—	—	—	—	—	—	—	—
PSPTO1453	C	levansucrase	1	56.3	—	—	—	—	—	—	—	—	—
PSPTO2305	C	levansucrase	1	57.3	—	—	—	—	—	—	—	—	—
PSPTOA0032	pA	levansucrase	1	55.6	—	—	—	—	—	—	—	—	—
PSPTO1378	C	membrane-bound lytic murein transglycosylase D, putative (CEL ORF1)	2	62.1	CAD15218	NP_636362	AAM35961	—	AAF84172	AAG05201	NP_404692	AAG54507	AAL19217
PSPTO0405	C	methionine sulfoxide reductase	2	58.8	CAD14294	NP_636219	AAM35788	AAL41149	AAF84742	AAG08403	NP_406984	AAG59417	AAL23228
PSPTO2829	C	non-ribosomal peptide synthase	5	63.5	CAD17793	NP_639207	AAM36952	AAL44495	AAF83100	AAG05787	NP_405472	AAG54921	AAL19539
PSPTO2830	C	non-ribosomal peptide synthase	5	64.4	CAD17793	NP_639207	AAM36952	AAL44495	AAF83100	AAG05787	NP_405472	AAG54921	AAL19539
PSPTO4518	C	non-ribosomal peptide synthase	5	58	CAD17792	NP_639207	AAM36951	AAL44495	AAF83100	AAG05812	NP_405472	AAG54921	AAL19539
PSPTO4519	C	non-ribosomal peptide synthase	5	59	CAD17792	NP_639207	AAM36951	AAL44495	AAF83100	AAG05787	NP_405472	AAG54921	AAL19539
PSPTO4699	C	non-ribosomal peptide synthase	5	59.3	CAD17793	NP_639207	AAM36952	AAL44495	AAF83100	AAG05790	NP_405472	AAG54921	AAL19539
PSPTO4283	C	pectin lyase	2	59.5	—	—	—	—	—	—	—	—	—
PSPTO1943	C	peptidoglycan hydrolase FlgJ	1	61.6	CAD17501	NP_637309	AAM36840	—	—	AAG04474	NP_405372	AAG55827	AAL20112

Table 2, page 5

PSPTO5162	C	periplasmic glucan biosynthesis protein MdoG	1	58.2	CAD16615	NP_639484	AAM39119	–	AAF85479	AAG08463	–	AAG55794	AAL20080
PSPTO5161	C	periplasmic glucan biosynthesis protein MdoH	1	60.3	CAD16616	NP_638861	AAM35507	AAL44122	AAF84432	AAG08462	–	AAG55795	AAL20081
PSPTO3440	C	polysaccharide biosynthesis protein, putative	1	61	–	NP_637799	–	AAL44858	–	–	–	–	–
PSPTO2153	C	pyoverdine ABC transporter, ATP-binding/permease protein PvdE	2	62	CAD17789	NP_636854	AAM36399	AAL44554	AAF85379	AAG05785	NP_405474	AAG57346	AAL21165
PSPTO2146	C	pyoverdine biosynthesis regulatory gene, putative	2	60.7	–	–	–	–	–	AAG05691	–	–	–
PSPTO2135	C	pyoverdine chromophore precursor synthetase PvsA	2	63.4	CAD17792	NP_639207	AAM36951	AAL44495	AAF83100	AAG05812	NP_405472	AAG54921	AAL19539
PSPTO2147	C	pyoverdine sidechain peptide synthetase I, epsilon-Lys module	2	64.5	CAD17792	NP_639207	AAM36951	AAL44494	AAF85054	AAG05788	NP_405472	AAG54921	AAL19539
PSPTO2148	C	pyoverdine sidechain peptide synthetase II, D-Asp-L-Thr component	2	61.4	CAD17792	NP_639207	AAM36952	AAL44495	AAF83100	AAG05790	NP_404403	AAG54921	AAL19539
PSPTO2149	C	pyoverdine sidechain peptide synthetase III, L-Thr-L-Ser component	2	61.9	CAD17792	NP_639207	AAM36951	AAL44495	AAF83100	AAG05787	NP_405472	AAG54921	AAL19539
PSPTO2150	C	pyoverdine sidechain peptide synthetase IV, D-Asp-L-Ser component	2	62.2	CAD17792	NP_639207	AAM36952	AAL44495	AAF83100	AAG05790	NP_404403	AAG54921	AAL19539
PSPTO2134	C	pyoverdine synthetase, thioesterase component	2	63.3	CAD15506	–	–	AAL44497	–	AAG05813	NP_404399	–	–
PSPTO3024	C	response regulator GacA	2	56.2	CAD13820	NP_637450	AAM36963	AAL45956	AAF85405	AAG05974	NP_405429	AAG56929	AAL20859
PSPTO4224	C	RNA polymerase sigma-24 factor AlgT	1	54.1	CAD14757	NP_636641	AAM36190	AAL43900	AAF85038	AAG04151	NP_406228	AAG57689	AAL21534
PSPTO1565	C	RNA polymerase sigma-38 factor RpoS	2	56.4	CAD14909	NP_639081	AAM38630	AAL43156	AAF84159	AAG07010	NP_406818	AAG57848	AAL21804
PSPTO4453	C	RNA polymerase sigma-54 factor RpoN	2	58.2	CAD13936	NP_638150	AAM37817	AAL41354	AAF84217	AAG07850	NP_407039	AAG58336	AAL22189
PSPTO1691	C	sensor histidine kinase/response regulator GacS	2	62.6	CAD16788	NP_635875	AAM35383	AAL41365	AAF83924	AAG04317	NP_407012	AAG57899	AAL21838
PSPTO1675	C	siderophore biosynthesis protein, putative	2	63.7	–	–	–	AAL45210	–	AAG04554	NP_405163	AAG54917	AAL19535
PSPTO4156	C	sodium-type flagellar protein MotY, putative	1	62.2	CAD17912	NP_636809	AAM36349	AAL43703	AAF83173	AAG06914	NP_404090	AAG57716	AAL19693
PSPTO1338	C	superoxide dismutase, Cu-Zn	2	59.9	CAD16075	NP_635586	AAM35102	–	–	–	NP_406837	AAG57201	AAL19860
PSPTO4363	C	superoxide dismutase, Fe	2	59.2	CAD16233	NP_637851	AAM37524	AAL41890	AAF85411	AAG07754	NP_405922	AAG56645	AAL20353
PSPTO4459	C	superoxide dismutase, Mn	2	56.3	CAD16233	NP_637633	AAM37238	AAL41890	AAF85411	AAG07856	NP_407488	AAG59102	AAL22895
PSPTO2752	C	thiol peroxidase	2	58.4	CAD16497	–	–	AAL42826	–	AAG05920	NP_405879	AAG56478	AAL20599

PSPTO0341	C	thiol:disulfide interchange protein DsbA	2	59.3	CAD13813	NP_638746	AAM35611	—	AAF84246	AAG08874	NP_403682	AAG59049	AAL22836
PSPTO0784	C	TonB dependent siderophore receptor, putative	2	60.3	CAD17962	NP_638414	AAM38463	AAL44724	AAF84933	AAG05310	NP_404908	AAG57293	AAL21103
PSPTO2151	C	TonB dependent siderophore receptor, putative	2	61	CAD16627	NP_638941	AAM35068	AAL45999	AAF83409	AAG05786	NP_405123	AAG55848	AAL20133
PSPTO2152	C	TonB dependent siderophore receptor, putative	2	61.3	CAD18213	NP_638941	AAM35068	AAL44823	AAF83409	AAG07609	NP_405123	AAG55848	AAL20133
PSPTO3462	C	TonB-dependent siderophore receptor	2	61.6	CAD17565	NP_636066	AAM38373	AAL44823	AAF83409	AAG05298	NP_405322	AAG58168	AAL19318
PSPTO4128	C	TonB-dependent siderophore receptor	2	59.1	CAD17565	NP_636066	AAM38373	AAL45999	AAF83409	AAG05298	NP_405322	AAG58168	AAL20133
PSPTO1206	C	TonB-dependent siderophore receptor, putative	2	60.8	CAD16627	NP_638941	AAM37038	AAL45999	AAF83409	AAG03859	NP_404572	AAG54454	AAL19155
PSPTO1610	C	TonB-dependent siderophore receptor, putative	2	61.9	CAD17565	NP_636066	AAM38372	AAL45999	AAF83409	AAG05298	NP_405322	AAG54454	AAL20133
PSPTO2463	C	TonB-dependent siderophore receptor, putative	2	60.5	CAD17251	NP_639060	AAM38599	AAL45999	AAF83409	AAG08222	NP_404572	AAG58168	AAL19155
PSPTO2484	C	TonB-dependent siderophore receptor, putative	2	60	CAD16107	NP_638390	AAM38010	AAL44724	AAF84936	AAG05978	NP_407355	AAG54918	AAL19536
PSPTO2605	C	TonB-dependent siderophore receptor, putative	2	60.6	CAD15509	NP_635493	AAM38045	AAL44499	AAF83409	AAG07543	NP_405467	AAG55323	AAL20505
PSPTO2607	C	TonB-dependent siderophore receptor, putative	2	63.3	CAD17962	NP_638426	AAM38052	AAL44724	AAF84933	AAG05477	NP_404908	AAG54918	AAL21663
PSPTO2846	C	TonB-dependent siderophore receptor, putative	2	56.5	CAD18213	NP_639060	AAM38599	AAL45999	AAF83409	AAG08222	NP_405123	AAG58168	AAL19155
PSPTO3574	C	TonB-dependent siderophore receptor, putative	2	59.8	CAD16436	NP_639060	AAM38599	AAL45999	AAF83409	AAG03541	NP_404572	AAG58168	AAL19318
PSPTO3692	C	TonB-dependent siderophore receptor, putative	2	59.9	CAD18213	NP_638864	AAM38341	AAL44823	AAF83409	AAG07609	NP_405123	AAG55848	AAL20133
PSPTO0518	C	tryptophan 2-monooxygenase IaaM, putative	5	60.9	—	NP_638624	AAM38269	AAL46247	—	—	—	—	—
PSPTO4589	C	TTSS chaperone candidate protein	4	59.2	—	—	—	—	—	—	—	—	—
PSPTO4599	C	TTSS chaperone candidate protein	4	58	—	—	—	—	—	—	—	—	—
PSPTOA0017	pA	TTSS chaperone candidate protein	4	54.9	—	—	—	—	—	—	—	—	—
PSPTO1376	C	TTSS chaperone candidate protein AvrF	4	56.7	—	—	—	—	—	—	—	—	—
PSPTO5353	C	TTSS chaperone protein ShcA(Pto)	4	54.9	—	—	—	—	—	—	NP_395207	—	—
PSPTO0503	C	TTSS chaperone protein ShcF	4	43.3	—	—	—	—	—	—	—	—	—
PSPTO1374	C	TTSS chaperone protein ShcM	4	57.3	—	—	—	—	—	—	—	—	—
PSPTO1369	C	TTSS chaperone protein ShcN	4	55.1	—	—	—	—	—	—	—	—	—
PSPTO4721	C	TTSS chaperone protein ShcV	4	50	—	—	—	—	—	—	—	—	—

