

## A Tale of Effectors; Their Secretory Mechanisms and Computational Discovery in Pathogenic, Non-Pathogenic and Commensal Microbes

Swati Bhowmick<sup>1</sup>, Sucheta Tripathy<sup>2\*</sup>

<sup>1</sup>National Institute of Cholera and Enteric Diseases; P-33, C.I.T. Road, Scheme XM, Beliaghata, Kolkata 700 010, India

<sup>2</sup>Indian Institute of Chemical Biology (Council of Scientific & Industrial Research); 4, Raja S. C. Mullick Road, Kolkata-32, India

\*Corresponding author: Sucheta Tripathy, Indian Institute of Chemical Biology (Council of Scientific & Industrial Research); 4, Raja S. C. Mullick Road, Kolkata-32, India, Tel: +91 33 2473 0492; Fax: +91 33 2473 5197, +91 33 2472 3967, E-mail: [tsucheta@gmail.com](mailto:tsucheta@gmail.com)

Rec date: May 09, 2014; Acc date: Jul 24, 2014; Pub date: Jul 27, 2014

Copyright: © 2014 Swati B et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

Secretory proteins that are involved in modulating hosts are called as effectors. Lately, finding these important proteins from a large array of other gene products have been a focus area in many high funding research programs. However, since the biological data is accumulating at a much faster rate now than ever before, this search process can be compared to finding a needle in the haystack. Conventional laboratory-based methods require critical experiments, extended time and high cost which in many cases result in failure of testing hypothesis. Using high throughput sequencing technologies, whole genome sequences are generated much more quickly and efficiently. The avalanche of genomics data has ushered new opportunities into discovery of large number of novel extra cellular secretory proteins that usually lie undetected with conventional methods. Recently powerful bioinformatics methods have emerged that can predict effectors from whole genome data of pathogens, commensal, symbiotic and environmental microorganisms much easily. In this review, we present a broad overview of these biological molecules that modulates host response in different ways in many organisms – pathogenic, non-pathogenic and commensal. We also catalogue the motifs associated with many secretory mechanisms and their prediction algorithms.

**Keywords:** Effectors; Secretory proteins; Algorithm; Motifs

### Abbreviations

Sec: Secretion, Tat: Twin arginine translocation, T1SS: Type I secretory system, T2SS: Type II secretory system, T3SS: Type III secretory system, T4SS: Type IV secretory system, T5SS: Type V secretory system, T6SS: Type VI secretory system, T7SS: Type VII secretory system, PV: Parasitophorous vacuole, NLS: Nuclear localization signal

### Introduction

Secretory mechanisms are highly specialized tools for microbes to interact with its host system. Certain specialized secretory proteins called as effectors play a very crucial role in host-microbe crosstalk. Effectors could be delivered into the host cells via varied mechanisms, including specialized secretion systems, physical injection and protein translocation via signal sequences. The common attribute each effector possesses is the ability to trespass the cell membrane of microbial cell and move into the extra cellular space. Once inside the external medium, be it in the host cell or outside of the host cell, the job of the effector is to modulate the host system.

Latest developments in genome sequencing technologies have advanced our knowledge about genomes of a broad diversity of microbes, including bacteria, apicomplexa, fungi, and oomycetes, as well as nematodes. Sequence analysis studies have revealed that effectors show very poor sequence similarity among themselves. Thus similarity based methods that are often used for finding such molecules have failed completely. In order to circumvent this issue,

many novel bioinformatics strategies have been devised from time to time for predicting them from the whole genome. A large number of effectors have been predicted using various algorithms that were earlier not known. Many of these effectors have now been characterized in the laboratory to prove the efficacy of these methods [1]. While there are major advances made in this area to generalize the algorithms for effector discovery in silico, there are no comprehensive review on classification and prediction of effectors in different microbial systems.

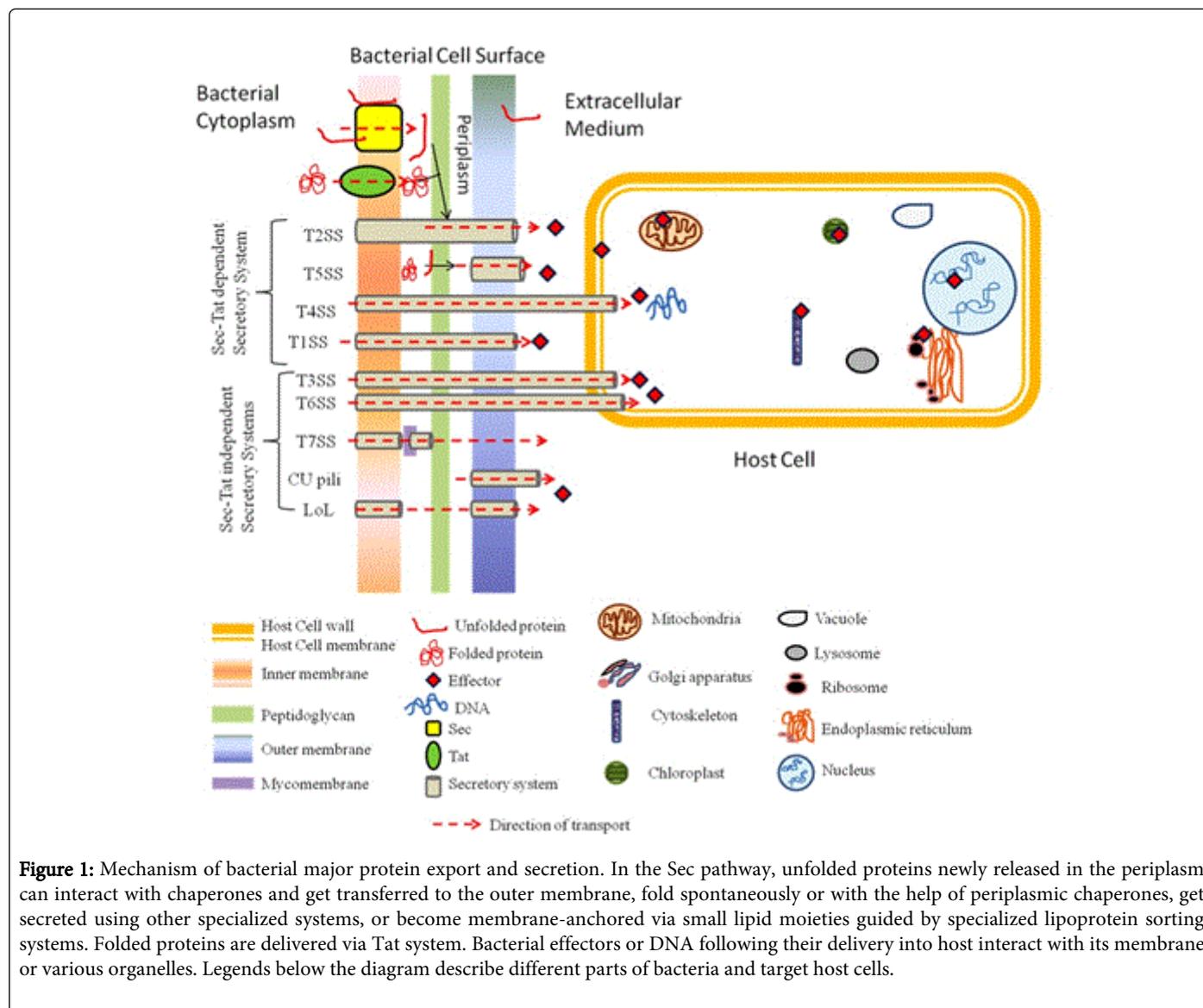
In this review article, we attempt to give an over view of secretory mechanisms in commensal, parasitic, saprophytic microbes. We discuss the Sec and Tat dependent and independent secretion systems in bacterial and other microbes. The different types of secretion systems such as Type I through VII is discussed with examples. Other not-so-common secretion mechanisms such as Chaperone Usher pathway and the LOL systems are also discussed. We have elaborated motifs involved in secretion mechanisms and the available computational methods to predict effectors from the genome sequences. This article summarizes many essential aspects of effector and secretion systems in microbes that may be of great help for researchers associated in this area.

