OPEN ACCESS Freely available online

PLoS PATHOGENS

Terminal Reassortment Drives the Quantum Evolution of Type III Effectors in Bacterial Pathogens

John Stavrinides[1☯*], Wenbo Ma[1☯], David S. Guttman[1,2]

1 Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada, 2 Center for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, Canada

Many bacterial pathogens employ a type III secretion system to deliver type III secreted effectors (T3SEs) into host cells, where they interact directly with host substrates to modulate defense pathways and promote disease. This interaction creates intense selective pressures on these secreted effectors, necessitating rapid evolution to overcome host surveillance systems and defenses. Using computational and evolutionary approaches, we have identified numerous mosaic and truncated T3SEs among animal and plant pathogens. We propose that these secreted virulence genes have evolved through a shuffling process we have called "terminal reassortment." In terminal reassortment, existing T3SE termini are mobilized within the genome, creating random genetic fusions that result in chimeric genes. Up to 32% of T3SE families in species with relatively large and well-characterized T3SE repertoires show evidence of terminal reassortment, as compared to only 7% of non-T3SE families. Terminal reassortment may permit the near instantaneous evolution of new T3SEs and appears responsible for major modifications to effector activity and function. Because this process plays a more significant role in the evolution of T3SEs than non-effectors, it provides insight into the evolutionary origins of T3SEs and may also help explain the rapid emergence of new infectious agents.

Citation: Stavrinides J, Ma W, Guttman DS (2006) Terminal reassortment drives the quantum evolution of type III effectors in bacterial pathogens. PLoS Pathog 2(10): e104. DOI: 10.1371/journal.ppat.0020104

Introduction

The type III secretion system (T3SS) and the proteins that traverse it are essential components of the virulence arsenal of many destructive bacterial pathogens. Pathogens utilize the T3SS to inject type III secreted effectors (T3SEs) into the host cell cytosol where they function to promote disease by facilitating cell attachment and entry, suppressing the host defense response, and modulating vesicular traffic, the host cytoskeleton, and hormones [1–3]. Consequently, T3SEs play a prominent role in bacterial pathogenesis and host-association [4] .

T3SEs are modular proteins, with the signals required to direct secretion from the bacterial cell and translocation into the host cell generally localized to the N terminus of the protein, and the functional domains typically localized to the central and C-terminal portions [5,6]. This modularity has been exploited in genetic screens designed to identify new T3SEs [7–9] and for creating reporters to monitor T3SS-dependent secretion [6,10]. Another feature common to the vast majority of T3SEs is that they are co-regulated with the assembly of the T3SS. This is achieved through T3SS-specific transcription factors that bind regulatory motifs found immediately upstream of nearly every T3SE and T3SS structural gene [11]. Motifs include the *Pseudomonas* hrp box [12], the *Salmonella* ssrAB box [13], and the *Xanthomonas* pip box [14]. An important functional consequence of the respective locations of these two T3SE features is that the regulatory motifs required for transcriptional activation of T3SEs are tightly linked to the signals required for secretion and translocation.

Despite several commonalities, T3SEs are evolutionarily diverse and highly variable in their distribution, both within and among species [9]. Their intimate interactions with host factors expose them to very strong selective pressures [15,16] resulting in their rapid evolutionary turnover [17,18]. Given their high degree of genetic diversity, it is perhaps not surprising that a number of studies have identified T3SEs that are truncated or chimeras of other sequences [10,19–22]; nevertheless, no study has recognized this particular form of variation as being significant from the perspective of either evolution or function. Here, we report on an evolutionary process that plays an extremely important role in the evolution of new T3SEs. Unlike previous studies that have reported the introduction of genetic variation through homologous intragenic recombination [23–28], the evolution of T3SEs is strongly influenced by a non-homologous recombinational process that is analogous to exon shuffling seen in eukaryotes. This process not only explains the high

Editor: Jeffrey Dangl, University of North Carolina, United States of America

Received June 29, 2006; **Accepted** August 28, 2006; **Published** October 13, 2006

DOI: 10.1371/journal.ppat.0020104

Abbreviations: ORPHET, orphaned effector terminus; T3SE, type III secreted effector; T3SS, type III secretion system

* To whom correspondence should be addressed. E-mail: johnstav@botany.utoronto.ca

☯ These authors contributed equally to this work.

Synopsis

Many pathogenic bacteria rely on specialized virulence proteins to cause disease. These proteins, known as type III secreted effectors (T3SEs), are directly injected into the host's cells and facilitate the disease process by interacting with host proteins and interfering with the defense response. Although most T3SEs lack any sequence similarity, several T3SEs share a common terminus, suggesting that part of these proteins was derived from the same sequence. The authors propose an evolutionary mechanism, called "terminal reassortment," in which the termini of T3SEs reassort with other genetic information to create new chimeric proteins. This study shows that this process has given rise to T3SEs with new virulence functions and that it may influence bacterial host specificity. Chimeric T3SEs are present in eight different genera and in some cases are present in as many as 32% of known T3SE families. This is significantly more than what is found in other protein families, suggesting that terminal reassortment plays a disproportionately important role in the evolution of T3SE. Terminal reassortment may lead to the very rapid evolution of new T3SEs, thereby contributing to the emergence of new infectious diseases.

frequency of truncated and chimeric T3SEs, but also provides insight into how the particular structure of these genes contributes to dramatic evolutionary and functional changes that may play a key role in the ongoing arms race between pathogen and host.

Results/Discussion

The complete nucleotide and protein sequences of all experimentally confirmed T3SEs were collected from the 23 species having at least one characterized T3SE (Table S1). Using a combination of BLASTP, TBLASTX, and pair-wise BLAST (BL2SEQ), we identified two common and interrelated features among T3SEs from all species. First, the N or C terminus of many T3SEs are homologous (E value $<10^{-5}$) to other loci or open reading frames. We will refer to these homologs as orphan effector termini (ORPHETs). Second, many T3SEs are chimeras of either two known T3SEs, or a T3SE and another gene.

ORPHETs and chimeric T3SEs were found in nine species, representing plant pathogens, animal pathogens, and mutualists. The largest number of chimeric T3SEs and ORPHETs were identified in the species with the largest T3SE repertoires, which included *Pseudomonas syringae* (12 of 56 T3SE families, 21%), *Salmonella enterica* (5/28, 18%), and *Xanthomonas campestris* (7/22, 32%) (Table 1). The high frequency of ORPHETs and chimeric effectors has led us to propose that the evolution of new T3SEs frequently occurs by the reassortment of termini from preexisting T3SEs. This stochastic process, which we will refer to as "terminal reassortment," involves the fusion of an existing T3SE N terminus to another T3SE, or an unrelated coding or non-coding sequence (Figure 1). Alternatively, terminal reassortment may occur via a large deletion that brings the N terminus of a T3SE into contact with a region downstream of the effector (Figure 1). Terminal reassortment does not describe or rely on a specific recombinational mechanism, but rather describes an evolutionary process that results in the rapid formation of new T3SEs via a single quantum evolutionary step in which the regulatory and secretion/translocation signals are coupled to new sequences.