Table 2, page 7

PSPTO1384	C	TTSS component protein HrcJ	4	59.3	CAD18018	NP_636607	AAM35300	AAL41541	—	AAG05112	NP_395193	AAG57972	AAL20333
PSPTO1400	C	TTSS component protein HrcN	4	65.5	CAD18021	NP_636610	AAM35303	AAL41574	AAF83953	AAG05086	NP_395174	AAG58832	AAL20339
PSPTO1397 ^c	C	TTSS component protein HrcQa	4	64.1	—	—	—	—	—	—	—	—	—
PSPTO1396 ^c	C	TTSS component protein HrcQb	4	60.8	—	—	—	—	—	—	NP_405388	—	—
PSPTO1395	C	TTSS component protein HrcR	4	58.2	CAD18011	NP_636600	AAM35293	AAL41563	—	AAG05082	NP_395178	AAG58847	AAL20343
PSPTO1394	C	TTSS component protein HrcS	4	58.3	CAD18010	NP_636599	AAM35292	AAL41597	—	AAG05081	NP_395179	AAG58846	AAL20344
PSPTO1393	C	TTSS component protein HrcT	4	61	CAD18023	NP_636612	AAM35305	AAL41599	—	AAG05080	NP_403924	AAG57979	AAL21768
PSPTO1392	C	TTSS component protein HrcU	4	57.2	CAD18015	NP_636604	AAM35297	AAL41581	—	AAG05079	NP_403925	AAG58844	AAL20346
PSPTO1402	C	TTSS component protein HrcV	4	59.8	CAD18014	NP_637274	AAM36798	AAL41598	—	AAG05092	NP_403917	AAG58833	AAL21776
PSPTO1383	C	TTSS component protein HrpB	4	58.3	—	—	—	—	—	—	—	—	—
PSPTO1385	C	TTSS component protein HrpD	4	63.6	—	—	—	—	—	—	—	—	—
PSPTO1386	C	TTSS component protein HrpE	4	63.6	—	—	—	—	—	—	—	—	—
PSPTO1387	C	TTSS component protein HrpF	4	54.5	—	—	—	—	—	—	—	—	—
PSPTO1388	C	TTSS component protein HrpG	4	63.9	—	—	—	—	—	—	—	—	—
PSPTO1403	C	TTSS component protein HrpJ	4	59.4	—	—	—	—	—	—	—	—	—
PSPTO1399	C	TTSS component protein HrpO	4	58.1	—	—	—	—	—	—	—	—	—
PSPTO1398	C	TTSS component protein HrpP	4	62.3	—	—	—	—	—	—	—	—	—
PSPTO1401	C	TTSS component protein HrpQ	4	62.5	—	—	—	—	—	—	—	—	—
PSPTO1390	C	TTSS component protein HrpT	4	58.7	—	—	—	—	—	—	—	—	—
PSPTO1377	C	TTSS effector avirulence protein AvrE(Pto)	3	57.6	CAD18432	—	—	—	—	—	—	—	—
PSPTOA0012	pA	TTSS effector avirulence protein AvrPphE(Pto)	3	53.8	CAD16866	NP_636999	AAM38068	—	—	—	—	—	—
PSPTO1022	C	TTSS effector avirulence protein AvrPpiB1(Pto)	3	39.9	—	—	—	—	—	—	—	—	—
PSPTOA0005	pA	TTSS effector avirulence protein AvrPpiB2(Pto)	3	39.9	—	—	—	—	—	—	—	—	—
PSPTO4001	C	TTSS effector avirulence protein AvrPto(DC3000)	3	49.6	—	—	—	—	—	—	—	—	—
PSPTO3087	C	TTSS effector avirulence protein AvrPtoB	3	59.2	—	—	—	—	—	—	—	—	—
PSPTO0901 ^d	C	TTSS effector candidate HolPsyAG(Pto)	3	53.1	—	—	—	—	—	—	—	—	—
PSPTO0905	C	TTSS effector candidate HolPsyAH(Pto) (glycosyl hydrolase, family 5 (cellulolytic activity))	3	58.5	—	—	—	—	—	—	—	—	—
PSPTO0906	C	TTSS effector candidate HolPsyAI(Pto)	3	54	—	—	—	—	—	—	—	—	—
PSPTO4703	C	TTSS effector candidate HolPtoAA	3	50.6	—	—	—	—	—	—	—	—	—

PSPTO4993 ^d , 4996 ^d	C	TTSS effector candidate HolPtoAC	3	61.1	—	—	—	—	—	—	—	—	—
PSPTO4732	C	TTSS effector candidate HolPtoQ2	3	48.8	CAD13773	NP_636447	AAM39163	—	—	—	—	—	—
PSPTO0883	C	TTSS effector candidate HolPtoR	3	59.4	CAD18432	NP_636464	—	—	—	—	—	—	—
PSPTO4593	C	TTSS effector candidate HolPtoU2	3	55.4	—	—	—	—	—	—	—	—	—
PSPTO4590	C	TTSS effector candidate HolPtoU-related	3	54.6	—	—	—	—	—	—	—	—	—
PSPTO0061	C	TTSS effector candidate HolPtoY	3	58.5	—	—	—	—	—	—	—	—	—
PSPTO4597 ^d , 4595 ^d	C	TTSS effector candidate HolPtoZ	3	54.2	—	—	—	—	—	—	—	—	—
PSPTO4588	C	TTSS effector candidate HolPtoZ2	3	59.2	—	—	—	—	—	—	—	—	—
PSPTO4724 ^{bd} , 4726 ^{bd}	C	TTSS effector candidate HopPtoD1-related protein (interruption)	3	54.3	—	—	—	—	—	—	—	—	—
PSPTO4592	C	TTSS effector candidate HopPtoO-related	3	54.2	—	—	—	—	—	—	—	—	—
PSPTO0875	C	TTSS effector candidate Ipx47	3	48.3	—	—	—	—	—	—	—	—	—
PSPTO5061	C	TTSS effector candidate Ipx53	3	64.3	—	—	—	—	—	—	NP_404834	—	AAL21797
PSPTO0474	C	TTSS effector candidate ORF01152	3	59.5	—	—	—	—	—	—	—	—	—
PSPTO0869	C	TTSS effector candidate ORF20	3	61.7	—	—	AAM37122	—	—	—	—	—	—
PSPTO4188	C	TTSS effector candidate ORF36	3	49.1	—	—	—	—	—	—	—	—	—
PSPTO4776	C	TTSS effector HopPmaI(Pto)	3	63.2	CAD18035	NP_635901	AAM36964	AAL42873	—	AAG08638	NP_406242	AAG54315	AAL19895
PSPTO1179	C	TTSS effector HopPmaJ(Pto)	3	58.7	—	—	—	—	—	—	—	—	—
PSPTO5354	C	TTSS effector HopPsyA(Pto)	3	53.3	—	—	—	—	—	—	—	—	—
PSPTO1372	C	TTSS effector HopPtoA1	3	61.1	CAD18428	—	—	—	—	—	—	—	—
PSPTO4718	C	TTSS effector HopPtoA2	3	57.1	CAD18428	—	—	—	—	—	—	—	—
PSPTO1406	C	TTSS effector HopPtoB1	3	55.5	—	—	—	—	—	—	—	—	—
PSPTO0589	C	TTSS effector HopPtoC	3	47	CAD17000	—	—	—	—	—	—	—	—
PSPTO0876	C	TTSS effector HopPtoD1	3	55.2	CAD17455	—	—	—	—	—	—	—	—
PSPTO4722	C	TTSS effector HopPtoD2	3	51.7	—	NP_637463	—	—	—	—	—	—	—
PSPTO4331	C	TTSS effector HopPtoE	3	50.7	—	—	—	—	—	—	—	—	—
PSPTO0502	C	TTSS effector HopPtoF	3	48	CAD17973	—	—	—	—	—	—	—	—
PSPTO4727	C	TTSS effector HopPtoG	3	43.7	CAD17474	NP_638946	—	—	—	—	—	—	—
PSPTO0588	C	TTSS effector HopPtoH	3	47.4	CAD17723	NP_638604	—	—	—	—	—	—	—
PSPTO4691	C	TTSS effector HopPtoI	3	48.9	—	—	—	—	—	—	—	—	—
PSPTO1568	C	TTSS effector HopPtoJ	3	49.4	—	—	—	—	—	—	—	—	—
PSPTO0044	C	TTSS effector HopPtoK	3	54.6	—	—	AAM38509	—	—	—	—	—	—
PSPTO2872	C	TTSS effector HopPtoL	3	61	—	—	—	—	—	—	NP_405828	—	AAL20513

PSPTO1375	C	TTSS effector HopPtoM	3	58.5	—	—	—	—	—	—	—	—	—
PSPTO1370	C	TTSS effector HopPtoN	3	53.6	—	—	—	—	—	—	—	—	—
PSPTO0877	C	TTSS effector HopPtoQ	3	51.1	CAD13773	NP_636447	AAM39163	—	—	—	—	—	—
PSPTOA0018	pA	TTSS effector HopPtoS1	3	46.5	—	—	—	—	—	—	—	—	—
PSPTO0501	C	TTSS effector HopPtoS2	3	46.5	—	—	—	—	—	—	—	—	—
PSPTO4594	C	TTSS effector HopPtoS3	3	49.8	—	—	—	—	—	—	—	—	—
PSPTOA0019	pA	TTSS effector HopPtoT1	3	51.9	—	—	—	—	—	—	—	—	—
PSPTO4720	C	TTSS effector HopPtoV	3	51.5	—	—	—	—	—	—	—	—	—
PSPTO0852	C	TTSS helper protein HopPmaG(Pto)	3	55.8	CAD18391	NP_639491	AAM38069	AAL44589	AAF84983	AAG07379	NP_405633	AAG57806	AAL21711
PSPTO4101	C	TTSS helper protein HopPmaH(Pto) (harpin-like protein)	3	56.8	—	NP_636036	AAM38405	—	—	AAG08445	—	—	—
PSPTO2678	C	TTSS helper protein HopPtoP (harpin-like protein)	3	59.7	CAD16603	NP_636615	AAM35308	—	—	—	—	AAG58843	—
PSPTO1381	C	TTSS helper protein HrpA(Pto) (pilin subunit)	3	54.9	—	—	—	—	—	—	—	—	—
PSPTO1405	C	TTSS helper protein HrpK(Pto)	3	58.4	—	—	—	—	—	—	—	—	—
PSPTO1373	C	TTSS helper protein HrpW(Pto) (harpin-like protein)	3	58	CAD16482	NP_636593	AAM37767	—	AAF83628	AAG04113	—	—	—
PSPTO1382	C	TTSS helper protein HrpZ(Pto) (harpin)	3	60.2	CAD18028	—	AAM35307	—	—	—	—	—	—
PSPTO1391	C	TTSS negative regulator of hrp expression HrpV	4	56.6	—	—	—	—	—	—	—	—	—
PSPTO1389	C	TTSS outer-membrane protein HrcC	4	57.9	CAD18025	NP_636613	AAM35306	—	AAF83183	AAG05105	NP_395186	AAG58839	AAL20318
PSPTO1404	C	TTSS RNA polymerase sigma factor HrpL	4	55.4	CAD14757	NP_636641	AAM36190	AAL43022	AAF85038	AAG06284	NP_406228	AAG57689	AAL21534
PSPTO1379	C	TTSS transcriptional regulator HrpR	4	59.4	CAD13860	NP_636405	AAM35118	AAL44112	AAF84654	AAG04724	NP_405881	AAG57601	AAL21739
PSPTO1380	C	TTSS transcriptional regulator HrpS	4	59.1	CAD17160	NP_636405	AAM36008	AAL44112	AAF84654	AAG08551	NP_405881	AAG57601	AAL21739
PSPTO4851	C	type II/III secretion system protein homolog of GspD/HrcC	4	62.5	CAD18238	NP_638574	AAM38224	AAL41244	AAF83183	AAG07692	NP_403806	AAG58491	AAL22350
PSPTO4853	C	type II/IV secretion system protein, putative (GspE/VirB-11)	2	62	CAD18236	NP_637830	AAM37467	AAL41241	NP_061671	AAG07690	NP_404326	AAG54411	AAL19107
PSPTO0811	C	type IV pillin PilA, putative	1	57.3	CAD18443	NP_637845	AAM37516	—	AAF82841	AAG07938	NP_404438	—	—
PSPTO0319	C	type IV pilus biogenesis protein	1	61.2	CAD16824	NP_639427	AAM39047	AAL46277	AAF85341	AAG08595	NP_404439	AAG54411	AAL19107
PSPTO0810	C	type IV pilus biogenesis protein	1	57.4	CAD14252	NP_637837	AAM37474	—	AAF83283	AAG07938	—	—	—
PSPTO0816	C	type IV pilus biogenesis protein	1	56.6	CAD16387	NP_637838	AAM37511	—	AAF82846	AAG07944	NP_404438	AAG54412	AAL19108
PSPTO0927	C	type IV pilus biogenesis protein	1	47.4	CAD14088	NP_638446	AAM38085	—	AAF85339	AAG07913	NP_406887	AAG54412	AAL19108