### An Overview of Microbial Secretion mechanisms

The secretion process in bacteria involves two important pathways e.g; the secretion (Sec) pathway and the two-arginine translocation (Tat) pathways that are universal across the tree of life. The Sec and Tat pathways are responsible for secreting proteins into the periplasmic regions of Gram negative bacteria that has a bilayer cell membrane, and to the exterior in Gram positive bacteria having a

mono layer cell membrane. Secretion in Gram negative bacteria generally involves transfer of proteins to the periplasmic space via Sec or Tat pathway and then across the outer membrane through Type II, Type V or less commonly via Type I or Type IV pathways. However, some secreted proteins in Gram negative bacteria are directly

transported into target cell in a single step across the bilayer membranes using Type I, III, IV and VI pathways [2,3]. Gram positive bacteria *Mycobacterium* spp utilize special type VII secretion machinery to transport proteins across their hydrophobic cell membrane and impermeable cell wall [Figure 1].



### The Universal Sec and Tat Pathways

Sec pathway carries out most of the protein export from bacterial cytoplasm to cell membrane and extracellular environment. These proteins have roles in nutrient acquisition, cell wall synthesis and virulence and viability of bacteria [4,5]. The key features of the proteins using Sec pathway are almost always synthesized as unfolded preproteins carrying an N-terminal signal sequence that often acts as an address tag [6]. Sec system relies on proton gradient and ATP hydrolysis powered by SecA motor protein as energy source for efficient preprotein translocation. SecA, in some cases are implicated in virulence of bacteria [5]. In contrast to Sec pathway, Tat secretion pathway is responsible for translocation of folded proteins in cytoplasmic membrane of bacteria [7]. This secretion machinery recognizes a distinct motif rich in basic amino acids (Ser/Thr-Arg-

Arg-X-Phe-Leu-Lys) in the N-terminal region of signal peptide. The proteins translocated to Tat translocon consisting of three membrane proteins TatA, TatB and TatC using proton gradient only as energy source [7,8]. In order to prevent mis-firing of tat peptides into sec pathways, there are subtle differences between the protein properties. The hydrophobicity in case of Tat proteins are slightly less than that of sec proteins [9]. Tat proteins carry out multiple cellular activities including anaerobic metabolism, cell envelope biogenesis, metal acquisition and detoxification, and virulence [7,9,10].

### The Sec and Tat dependent secretory systems

The type I, II, IV and V secretory systems require universal secretory mechanisms such as Sec and Tat pathways for secretion into the exterior. However, Type 1 secretory system (T1SS) is also capable

of exporting secretory molecules outside, independent of the Sec and Tat system [1]. For T1SS, ATP binding cassettes (ABC Transporters), Outer Membrane Factors (OMFs) and membrane fusion proteins (MFP) work in tandem for secretion. Several virulence factors such as metalloproteases, adhesins and glycanases of plant pathogenic bacteria and rhizobial proteins engaged in legume symbiosis are secreted through T1SS [2].

In case of type II secretory system (T2SS), the proteins that are already secreted into periplasmic space (via sec or tat pathway) are excreted out [11,12]. T2SS is conserved in Gram negative bacteria and consists of a set of 12-16 conserved proteins that are assembled into a supramolecular complex spanning the bacterial envelope, called the secreton [13]. T2SS is required for several plant and animal pathogens including *Vibrio cholerae*, *enterotoxigenic* and *enterohaemorrhagic Escherichia coli* (ETEC and EHEC, respectively), *Pseudomonas aeruginosa*, *Klebsiella spp.*, *Legionella pneumophila* and *Yersinia enterocolitica*. Virulence factors secreted via T2SS include exotoxin A, cholera toxin, pectinase and pectate lyase of plant pathogens *Erwinia carotovora* and *Xanthomonas campestris* [12].

Type IV secretion system (T4SS) is very unique in the sense that it can transport nucleic acids into several hosts including plants, animals, yeast or other bacteria when in direct contact. Three types of T4SS are known, (i) conjugation system that translocates DNA intercellularly by a contact dependent process, (ii) effector translocator system that is set to transfer proteins or other macromolecules to recipient cell, (iii) DNA release/uptake system that translocates DNA to or from the extracellular milieu [14] as in *Helicobacter pylori* ComB system. In case of *Neisseria gonorrhoeae*, DNA is secreted into the extracellular milieu. *Brucella*, *Bartonella*, *Rickettsia*, and *Anaplasma* release virulence factors through T4SS that consist of 12 proteins collectively termed as VirB/D4 family, similar in structure to *Agrobacterium tumefaciens* Vir B/D4 system [15]. *Legionella* and *Coxiella sp.* translocate their virulence factors through T4SS that consists of two groups of proteins, Dot (Defective in organelle trafficking) or Icm (Intracellular multiplication) proteins [15]. Recently many new type IV secretory systems have been discussed. The novel GI type IV secretory system is responsible for formation of wide variety of genome islands as in gonococcal genetic island (GGI) [16].

Interestingly, it has become apparent that T2SS of Gram-negative bacteria, the type IV pilus system (T4PS) of Gram-positive and Gram-negative bacteria, the archaeal flagellum synthesis system and the transformation system of Gram-positive bacteria are evolutionarily related and share several structural and functional features [12].

Type V secretory systems (T5SS) are autotransporters [17]. Paradoxically this is the simplest, largest and most recently discovered secretory system [11]. This system is also dependent on sec machinery for secretion of proteins into the periplasmic region via a signal peptide. After the signal sequence is cleaved, the passenger domain is ported out using the C terminal beta barrel that forms a pore like structure on the outer membrane. In case of toxins, the passenger domain is cleaved from the beta barrel forming a soluble component. These proteins are called as auto-transporters. Another subclass of Type V SS is T5cSS, where it is composed of a trimeric component, where each of the trimer contributes towards the formation of a beta barrel. There is another subclass called as two partner secretion (TPS), where it is composed of pairs of proteins and one partner carries the beta barrel and the other one carries the secretory protein.

Over 500 types of proteins are secreted via T5aSS class alone. Proteins secreted via the T5SS include adhesins such as AIDA-I and Ag43 of *E. coli*, Hia of *Haemophilus influenzae*, YadA of *Yersinia enterocolitica* and Prn of *Bordetella pertussis*. Toxins such as VacA of *Helicobacter pylori*; proteases such as IgA proteases of *Neisseria gonorrhoeae* and *Neisseria meningitidis*, SepA of *Shigella flexneri* and PrtS of *Serratia marcescens*; and S-layer proteins such as rOmpB of *Rickettsia sp.* and Hsr of *Helicobacter pylori* are released through this secretory mechanism. T5bSS (TPS) secreted proteins including adhesins such as HecA/HecB of the plant pathogen *Dickeya dadantii* (*Erwinia chrysanthemii*) and cytolysins such as ShlA/ShlB of *Serratia marcescens*, HpmA/HpmB of *Proteus mirabilis* and EthA/EthB of *Edwardsiella tarda* [11].

### The Sec/Tat Independent secretory systems

While T1SS, T4SS and can also operate with or without Sec/Tat secretory mechanisms, Type III and Type VI secretion systems are the classical independent mechanisms.