P. syringae, which has approximately 190 T3SE homologs and derivatives distributed among 56 T3SE families [10], provides the most interesting examples of ORPHETs and ORPHET-derived T3SEs. For example, one group of related T3SEs in *P. syringae* contains two ORPHETs, HopW1–$2_{PmaES4326}$ and HopW1–$2_{Pph1448A}$, and two larger T3SEs, HopW1–$1_{PmaES4326}$ and HopAE$1_{PsyB728a}$, all of which are homologous along their first 85 amino acids (Figure 2A). The two ORPHETs are clearly related to the N terminus of the larger T3SEs, while the two larger T3SEs are chimeras themselves, with HopW1–$1_{PmaES4326}$ sharing C-terminal homology to a prophage-related sequence from *Escherichia coli* O157:H7 EDL933, and HopAE$1_{PsyB728a}$ [29] sharing C-terminal homology to several putative *Xanthomonas* virulence genes. These data suggest that all four of these loci derived from a common N-terminal ORPHET. Indeed, *hopW1–1*$_{PmaES4326}$ is situated near its ORPHET on plasmid pPMA4326B and is flanked by numerous IS elements [30]. All of these related T3SEs except *hopW1–1*$_{PmaES4326}$ also have a highly conserved ribosome-binding site and hrp box. Surprisingly, the upstream region of *hopW1–1*$_{PmaES4326}$ is closely related to the upstream region of a completely unrelated *P. syringae* pv. tomato DC3000 T3SE, *hopD1* (64% nucleotide identity over 300 bp) (Figure 2A). This exchange of promoters is functionally significant since changes to the hrp box sequence and its relative distance from the start codon has been shown to alter T3SE expression and bacterial virulence [31].

An analysis of the flanking regions of ORPHETs indicates that they are not merely misannotations due to frameshift mutations in "full-length" T3SEs, nor are they incomplete sequence submissions. ORPHETs are homologous to their respective T3SEs and exhibit both conservation in either their 5′ or 3′ flanking regions (Table S2), and variation at the end of their protein-coding region, which is consistent with random mobilization and insertion, or large-scale deletion downstream of the T3SE secretion signal. Furthermore, because a single ORPHET can constitute the N terminus of multiple effectors, as with the HopW1/HopAE1 (Figure 2A) and HopAB3 effector families (Table 1), ORPHETs likely serve as precursors for the formation of new T3SEs.

Xanthomonas provides an excellent example of chimeric T3SEs. T3SEs XopJ [32], AvrXccE1, and hypothetical protein XAC3230 [33] share a common N terminus, but show absolutely no similarity throughout their C terminus (Figure 2B). The 100-bp region immediately upstream of all three effectors shares 70% nucleotide identity and contains a highly conserved ribosome-binding site and pip promoter box.

Seven *S. enterica* effectors, *srfH/sseI, sseJ, sspH1, sspH2, slrP, sifA,* and *sifB* share a common N terminus [21,34,35] and provide some of the most complex and interesting examples of terminal reassortment. This group comprises four distinct homology groups, with group I containing the full-length homologs SspH1, SspH2, and SlrP; group II containing SseJ; group III containing the relatively divergent full-length homologs SifA and SifB; and group IV containing only SrfH/SseI (Figure 3A). There is no sequence similarity among the C termini of those T3SEs from the different groups. One example of terminal reassortment within this group involves SspH2 and SrfH, which share 88% amino acid identity in

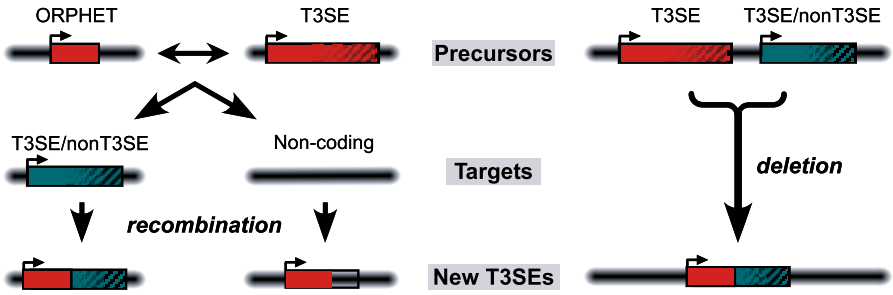

Figure 1. Proposed Mechanisms of Terminal Reassortment

ORPHETs containing the appropriate regulatory and secretion signals may evolve independently or be derived from another effector. ORPHETs can be mobilized and integrated randomly into the genome, capturing either an existing gene or a non-coding region and thereby instantaneously creating a new chimeric effector. Alternatively, deletions downstream of an N-terminal secretion signal may also result in the formation of a new fusion protein.

DOI: 10.1371/journal.ppat.0020104.g001

their N terminus and 98% identity in the 257 bp immediately upstream of their coding region that includes the ssrAB promoter box (Figure 3B). A second, yet more complex example, was identified within the full-length group I homologs, using the phylogenetic-based method of bootscanning on a nucleotide alignment of *sspH1, sspH2, slrP,* and the group IV effector *sseI/srfH*. Bootscanning is a general method for identifying recombination breakpoints based on incongruence among genealogies constructed from subregions within a locus. The bootscan revealed that SspH1 is a recombinant protein composed of an SlrP N terminus and an SspH2 C terminus, with a clear breakpoint near base 400 (Figure 3C). Interestingly, SlrP and SspH1, which share the most similar N terminus, are translocated by both *S. enterica* T3SS secretion systems SPI-1 and SPI-2, while all others are translocated by only SPI-2 [36]. This suggests the presence of a common chaperone binding domain or motif within the N terminus of these two functionally distinct proteins, in addition to the already identified WEK(I/M)XXFF motif [35], which permits translocation by both SPI-1 and SPI-2.

Terminal reassortment has been demonstrated experimentally in genetic screens designed to identify T3SS-specific substrates [7–9]. In these screens, an effector C terminus (reporter) is linked to a transposon, which integrates randomly in the bacterial genome, occasionally creating a functional fusion with an appropriate T3SS regulatory and secretion signal. Using these and related techniques, secretion and translocation into the host has been experimentally demonstrated for many ORPHETs, including HopAT1 [42] and HopS1 [43] from *P. syringae,* and OspD1 [44] and OspG