PSPTOA0045	pA	type IV pilus biogenesis protein	1	58.7	CAD16388	NP_635985	AAM38450	AAL46277	AAF84441	AAG03785	NP_404557	AAG58081	AAL21974
PSPTOB0052	pB	type IV pilus biogenesis protein	1	58.8	CAD16388	NP_635985	AAM38450	AAL46277	AAF84441	AAG03785	NP_404557	AAG58081	AAL21974
PSPTO0926	C	type IV pilus biogenesis protein PilB	1	54.4	CAD16532	NP_638444	AAM38083	AAL46277	AAF85341	AAG07914	NP_406888	AAG54411	AAL19107
PSPTO0925	C	type IV pilus biogenesis protein PilC	1	47.9	CAD16533	NP_638447	AAM38086	–	AAF85335	AAG07915	NP_406889	AAG54410	AAL19106
PSPTO1432	C	type IV pilus biogenesis protein PilF	1	56.2	CAD14915	NP_637348	AAM36879	–	AAF83270	AAG07192	NP_406386	–	–
PSPTO5031	C	type IV pilus biogenesis protein PilJ	1	61	CAD14201	NP_638270	AAM37944	AAL42908	AAF84755	AAG03800	NP_405252	AAG59538	AAL22437
PSPTO5132	C	type IV pilus biogenesis protein PilM	1	58.9	CAD16684	NP_638578	AAM38228	–	AAF83179	AAG08429	–	–	–
PSPTOB0007	pB	type IV pilus biogenesis protein PilM, putative	1	54.1	–	–	–	–	–	–	–	–	–
PSPTO5131	C	type IV pilus biogenesis protein PilN	1	59.8	CAD16683	NP_638577	AAM38227	–	AAF83180	AAG08428	–	–	–
PSPTO5130	C	type IV pilus biogenesis protein PilO	1	56.5	CAD16682	NP_638576	AAM38226	–	AAF83181	AAG08427	–	–	–
PSPTO5129	C	type IV pilus biogenesis protein PilP	1	60.2	CAD16681	NP_638575	AAM38225	–	AAF83182	AAG08426	–	–	–
PSPTO5128	C	type IV pilus biogenesis protein PilQ	1	59	CAD16680	NP_638574	AAM38224	AAL41244	AAF83183	AAG08425	NP_403806	AAG58491	AAL22350
PSPTOA0042	pA	type IV pilus biogenesis protein PilT	1	59.3	CAD16603	NP_636615	AAM35308	–	–	–	–	AAG57967	AAL21757
PSPTOB0049	pB	type IV pilus biogenesis protein PilT, putative	1	59.3	CAD16603	NP_636615	AAM35308	–	–	–	–	AAG57967	AAL21757
PSPTO3825	C	type IV pilus biogenesis protein PilZ	1	59.6	CAD15488	NP_636400	AAM36006	–	AAF83487	AAG06348	–	–	–
PSPTO0924	C	type IV pilus prepilin peptidase PilD	1	59.3	CAD16534	NP_638448	AAM38087	–	AAF85334	AAG07916	NP_404431	AAG58443	AAL22305
PSPTO5032	C	type IV pilus protein PilI	1	61.6	–	NP_638271	AAM37945	AAL43066	AAF84756	AAG03799	–	–	–
PSPTO0823	C	type IV pilus regulatory protein pilR	1	60.3	CAD16514	NP_638443	AAM38082	AAL42452	AAF85342	AAG07935	NP_406415	AAG59201	AAL23002
PSPTO0824	C	type IV pilus regulatory sensor protein PilS	1	60.1	CAD16515	NP_638442	AAM38081	AAL45099	AAF85343	AAG07934	NP_403906	AAG59200	AAL23001
PSPTO5034	C	type IV pilus response regulator PilG	1	52.9	CAD14198	NP_638273	AAM37947	AAL41534	AAF84757	AAG03797	NP_403738	AAG56982	AAL20028
PSPTO5033	C	type IV pilus response regulator PilH	1	58.1	CAD14199	NP_636424	AAM36026	AAL42303	AAF83260	AAG03798	NP_406679	AAG54745	AAL19351
PSPTO0531	C	type IV secretion system protein, putative	1	52.4	–	–	–	AAL45856	NP_061664	–	–	–	–
PSPTO0533	C	type IV secretion system protein, truncation	1	51.6	–	–	–	AAL45857	NP_061667	–	–	–	–

Table 2, page 11

PSPTO2602	C	yersiniabactin non-ribosomal peptide synthetase	2	63.4	CAD15508	NP_639207	AAM38759	AAL44494	AAF83100	AAG07614	NP_405472	AAG54921	AAL19539
PSPTO2600	C	yersiniabactin polyketide/non-ribosomal peptide hybrid synthetase	2	64.3	CAD17792	NP_636396	AAM36002	AAL44493	AAF83483	AAG07613	NP_405471	AAG55841	AAL20126
PSPTO2597	C	yersiniabactin synthetase, salicylate ligase component	2	65.8	CAD17188	NP_636198	AAM35769	AAL45616	AAF83100	AAG07616	NP_405468	AAG54929	AAL19547
PSPTO2599	C	yersiniabactin synthetase, thiazoliny l reductase component	2	65.7	CAD15507	–	–	–	–	AAG07612	NP_405470	–	–
PSPTO2598	C	yersiniabactin synthetase, thioesterase component	2	64.1	CAD15506	–	–	AAL44497	–	AAG07617	NP_405469	–	–

The inferred proteomes of nine pathogenic bacteria were searched with the 298 DC3000 proteins implicated in virulence by using a BLASTP cutoff criterion of

$E < 10^{-5}$, which will identify putative homologs as well as paralogous gene family members (note, for example, that Hrc proteins and HrpR show similarity to flagellar proteins and transcriptional regulators in *A. tumefaciens*, a pathogen that lacks a TTSS). Rs = *Ralstonia solanacearum*, Xac = *Xanthomonas axonopodis* pv *citri*, Xcc = *Xanthomonas campestris* pv *campestris*, Xf = *Xylella fastidiosa*, At = *Agrobacterium tumefaciens* (U. Wash seq.), Pa = *Pseudomonas aeruginosa* PAO1, St = *Salmonella typhimurium* LT2, Yp = *Yersinia pestis* CO92, and Ec = *Escherichia coli* O157:H7 EDL933. Categories are as follows: 1, Adhesins and other cell surface associated factors; 2, Miscellaneous virulence factors; 3, TTSS effectors, helpers, and candidates; 4, TTSS secretion and regulatory proteins; and 5, low molecular weight diffusible factors.

^a C, chromosome; pA, plasmid pDC3000A; pB, plasmid pDC3000B.

^b This gene does not have a protein record in GenBank.

^c HrcQa and HrcQb are present as one protein in other TTSS pathogens.

^d Indicates a gene disrupted by IS elements, with PSPTO numbers given for both the N-terminal and C-terminal regions. Only the N-terminal part is proposed as a candidate virulence factor (see Table 3).

Table 3. The chromosomal and plasmid genes disrupted by IS elements or phage genes in *P. syringae* pv. *tomato* DC3000

	N-terminus	C-terminus	Gene assignment
1	PSPTO0198	PSPTO0195	ISPsy5, Orf1
2	PSPTO0843	PSPTO0845	ISPsy5, transposase
3	PSPTO0846	PSPTO0841	membrane protein, putative
4	PSPTO1018	PSPTO1021	hypothetical protein
5	PSPTO1225	PSPTO1228	insecticidal toxin protein
6	PSPTO1314	PSPTO1316	fimbrial biogenesis usher protein
7	PSPTO0901	-	TTSS effector candidate HolPsyAG(Pto)
8	PSPTO2329	PSPTO2327	response regulator
9	PSPTO2462	PSPTO2459	conserved hypothetical protein
10	PSPTO2540	PSPTO2542	conserved hypothetical protein
11	PSPTO3015	PSPTO3012	conserved hypothetical protein
12	PSPTO3431	PSPTO3434	ISPsy6, transposase
13	PSPTO3611	PSPTO3614	IS52, transposase
14	PSPTO4268	PSPTO4271	sulfate permease family protein
15	PSPTO4565	PSPTO4568	peptide ABC transporter, ATP-binding protein
16	PSPTO4597	PSPTO4595	TTSS effector candidate HolPtoZ
17	PSPTO4724	PSPTO4726	TTSS effector candidate HopPtoD1-related protein
18	PSPTO4728	PSPTO4731	ISPsy10, transposase
19	PSPTO4993	PSPTO4996	TTSS effector candidate HolPtoAC
20	PSPTO5366	PSPTO5369	ISPsy4, transposase
21	PSPTO5572	PSPTO5570	membrane protein
22	PSPTO5574	PSPTO5577	response regulator
23	PSPTOB0038	PSPTOB0041	transcriptional regulator, GntR family

24	PSPTOA0003	PSPTOA0006	ultraviolet light resistance protein B
----	------------	------------	--

Table 4. IS elements present in *P. syringae* pv. *tomato* DC3000

Element	Family	Copies ^a	Length / DR ^b		Structure	Left and right ends of newly identified elements (5'-3')
IS52	IS5	3(2)	ND ^c	ND	Orf A	previously identified: M14366
ISPs1	IS5	1(2)	ND	ND	Orf AB	previously identified: M11035
ISPssy	IS5	36(8)	ND	ND	Orf A	previously identified: AB023075
ISPsy3	IS91	2(4)	ND	ND	Orf A	previously identified: AB063176
ISPsy4	IS21	23(1)	1962	6	Orf AB	TGTCACCGCCACTGTAAAAATGACCCCCTAACGCCAAC TGTC AACGCCAACTAAAAAGTGACCCCCTTCCGTGCTA
ISPsy5	IS66	35(13)	2059	8	Orf AB	GTATCCGATCCACGAACCCACATTTCAAGCCACTGGC GTAAGCGATCCACCAACCACACCTACCATCAGATGCGG
ISPsy6	unclassified	14(1)	1603	4(5)	Orf A	CAGACTGTGTGAAAACCTAGCAATCTGCGCAGCGCTCA CAGACTGTGTGAAAACACACTGATGAAGGCCCAAGCAA
ISPsy7	IS110	10	1356	0	Orf A	ATAATGAGATGGTCAGCCCGATCCGAGAAACCCGGAGT TAGGCAAATGTATGGTCAGCCCACCTCCTGCAATCGCA
ISPsy8	IS3	5	1432	3	Orf AB	TGGACTGCCCCCAAGAAGTTGGACACCAATCCGACCTA TGGACTGCCCCCGGAAGTTGGACAGTTTTAGCTTGCC
ISPsy9	IS3	1	1452	3	Orf AB	TGGTGTGACCCCGAAGGTTGGATGCAACCTTTGGAGG TGGTGTGACCCCGAAGGTTGGACAATTATCCACCTAC
ISPsy10	IS5	(3)	ND	ND	Orf A	ND ND
ISPsy11	IS3	3	1313	3	Orf AB	TGTCCCGTCCTGAATAAGGTTTACACCTTCTGACCTTT

						TGTCTAGTCCTGAAATAGGTTTACACCTGTTTCACGCT
ISPsy12 IS3	1	ND	ND	Orf AB	ND	ND
ISPsy13 IS3	1	ND	ND	Orf AB	ND	ND
ISPsy14 IS21	1	ND	ND	Orf AB	ND	ND
ISPsy15 IS5	(2)	ND	ND	Orf A	ND	ND

^a The number of full-length copies and the number of incomplete copies in parentheses.

^b Size in base pairs of the element (length) and the direct repeat (DR) generated by insertion into the target site.

^c ND = not determined.

Table 4. IS elements present in *P. syringae* pv. *tomato* DC3000

Element	Family	Copies ^a	Length / DR ^b		Structure	Left and right ends of newly identified elements (5'-3')
IS52	IS5	3(2)	ND ^c	ND	Orf A	previously identified: M14366
ISPs1	IS5	1(2)	ND	ND	Orf AB	previously identified: M11035
ISPssy	IS5	36(8)	ND	ND	Orf A	previously identified: AB023075
ISPsy3	IS91	2(4)	ND	ND	Orf A	previously identified: AB063176
ISPsy4	IS21	23(1)	1962	6	Orf AB	TGTCACCGCCACTGTAAAAATGACCCCCTAACGCCAAC TGTC AACGCCAACTAAAAAGTGACCCCCTTCCGTGCTA
ISPsy5	IS66	35(13)	2059	8	Orf AB	GTATCCGATCCACGAACCCACATTTCAAGCCACTGGC GTAAGCGATCCACCAACCACACCTACCATCAGATGCGG
ISPsy6	unclassified	14(1)	1603	4(5)	Orf A	CAGACTGTGTGAAAACCTAGCAATCTGCGCAGCGCTCA CAGACTGTGTGAAAACACACTGATGAAGGCCCAAGCAA
ISPsy7	IS110	10	1356	0	Orf A	ATAATGAGATGGTCAGCCCGATCCGAGAAACCCGGAGT TAGGCAAATGTATGGTCAGCCCACCTCCTGCAATCGCA
ISPsy8	IS3	5	1432	3	Orf AB	TGGACTGCCCCCAAGAAGTTGGACACCAATCCGACCTA TGGACTGCCCCCGGAAGTTGGACAGTTTTAGCTTGCC
ISPsy9	IS3	1	1452	3	Orf AB	TGGTGTGACCCCGAAGGTTGGATGCAACCTTTGGAGG TGGTGTGACCCCGAAGGTTGGACAATTATCCACCTAC
ISPsy10	IS5	(3)	ND	ND	Orf A	ND ND
ISPsy11	IS3	3	1313	3	Orf AB	TGTCCCGTCCTGAATAAGGTTTACACCTTCTGACCTTT

						TGTCTAGTCCTGAAATAGGTTTACACCTGTTTCACGCT
ISPsy12 IS3	1	ND	ND	Orf AB	ND	ND
ISPsy13 IS3	1	ND	ND	Orf AB	ND	ND
ISPsy14 IS21	1	ND	ND	Orf AB	ND	ND
ISPsy15 IS5	(2)	ND	ND	Orf A	ND	ND

^a The number of full-length copies and the number of incomplete copies in parentheses.