Type III secretion system (T3SS): Some of the world's most important diseases of plants, animals and humans are caused by effector proteins delivered via type T3SS. T3SS is a needle like nanomachine that injects effectors directly into cytoplasm of eukaryotic host cells to initiate infection. It consists of bacterial membrane embedded basal apparatus, an external needle which protrudes from the cell surface and a tip complex which caps the needle. A translocon is assembled between the needle tip complex and host cells creating a pore in host cell membrane [18]. Bacterial flagellum, which is a key motility organelle also secretes virulence factors through extracellular filament [19]. Following entry into host cells, T3SS effectors manipulate host cellular compartments and molecular pathways. Modification of cell signaling pathways, destruction of cell membrane and mimicry of structure/function of eukaryotic proteins are some of the strategies employed by the effectors to help bacteria in colonization, invasion and pathogenesis [20,21]. There are distinct regions, domains or motifs in the effectors which function in infection process through protein-protein interaction, organelle targeting and immune modulation [22].

Type VI secretion system (T6SS): T6SS of Gram negative bacteria perforate prokaryotic or eukaryotic cells and release toxic effector proteins directly into the target cells in a single cell-contact dependent step. T6SSs are syringe-like contractile injection system which are similar in structure and function to cell-puncturing device of tailed bacteriophages [23]. The system was initially found to deliver effectors into eukaryotic cell, however, recent studies suggest T6SSs are more important to mediate interbacterial interactions [24]. T6SS *Pseudomonas aeruginosa* effectors T6S exported 1 (Tse1) and Tse3 degrade peptidoglycan of other Gram negative bacteria into small soluble fragments [25]. Several pathogenic genera *Burkholderia*, *Pseudomonas*, *Yersinia* and *Vibrio* secrete T6SS effector valine glycine repeat protein G (VgrG), which have dual functions. VgrG serves as integral structural component of T6SS and performs distinct effector activities that include cell adhesion, chitosan degradation and actin filament binding and modification. T6SS also makes cell membrane as a target [23]. T6SS phospholipase effector, known as type 6 lipase effector (Tle) proteins hydrolyse component lipids of cell membrane [26]. T6SS effectors also act as nucleases, which under experimental condition degrade plasmid and chromosomal DNA [27]. T6SS plays important role in mediating interbacterial antagonism, which in turn is important for bacterial pathogenicity. Enteric pathogens that have

T6SS, including *Salmonella enterica*, *Aeromonas hydrophila*, *Citrobacter rodentium* establish infection by disruption of the colonization barrier of native microflora of the gut [24].

Type VII secretion system (T7SS): Secretory proteins in Gram positive bacteria have much shorter path to cover to get excreted. However, in organisms like *Mycobacterium* where the cell wall core consists of three covalently linked structures, peptidoglycan, arabinoglycan and mycolic acids, where the mycolic acids forms hydrophobic layer with extreme low fluidity, secretion requires specialized T7SS [28]. T7SS or ESX (Early secreted antigenic target protein) transports proteins across the hydrophobic mycomembrane. Pathogenic mycobacteria have 5 different T7SS, designated as ESX-1 to ESX-5, some of which are important for mycobacterial virulence (ESX-1, ESX-5) and some for viability (ESX-3, ESX-5) [29]. Proteins secreted through T7SS apart from ESX-4 have characteristic N-terminal proline-glutamic acid (PE) or proline-proline-glutamic acid (PPE) motifs [29]. In the extracellular milieu, mycobacteria depend on ESX-3 for metal acquisition and therefore, viability. Following internalization into macrophages, ESX-1 secreted factors acting as pore forming toxins circumvent phagolysosomal degradation and facilitate translocation of pathogenic bacteria into host cytosol. The cytosolic localization promotes mycobacterial replication and ESX-5 secreted proteins manipulate immune response of macrophages inducing cell death and promote bacterial dissemination.

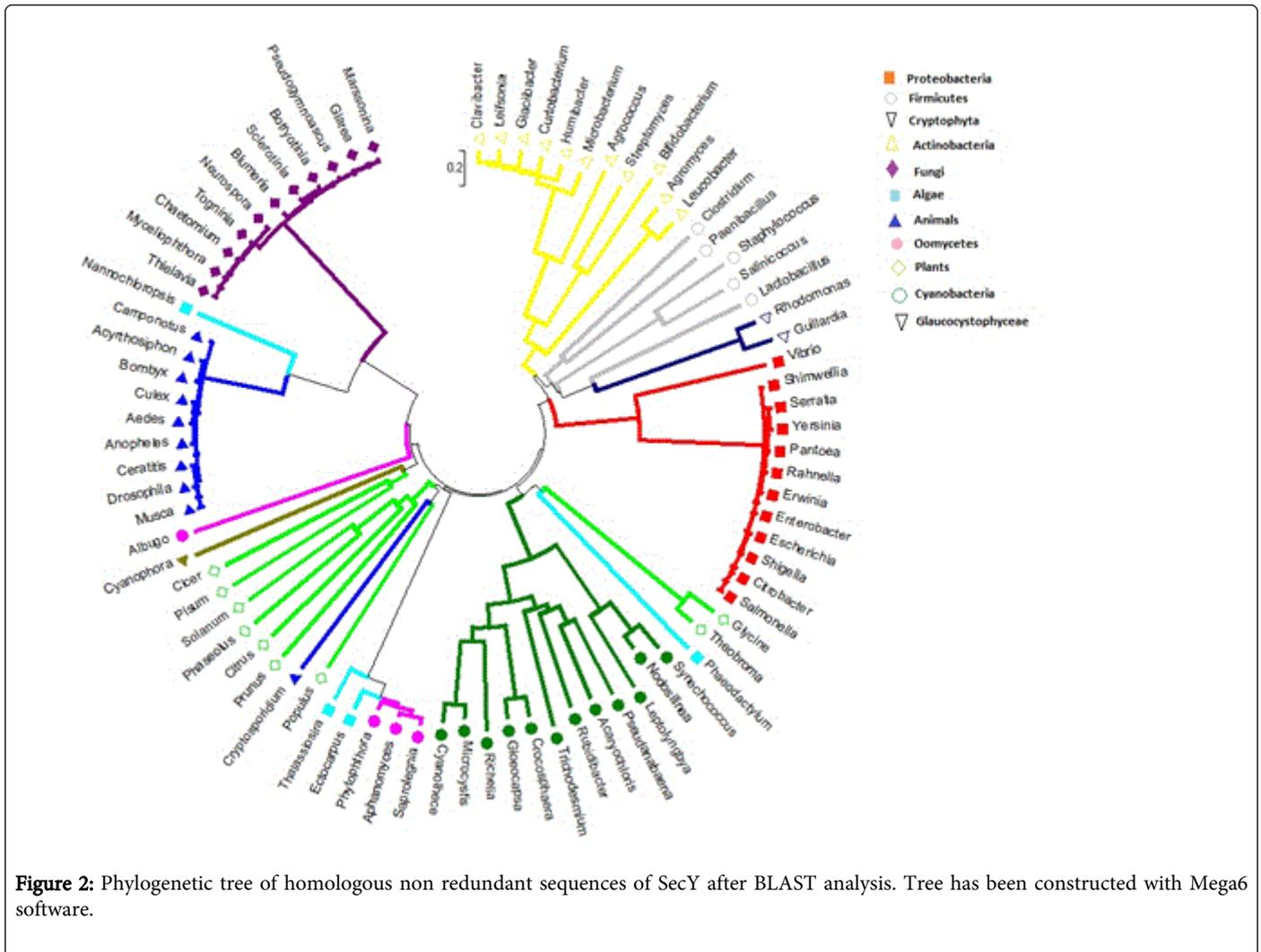
Chaperone usher pathway: Extracellular proteinaceous fibres are critical virulence factors of many Gram negative bacteria, e.g. the p and type 1 pili are virulence factors of uropathogenic *Escherichia coli* (UPEC), which cause majority of urinary tract infection (UTI). Pili, also termed as fimbriae are non-flagellar proteinaceous appendages composed of multiple pilin units. They are involved in host attachment and invasion, biofilm formation, cell motility and transport of proteins and DNA across membranes [30]. The chaperone-usher (CU) pathway is the most widespread among several pathways that assemble adhesive pili at the surface of Gram negative bacteria. Pili are assembled by two units, periplasmic chaperone and outer membrane pore-forming protein, usher. Chaperone facilitates folding of pilus subunits in periplasm and targets them to usher. Usher is assembly platform where pilus subunits are coordinated to form