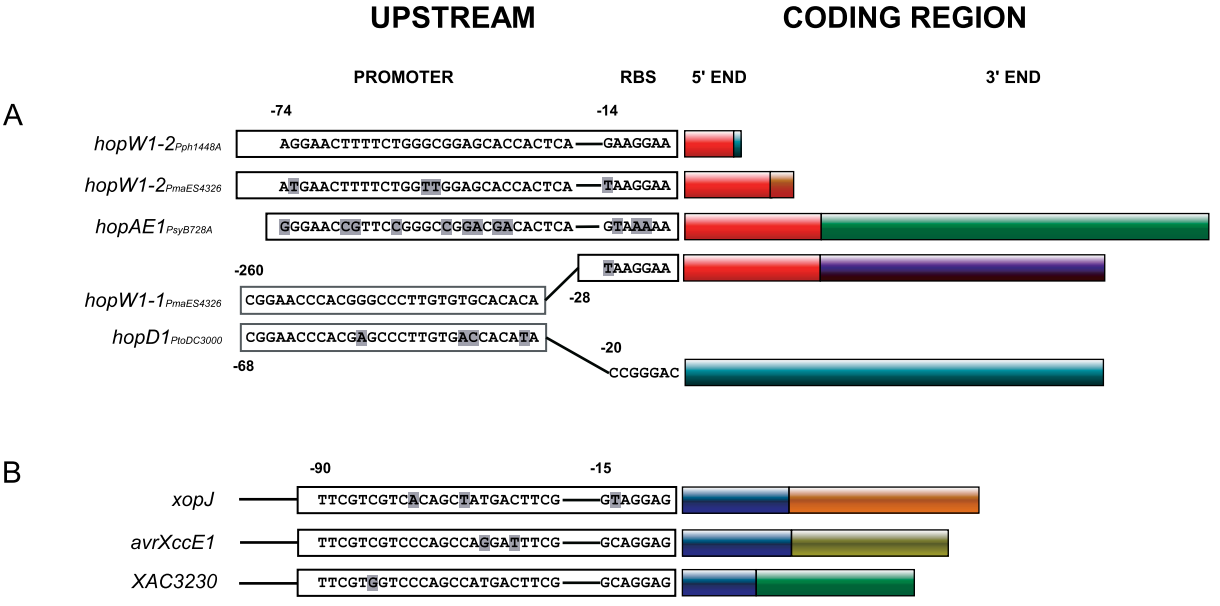

Figure 2. Genetic Organization of ORPHETs and Chimeras Illustrating Conserved Regulatory and Protein-Coding Regions

Homologous gene regions within each panel share the same coloring. The coding region of each T3SE starts at the beginning of the solid black bar, although the length of the upstream region is not to scale with the coding region. The promoter sequences and ribosome-binding sites are indicated, but not shown to scale. Variable residues in these sequences have been highlighted in gray.

(A) The three HopW1 family members along with HopAE1$_{PsyB278a}$ in *P. syringae* share N-terminal homology that extends to the upstream region and includes the ribosome-binding site and hrp box promoter. The hrp box promoter of HopW1–1$_{PmaES4326}$ and flanking region, however, is more closely related to that of HopD1$_{PtoDC3000}$, although these two proteins are not homologous.

(B) *Xanthomonas* spp. effectors XopJ and AvrXccE1 along with hypothetical protein XAC3230 share a common N terminus and display conservation upstream of the protein-coding region.

DOI: 10.1371/journal.ppat.0020104.g002

Table 1. Examples of Terminal Reassortment from Eight Genera, Compiled from Pair-Wise Amino Acid Comparisons of Known T3SEs

ORPHET/Chimera					Homolog				Similarity		
Effector	Species (Strain)[a]	Size (Amino Acid)	Linkage[c]	O/C[b]	Effector	Species (Strain)[a]	Size (Amino Acid)	Linkage[c]	Shared Terminal[d]	Percent aa Identity	Conserved Flanking[e]
Bradyrhizobium											
ID322	Bja (110spc4)	127	tn, T3SS	O	NopP	Bja (110spc4)	277	*hrcU*	C	74/121(60%)	3′
Escherichia											
ECs1566	Eco (O157:H7)	57	tn	O	NleB/Z4328	Cro (DBS100)	328	—	C	32/37(86%)	3′
Z1385	Eco (O157:H7)	250	ph	O	EspFu	Eco (O157:H7)	337	Ph	C	247/250(98%)	3′
Photorhabdus											
Plu2401	Plu (TTO1)	323	tn	C	LopT	Plu (W14)	311	—	C	210/278(75%)	3′
Pseudomonas											
AvrA	Pgy (R6)	907	N/A	C	PsvA	Per (NAE6)	731	tn, pl	N	72/161(44%)	5′
AvrA	Pgy (R6)	907	N/A	C	HopQ1–1	Pto (DC3000)	447	tn, pl	N	—	[g]5′
AvrRps4	Pph (1448A)	221	tn, pl	C	HopAQ1	Pto (DC3000)	84	tn	N	36/44(81%)	5′
HopAB3–2	Pma (ES4326)	155	N/A	O	HopAB3–1	Pma (ES4326)	384	N/A	N	103/176(58%)	5′
HopAB3–2	Pma (ES4326)	155	N/A	O	ORF4	Pph (1449B)	323	tn, pl	N	46/79(57%)	5′
HopAB3–1	Pma (ES4326)	384	N/A	C	ORF4	Pph (1449B)	323	tn, pl	N	51/87(58%)	5′
HopAE1	Psy (B728A)	914	—	C	HopW1–1	Pma (ES4326)	774	tn, pl	N	197/323(60%)	[f]5′
HopAT1	Pph (1448A)	87	hopAJ1	O	HopAV1	Pph (1448A)	807	tn, pl	N	53/66(80%)	5′
HopAT1	Pph (1448A)	87	hopAJ1	O	PsvA	Per (NAE6)	731	tn, pl	N	32/64(49%)	5′
HopAV1	Pph (1448A)	807	tn, pl	C	HopAU1	Pph (1448A)	731	—	N/A	N/A	5′
HopD1	Pto (DC3000)	705	pl	C	HopAO1	Pto (DC3000)	468	—	N	141/198(70%)	5′
HopH1	Pto (DC3000)	218	tn, hopC1	C	HopAP1	Psy (B728A)	176	hopH1	C	29/70(41%)	—
HopK1	Pto (DC3000)	338	—	C	AvrRps4	Pph (1448A)	221	tn, pl	N	127/151(83%)	5′
HopK1	Pto (DC3000)	338	—	C	HopAQ1	Pto (DC3000)	84	tn	N	46/53(86%)	5′
HopK1	Pto (DC3000)	338	—	C	XopO	Xcv (85–10)	220	tn	N	82/147(55%)	—
HopK1	Pto (DC3000)	338	—	C	XAC3666	Xac (306)	221	—	C	65/174 (37%)	—
HopS1	Pto (DC3000)	119	tn	O	HopS2	Pto (DC3000)	177	—	N	61/119(50%)	[h]5′
HopT2	Pto (DC3000)	127	—	O	HopT1–1	Pto (DC3000)	379	—	N	73/106(68%)	[h]5′
HopW1–2	Pph (1448A)	94	tn, pl	O	HopW1–1	Pma (ES4326)	774	tn, pl	N	72/86(82%)	[f]5′
HopW1–2	Pph (1448A)	94	tn, pl	O	HopAE1	Psy (B728A)	914	—	N	50/85(58%)	5′
HopW1–2	Pma (ES4326)	240	tn, pl	O	HopW1–1	Pma (ES4326)	774	tn, pl	N	217/229(94%)	5′
HopW1–2	Pma (ES4326)	240	tn, pl	O	HopAE1	Psy (B728A)	914	—	N	145/230(62%)	5′
Salmonella											
SrfH	Set (14028s)	322	ph	C	SspH2	Set (14028s)	788	Ph	N	130/142(90%)	5′
SrfH	Set (14028s)	322	ph	C	SlrP	Set (LT2)	765	tn	N	48/114 (42%)	—
SrfH	Set (14028s)	322	ph	C	SspH1	Set (14028s)	700	tn, ph	N	63/130 (48%)	—
SseJ	Set (LT2)	408	—	C	SrfH	Set (14028s)	322	Ph	N	60/116(51%)	5′
SseJ	Set (LT2)	408	—	C	SspH1	Set (14028s)	700	tn, ph	N	49/103(47%)	5′
SseJ	Set (LT2)	408	—	C	SspH2	Set (14028s)	788	Ph	N	62/116(53%)	5′
SseJ	Set (LT2)	408	—	C	SlrP	Set (LT2)	765	tn	N	41/80(51%)	5′
SseJ	Set (LT2)	408	—	C	SifB	Set (LT2)	316	—	N	55/115(47%)	5′
SseJpseudo	Set (CT18)	46	—	O	SseJ	Set (LT2)	408	—	C	44/46(95%)	3′
Shigella											
IpgD	Sfl (M90T)	538	—	C	SopB	Set (LT2)	561	Pl	C	187/396(47%)	3′
OspD1	Sfl (M90T)	225	tn, pl	O	OspD2	Sfl (M90T)	569	tn, pl	C	68/133(51%)	—
OspD1	Sfl (M90T)	225	tn, pl	O	OspD3	Sfl (M90T)	565	tn, pl	C	69/130 (52%)	—
OspG	Sfl (M90T)	196	tn, pl	O	NleH/Z6021	Eco (O157:H7)	293	tn, ph	C	63/133(46%)	—