^b Size in base pairs of the element (length) and the direct repeat (DR) generated by insertion into the target site.

^c ND = not determined.

Table 5. Phage and bacteriocin regions of *P. syringae* pv. *tomato* DC3000

Feature	5' end	3' end	Size (bp)	Type of Feature
PYOCIN001	625,118	647,483	22,365	Putative F-type pyocin
PHAGE02	2,252,628	2,264,996	12,368	Defective prophage region
PHAGE03	3,824,683	3,865,352	40,669	Putative P2-like integrated prophage
PHAGE04	4,445,530	4,453,456	7,926	Defective prophage region
PHAGE05	4,515,380	4,530,373	14,993	Defective prophage region
PHAGE06	4,543,843	4,563,492	19,649	Putative lysogenic prophage region

Table 6. Genes with unknown function unique to DC3000 identified from the comparative analysis with the *P. aeruginosa* PAO1 and *P. putida* KT2440 genomes and the identification of putative homologs (BLASTP $E < 10^{-5}$) in pathogenic bacteria

Locus	Location	Clustering		Predicted product ^c	GC%	Length (aa)	Plant pathogens					Animal pathogens		
		Genes ^a	TTSS ^b				Rs	Xcc	Xac	At	Xf	Yp	Ec	St
PSPTO0007	C	+	+	hypothetical protein	55.2	294								
PSPTO0011	C	+	+	conserved hypothetical protein	55.9	236	Rs	Xcc	Xac	At	Xf			
PSPTO0012	C	+	+	hypothetical protein	54.9	167								
PSPTO0013	C	+	+	hypothetical protein	58	189								
PSPTO0014	C	+	+	hypothetical protein	66.3	96								
PSPTO0015	C	+	+	hypothetical protein	55.7	203								
PSPTO0016	C	+	+	hypothetical protein	53.5	155								
PSPTO0020	C	+	+	hypothetical protein	54.9	116								
PSPTO0021	C	+	+	hypothetical protein	51.2	746								
PSPTO0022	C	+	+	DNA-binding protein	56.2	281								
PSPTO0023	C	+	+	hypothetical protein	53.9	199								
PSPTO0024	C	+	+	hypothetical protein	60	45								
PSPTO0025	C	+	+	hydrolase, haloacid dehalogenase-like family	56.6	209								
PSPTO0026	C	+	+	hypothetical protein	53	348								
PSPTO0027	C	+	+	hypothetical protein	51.7	240								
PSPTO0032	C	+	+	conserved hypothetical protein	48.9	322				At				
PSPTO0034	C	+	+	recombinase, putative	50.3	518	Rs			At				
PSPTO0037	C	+	+	helicase domain protein	53.4	1636						Yp		
PSPTO0038	C	+	+	conserved domain protein	53.4	696								

PSPTO0041	C	+	+	hypothetical protein	54.2	128								
PSPTO0043	C	+	+	cytidine/deoxycytidylate deaminase family protein	51.4	520								
PSPTO0045	C	+	+	hypothetical protein	53.2	52								
PSPTO0048	C	+	+	hypothetical protein	50	388								
PSPTO0049	C	+	+	conserved hypothetical protein	49.8	247								
PSPTO0050	C	+	+	conserved hypothetical protein	47.8	443								
PSPTO0051	C	+	+	conserved hypothetical protein	53.7	147	Rs							
PSPTO0052	C	+	+	hypothetical protein	53.1	589								
PSPTO0053	C	+	+	hypothetical protein	55.6	481								
PSPTO0054	C	+	+	hypothetical protein	57	203								
PSPTO0059	C	+	+	hypothetical protein	54.2	206								
PSPTO0088	C			hypothetical protein	58.3	44								
PSPTO0093	C			hypothetical protein	51.8	83								
PSPTO0104	C			hypothetical protein	53.9	34								
PSPTO0122	C			hypothetical protein	59.9	54								
PSPTO0133	C			hypothetical protein	54	50								
PSPTO0142	C			hypothetical protein	49.1	76								
PSPTO0149	C			conserved hypothetical protein	48.9	225								
PSPTO0160	C			hypothetical protein	60.3	125								
PSPTO0167	C			hypothetical protein	53.8	140								
PSPTO0172	C			hypothetical protein	54	97								
PSPTO0192	C			recombinase, putative	50.1	518	Rs			At				
PSPTO0205	C	+		hypothetical protein	48.4	331								
PSPTO0206	C	+		conserved hypothetical protein	55.3	158								

PSPTO0208	C	+		hypothetical protein	58	207								
PSPTO0216	C	+		hypothetical protein	50.7	46								
PSPTO0240	C	+		prevent-host-death family protein	51.8	83			Xac			Yp		
PSPTO0242	C	+		hypothetical protein	57.7	190								
PSPTO0243	C	+		conserved hypothetical protein	49	273								
PSPTO0244	C	+		hypothetical protein	62.3	237				At				
PSPTO0245	C	+		acetyltransferase, GNAT family	62.5	285				At				
PSPTO0246	C	+		hypothetical protein	54.7	137								
PSPTO0260	C			conserved hypothetical protein	62.8	218								
PSPTO0274	C	+		hypothetical protein	61.9	84								
PSPTO0275	C	+		DNA-binding protein	56.1	377	Rs	Xcc			Xf			
PSPTO0277	C	+		hypothetical protein	51.6	162								
PSPTO0279	C	+		conserved hypothetical protein	54	118								
PSPTO0283	C	+		hypothetical protein	54.4	147								
PSPTO0285	C	+		conserved domain protein	53.3	360								
PSPTO0289	C	+		hypothetical protein	56.2	51								
PSPTO0290	C	+		conserved hypothetical protein	59.8	78								
PSPTO0293	C	+		hypothetical protein	54.2	102								
PSPTO0294	C	+		hypothetical protein	62.6	66								
PSPTO0295	C	+		hypothetical protein	55.3	73								
PSPTO0296	C	+		hypothetical protein	58.3	107								
PSPTO0297	C	+		hypothetical protein	56.9	96								
PSPTO0302	C	+		hypothetical protein	49.4	58								
PSPTO0331	C			hypothetical protein	50.5	74								

PSPTO0332	C			hypothetical protein	54.4	95								
PSPTO0342	C			hypothetical protein	52.2	46								
PSPTO0354	C			hypothetical protein	52.6	52								
PSPTO0368	C			hypothetical protein	56.5	111								
PSPTO0374	C			hypothetical protein	61.2	195								
PSPTO0391	C			hypothetical protein	53.8	83								
PSPTO0392	C			hypothetical protein	50.8	63								
PSPTO0442	C			hypothetical protein	58.6	33								
PSPTO0447	C			hypothetical protein	68.3	60								
PSPTO0467	C	+	+	hypothetical protein	51.6	51								
PSPTO0473	C	+	+	hypothetical protein	59.8	1073		Xcc	Xac					
PSPTO0475	C	+	+	hypothetical protein	55.2	143								
PSPTO0476	C	+	+	hypothetical protein	53.2	148								
PSPTO0477	C	+	+	hypothetical protein	49.2	147								
PSPTO0478	C	+	+	hypothetical protein	50.6	308								
PSPTO0482	C	+	+	hypothetical protein	56.3	116								
PSPTO0521	C	+	+	DNA-binding protein	63.8	94								
PSPTO0522	C	+	+	conserved domain protein	58.5	408	Rs	Xcc		At				
PSPTO0526	C	+	+	hypothetical protein	54.3	70								
PSPTO0527	C	+	+	hypothetical protein	52.4	650								
PSPTO0528	C	+	+	hypothetical protein	50.7	295								
PSPTO0529	C	+	+	hypothetical protein	51	266								
PSPTO0530	C	+	+	conserved hypothetical protein	53.9	173					Xf			
PSPTO0561	C			hypothetical protein	61.4	57								

PSPTO0582	C	+	+	hypothetical protein	51.2	213								
PSPTO0583	C	+	+	hypothetical protein	47.2	229								
PSPTO0585	C	+	+	hypothetical protein	50.5	155								
PSPTO0590	C	+	+	hypothetical protein	45.4	47								
PSPTO0608	C			hypothetical protein	30.3	33								
PSPTO0657	C	+		hypothetical protein	55.1	196								
PSPTO0658	C	+		hypothetical protein	49.7	181								
PSPTO0665	C	+		ROK family protein	60.1	376		Xcc	Xac	At		Yp	Ec	St
PSPTO0666	C	+		hypothetical protein	53	173								
PSPTO0684	C			hypothetical protein	56.3	183								
PSPTO0702	C	+		conserved hypothetical protein	56.8	339	Rs		Xac				Ec	St
PSPTO0707	C	+		hypothetical protein	56	204								
PSPTO0709	C	+		hypothetical protein	54.9	240								
PSPTO0710	C	+		conserved domain protein	56.1	98								
PSPTO0711	C	+		hypothetical protein	54.5	591								
PSPTO0712	C	+		hypothetical protein	52.4	201								
PSPTO0713	C	+		conserved hypothetical protein	55.3	347								
PSPTO0715	C	+		conserved hypothetical protein	58.5	265	Rs							
PSPTO0717	C	+		DNA-binding protein	54.2	80			Xac					
PSPTO0718	C	+		hypothetical protein	55.8	132								
PSPTO0719	C	+		hypothetical protein	48.9	58								
PSPTO0745	C			hypothetical protein	54.8	76								
PSPTO0746	C			hypothetical protein	62.4	465								
PSPTO0792	C			hypothetical protein	65.2	44								

PSPTO0803	C			hypothetical protein	53.1	32								
PSPTO0819	C			conserved hypothetical protein	58.3	248	Rs	Xcc		At	Xf			
PSPTO0831	C	+	+	conserved hypothetical protein	56	642		Xcc	Xac		Xf			
PSPTO0837	C	+	+	conserved hypothetical protein	44.2	200			Xac					
PSPTO0840	C	+	+	hypothetical protein	50.2	138								
PSPTO0847	C	+	+	hypothetical protein	50	122								
PSPTO0848	C	+	+	conserved hypothetical protein	57.8	457		Xcc	Xac					
PSPTO0849	C	+	+	conserved hypothetical protein	59.5	312		Xcc	Xac					
PSPTO0850	C	+	+	conserved hypothetical protein	58.5	138			Xac					
PSPTO0851	C	+	+	conserved hypothetical protein	51.6	162						Yp	Ec	
PSPTO0854	C	+	+	hypothetical protein	51.1	73								
PSPTO0856	C	+	+	conserved hypothetical protein	52	272								
PSPTO0858	C	+	+	hypothetical protein	53.2	99								
PSPTO0859	C	+	+	conserved hypothetical protein	56.1	889			Xac					
PSPTO0860	C	+	+	conserved hypothetical protein	58.1	151								
PSPTO0861	C	+	+	conserved hypothetical protein	58.2	498			Xac					
PSPTO0862	C	+	+	conserved hypothetical protein	58.5	307			Xac					
PSPTO0863	C	+	+	conserved hypothetical protein	59.6	222								
PSPTO0866	C	+	+	hypothetical protein	59.8	121			Xac					
PSPTO0867	C	+	+	conserved hypothetical protein	59	122			Xac					
PSPTO0868	C	+	+	hypothetical protein	61.2	49								
PSPTO0870	C	+	+	hypothetical protein	49	66								
PSPTO0874	C	+	+	nikkomycin biosynthesis domain protein	49.2	417								
PSPTO0878	C	+	+	conserved domain protein	59.4	105								