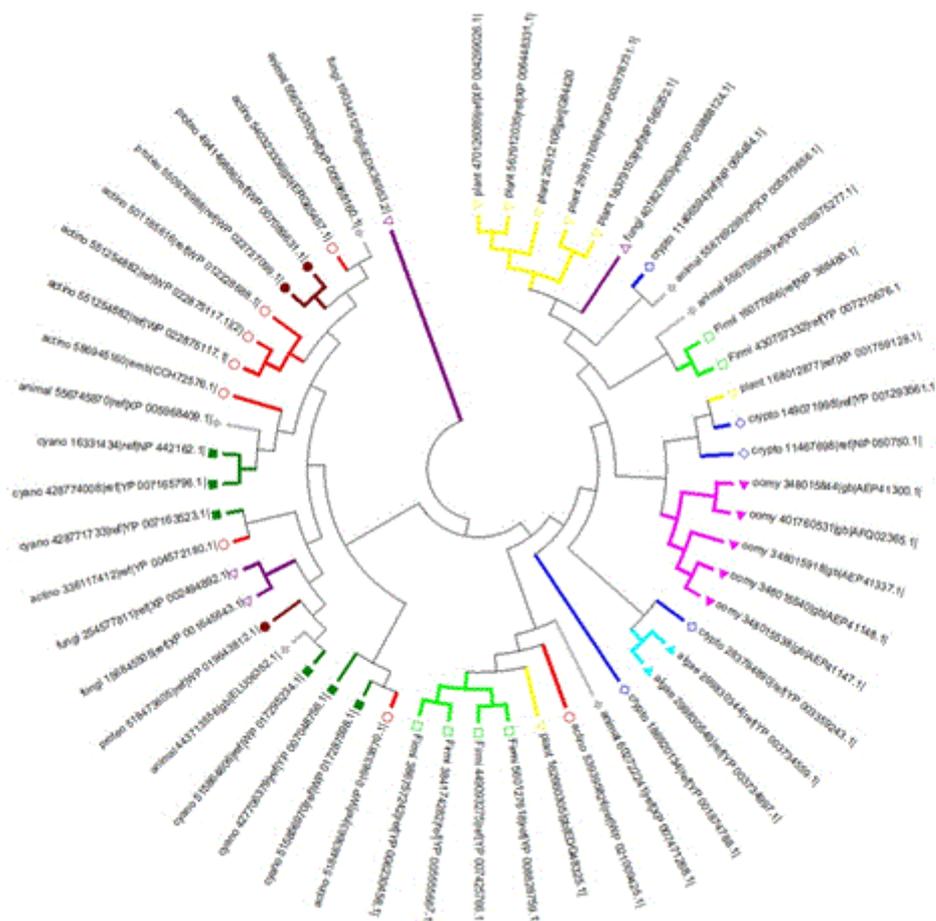
pilus and released into extracellular environment through usher pore [30]. Adhesions of pili (FimH) interact with mannose residues on host epithelial cells to initiate pathogenesis [31]. Because of their important role in virulence, several antibacterial agents have been designed which disturb pilus biogenesis or block bacterial adhesion to host cells [32].

Localization of lipoprotein (LoL) system: Bacterial lipoproteins are synthesized as precursors in the cytoplasm and processed into mature forms on the cytoplasmic membrane. Some lipoproteins play vital roles in the sorting of other lipoproteins, lipopolysaccharides, and  $\beta$ -barrel proteins to the outer membrane. *Helicobacter pylori* colonize the gastric mucosa of the human stomach with a variety of factors secreted from its outer membrane. Lipopolysaccharide and numerous outer membrane proteins are involved in adhesion and immune stimulation [4].

### Evolutionary Patterns of Sec and Tat mechanisms

Although some information is available on Sec and Tat proteins evolutionary relationships, it does not include life from different clades. In order to investigate the relationship between Sec and Tat machineries across the kingdoms, we did a fresh phylogenetic analysis on Sec Y and TatC proteins across the tree of life. We studied the evolutionary patterns of secY and TatC proteins across 11 clades e.g.; Proteobacteria, firmicutes, Cyanobacteria, cryptophyta, actinobacteria, fungi, algae, oomycetes, plants Glaucocystophyceae, animalia. Glaucocystophyceae belongs to a plant taxon that has reduced secretion machinery [33]. In secY protein cluster, substantial amount of sequence conservation was observed. Some members of animal secY grouped with algae and vice versa indicating acquisition of these genes via lateral gene transfer in these organisms [Figure. 2]. The TatC system however appeared to be much diverse and much disperse in their evolutionary patterns as was already discussed [34]. For instance the oomycetes TatC system showed very little sequence similarity with other groups and all other groups indicating a much more polyphyletic origin. Members of different groups shared much less identity with each other's (< 40%) and oomycetes groups had the least identity with the remaining taxonomic groups (<=25%) [Figure. 3]. The red alga Glaucophyta lacked any tatC homologue.





**Figure 3:** Phylogenetic tree of homologous non redundant sequences of TatC after BLAST analysis. Tree has been constructed with Mega6 software.

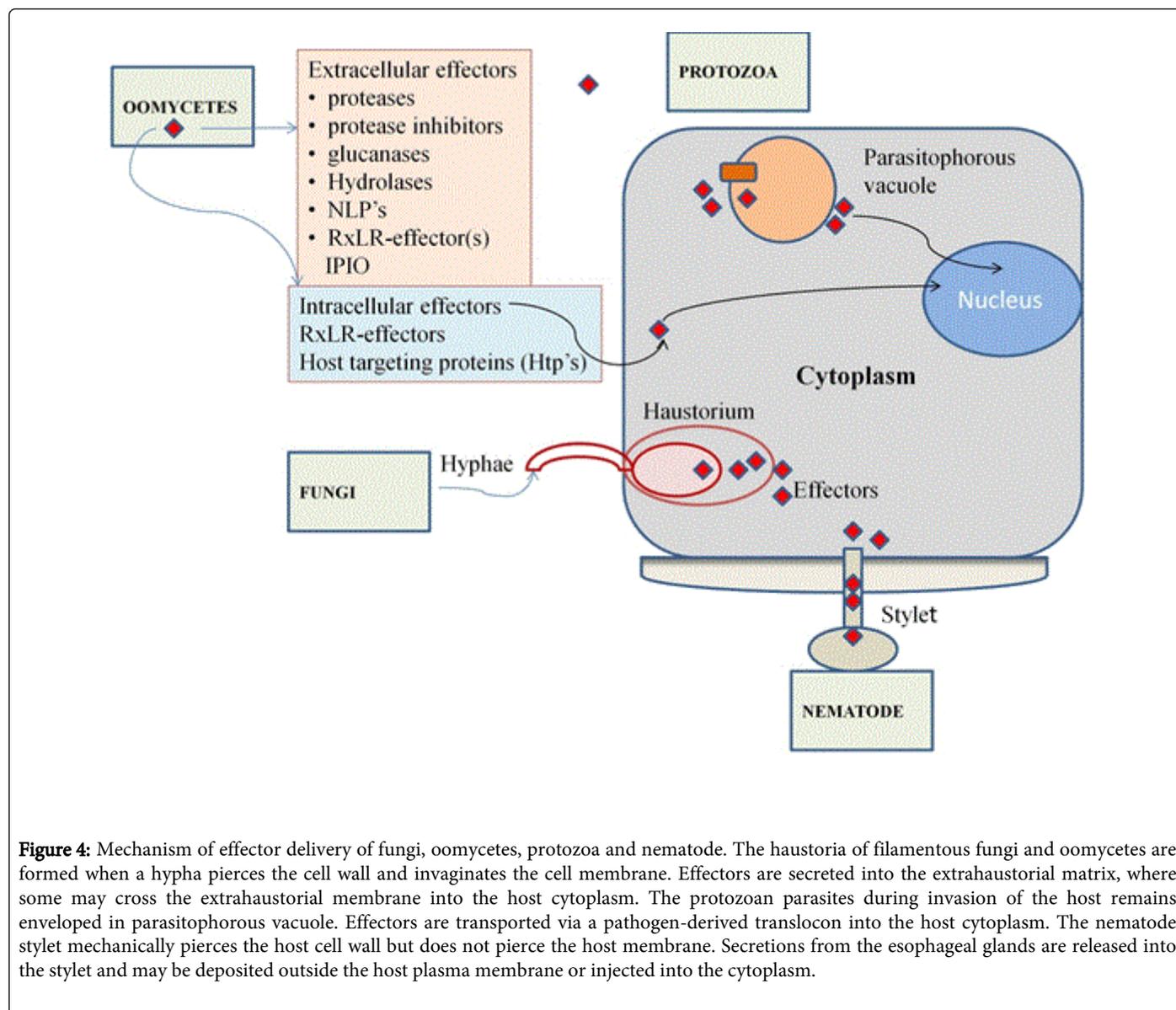
### Secretory systems in other microorganisms