Table 1. continued

ORPHET/Chimera					Homolog				Similarity		
Effector	**Species (Strain)[a]**	**Size (Amino Acid)**	**Linkage[c]**	**O/C[b]**	**Effector**	**Species (Strain)[a]**	**Size (Amino Acid)**	**Linkage[c]**	**Shared Terminal[d]**	**Percent aa Identity**	**Conserved Flanking[e]**
Xanthomonas											
PsvA	Xcc (33913)	442	tn	O	XopD	Xcv (75–3)	545	tn	C	340/442(75%)	3′
XAC2785	Xac (306)	254	—	O	XopF1	Xcv (85–10)	670	—	C	73/158(45%)	—
XAC2785	Xac (306)	254	—	O	XopF2	Xcv (85–10)	667	—	C	118/159(73%)	3′
XAC3230	Xac (306)	296	—	C	AvrXccE1	Xcc (33913)	355	tn	N	44/57(76%)	5′
XAC3230	Xac (306)	296	—	C	XopJ	Xcv (75–3)	373	N/A	N	35/61(56%)	5′
XCC1218	Xcc (33913)	288	hrpF/hrpW	O	XopF1	Xcv (85–10)	670	—	C	230/283(80%)	3′
XCC1218	Xcc (33913)	288	hrpF/hrpW	O	XopF2	Xcv (85–10)	667	—	C	198/287(68%)	—
XopJ	Xcv (75–3)	373	N/A	C	AvrXccE1	Xcc (33913)	355	tn	N	46/87(52%)	5′

The ORPHET/Chimera columns indicate the name, origin, size, and linkage to mobile elements for the observed N- or C-terminal ORPHET or chimeric T3SE. Homolog columns indicate the same information for the full-length T3SE related to the specified ORHPET or chimera. Similarity columns indicate the degree of similarity between the specific ORPHET or chimera and the related T3SE. Further information on each ORHPET and chimera is presented in Table S1.

[a]Abbreviations used: Bja, *Bradyrhizobium japonicum;* Eco, *Escherichia coli;* Cro, *Citrobacter rodentium;* Plu, *Photorhabdus luminescens;* Pma, *Pseudomonas syringae* pv. *maculicola;* Pto, *Pseudomonas syringae* pv. *tomato;* Psy, *Pseudomonas syringae* pv. *syringae;* Pph, *Pseudomonas syringae* pv. *phaseolicola;* Pgy, *Pseudomonas syringae* pv. *glycinea;* Per, *Pseudomonas syringae* pv. *eriobotryae;* Set, *Salmonella enterica;* Sfl, *Shigella flexneri;* Xcc, *Xanthomonas campestris* pv. *campestris;* Xac, *Xanthomonas axonopodis* pv. *citri;* Xcv, *Xanthomonas campestris* pv. *vesicatoria.*

[b]O, orphet; C, chimera.

[c]Linkage to mobile elements. tn, transposable element or fragment; pl, plasmid; ph, prophage; N/A, not available; –, none.

[d]N or C terminus.

[e]Conserved 5′ (upstream) or 3′ (downstream).

[f]Region 30 bp immediately upstream of *hopW1–1* is homologous to related effectors. Region further upstream encompassing promoter is homologous to that of *hopD1–1*.

[g]Conserved region is upstream of *avrA* coding region, but within the *hopQ1* coding region.

[h]Conserved upstream region is within a coding region (effector is in an operon with a chaperone or another effector).

DOI: 10.1371/journal.ppat.0020104.t001

[45] from *Shigella flexneri*. As with these contrived screens, the natural evolutionary process of terminal reassortment likely involves mobile genetic elements such as insertion sequences, integrative conjugative elements, phage, and plasmids. Indeed, over 50% of ORPHETs or chimeric T3SEs are associated with mobile elements (Table 1). Many effectors are plasmid-borne, placing them on a genome that is typically in genetic flux and susceptible to frequent rearrangements. Prophage are also implicated in terminal reassortment, illustrated by C-terminal homology of effectors EspFu and HopW1–1_{ES4326} to prophage-related genes. Phage may serve as mixing vessels for T3SEs, or provide the novel genetic variation and homologous sequence for recombination. Both phage and plasmids, as well as transposable elements, ultimately provide multiple copies of ORPHETs or T3SE precursors. Consequently, terminal reassortment can occur without loss of the original effector or ORPHET.