PSPTO0880	C	+	+	conserved hypothetical protein	56.9	249			Xac					
PSPTO0881	C	+	+	conserved domain protein	57.9	293			Xac					
PSPTO0882	C	+	+	hypothetical protein	50.8	177								
PSPTO0891	C		+	hypothetical protein	57	76								
PSPTO0895	C	+	+	hypothetical protein	50.4	168								
PSPTO0899	C	+	+	conserved domain protein	55.3	155	Rs							
PSPTO0900	C	+	+	hypothetical protein	63.4	93								
PSPTO0904	C	+	+	mutT/nudix family protein	51.9	504								
PSPTO0907	C	+	+	hypothetical protein	53.1	403								
PSPTO0914	C	+	+	STAS domain protein	63.9	96		Xcc	Xac					
PSPTO0918	C	+	+	conserved hypothetical protein	60	35								
PSPTO0921	C	+	+	hypothetical protein	44.2	43								
PSPTO0930	C	+		conserved hypothetical protein	55.8	230								
PSPTO0931	C	+		hypothetical protein	51.9	135								
PSPTO0932	C	+		conserved hypothetical protein	56.2	181								
PSPTO0933	C	+		conserved hypothetical protein	60.1	332						Yp	Ec	
PSPTO0935	C	+		conserved hypothetical protein	62.7	377						Yp	Ec	
PSPTO0936	C	+		conserved hypothetical protein	59.5	252						Yp	Ec	
PSPTO0937	C	+		conserved hypothetical protein	62.5	364						Yp	Ec	
PSPTO0938	C	+		conserved hypothetical protein	60.1	316						Yp	Ec	
PSPTO0966	C			hypothetical protein	53.4	58								
PSPTO0974	C			hypothetical protein	54.3	35								
PSPTO0987	C			conserved domain protein	50.4	76								
PSPTO1003	C			acetyltransferase, GNAT family	58.2	145				At				St

PSPTO1006	C			hypothetical protein	57.6	33								
PSPTO1016	C			hypothetical protein	53.2	534								
PSPTO1036	C	+		hypothetical protein	47.7	153								
PSPTO1038	C	+		conserved hypothetical protein	63.5	149						Yp		
PSPTO1044	C	+		hypothetical protein	55.9	59								
PSPTO1045	C	+		PIN domain protein	54.2	134				At	Xf			St
PSPTO1046	C	+		conserved hypothetical protein	52.2	76								St
PSPTO1055	C	+		hypothetical protein	53.7	72								
PSPTO1057	C	+		conserved hypothetical protein TIGR00305	56.4	136					Xf			
PSPTO1058	C	+		conserved hypothetical protein	53.2	79					Xf			
PSPTO1060	C	+		hypothetical protein	52.5	47								
PSPTO1062	C	+		hypothetical protein	54	87								
PSPTO1082	C			hypothetical protein	54.2	72								
PSPTO1088	C	+		conserved hypothetical protein	58.6	350	Rs		Xac				Ec	St
PSPTO1090	C	+		hypothetical protein	38.1	325								
PSPTO1091	C	+		conserved hypothetical protein	39.1	138								
PSPTO1093	C	+		relaxase/mobilization nuclease domain protein	52.8	805								
PSPTO1094	C	+		conserved hypothetical protein	53.1	125								
PSPTO1113	C			hypothetical protein	54.1	194								
PSPTO1114	C			hypothetical protein	55.9	74		Xcc	Xac					
PSPTO1124	C			hypothetical protein	56.7	40								
PSPTO1125	C			hypothetical protein	58.7	71								
PSPTO1137	C			hypothetical protein	62	36								
PSPTO1142	C			hypothetical protein	60.5	86		Xcc	Xac					

PSPTO1161	C			hypothetical protein	60.4	64								
PSPTO1176	C		+	conserved domain protein	51.8	114								
PSPTO1185	C	+	+	hypothetical protein	56.2	201								
PSPTO1193	C	+	+	conserved hypothetical protein	62.8	654	Rs							
PSPTO1195	C	+	+	hypothetical protein	63.7	713								
PSPTO1198	C	+	+	hypothetical protein	57.5	40								
PSPTO1199	C			hypothetical protein	46.7	45								
PSPTO1259	C			hypothetical protein	50.6	58								
PSPTO1297	C			hypothetical protein	55.3	44								
PSPTO1311	C	+		hypothetical protein	55.7	125								
PSPTO1319	C	+		hypothetical protein	59.6	95								
PSPTO1321	C	+		hypothetical protein	61.4	51								
PSPTO1322	C	+		acetyltransferase, GNAT family	62.8	157								
PSPTO1341	C	+		hypothetical protein	58	322								
PSPTO1342	C	+		conserved domain protein	53.6	352								St
PSPTO1344	C	+		hypothetical protein	59.1	190		Xcc	Xac					
PSPTO1350	C	+	+	hypothetical protein	61.4	400								
PSPTO1351	C	+	+	hypothetical protein	55.8	98								
PSPTO1353	C	+	+	conserved hypothetical protein	61.1	131		Xcc	Xac			Yp		
PSPTO1367	C	+	+	acetyltransferase, GNAT family	53.4	148								
PSPTO1408	C	+	+	hypothetical protein	64.5	61								
PSPTO1409	C	+	+	conserved hypothetical protein	49.3	281						Yp		St
PSPTO1410	C	+	+	exchangeable effector locus protein	53.6	138								
PSPTO1411	C	+	+	hypothetical protein	49.6	191								

PSPTO1442	C			hypothetical protein	58.2	178								
PSPTO1451	C			conserved hypothetical protein	51.3	393	Rs	Xcc				Yp	Ec	
PSPTO1452	C			hypothetical protein	57.9	38								
PSPTO1463	C			hypothetical protein	57.6	48								
PSPTO1530	C			hypothetical protein	54.6	282								
PSPTO1566	C	+	+	hypothetical protein	62.5	32								
PSPTO1569	C	+	+	conserved domain protein	50	88								
PSPTO1570	C	+	+	hypothetical protein	56.2	51								
PSPTO1571	C	+	+	hypothetical protein	52.6	523								
PSPTO1572	C	+	+	hypothetical protein	47.4	104								
PSPTO1573	C	+	+	conserved hypothetical protein	52.8	60					Xf		Ec	
PSPTO1575	C	+	+	hypothetical protein	56.6	103								
PSPTO1577	C	+	+	hypothetical protein	56.9	370			Xac					
PSPTO1581	C	+	+	conserved domain protein	55.2	212	Rs				Xf			
PSPTO1586	C	+	+	hypothetical protein	56.1	123								
PSPTO1591	C	+	+	hypothetical protein	51.7	40								
PSPTO1594	C	+	+	conserved hypothetical protein	63.9	134	Rs			At				
PSPTO1595	C	+	+	hypothetical protein	57.8	225								
PSPTO1601	C	+		hypothetical protein	59.6	166								
PSPTO1603	C	+		hypothetical protein	58.1	39								
PSPTO1611	C	+		hypothetical protein	65.8	37								
PSPTO1614	C	+		hypothetical protein	56.6	179								
PSPTO1615	C	+		hypothetical protein	48.7	290								
PSPTO1619	C	+		hypothetical protein	54.9	226								

PSPTO1634	C			conserved hypothetical protein	50.6	60						Yp		
PSPTO1635	C			conserved hypothetical protein	53.5	134						Yp		
PSPTO1641	C			hypothetical protein	57.7	93								
PSPTO1651	C	+		conserved hypothetical protein	58.7	96							Ec	
PSPTO1653	C	+		conserved domain protein	53.4	517		Xcc				Yp		
PSPTO1654	C	+		conserved hypothetical protein	54.9	424		Xcc				Yp		
PSPTO1655	C	+		hypothetical protein	51.7	158								
PSPTO1656	C	+		conserved hypothetical protein	53.5	629							Ec	
PSPTO1657	C	+		hypothetical protein	51.4	556								
PSPTO1662	C			hypothetical protein	55.9	65								
PSPTO1668	C			hypothetical protein	57.3	82								
PSPTO1676	C			conserved hypothetical protein	53.1	157								
PSPTO1678	C			hypothetical protein	58.9	116								
PSPTO1681	C			hypothetical protein	54.3	94								
PSPTO1722	C	+		hypothetical protein	63.7	101								
PSPTO1723	C	+		hypothetical protein	48.2	47								
PSPTO1726	C	+		hypothetical protein	50	104								
PSPTO1727	C	+		hypothetical protein	59.3	50								
PSPTO1761	C			hypothetical protein	54.9	68								
PSPTO1774	C			hypothetical protein	57.6	59								
PSPTO1804	C			hypothetical protein	57.8	30								
PSPTO1816	C			hypothetical protein	51	70								
PSPTO1820	C			hypothetical protein	57.6	59								
PSPTO1837	C			hypothetical protein	58.8	119								

PSPTO1840	C			hypothetical protein	50	52								
PSPTO1848	C	+		hypothetical protein	53.2	77								
PSPTO1849	C	+		hypothetical protein	49.5	171								
PSPTO1850	C	+		hypothetical protein	54.8	532		Xcc						
PSPTO1851	C	+		hypothetical protein	54.9	720		Xcc						
PSPTO1853	C	+		hypothetical protein	50.5	35								
PSPTO1858	C	+		conserved hypothetical protein	53.4	293							Ec	
PSPTO1869	C	+		conserved hypothetical protein	60.8	96			Xac					
PSPTO1872	C	+		hypothetical protein	57	31								
PSPTO1876	C	+		hypothetical protein	60.2	67								
PSPTO1895	C	+		conserved hypothetical protein	55.6	87	Rs							
PSPTO1897	C	+		BNR/Asp-box repeat protein	61	432		Xcc	Xac	At				St
PSPTO1900	C	+		hypothetical protein	52.7	31								
PSPTO1903	C	+		hypothetical protein	49.3	228								
PSPTO1905	C	+		hypothetical protein	58.6	79								
PSPTO1913	C			conserved hypothetical protein	53.7	98								
PSPTO1932	C			hypothetical protein	56.4	52								
PSPTO1938	C			hypothetical protein	61.7	47								
PSPTO1967	C			hypothetical protein	53.1	59								
PSPTO1999	C	+		PIN domain protein	58.4	150	Rs		Xac	At	Xf			
PSPTO2000	C	+		prevent-host-death family protein	57.8	83								
PSPTO2006	C	+		hypothetical protein	55.4	157								
PSPTO2012	C	+		hypothetical protein	52.8	178								
PSPTO2013	C	+		hypothetical protein	53.5	220								

PSPTO2027	C			hypothetical protein	61.4	38								
PSPTO2058	C	+		hypothetical protein	52.7	62								
PSPTO2062	C	+		conserved hypothetical protein	56.9	123								
PSPTO2067	C	+		hypothetical protein	58.1	39								
PSPTO2068	C	+		conserved hypothetical protein	54.4	109		Xcc	Xac					
PSPTO2069	C	+		hypothetical protein	57.9	61								
PSPTO2073	C	+		hypothetical protein	54	182								
PSPTO2074	C	+		hypothetical protein	54.3	97								
PSPTO2076	C	+		hypothetical protein	46.6	153								
PSPTO2077	C	+		conserved domain protein	59.5	121								
PSPTO2078	C	+		hypothetical protein	57.3	125								
PSPTO2079	C	+		hypothetical protein	57.8	120								
PSPTO2080	C	+		conserved hypothetical protein, internal deletion	56.3	126						Yp		
PSPTO2081	C	+		hypothetical protein	61.6	150								
PSPTO2083	C	+		hypothetical protein	47.4	175								
PSPTO2084	C	+		conserved hypothetical protein	58.6	83								
PSPTO2085	C	+		conserved hypothetical protein	59.5	131								
PSPTO2086	C	+		hypothetical protein	61.1	60								
PSPTO2087	C	+		hypothetical protein	62.4	102								
PSPTO2088	C	+		conserved hypothetical protein	60.9	132								
PSPTO2090	C	+		hypothetical protein	49.5	122								
PSPTO2095	C	+		conserved domain protein	56.7	70								
PSPTO2096	C	+		hypothetical protein	60.5	414								
PSPTO2122	C			hypothetical protein	57.4	97								