In addition to bacterial effectors, there are a number of organisms such as oomycetes, fungi, nematodes and protozoa which cause devastating diseases by secretion of large number of effector proteins. Figure. 4 summarizes the mechanisms by which these effectors modulate host system.

In the case of many pathogenic fungi (e.g., powdery mildew and rust fungi, *Magnaporthe oryzae* and *Colletotrichum higginsianum*) and oomycetes (*Hyaloperonospora arabidopsidis*, *Phytophthora infestans* and *Phytophthora sojae*) penetration of the host cell wall is accomplished via a hypha that differentiates into a specialized feeding structure called as haustorium. Effectors are secreted in the extrahaustorial matrix and further translocated to host cytoplasm [35-37]. Several haustorially expressed secreted proteins (HESPs; effectors) interact with cytoplasmic nucleotide-binding site and leucine-rich repeat (NBS-LRR) resistance proteins in the host. The

intracellular protozoan parasites of humans (e.g. *P. falciparum* causing malaria, *Toxoplasma* spp. causing sleeping sickness) forms a specialized structure when entering a host cell during the blood stage of infection. This structure, the parasitophorous vacuole (PV), is functionally analogous to the haustorium of fungi and oomycetes. Effectors, similar to oomycetes and fungi, are translocated across the PV to host cytoplasm [38,39].

Plant parasitic nematodes (PPNs) are small roundworms infesting roots of thousands of plant species. PPNs have hollow protrusible syringe-like stylet, which mechanically pierces the host cell wall and injects gland secretions into the host cell cytoplasm [40]. The injectisomes of bacterial T3SS and T4SS are analogous to the stylets of PPNs [37,41]. A large number of nematode effectors are cell wall degrading or cell wall modifying enzymes and also affect plant signaling, hormone balance, and cell morphogenesis [42].



### Targeting of effectors to host cell, role of conserved motifs

Pathogenic microorganisms secrete various effector proteins to manipulate host cell machinery for virulence and survival. Effectors are composed of functionally distinct domains or motifs that are responsible for protein-protein interaction, enzymatic action or organelle targeting. Identification of these motifs are important leads for designing antimicrobial peptides or drugs. Here we summarize some well recognized motifs of bacterial effectors and their targeting mechanisms [43]. A large number of toxins secreted through T1SS mostly exhibiting cytotoxic pore forming activity belong to RTX family of proteins characterized by the presence of arrays of glycine- and aspartate-rich repeats [44,45]. Following entry into host cells, bacterial effectors reach their specific targets including proteins in the nucleus, mitochondria, endoplasmic reticulum, cytoskeleton, or plasma membrane utilizing specific motifs in effectors and carry out their activities [43,45]. A class of T3SS and T4SS effectors bearing Glu-

Pro-Ile-Tyr-Ala (EPIYA) sequence (EPIYA motif) elicit pathogenic response by manipulating host cell signaling [46] e.g. CagA protein of *Helicobacter pylori* is the archetypal EPIYA effector causing gastric carcinoma [47]. A number of T3SS effectors with their motifs playing roles during infection has been demonstrated [22]. *Salmonella enterica* serovar Typhimurium utilize distinct motifs in their T3SS effector to advance infection following host cell function [48]. The tetratricopeptide repeat (TPR) motifs containing proteins are involved in virulence mechanisms of large number of T3SS effectors [49]. C-terminal amino acid composition and possible motifs have been exploited to predict T4SS effectors causing pathogenesis [50]. Arg-Gly-Asp (RGD) motif present in extracellular host adhesion factor is used by some bacteria to attach onto the host cells. IPIO effector or some proteins/peptides with an RGD motif interact with adhesion between cell wall and plasma membrane leading to disruption of plant cells. The motifs composed of less than 10 amino acids, called short linear motifs (SliMs) enable pathogenic bacteria to control intracellular process of host [51]. The cell wall peptidoglycan of Gram positive

bacteria act as surface organelle for transport and assembly of many proteins which interact with hosts. Proteins secreted and released into the environment by living Gram positive beneficial bacteria (probiotics) interact directly with mucosal cells, such as epithelial and immune cells [52]. Identification of these proteins and their motifs (Ala-X-Ala, signal peptide type I and the Leu-Ala-Gly-Cys, signal peptide type II or lipobox sequence) are important to design probiotic effector molecules [53-55]. Extracellular pilus of several probiotic bacteria induce immunomodulatory activity [56-58]. Bacterium *Xanthomonas* inject TAL effector proteins into plant cells which recognize effector specific promoter sequence in nuclear DNA via unique repeat-variable diresidues (RVDs) and activate plant genes that either benefit the bacterium or trigger host defense response [59]. Table 1 presents different examples of effectors with their targeting mechanisms. The availability of genome sequences has led to identification of plethora of putative translocated effector proteins in oomycetes [35,60]. Effectors of this pathogen are categorized into extracellular and intracellular types. Extracellular effectors mediate protection against host defense which include extracellular protease inhibitor EPI1, EPI10, cysteine protease inhibitors EPIC1, EPIC2B and glucanase inhibitors which all prevent degradation of pathogen cell wall components and subsequent release of elicitors to control host defence. The effectors which aid in invasion include some hydrolytic enzymes and toxins.

Intracellular effector proteins act by suppressing host defense responses. However, some effectors can be recognized by host resistance (R) proteins and, as a result, initiate specific gene-for-gene defense responses. In such cases, these effectors are termed avirulence (Avr) proteins. A large number of Avr proteins show presence of a highly conserved amino acid motif, RXLR (Arg-any amino acid-Leu-Arg) which is often followed by an EER-motif (Glu-Glu-Arg) [35,60] in the C terminus. RXLR motif and its variants in both oomycetes and fungi enable the effectors to bind phosphatidylinositol-3-phosphate on the outer surface of plant and human cell plasma membranes, and this binding mediates the effector entry through lipid raft mediated endocytosis [61]. *Pseudoperonospora cubensis* has a nuclear-localized effector (PcQNE) with a QXLR motif, where the R of RXLR is substituted with a Q in the N terminus; plays a pivotal role in pathogenicity [62]. Another group of cytoplasmic effectors of oomycetes, known as crinklers (CRN) cause leaf crinkling and necrosis of host plants [35,60]. They exhibit highly conserved N-terminal Leu-Xaa-Leu-Phe-Leu-Ala-Lys (LxLFLAK) domain and a tri-peptide signature (Asp-Trp-Leu, DWL) at C-terminal region that ends with a

conserved His-Val-Leu-Val-Xaa-Xaa-Pro motif. Crinklers have a nuclear localization signal motif that enables them to get targeted to the nucleus of the host cells.