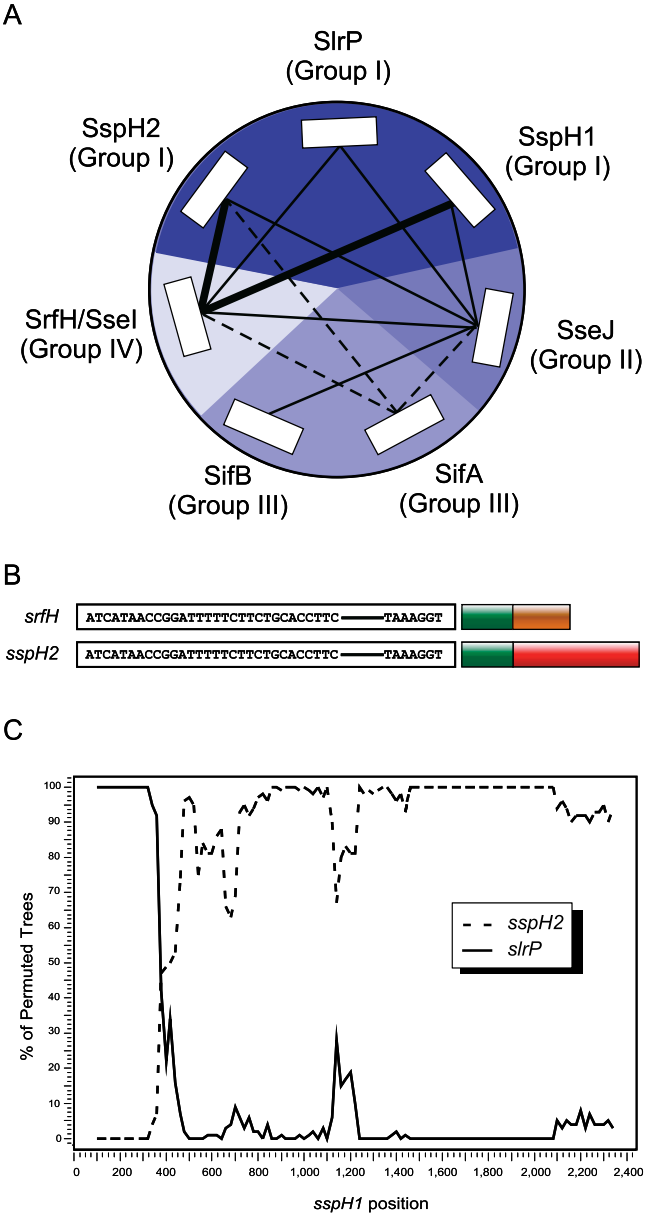

Figure 3. Terminal Reassortment among Effectors in *Salmonella enterica*

(A) N-terminal homology among seven effectors of *S. enterica* representing four distinct homology groups. Each homology group has been encapsulated in a shade of blue and contains proteins that share full-length homology. The interconnecting lines indicate the degree of N-terminal relatedness, with dashed lines indicating <40% similarity, thin, solid lines 40%–60% similarity, and thick, solid lines >60% similarity.

(B) *S. enterica* effectors *sspH2* and *srfH* share a highly conserved upstream region, ssrAB promoter box, and N-terminal region, while the C terminus of these proteins is non-homologous.

(C) Bootscan of the homologous *S. enterica* effectors *sspH1, sspH2, slrP,* and N-terminal homolog *ssel/srfH* to detect recombination, with the nucleotide sequence of *sspH1* as the query sequence. *ssel/srfH* had <35% permuted trees and was omitted from the figure. Bootscanning was performed with Simplot v 3.5.1 using neighbor joining with the F84 model, gapstrip off, and window and step sizes of 200 and 20, respectively. Other members of this effector family were excluded to facilitate optimal nucleotide alignment.

DOI: 10.1371/journal.ppat.0020104.g003

An analysis of the relationship between the number of ORPHETs and chimeras, and the number of T3SE homology families identified for each species (Figure 4A) reveals a very strong linear relationship ($R^2 = 0.91$, $p < 0.0001$, Figure 4B), reinforcing the significance of terminal reassortment on the evolution of T3SEs. Approximately one out of every 4.2 T3SE families (weighted mean = 24% of T3SE families) carries an ORPHET or chimeric T3SE. This is significant because it indicates that terminal reassortment is strongly influencing the evolution of T3SEs. Additional support for the importance of terminal reassortment in the evolution of T3SE is apparent when the frequency of ORPHETs and chimeras in T3SE homology families is compared to the frequency of truncated and chimeric loci among families of non-T3SE proteins. To reduce bias introduced by incomplete T3SE inventories, we conducted this analysis on *P. syringae, X. campestris,* and *S. enterica*, which have the largest, best-characterized effector complements. While, on average, 24% of T3SE families have at least one ORPHET or chimera, only 7% of non-T3SE protein families were found to carry a truncated or chimeric locus among the 2,943 protein families of *P. syringae* (204 gene families, $p = 3 \times 10^{-5}$, two-tailed chi-squared test), the 2,760 non-T3SE protein families of *S. enterica* (198 gene families, $p = 0.03$), and the 2,688 non-T3SE protein families of *X. campestris* (195 genes, $p = 1 \times 10^{-5}$). Transporters (18%) and regulators (10%) are among the most highly represented functional groups among the 204 non-T3SE truncated or chimeric loci in *P. syringae*. Interestingly, proteins that are expected to be under strong selective pressures (e.g., outer membrane proteins, alginate biosynthetic genes, type IV pilus subunits) only represent 4% of these chimeras.

T3SEs are assumed to be predominantly associated with the flexible genome—the part of the genetic compliment that varies among strains within a species. Given this, it is possible that terminal reassortment simply occurs more frequently in the flexible genome, and therefore, our observations are a consequence of the genomic context of T3SEs, rather than something inherent to T3SEs themselves. To address this, we determined that 28.3% of all *P. syringae* coding sequences are absent from one or two of the three sequenced *P. syringae* genomes (Nahal and Guttman, unpublished data) and are therefore putatively components of the flexible genome. In comparison, only 9.3% of the 204 non-T3SE truncated and chimeric loci were absent from one or two of the three sequenced *P. syringae* strains. One would expect a larger fraction of the non-T3SE truncated and chimeric loci to be in the flexible genome if terminal reassortment were simply a byproduct of those processes acting on the flexible genome.

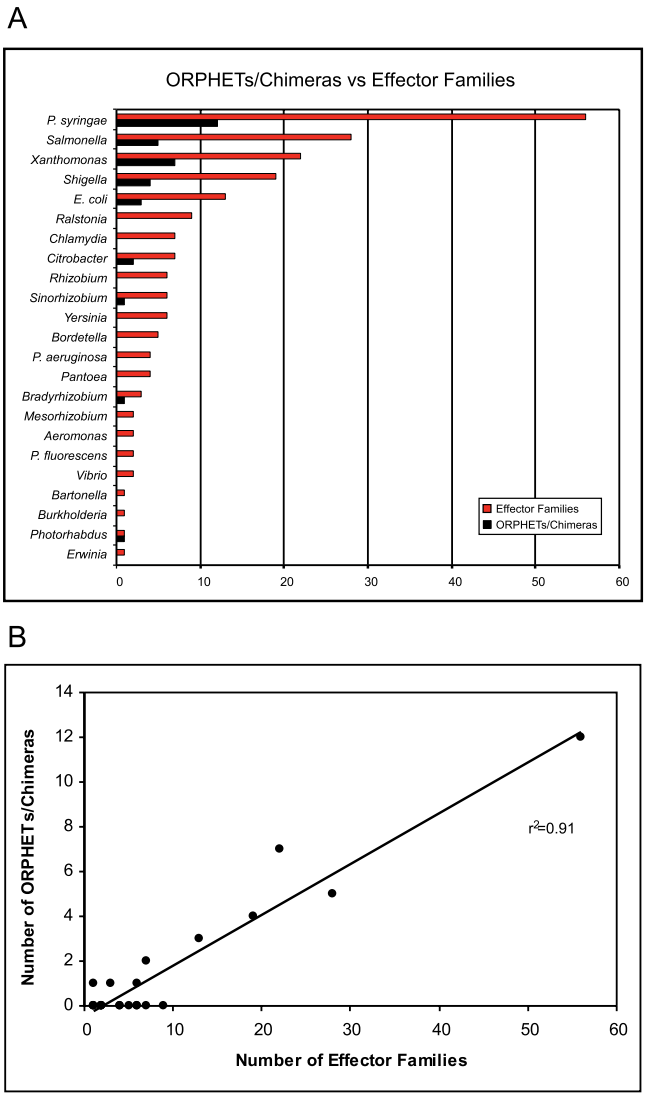

Figure 4. Frequency of ORPHETs and Chimeras among Bacterial Pathogens

(A) Histogram illustrating the number of effector families analyzed for each species and the number of ORPHETs/chimeras identified for each. (B) Regression analysis of the number of effector families versus the number of ORPHETs/chimeras identified for each given species, $r^2 = 0.91$.