PSPTO2124	C			hypothetical protein	54.5	85								
PSPTO2157	C			hypothetical protein	64.4	568								
PSPTO2162	C			hypothetical protein	55.5	164								
PSPTO2219	C			hypothetical protein	56.8	54								
PSPTO2260	C	+		conserved hypothetical protein	59.8	409								
PSPTO2262	C	+		hypothetical protein	52.9	347								
PSPTO2263	C	+		hypothetical protein	53.8	145								
PSPTO2264	C	+		hypothetical protein	59	109								
PSPTO2271	C	+		hypothetical protein	58.6	37								
PSPTO2273	C	+		hypothetical protein	51.5	57								
PSPTO2308	C			hypothetical protein	55.4	65								
PSPTO2318	C			hypothetical protein	46.5	104								
PSPTO2323	C			hypothetical protein	51.3	596								
PSPTO2325	C			hypothetical protein	46.9	86								
PSPTO2332	C			hypothetical protein	61.1	155								
PSPTO2356	C	+		hypothetical protein	55.6	127								
PSPTO2357	C	+		hypothetical protein	58.6	141								
PSPTO2360	C	+		hypothetical protein	56.2	89								
PSPTO2361	C	+		conserved domain protein	61.8	76			Xac			Yp		
PSPTO2363	C	+		hypothetical protein	54.4	79								
PSPTO2377	C			hypothetical protein	59.7	101								
PSPTO2389	C	+		hypothetical protein	57.1	260								
PSPTO2390	C	+		hypothetical protein	46.1	256								
PSPTO2391	C	+		hypothetical protein	55.7	106								

PSPTO2394	C	+		radical SAM domain protein	46.3	306								
PSPTO2407	C			conserved hypothetical protein	58.3	214	Rs	Xcc		At	Xf			
PSPTO2431	C			conserved hypothetical protein	43.5	305								
PSPTO2433	C			hypothetical protein	45.7	43								
PSPTO2442	C			CheW domain protein	63.3	487						Yp		
PSPTO2443	C			hypothetical protein	54.5	41								
PSPTO2455	C			hypothetical protein	57.3	78								
PSPTO2456	C			hypothetical protein	56.1	38								
PSPTO2497	C	+		hypothetical protein	50.4	41								
PSPTO2498	C	+		hypothetical protein	65.8	40								
PSPTO2501	C	+		hypothetical protein	58.1	89								
PSPTO2502	C	+		hypothetical protein	60.2	67								
PSPTO2503	C	+		hypothetical protein	58.7	189								
PSPTO2512	C			hypothetical protein	53.2	173								
PSPTO2513	C			hypothetical protein	54.3	35								
PSPTO2530	C	+		hypothetical protein	54.8	42								
PSPTO2531	C	+		hypothetical protein	57.1	94								
PSPTO2534	C	+		hypothetical protein	49.6	226								
PSPTO2536	C	+		hypothetical protein	56	147								
PSPTO2537	C	+		conserved domain protein	52.5	155								
PSPTO2584	C	+		conserved hypothetical protein	61.1	227				At				
PSPTO2587	C	+		hypothetical protein	52.5	40								
PSPTO2589	C	+		hypothetical protein	58.9	111								
PSPTO2613	C	+		hypothetical protein	54.8	191								

PSPTO2615	C	+		GAF domain protein	54.3	165		Xcc	Xac	At				
PSPTO2619	C	+		hypothetical protein	60.6	218								
PSPTO2631	C			hypothetical protein	59.6	197								
PSPTO2647	C			hypothetical protein	57	246								
PSPTO2648	C			hypothetical protein	57	155								
PSPTO2669	C	+	+	conserved hypothetical protein	58.3	96	Rs			At				
PSPTO2675	C	+	+	conserved hypothetical protein	58.2	196								
PSPTO2679	C	+	+	hypothetical protein	60.8	40								
PSPTO2682	C	+	+	hypothetical protein	59.5	199			Xac	At				
PSPTO2695	C	+	+	conserved hypothetical protein	60.9	353		Xcc	Xac					
PSPTO2698	C	+	+	hypothetical protein	59.6	104								
PSPTO2700	C	+	+	hypothetical protein	56.9	184								
PSPTO2727	C			hypothetical protein	49.5	31								
PSPTO2768	C	+		hypothetical protein	50.6	58								
PSPTO2770	C	+		hypothetical protein	60	35								
PSPTO2772	C	+		hypothetical protein	55.1	141								
PSPTO2774	C	+		conserved hypothetical protein	57.5	131						Yp		
PSPTO2781	C	+		conserved hypothetical protein	62.8	411	Rs			At				
PSPTO2789	C			hypothetical protein	43.1	48								
PSPTO2796	C			conserved hypothetical protein	60.2	283				At				
PSPTO2797	C			hypothetical protein	63.1	56								
PSPTO2809	C	+		conserved hypothetical protein	64.6	486	Rs	Xcc	Xac	At			Ec	St
PSPTO2815	C	+		conserved hypothetical protein	61.3	173	Rs			At		Yp		
PSPTO2817	C	+		conserved domain protein	62.8	61		Xcc	Xac	At		Yp		

PSPTO2819	C	+		conserved hypothetical protein	58.9	262								
PSPTO2820	C	+		hypothetical protein	63.8	371								
PSPTO2821	C	+		hypothetical protein	65.6	302								
PSPTO2823	C	+		hypothetical protein	59.3	249								
PSPTO2824	C	+		auxin-responsive GH3-related protein	61.8	487								
PSPTO2826	C	+		conserved domain protein	58.9	99				At		Yp	Ec	
PSPTO2827	C	+		conserved domain protein	60.8	80				At		Yp		St
PSPTO2835	C			conserved domain protein	58.7	50				At				
PSPTO2858	C	+	+	hypothetical protein	51.7	118								
PSPTO2859	C	+	+	conserved hypothetical protein	56.2	866						Yp		
PSPTO2860	C	+	+	helicase domain protein	57.5	459							Ec	
PSPTO2866	C	+	+	cupin family protein	57.1	133		Xcc	Xac	At		Yp		
PSPTO2867	C	+	+	TspO/MBR family protein	56.8	146		Xcc	Xac	At				
PSPTO2868	C	+	+	conserved hypothetical protein	59.4	392	Rs				Xf			
PSPTO2869	C	+	+	hypothetical protein	43.3	40								
PSPTO2871	C	+	+	conserved hypothetical protein	59.7	1044						Yp		St
PSPTO2873	C	+	+	conserved hypothetical protein	61.2	345								
PSPTO2880	C	+	+	conserved hypothetical protein	61.6	423								
PSPTO2881	C	+	+	hypothetical protein	61.5	32								
PSPTO2888	C	+	+	conserved domain protein	45.1	48								
PSPTO2889	C	+	+	hypothetical protein	55	37								
PSPTO2892	C	+	+	hypothetical protein	52.1	32								
PSPTO2894	C	+	+	lectin repeat domain protein	50	801								
PSPTO2895	C	+	+	hypothetical protein	57.1	339				At				

PSPTO2897	C	+	+	conserved hypothetical protein	58.9	380		Xcc	Xac					
PSPTO2899	C	+	+	conserved hypothetical protein	59.1	394		Xcc	Xac					
PSPTO2908	C	+		hypothetical protein	56.8	230								
PSPTO2918	C	+		conserved hypothetical protein	61.2	135	Rs	Xcc	Xac	At				
PSPTO2924	C	+		hypothetical protein	54.7	53								
PSPTO2929	C	+		hypothetical protein	53.5	320								
PSPTO2931	C	+		conserved hypothetical protein	59.3	394								
PSPTO2939	C	+		hypothetical protein	54.5	170								
PSPTO2942	C	+		hypothetical protein	57.8	34								
PSPTO2959	C	+		conserved hypothetical protein	64.2	551								
PSPTO2963	C	+		oxidoreductase, FAD-binding	65	379	Rs	Xcc	Xac	At				
PSPTO2965	C	+		conserved hypothetical protein	57.9	130	Rs							
PSPTO2977	C	+		conserved hypothetical protein	51.7	211								
PSPTO2979	C	+		conserved hypothetical protein	52.3	128								
PSPTO2982	C	+		conserved hypothetical protein	62.3	576		Xcc		At				
PSPTO2998	C			conserved hypothetical protein	58.7	142	Rs	Xcc	Xac	At	Xf			
PSPTO2999	C			hypothetical protein	54.4	250								
PSPTO3016	C			conserved hypothetical protein	48.4	117	Rs					Yp	Ec	St
PSPTO3018	C			conserved hypothetical protein	58.6	307	Rs			At		Yp		St
PSPTO3026	C			conserved hypothetical protein	55.1	81		Xcc						
PSPTO3055	C			hypothetical protein	54.4	49								
PSPTO3067	C		+	hypothetical protein	59.1	230								
PSPTO3078	C		+	conserved hypothetical protein	52.9	128								
PSPTO3092	C		+	conserved hypothetical protein	60.1	278		Xcc	Xac					

PSPTO3104	C			hypothetical protein	59	48								
PSPTO3117	C			hypothetical protein	57.2	60								
PSPTO3120	C			Cof-like hydrolase family protein	60.4	272						Yp	Ec	St
PSPTO3151	C			conserved hypothetical protein	64.5	61				At				
PSPTO3166	C	+		conserved domain protein	58.5	320								
PSPTO3167	C	+		hypothetical protein	60.7	101								
PSPTO3168	C	+		hypothetical protein	39.1	46								
PSPTO3169	C	+		hypothetical protein	62.1	87								
PSPTO3177	C			conserved hypothetical protein	57.6	99								
PSPTO3187	C			hypothetical protein	53.9	86								
PSPTO3189	C			conserved hypothetical protein	62.1	317								
PSPTO3200	C	+		hypothetical protein	60.5	290								
PSPTO3201	C	+		hypothetical protein	60.6	259								
PSPTO3203	C	+		conserved hypothetical protein	51.7	154								
PSPTO3211	C			conserved hypothetical protein	40.5	171						Yp		
PSPTO3217	C	+		hypothetical protein	50.3	55								
PSPTO3218	C	+		hypothetical protein	52.6	398								
PSPTO3219	C	+		hypothetical protein	57.7	41								
PSPTO3222	C	+		hypothetical protein	53.1	71								
PSPTO3228	C	+		hypothetical protein	53	56								
PSPTO3236	C			hypothetical protein	51.9	131								
PSPTO3241	C			phytase domain protein	62.9	640								
PSPTO3246	C			hypothetical protein	59.3	232							Ec	St
PSPTO3253	C			conserved hypothetical protein	55	240						Yp	Ec	St

PSPTO3261	C			hypothetical protein	58.3	84								
PSPTO3270	C	+		conserved hypothetical protein	55.7	94								
PSPTO3271	C	+		hypothetical protein	60.3	110								
PSPTO3276	C	+		hypothetical protein	44.4	39								
PSPTO3289	C	+		hypothetical protein	54.6	72								
PSPTO3292	C	+		hypothetical protein	53.1	404	Rs							
PSPTO3293	C	+		hypothetical protein	57.2	416								
PSPTO3303	C			hypothetical protein	55.4	62								
PSPTO3305	C			hypothetical protein	54.2	280								
PSPTO3324	C			hypothetical protein	50.5	35								
PSPTO3337	C			hypothetical protein	53.2	195								
PSPTO3384	C	+		hypothetical protein	52.9	68								
PSPTO3386	C	+		hypothetical protein	50.4	784				At				
PSPTO3387	C	+		conserved hypothetical protein	41.6	109							Ec	
PSPTO3402	C			hypothetical protein	62.7	92								
PSPTO3414	C	+		hypothetical protein	61.1	48								
PSPTO3415	C	+		conserved domain protein	57.6	128								
PSPTO3420	C	+		hypothetical protein	56.8	199	Rs							
PSPTO3422	C	+		hypothetical protein	55	191								
PSPTO3424	C	+		hypothetical protein	54.6	124								
PSPTO3426	C	+		hypothetical protein	53.3	140								
PSPTO3428	C	+		conserved domain protein	55.7	152	Rs							
PSPTO3429	C	+		DNA-binding protein, putative	53.8	80	Rs		Xac		Xf	Yp	Ec	
PSPTO3468	C			conserved hypothetical protein	65.2	467		Xcc	Xac	At				