Transport across the PV membrane by *P. falciparum* effectors into the host cytosol requires an 11-amino-acid host cell-targeting (HT) motif (Rx1SRxLxE/D/Qx2x3x4) with a 5-amino-acid PEXEL [alternatively known as Plasmodium Export Element (PEXEL)] core (RxLxE/D/Q) that is conserved among diverse proteins [63]. The HT-motif and PEXEL are identified by different algorithms and have slightly different specificities, but recognize the same core sequence (RxLxE/Q/D). Proteins with HT motif cross PV and enter into red blood cell, whereas proteins without HT are released into PV. Recent studies show the export mechanism is due to the HT signal binding to the lipid phosphatidylinositol-3-phosphate (PI (3) P) in the parasite endoplasmic reticulum (ER). The protease in parasite ER, plasmepsin V cleave HT signal and thus proteins are released into erythrocyte cytoplasm [64], however, PI(3)P independent transport pathway is also reported [65]. HT signal was originally identified as a highly conserved motif by pattern recognition programs in five functionally equivalent~40 amino acid, vacuolar translocation sequence (VTS) [AFNNLCSKNAKGLNLNKRLLYETQAHVDDVHHAHHADV]. The R, L and E region (shown in bold in the pattern describe above), is known to contribute to both PI (3) P binding and protein export to the erythrocyte. Substitution of sequences KNAKGLN upstream of the HT motif (underlined in the pattern describe above), also blocked export but have not been tested for PI (3) P binding. The italicized sequence in the pattern (QAHVDDVHHAHHADV) indicates charged amino acids downstream of the HT motif enable export independent of PI (3)P. Substitution of HHAHHA sequence in repeat sequences does not block export [63,66].

*Toxoplasma* parasite secrete a number of effectors to manipulate host immune system, these are rhoptyr protein kinases (the ROP class); ROP18, ROP5, ROP16, ROP38, which accumulate in PV. Conserved HGB motif was identified in ROP5, that is a major virulence determinant of this protozoa [67].

The cyst nematodes secrete active CLAVATA3/ESR (CLE)-like proteins that facilitate nematode parasitism by interacting with several plant signaling pathways [42]. Nuclear localization signals (NLSs) have been predicted in several secreted proteins. During parasitism, these effectors interact with nuclear proteins leading to suppression of plant immunity.

Organisms	Effectors	Conserved sequence/ Motifs	Targeting mechanism
Salmonella spp.	T3SS effector SspH2 SseI (P)	C-terminal NEL (new E3 ligase) domain C-terminal bacterial deamidases	Effectors targeted to host plasma membrane through Spalmitoylation: (Addition of 16-carbon palmitic acid to a specific Cys residue of effector protein) [86,87]
Pseudomonas spp.	T3SS AvrB, AvrRpm1, HopF2, HopZ, AvrPphB, ORF4, AvrPto, (M+ P) ExoU	GXXX(S/T/C)	Effectors targeted to host plasma membrane through Nmyristoylation (addition of 14-carbon myristic acid to effector protein) [88-90]
Legionella spp. Salmonella spp	T4SS effector AnkB T3SS effector SifA	CaaX motif (a, aliphatic residue; X is Ser, Met or Gln) CLCCFL motif	Targeted to LCV through prenylation (addition of a 15-carbon farnesyl or a 20-carbon geranyl geranyl isoprenoid group to a Cys residue in CaaX motif) [91] Targeted to SCV and SIF through prenylation [92]

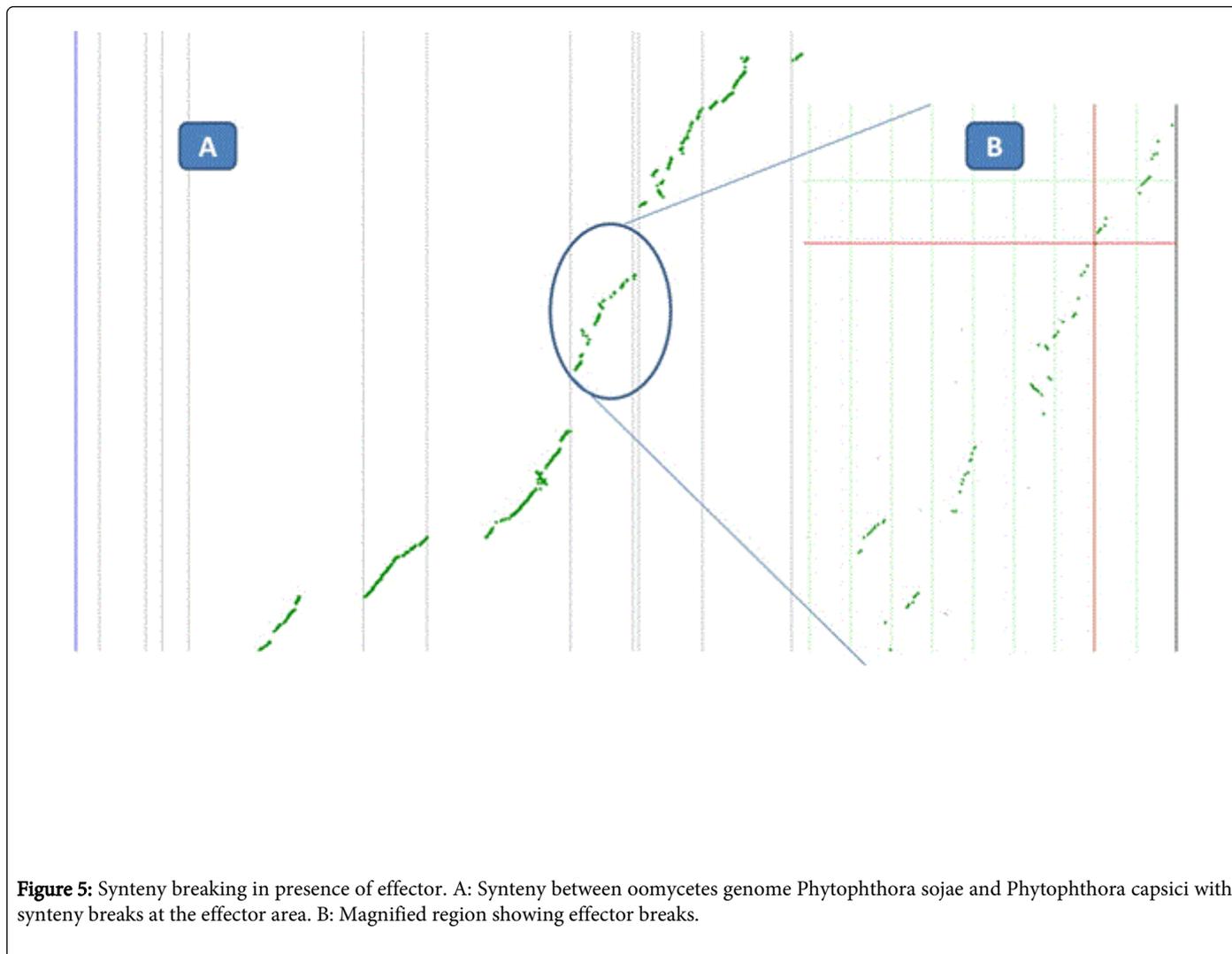
Salmonella spp	T3SS effector SopB	-	Targeted to SCV through ubiquitylation (covalent attachment of ubiquitin to a Lys residue on a target protein) [43]
Legionella spp. Pseudomonas spp. Helicobacter pylori	T3SS effector SidC, SidM T3SS effector ExoU T4SS effector CagA	PtdIns4P binding domain C-terminal phosphoinositide (PtdIns(4,5)P <sub>2</sub> ) binding domain EPIYA (Glu-Pro-Ile-Tyr-Ala)	Targeted to LCV by binding with PtdIns4P [93,94] Affinity for PtdIns-4,5-bisphosphate which is abundant at the cytoplasmic side of the plasma membrane Phosphatidylserine in plasma membrane [46,47]
Pseudomonas spp. Escherichia coli	T3SS effector HopL1 and HopG1 T3SS effectors Map EspF, EspZ	Chloroplast and mitochondria targeting N-terminal pre sequence Mitochondria targeting N-terminal pre sequence along with C-terminal Thr-Arg-Leu (TRL) motif Mitochondria targeting N-terminal pre sequence	Alters thylakoid structures and suppresses host defense [95] Disruption of mitochondrial function and fragmentation of mitochondria [96-98]
Xanthomonas spp. Agrobacterium tumefaciens	T3SS TAL effector AvrBs3, AvrXa5, AvrXa7, AvrXa10 and PthA . T4SS effectors VirD2 and VirE2	A central DNA-binding region consisting of a nearly identical tandem of 34 amino acid repeats, followed by a classic NLS and an acidic transcriptional activation domain (AAD) NLS: one or two short amino acid sequences rich in the positively charged amino acids Lys and Arg. NLS	Control gene expression leading to host cell death [99] Transfer of bacterial T-DNA into the host plant nucleus leads to tumor formation [100]
Oomycetes Extracellular effectors	EPI1, EPI10, EPIC1 IPI-O Toxins NLP	NA RGD NA	Prevents degradation of cell wall [35] Disrupts adhesion between cell wall and plasma membrane [60] Disrupts adhesion between cell wall and plasma membrane [60]
Intracellular effectors	Avr PcQNE CRN	RXLR QXLR N-terminal LXLFLAK, DWL at the beginning of C-terminal which end with His-Val-Leu-Val-Xaa-Xaa-Pro motif	Suppression of plant immunity [35,60] Targeted to nucleus by NLS [62] Targeted to nucleus by NLS [35,60]
Protozoa Plasmodium Toxoplasma	PfEMP1, RIFIN, STEVOR ROP5	Rx1SRxLxE/D/Qx2x3x4 RGB	Target protein from parasite PV to host cytoplasm [63,66] Virulence [67]