DOI: 10.1371/journal.ppat.0020104.g004

Furthermore, 69.7% of the 204 non-T3SE truncated or chimeric loci have homologs in *Pseudomonas putida, Pseudomonas aeruginosa,* and *Pseudomonas fluorescens,* further supporting the contention that these loci are not components of the flexible genome. Clearly, the processes that are responsible for the evolutionary dynamics of the flexible genome are not strictly responsible for the formation of chimeras and ORPHETs.

We believe that the high frequency of ORPHETs and chimeras among T3SE is due to three key features that are collectively unique to T3SEs. First and foremost, T3SEs are modular [5,6], with all of the signals required for expression, secretion, and translocation located in close proximity. Second, T3SEs are under intense selective pressures due to their central roles in pathogenesis and host adaptation [16]. Third, T3SEs are commonly associated with mobile elements [40,41], which can act as catalysts for the reassortment of linked loci.

Horizontal gene transfer also likely plays an important role at all levels of the terminal reassortment process. Horizontal gene transfer may be responsible for the supply of precursor ORPHETs and T3SEs and the introduction of novel genetic variation that will eventually serve as the target. Horizontal gene transfer may even lead to the intra- or inter-specific dissemination of the new chimera following terminal reassortment. The association of ORPHETs and T3SEs with multi-copy mobile elements provides a substantial pool of these elements, and as a result, terminal reassortment does not necessarily lead to loss of the original T3SE or ORPHET precursor.

Terminal reassortment may fundamentally affect bacterial virulence and host specificity. For example, the *P. syringae* T3SEs HopD1 and HopAO1 are N-terminal chimeras that have different biological functions due to their unique C termini. Unlike HopD1, the C terminus of HopAO1 has tyrosine phosphatase activity, which is responsible for the suppression of the innate immune response of *Nicotiana benthamiana* [22,37]. Since only one of these two related T3SEs has tyrosine phosphatase activity, terminal reassortment has in this case given rise to T3SEs that may act in a host-specific manner. A similar situation is found with the *Xanthomonas* chimeras XopJ and AvrXccE1. XopJ and other members of the YopJ family have SUMO isopeptidase activity governed by a catalytic triad of histidine, glutamic acid, and cysteine residues [3,38]. This catalytic triad is present in XopJ, but absent from its N-terminal homolog AvrXccE1, making it very likely that these two effectors have different biological functions. Another example comes from the chimeric T3SEs SspH2, SspH1, and SlrP, which have been shown to differentially modulate *Salmonella* host specificity. SspH1 and SspH2 contribute to *Salmonella* virulence in calves, while SlrP plays a similar role during infection of mice [35,39]. Since these chimeric T3SEs have differential effects in different hosts, they can be considered host-specificity factors.

Because ORPHETs are often mobilized with their upstream regulatory and binding motifs, virtually any coding or non-coding gene segment can be instantly recruited by an N-terminal secretion domain to serve as part of a new T3SE with potentially new virulence function, effectively making a novel virulence protein in a single quantum step. The random shuffling of any one ORPHET or T3SE has only a small probability of giving rise to a functional protein that is of use to the bacterium. Nevertheless, over evolutionary time, terminal reassortment will produce a wide spectrum of novel proteins, some of which will confer a selective advantage to their bacterial host. These proteins may persist and even become fixed in the population. This evolutionary process, a bacterial version of Goldschmidt's hopeful monster, has significant implications for pathogenicity and virulence as it enables the near instantaneous evolution of novel genes or altered expression patterns that can permit the bacterium to evade host defenses or exploit new eukaryotic targets. It will be interesting to determine if this remarkably simple and efficient evolutionary process is responsible for some of the dramatic host-specificity shifts that underlie the emergence of new infectious agents.

Materials and Methods

Bioinformatic identification of chimeric effectors. Nucleotide and protein sequences of all known and experimentally confirmed T3SEs were collected from NCBI following extensive literature and database searches. Genera analyzed include *Aeromonas, Bartonella, Bordetella, Bradyrhizobium, Brucella, Burkholderia, Citrobacter, Chromobacterium, Chlamydia, Escherichia, Edwardsiella, Haemophilus, Mesorhizobium, Pantoea, Photorhabdus, Pseudomonas, Ralstonia, Rhizobium, Salmonella, Shewanella, Shigella, Sinorhizobium, Sodalis, Vibrio, Xanthomonas,* and *Yersinia.* The NCBI nr database was searched with BLASTN, BLASTP, and TBLASTX to identify candidate chimeras and ORPHETs, with subsequent verification achieved with BL2SEQ and DIALIGN2. A protein was considered an ORPHET if its full-length sequence had >45% similarity to the N or C terminus of a different protein. Similarly, two proteins were considered chimeras if they shared >45% similarity in either terminus. In both cases, similarity was required to begin within 30 amino acids of their start position for N-terminal homologs and 30 amino acids of their stop position for C-terminal homologs. Sequences flanking effectors were extracted from NCBI manually and compared using BL2SEQ and ClustalX1.83 [46]. Upstream promoter sequences and ribosome-binding sites were identified and extracted manually from available sequence data, except for *hopAB3–2*$_{PmaES4326}$, whose upstream sequence was obtained by inverse PCR (see below). Hrp box promoter sequences for *Pseudomonas syringae* will be deposited at the PPI site: http://www.pseudomonas-syringae.org.

Phylogenetic analyses and bootscanning. Phylogenetic trees were constructed using alignments generated with ClustalX1.83 using MEGA3 [47]. The JTT amino acid substitution model was used with pair-wise deletion and 1,000 bootstrap replicates. Bootscanning was performed with Simplot v 3.5.1 [48] using neighbor joining with the F84 (maximum likelihood) model, gapstrip off, and window and step sizes of 200 and 20, respectively. The inclusion of the group IV *Salmonella* effectors was required for this analysis as bootscanning requires a minimum of four sequences. Because the other effector groups were highly divergent, they were excluded from the analysis to facilitate optimal nucleotide alignment. A significant recombination event is usually indicated by a percentage of permuted trees over 70%.