PSPTO3478	C	+		hypothetical protein	53.2	79		Xcc	Xac					
PSPTO3481	C	+		hypothetical protein	55.7	76								
PSPTO3483	C	+		hypothetical protein	52.7	282								
PSPTO3486	C	+		hypothetical protein	50.9	38								
PSPTO3491	C	+		hypothetical protein	58.2	122								
PSPTO3509	C			hypothetical protein	47.5	40								
PSPTO3518	C			hypothetical protein	60.2	41								
PSPTO3525	C			hypothetical protein	52.9	143								
PSPTO3542	C	+		conserved hypothetical protein	60.7	305								
PSPTO3544	C	+		conserved hypothetical protein	60.1	381				At				
PSPTO3546	C	+		hypothetical protein	51.8	47								
PSPTO3560	C	+		GDA1/CD39 family protein	50.7	402								
PSPTO3562	C	+		hypothetical protein	60.6	33								
PSPTO3567	C	+		hypothetical protein	60.4	466								
PSPTO3568	C	+		hypothetical protein	53.7	41								
PSPTO3571	C	+		hypothetical protein	57.9	38								
PSPTO3575	C	+		hypothetical protein	53.6	173								
PSPTO3578	C	+		hypothetical protein	53.3	162								
PSPTO3597	C	+		conserved hypothetical protein	63.2	402						Yp	Ec	
PSPTO3599	C	+		conserved domain protein	58.5	274						Yp	Ec	
PSPTO3606	C	+		hypothetical protein	47.4	52								
PSPTO3610	C	+		hypothetical protein	47.4	559								
PSPTO3616	C	+		hypothetical protein	45.1	328		Xcc	Xac		Xf			
PSPTO3622	C			hypothetical protein	60.7	112								

PSPTO3623	C			hypothetical protein	60.3	121								
PSPTO3640	C			conserved hypothetical protein	63.1	378	Rs			At				
PSPTO3650	C	+		hypothetical protein	42.3	426								
PSPTO3654	C	+		acetyltransferase, GNAT family	59.3	162								
PSPTO3655	C	+		conserved hypothetical protein	60.7	388	Rs			At				
PSPTO3673	C			hypothetical protein	53.3	40								
PSPTO3682	C	+		conserved hypothetical protein	63.4	375	Rs	Xcc	Xac					
PSPTO3683	C	+		hypothetical protein	54.6	61								
PSPTO3684	C	+		hypothetical protein	54.6	133								
PSPTO3690	C			conserved hypothetical protein	48.7	286			Xac					
PSPTO3693	C			hypothetical protein	60.4	95								
PSPTO3697	C			hypothetical protein	61.4	38								
PSPTO3710	C			acetyltransferase, GNAT family	57.4	184								
PSPTO3728	C			hypothetical protein	56.1	38								
PSPTO3732	C			hypothetical protein	52.3	79								
PSPTO3736	C			conserved hypothetical protein	57	90	Rs			At				
PSPTO3754	C			hypothetical protein	50.7	297								
PSPTO3767	C			hypothetical protein	52.7	86								
PSPTO3775	C	+		hypothetical protein	59.9	94								
PSPTO3776	C	+		hypothetical protein	47.5	148								
PSPTO3780	C	+		glutathionylspermidine synthase family protein	58.2	385		Xcc	Xac	At		Yp	Ec	St
PSPTO3781	C	+		conserved hypothetical protein	61.4	240		Xcc	Xac				Ec	
PSPTO3782	C	+		conserved hypothetical protein	59.3	135	Rs	Xcc	Xac	At			Ec	St
PSPTO3783	C	+		conserved hypothetical protein	59.5	218		Xcc	Xac				Ec	St

PSPTO3785	C	+		rhomboid family protein	58.4	185		Xcc	Xac					
PSPTO3794	C	+		hypothetical protein	53.8	137								
PSPTO3795	C	+		hypothetical protein	50.2	97								
PSPTO3851	C	+		hypothetical protein	58.9	1104								
PSPTO3852	C	+		hypothetical protein	55.2	306								
PSPTO3853	C	+		hypothetical protein	55	299								
PSPTO3865	C	+		conserved hypothetical protein	56.2	398								
PSPTO3866	C	+		hypothetical protein	57.6	177								
PSPTO3868	C	+		hypothetical protein	58	285								
PSPTO3891	C	+		hypothetical protein	59.4	471								
PSPTO3892	C	+		hypothetical protein	60	205								
PSPTO3895	C	+		hypothetical protein	51.9	131								
PSPTO3896	C	+		hypothetical protein	51.2	41								
PSPTO3904	C			hypothetical protein	52	50								
PSPTO3907	C			conserved hypothetical protein	54	250								
PSPTO3927	C	+		hypothetical protein	54.1	221								
PSPTO3931	C	+		hypothetical protein	60.3	47								
PSPTO3934	C	+		tail fiber assembly domain protein	54.7	92	Rs					Yp	Ec	St
PSPTO3938	C	+		hypothetical protein	49.5	122								
PSPTO3942	C	+		hypothetical protein	59.5	200								
PSPTO3944	C	+		hypothetical protein	58.8	225								
PSPTO3945	C	+		hypothetical protein	61.1	66								
PSPTO3946	C	+		hypothetical protein	56.8	155								
PSPTO3947	C	+		hypothetical protein	56.1	85								

PSPTO3948	C	+		hypothetical protein	61.9	127								
PSPTO3961	C	+		hypothetical protein	41.4	66								
PSPTO3962	C	+		hypothetical protein	43.1	34								
PSPTO3965	C	+		hypothetical protein	51.3	517								
PSPTO4003	C	+	+	conserved hypothetical protein	62.9	178								
PSPTO4004	C	+	+	conserved hypothetical protein	63.1	310								
PSPTO4005	C	+	+	hypothetical protein	58.5	142								
PSPTO4007	C	+	+	conserved hypothetical protein	61.9	174								
PSPTO4008	C	+	+	hypothetical protein	59.8	135								
PSPTO4011	C	+	+	hypothetical protein	57.1	94								
PSPTO4012	C	+	+	hypothetical protein	58.6	253								
PSPTO4015	C	+	+	hypothetical protein	55.3	82								
PSPTO4016	C	+	+	hypothetical protein	56.4	68								
PSPTO4017	C	+	+	hypothetical protein	53.8	179								
PSPTO4028	C			hypothetical protein	59.3	41								
PSPTO4035	C	+		hypothetical protein	57.7	208								
PSPTO4040	C	+		conserved domain protein	49.7	122								
PSPTO4042	C	+		conserved hypothetical protein	60.1	132								
PSPTO4043	C	+		hypothetical protein	61.3	200								
PSPTO4044	C	+		hypothetical protein	61.2	194								
PSPTO4045	C	+		hypothetical protein	51.2	99								
PSPTO4047	C	+		hypothetical protein	50.8	120								
PSPTO4048	C	+		conserved hypothetical protein	58.2	83								
PSPTO4049	C	+		conserved hypothetical protein	60.1	132								

PSPTO4050	C	+		hypothetical protein	62.1	250								
PSPTO4051	C	+		hypothetical protein	59.9	191								
PSPTO4052	C	+		hypothetical protein	56.8	288								
PSPTO4053	C	+		hypothetical protein	54.2	32								
PSPTO4054	C	+		hypothetical protein	59.5	270								
PSPTO4055	C	+		hypothetical protein	58.3	64								
PSPTO4056	C	+		hypothetical protein	53.2	195								
PSPTO4066	C	+		conserved domain protein	52.6	97								
PSPTO4069	C	+		hypothetical protein	46.9	54								
PSPTO4070	C	+		conserved hypothetical protein	50.3	145						Yp		
PSPTO4071	C	+		conserved hypothetical protein	57.4	191			Xac					
PSPTO4085	C			hypothetical protein	61.7	54								
PSPTO4086	C			hypothetical protein	61.1	54								
PSPTO4096	C	+	+	conserved domain protein	56.5	46								St
PSPTO4099	C	+	+	hypothetical protein	57.8	49								
PSPTO4102	C	+	+	hypothetical protein	54.9	65								
PSPTO4118	C			hypothetical protein	44.7	38								
PSPTO4150	C			hypothetical protein	56.7	120								
PSPTO4153	C			hypothetical protein	57.3	132								
PSPTO4177	C	+		2-hydroxychromene-2-carboxylate isomerase family protein	58.8	215	Rs	Xcc	Xac	At	Xf			
PSPTO4184	C	+		hypothetical protein	61.1	66								
PSPTO4186	C	+		hypothetical protein	43.5	226								
PSPTO4187	C	+		hypothetical protein	47.8	217								
PSPTO4189	C	+		hypothetical protein	43.4	229								

PSPTO4190	C	+		hypothetical protein	49.5	289								
PSPTO4197	C	+		conserved hypothetical protein	62.6	398				At				
PSPTO4199	C	+		conserved hypothetical protein	63.6	99								
PSPTO4207	C	+		hypothetical protein	61.5	291		Xcc	Xac					
PSPTO4209	C	+		hypothetical protein	58.9	223		Xcc	Xac					
PSPTO4241	C			conserved hypothetical protein	61.5	241				At				
PSPTO4242	C			conserved hypothetical protein	59.3	205				At				
PSPTO4250	C			DNA-binding protein	53.2	378								
PSPTO4261	C	+		isomerase, putative	50.4	232	Rs	Xcc	Xac	At	Xf			
PSPTO4264	C	+		conserved hypothetical protein	55.4	184								
PSPTO4272	C	+		hypothetical protein	47.8	494								
PSPTO4282	C	+		hypothetical protein	55.1	46								
PSPTO4284	C	+		acetyltransferase, GNAT family	60.1	148								
PSPTO4286	C	+		conserved hypothetical protein	57.1	406	Rs							
PSPTO4287	C	+		hypothetical protein	52	336								
PSPTO4311	C			hypothetical protein	53.6	74								
PSPTO4313	C			hypothetical protein	61.4	69								
PSPTO4321	C	+	+	hypothetical protein	55.1	330								
PSPTO4322	C	+	+	hypothetical protein	52.6	606								
PSPTO4323	C	+	+	conserved hypothetical protein	53.2	159								
PSPTO4324	C	+	+	hypothetical protein	54.4	98								
PSPTO4325	C	+	+	hypothetical protein	54.5	502								
PSPTO4326	C	+	+	conserved hypothetical protein	52.4	313			Xac					
PSPTO4327	C	+	+	hypothetical protein	51.4	109								

PSPTO4346	C			hypothetical protein	48.5	114							Yp		St
PSPTO4348	C			hypothetical protein	49.3	161									
PSPTO4364	C			hypothetical protein	57.3	32									
PSPTO4378	C	+		hypothetical protein	52.8	36									
PSPTO4384	C	+		hypothetical protein	56.4	107									
PSPTO4386	C	+		hypothetical protein	52	523									
PSPTO4387	C	+		hypothetical protein	52.1	455									
PSPTO4388	C	+		hypothetical protein	60.3	52									
PSPTO4440	C			hypothetical protein	60.3	42									
PSPTO4520	C	+		methyltransferase domain protein	57.8	783									
PSPTO4525	C	+		hypothetical protein	55.2	55									
PSPTO4529	C	+		hypothetical protein	50.2	162									
PSPTO4532	C	+		hypothetical protein	55.2	102									
PSPTO4542	C			GAF domain protein	55.8	166		Xcc	Xac						
PSPTO4552	C			hypothetical protein	47.2	36									
PSPTO4556	C			conserved domain protein	59.3	559									
PSPTO4570	C	+	+	hypothetical protein	53.5	337									
PSPTO4571	C	+	+	hypothetical protein	52.5	399									
PSPTO4573	C	+	+	hypothetical protein	48.1	36									
PSPTO4574	C	+	+	hypothetical protein	48	370									
PSPTO4591	C	+	+	hypothetical protein	50.3	96									
PSPTO4599	C	+	+	hypothetical protein	58	157									
PSPTO4605	C	+	+	conserved hypothetical protein	54.9	745			Xac					Ec	
PSPTO4606	C	+	+	conserved hypothetical protein	49	132									