**Table 1:** The effectors with their conserved motifs and targeting mechanism in host. Abbreviations: AnkB, ankyrin B; Avr, avirulence; LCV, Legionella-containing vacuole; NLS, nuclear localization signal; PtdIns4P, phosphatidylinositol-4-phosphate; PtdIns(4,5)P<sub>2</sub>, PtdIns-4,5-bisphosphate; SCV, Salmonella-containing vacuole; SIFs, Salmonella-induced filaments; TAL, Transcription activator-like; T-DNA, tumour-inducing DNA.

## Effectors localize themselves in Long Intergenic Regions

In eukaryotic pathogens, effector molecules are found to be localized in long intergenic regions [68]. The most plausible reason for this is when effectors try to integrate themselves in a tightly packed, more stable gene rich region, that region become unstable and hence undergoes negative selection. However, gene sparse long intergenic regions are more plastic and can tolerate transposon mediated gene expansion [69]. Many phytopathogenic ascomycetes such as *Leptosphaeria maculans* are known to have effectors localized in AT

rich region [70,61]. Recently a great deal of genomics work has been carried in effector systems in Oomycetes pathogens [71,76]. In oomycetes pathogens, there are pockets in the genome called as gated communities, where all the housekeeping and conserved genes reside [72]. Any effector integration event is rapidly flushed out in these area and hence the term gated community has been coined for such areas [73]. The point of effector insertion is known to cause a break of co-linearity between conserved genomes [Figure 5]. Recently chromosomal positions of effector genes have been built into algorithms for effector discovery.



**Figure 5:** Synteny breaking in presence of effector. A: Synteny between oomycetes genome *Phytophthora sojae* and *Phytophthora capsici* with synteny breaks at the effector area. B: Magnified region showing effector breaks.

### Algorithms for effector prediction

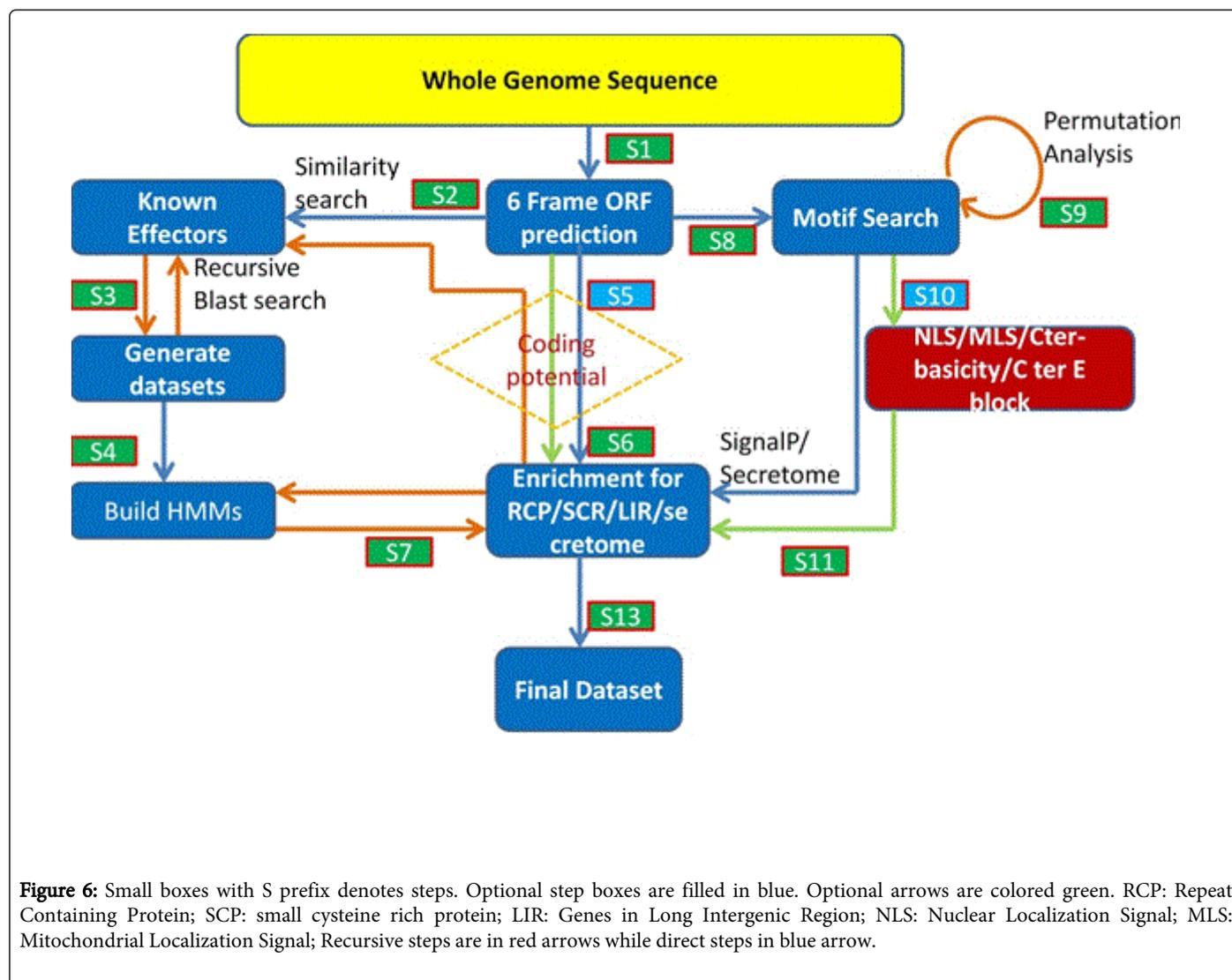
For development of algorithms to predict effectors, it is essential to understand few of their common properties [74,75]. Effectors almost always are small molecules, undergoing positive selection, expansion, rapid evolution [71,76]. They have a N terminus secretory signal 77 and a C terminal E rich basic motif. Interspersed in the protein sequence lies relatively larger number of cysteine residues for proper protein folding [78]. Several types of effectors have been studied to have a specific signature motif that distinguishes and defines its physiological role [Table 1]. These often lack introns and are located in long intergenic regions attached with a transposon. Other than this, they often are characterized as lacking a transmembrane domain [37]. Depending on their localization in the host cell, they are known to carry a NLS (Nuclear Localization Signal), MLS (mitochondrial Localization Signal) [1,79]. Since most of the effectors are located in Genome islands and are acquired through HGT events, they have subtle difference in GC content, coding potential and other genomic properties with the native genome. Based on these traits, many algorithms have been designed for efficient effector discovery [Figure. 6].