Proteomic comparison and analyses. The genome-wide identification of non-T3SE truncated and chimeric loci was performed with PERL scripts using the complete proteomes of *P. syringae* pv. tomato DC3000 (5,470 proteins) [49], *Xanthomonas campestris* pv. vesicatoria 85–10 (4,487 proteins) [50], and *Salmonella enterica* subsp. enterica serovar Typhi str. CT18 (4,395 proteins) [51]. Following the exclusion of transposable elements, type III effectors, pseudogenes, and leader peptides, each proteome was categorized into homology groups (2,943 for *P. syringae*, 2,688 for *X. campestris,* and 2,760 for *S. enterica*) based on COG designation. Proteins not belonging to an established COG were BLASTed against the entire proteome and grouped together if they shared 70% similarity over 70% of their length. All homology groups were examined for evidence of chimeras and ORPHET-like derivation by BLASTP against a custom database, yielding 204 chimeras/ORPHETs for *P. syringae,* 195 for *X. campestris,* and 198 for *S. enterica.* The custom *P. syringae* database included *P. syringae* pv. tomato DC3000, *P. syringae* pv. phaseolicola 1448A, and *P. syringae* pv syringae B728a, as well as proteins from fully sequenced plasmid genomes (pDC3000A, pDC3000B, p1448A-A, p1448A-B, pFKN, pPSR1, pPMA4326A, pPMA4326B). For *S. enterica,* the custom database included *S. enterica* ATCC9150, *S. enterica* SC-B67, *S. enterica* CT18, and *S. enterica* TY2, as well as fully sequenced plasmid genomes (pBERT, pCryptic, pHCM1, pHCM2, pKDSC50, pOU1113, pSC138, pSCV50). The custom *Xanthomonas* database included the proteomes of *X. campestris* pv. vesicatoria 85–10, *X. campestris* pv. campestris 8004, *X. campestris* pv. campestris ATCC33913, and the four plasmids, pXCV2, pXCV19, pXCV38, and pXCV183. A two-tailed chi-squared test was used to compare the relative impact of terminal reassortment on *P. syringae* effector families ($n = 56$), *S. enterica* effector families ($n = 28$), and *X. campestris* effector families ($n = 28$) versus the *P. syringae* ($n = 2{,}943$), the *S. enterica* ($n = 2{,}760$), or the *X. campestris* ($n = 2{,}688$) non-effector datasets, respectively. The classification of the 204 *P. syringae* non-T3SE truncated and chimeric loci into "core" and "flexible" was achieved by BLAST analysis of a single chimeric representative from each family against a custom database comprising the proteomes of the three sequenced *P. syringae* strains (*P. syringae* pv. tomato DC3000, *P. syringae* pv. syringae B728A, and *P. syringae* pv. phaseolicola 1448A), *Pseudomonas putida* KT2440, *Pseudomonas fluorescens* Pf-5, and *Pseudomonas aeruginosa* PAO1. Proteins were considered orthologous if they shared >45% amino acid similarity over their entire lengths and had a BLAST expect value less than 10^{-20}. Regression analyses were performed with StatView version 5.0.1 (SAS Institute) using data from all bacteria having at least one confirmed T3SE.

Inverse PCR and DNA sequencing. Genomic DNA of *P. syringae* pv. maculicola ES4326 was extracted using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, Minnesota, United States). Approximately 8 μg of genomic DNA was completely digested with BsaAI (New England Biolabs Incorporated, Beverly, Massachusetts, United States) and self-ligated at 16 °C overnight using T4 DNA ligase. The upstream region of *hopAB3–2*$_{PmaES4326}$ was obtained by inverse PCR with 20 ng of self-ligated DNA as template, using the primers 5′GGCTCTGTTGATACTACCCACCATG 3′ and 5′ GTGCCGCTACCGCCGTGCC 3′ at an annealing temperature of 57 °C. The DNA sequence of the single 420-bp PCR product was obtained with the same primers using a CEQ8000 sequencer (Beckman Coulter Incorporated, Fullerton, California, United States).

Supporting Information

Table S1. T3SE Proteins from 21 Different Genera

Found at DOI: 10.1371/journal.ppat.0020104.st001 (199 KB DOC).

Table S2. Examples of Terminal Reassortment Involving T3SEs from Eight Genera

Found at DOI: 10.1371/journal.ppat.0020104.st002 (209 KB DOC).

Acknowledgments

We would like to acknowledge S. Sarkar and P. Wang for discussions on effector evolution and input on the content of this manuscript.

Author contributions. JS conceived and designed the experiments. JS and WM performed the experiments. JS analyzed the data. JS, WM, and DSG wrote the paper.

Funding. DSG is supported by grants from the Canada Research Chairs Program, the National Science and Engineering Council of Canada (NSERC), and Performance Plants Incorporated of Kingston, Ontario, Canada. JS is supported by an NSERC Canada Graduate Scholarship and WM is supported by an NSERC Postdoctoral Fellowship.

Competing interests. The authors have declared that no competing interests exist.

References

1. Abe A, Matsuzawa T, Kuwae A (2005) Type-III effectors: Sophisticated bacterial virulence factors. C R Biol 328: 413–428.
2. Mota LJ, Cornelis GR (2005) The bacterial injection kit: Type III secretion systems. Ann Med 37: 234–249.
3. Mudgett MB (2005) New insights into the function of phytopathogenic bacterial type III effectors in plants. Annu Rev Plant Biol 56: 509–531.
4. Cornelis GR, Van Gijsegem F (2000) Assembly and function of type III secretory systems. Annu Rev Microbiol 54: 735–774.
5. Sory M, Boland A, Lambermont I, Cornelis GR (1995) Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the *cyaA* gene fusion approach. Proc Natl Acad Sci U S A 92: 11998–12002.
6. Guttman DS, Greenberg JT (2001) Functional analysis of the type III effectors AvrRpt2 and AvrRpm1 of *Pseudomonas syringae* with the use of a single-copy genomic integration system. Mol Plant Microbe Interact 14: 145–155.
7. Roden JA, Belt B, Ross JB, Tachibana T, Vargas J, et al. (2004) A genetic screen to isolate type III effectors translocated into pepper cells during *Xanthomonas* infection. Proc Natl Acad Sci U S A 101: 16624–16629.
8. Schechter LM, Roberts KA, Jamir Y, Alfano JR, Collmer A (2004) *Pseudomonas syringae* type III secretion system targeting signals and novel effectors studied with a Cya translocation reporter. J Bacteriol 186: 543–555.
9. Guttman DS, Vinatzer BA, Sarkar SF, Ranall MV, Kettler G, et al. (2002) A functional screen for the type III (Hrp) secretome of the plant pathogen *Pseudomonas syringae*. Science 295: 1722–1726.
10. Lindeberg M, Stavrinides J, Chang JH, Alfano JR, Collmer A, et al. (2005)

Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen *Pseudomonas syringae*. Mol Plant Microbe Interact 18: 275–282.
11. Hueck CJ (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 62: 379–423.
12. Innes RW, Bent AF, Kunkel BN, Bisgrove SR, Staskawicz BJ (1993) Molecular analysis of avirulence gene *avrRpt2* and identification of a putative regulatory sequence common to all known *Pseudomonas syringae* avirulence genes. J Bacteriol 175: 4859–4869.
13. Feng XH, Walthers D, Oropeza R, Kenney LJ (2004) The response regulator SsrB activates transcription and binds to a region overlapping OmpR binding sites at *Salmonella* pathogenicity island 2. Mol Microbiol 54: 823–835.
14. Fenselau S, Bonas U (1995) Sequence and expression analysis of the hrpB pathogenicity operon of *Xanthomonas campestris* pv vesicatoria which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems. Mol Plant Microbe Interact 8: 845–854.
15. Guttman D, Gropp S, Morgan R, Wang P (2006) Diversifying selection drives the evolution of the type III secretion system pilus of *Pseudomonas syringae*. Mol Biol Evol. E-pub 1 September 2006.
16. Rohmer L, Guttman DS, Dangl JL (2004) Diverse evolutionary mechanisms shape the type III effector virulence factor repertoire in the plant pathogen *Pseudomonas syringae*. Genetics 167: 1341–1360.
17. Pitman AR, Jackson RW, Mansfield JW, Kaitell V, Thwaites R, et al. (2005) Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Curr Biol 15: 2230–2235.
18. Wichmann G, Ritchie D, Kousik CS, Bergelson J (2005) Reduced genetic variation occurs among genes of the highly clonal plant pathogen *Xanthomonas axonopodis* pv. vesicatoria, including the effector gene *avrBs2*. Appl Environ Microbiol 71: 2418–2432.
19. Campellone KG, Robbins D, Leong JM (2004) EspFu is a translocated EHEC effector that interacts with Tir and N-WASP and promotes nck-independent actin assembly. Dev Cell 7: 217.
20. Noel L, Thieme F, Nennstiel D, Bonas U (2001) cDNA-AFLP analysis unravels a genome-wide hrpG-regulon in the plant pathogen *Xanthomonas campestris* pv. vesicatoria. Mol Microbiol 41: 1271–1281.
21. Brumell JH, Marcus SL, Finlay BB (2000) N-terminal conservation of putative type III secreted effectors of *Salmonella typhimurium*. Mol Microbiol 36: 773–774.
22. Espinosa A, Guo M, Tam VC, Fu ZQ, Alfano JR (2003) The *Pseudomonas syringae* type III-secreted protein HopPtoD2 possesses protein tyrosine phosphatase activity and suppresses programmed cell death in plants. Mol Microbiol 49: 377–387.
23. Davies RL, Campbell S, Whittam TS (2002) Mosaic structure and molecular evolution of the leukotoxin operon (lktCABD) in *Mannheimia (Pasteurella) haemolytica*, *Mannheimia glucosida*, and *Pasteurella trehalosi*. J Bacteriol 184: 266–277.
24. Ko KS, Hong SK, Lee HK, Park MY, Kook YH (2003) Molecular evolution of the *dotA* gene in *Legionella pneumophila*. J Bacteriol 185: 6269–6277.
25. Davies RL, Lee I (2004) Sequence diversity and molecular evolution of the heat-modifiable outer membrane protein gene *(ompA)* of *Mannheimia (Pasteurella) haemolytica*, *Mannheimia glucosida*, and *Pasteurella trehalosi*. J Bacteriol 186: 5741–5752.
26. Baldo L, Lo N, Werren JH (2005) Mosaic nature of the *Wolbachia* surface protein. J Bacteriol 187: 5406–5418.
27. Krzywinska E, Krzywinski J, Schorey JS (2004) Naturally occurring horizontal gene transfer and homologous recombination in *Mycobacterium*. Microbiology 150: 1707–1712.
28. Haake DA, Suchard MA, Kelley MM, Dundoo M, Alt DP, et al. (2004) Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer. J Bacteriol 186: 2818–2828.
29. Greenberg JT, Vinatzer BA (2003) Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. Curr Opin Microbiol 6: 20–28.
30. Stavrinides J, Guttman DS (2004) Nucleotide sequence and evolution of the five-plasmid complement of the phytopathogen *Pseudomonas syringae* pv. maculicola ES4326. J Bacteriol 186: 5101–5115.
31. Nissan G, Manulis S, Weinthal DM, Sessa G, Barash I (2005) Analysis of promoters recognized by HrpL, an alternative sigma-factor protein from *Pantoea agglomerans* pv. gypsophilae. Mol Plant Microbe Interact 18: 634–643.
32. Noel L, Thieme F, Gabler J, Buttner D, Bonas U (2003) XopC and XopJ, two novel type III effector proteins from *Xanthomonas campestris* pv. vesicatoria. J Bacteriol 185: 7092–7102.
33. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, et al. (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature 417: 459–463.
34. Miao EA, Miller SI (2000) A conserved amino acid sequence directing intracellular type III secretion by *Salmonella typhimurium*. Proc Natl Acad Sci U S A 97: 7539–7544.
35. Miao EA, Scherer CA, Tsolis RM, Kingsley RA, Adams LG, et al. (1999) *Salmonella typhimurium* leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. Mol Microbiol 34: 850–864.
36. Waterman SR, Holden DW (2003) Functions and effectors of the *Salmonella* pathogenicity island 2 type III secretion system. Cell Microbiol 5: 501–511.
37. Bretz JR, Mock NM, Charity JC, Zeyad S, Baker CJ, et al. (2003) A translocated protein tyrosine phosphatase of *Pseudomonas syringae* pv. tomato DC3000 modulates plant defense response to infection. Mol Microbiol 49: 389–400.
38. Hotson A, Mudgett MB (2004) Cysteine proteases in phytopathogenic bacteria: Identification of plant targets and activation of innate immunity. Curr Opin Plant Biol 7: 384–390.
39. Tsolis RM, Townsend SM, Miao EA, Miller SI, Ficht TA, et al. (1999) Identification of a putative *Salmonella enterica* serotype typhimurium host range factor with homology to IpaH and YopM by signature-tagged mutagenesis. Infect Immun 67: 6385–6393.
40. Arnold DL, Jackson RW, Fillingham AJ, Goss SC, Taylor JD, et al. (2001) Highly conserved sequences flank avirulence genes: Isolation of novel avirulence genes from *Pseudomonas syringae* pv. pisi. Microbiology 147: 1171–1182.
41. Kim JF, Charkowski AO, Alfano JR, Collmer A, Beer SV (1998) Sequences related to transposable elements and bacteriophages flank avirulence genes of *Pseudomonas syringae*. Mol Plant Microbe Interact 11: 1247–1252.
42. Chang JH, Urbach JM, Law TF, Arnold LW, Hu A, et al. (2005) A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae*. Proc Natl Acad Sci U S A 102: 2549–2554.
43. Guo M, Chancey ST, Tian F, Ge ZX, Jamir Y, et al. (2005) *Pseudomonas syringae* type III chaperones ShcO1, ShcS1, and ShcS2 facilitate translocation of their cognate effectors and can substitute for each other in the secretion of HopO1–1. J Bacteriol 187: 4257–4269.
44. Parsot C, Ageron E, Penno C, Mavris M, Jamoussi K, et al. (2005) A secreted anti-activator, OspD1, and its chaperone, Spa15, are involved in the control of transcription by the type III secretion apparatus activity in *Shigella flexneri*. Mol Microbiol 56: 1627–1635.
45. Buchrieser C, Glaser P, Rusniok C, Nedjari H, d'Hauteville H, et al. (2000) The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. Mol Microbiol 38: 760–771.
46. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882.
47. Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings In Bioinformatics 5: 150–163.
48. Salminen MO, Carr JK, Burke DS, McCutchan FE (1995) Identification of breakpoints in intergenotypic recombinants of HIV Type-1 by bootscanning. AIDS Res Hum Retrovir 11: 1423–1425.
49. Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, et al. (2003) The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. Proc Natl Acad Sci U S A 100: 10181–10186.
50. Thieme F, Koebnik R, Bekel T, Berger C, Boch J, et al. (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. J Bacteriol 187: 7254–7266.
51. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, et al. (2001) Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature 413: 848–852.