PSPTO4607	C	+	+	hypothetical protein	49	819								
PSPTO4609	C	+	+	hypothetical protein	56	81								
PSPTO4614	C	+	+	conserved hypothetical protein	51.7	200								
PSPTO4617	C	+	+	hypothetical protein	51.3	80								
PSPTO4618	C	+	+	hypothetical protein	54.5	107								
PSPTO4620	C	+	+	conserved hypothetical protein	49.2	107								
PSPTO4627	C	+	+	hypothetical protein	48.3	817								
PSPTO4628	C	+	+	hypothetical protein	55.4	68								
PSPTO4629	C	+		hypothetical protein	54.3	35								
PSPTO4649	C			hypothetical protein	49.7	57								
PSPTO4655	C			hypothetical protein	62.1	289								
PSPTO4658	C			conserved domain protein	58.4	329	Rs							
PSPTO4670	C	+		hypothetical protein	54.7	265		Xcc	Xac					
PSPTO4672	C	+		hypothetical protein	57.9	57								
PSPTO4674	C	+		hypothetical protein	51.7	58								
PSPTO4676	C	+		hypothetical protein	56	69								
PSPTO4677	C	+		hypothetical protein	57	213								
PSPTO4688	C		+	hypothetical protein	65.1	147								
PSPTO4716	C	+	+	hypothetical protein	54.7	165								
PSPTO4717	C	+	+	hypothetical protein	59.6	89								
PSPTO4719	C	+	+	hypothetical protein	60	55								
PSPTO4723	C	+	+	conserved hypothetical protein	48.4	102	Rs			At				
PSPTO4733	C	+	+	hypothetical protein	44.9	179								
PSPTO4734	C	+	+	conserved domain protein	58.7	75								

PSPTO4741	C	+	+	hypothetical protein	45	140								
PSPTO4743	C	+	+	hypothetical protein	44	633								
PSPTO4747	C	+	+	hypothetical protein	53	212								
PSPTO4750	C	+	+	hypothetical protein	54.7	159								
PSPTO4752	C	+	+	hypothetical protein	56.1	321								
PSPTO4753	C	+	+	hypothetical protein	56.2	363								
PSPTO4754	C	+	+	hypothetical protein	51.4	281								
PSPTO4756	C	+	+	hypothetical protein	53.8	44								
PSPTO4757	C	+	+	hypothetical protein	55.3	73								
PSPTO4758	C	+	+	hypothetical protein	51.5	44								
PSPTO4761	C	+	+	hypothetical protein	51.4	673							Ec	
PSPTO4762	C	+	+	von Willebrand factor type A domain protein	52.8	224						Yp	Ec	
PSPTO4763	C	+	+	hypothetical protein	55	260							Ec	
PSPTO4766	C	+	+	hypothetical protein	54.6	102								
PSPTO4767	C	+	+	pentapeptide repeat protein	45.4	327								
PSPTO4772	C	+	+	hypothetical protein	50.8	346								
PSPTO4780	C	+	+	hypothetical protein	59.2	179								
PSPTO4783	C	+	+	hypothetical protein	55.8	49								
PSPTO4785	C	+	+	hypothetical protein	51.9	36								
PSPTO4797	C			hypothetical protein	50	72								
PSPTO4801	C			hypothetical protein	45.8	51								
PSPTO4811	C			hypothetical protein	59	39								
PSPTO4815	C			hypothetical protein	51.5	68								
PSPTO4838	C			hypothetical protein	55.1	49								

PSPTO4870	C	+		hypothetical protein	55	211								
PSPTO4871	C	+		hypothetical protein	55	211								
PSPTO4872	C	+		hypothetical protein	57.3	232								
PSPTO4884	C			hypothetical protein	57	768								
PSPTO4902	C			hypothetical protein	55.8	77								
PSPTO4927	C			hypothetical protein	54.3	914								
PSPTO4966	C	+		hypothetical protein	52.9	128								
PSPTO4967	C	+		hypothetical protein	44.4	220								
PSPTO4968	C	+		hypothetical protein	54.6	213								
PSPTO5025	C			hypothetical protein	43.6	123								
PSPTO5026	C			hypothetical protein	55.6	36								
PSPTO5042	C			hypothetical protein	65.2	46								
PSPTO5059	C			hypothetical protein	56.6	189								
PSPTO5073	C			hypothetical protein	49.4	79								
PSPTO5088	C			hypothetical protein	52.5	33								
PSPTO5091	C	+		conserved hypothetical protein	60.1	239	Rs	Xcc	Xac				Ec	
PSPTO5098	C	+		conserved hypothetical protein	63.9	315	Rs	Xcc	Xac		Xf		Ec	
PSPTO5101	C	+		conserved hypothetical protein	59.8	204	Rs	Xcc	Xac				Ec	
PSPTO5108	C	+		conserved hypothetical protein	62.9	151	Rs	Xcc	Xac				Ec	
PSPTO5113	C	+		conserved hypothetical protein	52.6	422								St
PSPTO5114	C	+		conserved domain protein	53	214								
PSPTO5122	C			hypothetical protein	52.1	32								
PSPTO5124	C			hypothetical protein	56.3	32								
PSPTO5183	C			hypothetical protein	59.7	239								

PSPTO5199	C	+		hypothetical protein	56.8	54								
PSPTO5202	C	+		hypothetical protein	53.1	332								
PSPTO5203	C	+		hypothetical protein	52.6	359						Yp		
PSPTO5205	C	+		hypothetical protein	51.9	77								
PSPTO5206	C	+		EF hand domain protein	45.4	310								
PSPTO5209	C	+		hypothetical protein	45.2	392								
PSPTO5210	C	+		EF hand domain protein	44.3	310								
PSPTO5234	C	+		hypothetical protein	51.8	191								
PSPTO5235	C	+		hypothetical protein	56.5	118								
PSPTO5236	C	+		hypothetical protein	58.6	37								
PSPTO5257	C			hypothetical protein	59	216								
PSPTO5281	C			hypothetical protein	54.3	78								
PSPTO5298	C			hypothetical protein	58.7	63								
PSPTO5331	C			hypothetical protein	53.8	630								
PSPTO5332	C			trypsin domain protein	58.2	510								
PSPTO5345	C	+	+	hypothetical protein	53.8	318								
PSPTO5346	C	+	+	hypothetical protein	60.8	397								
PSPTO5347	C	+	+	hypothetical protein	59.4	83								
PSPTO5348	C	+	+	hypothetical protein	53.3	366								
PSPTO5349	C	+	+	prevent-host-death family protein	54.7	75								
PSPTO5355	C	+	+	hypothetical protein	56.7	161								
PSPTO5364	C	+	+	hypothetical protein	55	100								
PSPTO5365	C	+	+	conserved domain protein	41.5	345								
PSPTO5373	C	+	+	hypothetical protein	39.9	360								

PSPTO5374	C	+	+	conserved hypothetical protein	51.7	107	Rs				Xf	Yp		
PSPTO5384	C	+		conserved hypothetical protein	52.8	113								
PSPTO5385	C	+		hypothetical protein	57.3	245								
PSPTO5386	C	+		hypothetical protein	56.2	35								
PSPTO5392	C	+		hypothetical protein	54.2	75								
PSPTO5410	C			hypothetical protein	63.3	270	Rs							
PSPTO5413	C			EF hand domain protein	49.1	792								
PSPTO5429	C			hypothetical protein	43.9	38								
PSPTO5437	C			hypothetical protein	55.7	227								
PSPTO5442	C	+		hypothetical protein	61.1	119								
PSPTO5446	C	+		hypothetical protein	55.8	40								
PSPTO5447	C	+		conserved hypothetical protein	51.6	142		Xcc	Xac					St
PSPTO5449	C	+		hypothetical protein	61.4	166								
PSPTO5450	C	+		conserved hypothetical protein	59.7	101		Xcc						
PSPTO5455	C	+		conserved hypothetical protein	60.9	258								
PSPTO5456	C	+		hypothetical protein	58.6	309								
PSPTO5458	C	+		conserved hypothetical protein	53.4	158								
PSPTO5459	C			hypothetical protein	50.5	31								
PSPTO5471	C			hypothetical protein	59.6	142								
PSPTO5481	C			hypothetical protein	57.4	1550								
PSPTO5495	C	+		hypothetical protein	53.7	210								
PSPTO5496	C	+		hypothetical protein	52.3	51								
PSPTO5497	C	+		hypothetical protein	52.2	210								
PSPTO5513	C			hypothetical protein	58.4	73								

PSPTO5514	C			conserved domain protein	52.1	103								
PSPTO5531	C			hypothetical protein	61.7	346								
PSPTO5552	C			hypothetical protein	61.8	41								
PSPTO5555	C			hypothetical protein	54.1	77								
PSPTO5565	C			hypothetical protein	51.6	213								
PSPTO5568	C			hypothetical protein	52.4	49								
PSPTO5583	C			Ais protein, putative	52.6	228							Ec	St
PSPTO5592	C			hypothetical protein	53.5	293								
PSPTOA0015	pA	+	+	conserved hypothetical protein	50.5	72								
PSPTOA0022	pA	+	+	hypothetical protein	54.2	193								
PSPTOA0023	pA	+	+	hypothetical protein	65.5	116								
PSPTOA0024	pA	+	+	hypothetical protein	56	125								
PSPTOA0025	pA	+	+	conserved hypothetical protein	56.4	278	Rs						Ec	
PSPTOA0026	pA	+	+	conserved hypothetical protein	52.7	62								
PSPTOA0031	pA	+	+	conserved hypothetical protein	50.6	393	Rs	Xcc					Yp	Ec
PSPTOA0033	pA	+	+	hypothetical protein	56.1	132								
PSPTOA0038	pA	+	+	hypothetical protein	40.4	61								
PSPTOA0040	pA	+		conserved hypothetical protein	55	177							Yp	Ec
PSPTOA0046	pA	+		hypothetical protein	57.5	69								
PSPTOA0057	pA	+		conserved hypothetical protein	57.1	91		Xcc	Xac	At	Xf			
PSPTOA0058	pA	+		conserved hypothetical protein	57.1	66					Xf		Ec	
PSPTOA0061	pA	+		hypothetical protein	56.8	78								
PSPTOA0069	pA			hypothetical protein	51.7	140								
PSPTOA0070	pA			hypothetical protein	59.3	207								

PSPTOB0003	pB	+		conserved hypothetical protein	44.8	157		Xcc				Yp		
PSPTOB0006	pB	+		hypothetical protein	61.9	42								
PSPTOB0012	pB	+		hypothetical protein	56.8	352			Xac					
PSPTOB0015	pB	+		hypothetical protein	50	98								
PSPTOB0016	pB	+		conserved domain protein	57.3	117								
PSPTOB0020	pB	+		hypothetical protein	53.9	315								
PSPTOB0023	pB	+		hypothetical protein	52.5	73								
PSPTOB0027	pB	+		hypothetical protein	57.7	138								
PSPTOB0028	pB	+		hypothetical protein	68.1	91								
PSPTOB0029	pB	+		hypothetical protein	58.7	84								
PSPTOB0030	pB	+		conserved hypothetical protein	57.1	278	Rs						Ec	
PSPTOB0031	pB	+		conserved hypothetical protein	49.5	66								
PSPTOB0033	pB	+		hypothetical protein	48.4	93								
PSPTOB0034	pB	+		hypothetical protein	52.3	114								
PSPTOB0037	pB	+		hypothetical protein	50.7	69								
PSPTOB0044	pB	+		hypothetical protein	37.2	61								
PSPTOB0046	pB	+		conserved hypothetical protein	55.6	177						Yp	Ec	St
PSPTOB0047	pB	+		hypothetical protein	51.3	102								
PSPTOB0053	pB	+		hypothetical protein	55.7	231								
PSPTOB0055	pB	+		hypothetical protein	56.8	81								
PSPTOB0070	pB	+		hypothetical protein	54.1	90								
PSPTOB0072	pB	+		conserved hypothetical protein	59	400								
PSPTOB0073	pB	+		PIN domain protein	57.3	125								
PSPTOB0074	pB	+		conserved hypothetical protein	56.6	66								

The inferred proteomes of eight pathogenic bacteria were searched with the 811 genes unique to DC3000 by using a BLASTP cutoff criterion of $E < 10^{-5}$ which will identify putative homologs as well as paralogous gene family members. Rs = *Ralstonia solanacearum*, Xac = *Xanthomonas axonopodis* pv. *citri*, Xcc = *Xanthomonas campestris* pv. *campestris*, At = *Agrobacterium tumefaciens* (U. Wash seq.), Xf = *Xylella fastidiosa*, Yp = *Yersinia pestis* CO92, St = *Salmonella typhimurium* LT2, and Ec = *Escherichia coli* O157:H7 EDL933.

^a Individual unique genes that cluster with one another at a density of at least three genes within a 10-kb region.

^b Genes or clusters (as defined in ^a) that are present within 20 kb of TTSS components and effectors.

^c The annotation of some of the genes in the table is based on multiple alignments to a nonredundant amino acid database as well as Hidden Markov Models (HMMs), which identify domains in protein families (Pfam and TIGRFAMs). Additional information on the individual loci and protein families is available in The Institute for Genomic Research Comprehensive Microbial Resource (<http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl>).