The TATFIND algorithm was developed to identify putative Tat substrates by looking for the presence of this conserved motif in bacterial signal peptides [80].

Effectors lack significant sequence homology with each other, so prediction methods based on homology almost always fails. Most of the gene prediction algorithms miss out on finding effectors in a newly sequenced organism [71,81]. Effector prediction in a newly sequenced organism can be carried out with the following steps. Step-1: Genome wide 6 frame translation: The whole genome six frame translation can be carried out using getorf program of EMBOSS package [82]. Step-2: Filtration: Sequences smaller than 30 aa residues can be filtered out. Step-3: Recursive Blast search [83]: Perform blast against known effectors and merge the homologous sequences predicted by blast into the database and repeat blast till no more new blast hits are obtained. Step 4: Cluster proteins and build HMMs [84]: From the filtered blast output, cluster proteins using bbh (bi-directional best blast hit), followed by running clustalw [85]. Build Hmm using HMMER software [84]. Step 5: coding potential filtering: For the given organism create codon usage table using CUSP program from EMBOSS package and filter out sequences with coding potential < zero. Step 6: Filter RCPs, SCPs and LIRs with higher coding potential: The 6 frame ORF data that has Repeat containing proteins (RCP), Small Cysteine Rich

Proteins (SCRs), Long Intergenic Regions (LIRs) [78] can be further selected for the next process. Step 7: HMM search: Do HMM search of the dataset created from step 4 and continue a recursive search against by adding new HMM plus sequences into HMM datasets. Continue till no new HMMs are obtained. Step 8: Motif search: Further filter down the sequences obtained from Step 5 and look for the motifs that is typical for the kind of effectors you are looking for. Step 9: Permutation Analysis: Shuffle the nucleotides flanking the motif of interest and see if the motif sequence occurs by chance in how many

cases that lacks any biological significance. Step 10: Detecting Signals : From the sequences with true positive motifs, find out the proteins having Nuclear Localization Signal (NLS), Mitochondrial Localization Signal (MLS), C terminal basic region [55], C terminal E motifs, secretome positive proteins [77], transmembrane negative proteins and merge with dataset generated from step 6. Repeat step 7 till no more new proteins are obtained. This final dataset is now clustered using clustalw for grouping them into distinct clusters for evolutionary studies.



**Figure 6:** Small boxes with S prefix denotes steps. Optional step boxes are filled in blue. Optional arrows are colored green. RCP: Repeat Containing Protein; SCP: small cysteine rich protein; LIR: Genes in Long Intergenic Region; NLS: Nuclear Localization Signal; MLS: Mitochondrial Localization Signal; Recursive steps are in red arrows while direct steps in blue arrow.

## Conclusion

The avalanche of genomic information available over the last decade has enabled scientists in understanding the key mechanisms responsible for secretion process. Secretory proteins are no longer the elusive elements as they were before. From whole genome sequences, canonical and non-canonical secretomes are now predicted effortlessly using modern computational methods. We have taken the most up-to-date information available on secretory mechanisms and summarize them in this article. The Sec/Tat dependent and independent systems are discussed in detail with proper examples to serve as a catalogue for the researchers working in this area. We further corroborated the fact that Tat machineries are more diverse whereas Sec pathways are quite

conserved using diverse genomic data. We have compiled the secretion types with their associated functions and motifs involved in a great detail. In addition we have also summarized the software resources available for effector prediction for the benefit of wet lab researchers. In this article we made an attempt to compile relevant information on effector biology and prediction across the microbial world. This review will undoubtedly serve as a quick guide for wet lab and dry lab scientists in understanding effector biology, motif finding and prediction of novel effectors.

## Acknowledgement

ST thanks Council of Scientific and Industrial Research, Department of Biotechnology for Ramalingaswamy fellowship. SB thanks Department of Science & Technology, Government of India for financial support.

## References

1. Dou D, Kale SD, Wang X, Chen Y, Wang Q, et al. (2008) Conserved C-terminal motifs required for avirulence and suppression of cell death by *Phytophthora sojae* effector Avr1b. *Plant Cell* 20: 1118-1133.
2. Tseng TT, Tyler BM, Setubal JC (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol* 9 Suppl 1: S2.
3. Bleves S, Viarre V, Salacha R, Michel GP, Filloux A, et al. (2010) Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons. *Int J Med Microbiol* 300: 534-543.
4. Papanikou E, Karamanou S, Economou A (2007) Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol* 5: 839-851.
5. Feltcher ME, Braunstein M (2012) Emerging themes in SecA2-mediated protein export. *Nat Rev Microbiol* 10: 779-789.
6. Chatzi KE, Sardis MF, Karamanou S, Economou A (2013) Breaking on through to the other side: protein export through the bacterial Sec system. *Biochem J* 449: 25-37.
7. Lee PA, Tullman-Ercek D, Georgiou G (2006) The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* 60: 373-395.
8. Barnett JP, Robinson C, Scanlan DJ, Blindauer CA (2011) The Tat protein export pathway and its role in cyanobacterial metalloprotein biosynthesis. *FEMS Microbiol Lett* 325: 1-9.
9. Palmer T, Berks BC (2012) The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* 10: 483-496.
10. Natale P, Bruser T, Driessen A J (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* 1778, 1735-56.
11. Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 68: 692-744.
12. Korotkov KV, Sandkvist M, Hol WG (2012) The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* 10: 336-351.
13. Douzi B, Filloux A, Voulhoux R (2012) On the path to uncover the bacterial type II secretion system. *Philos Trans R Soc Lond B Biol Sci* 367: 1059-1072.
14. Alvarez-Martinez CE, Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* 73: 775-808.
15. Voth DE, Broderdorf LJ, Graham JG (2012) Bacterial Type IV secretion systems: versatile virulence machines. *Future Microbiol* 7: 241-257.
16. Juhas M, Crook DW, Hood DW (2008) Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol* 10: 2377-2386.
17. Leo JC, Grin I, Linke D (2012) Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* 367: 1088-1101.
18. Chatterjee S, Chaudhury S, McShan AC, Kaur K, De Guzman RN (2013) Structure and biophysics of type III secretion in bacteria. *Biochemistry* 52: 2508-2517.
19. Büttner D (2012) Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol Mol Biol Rev* 76: 262-310.
20. Tosi T, Pflug A, Discola KF, Neves D, Dessen A (2013) Structural basis of eukaryotic cell targeting by type III secretion system (T3SS) effectors. *Res Microbiol* 164: 605-619.
21. Deslandes L, Rivas S (2012) Catch me if you can: bacterial effectors and plant targets. *Trends Plant Sci* 17: 644-655.
22. Dean P (2011) Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol Rev* 35: 1100-1125.
23. Kapitein N, Mogk A (2013) Deadly syringes: type VI secretion system activities in pathogenicity and interbacterial competition. *Curr Opin Microbiol* 16: 52-58.
24. Russell AB, Peterson SB, Mougous JD (2014) Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* 12: 137-148.
25. Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W, et al. (2011) Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* 475: 343-347.
26. Russell AB, LeRoux M, Hathazi K, Agnello DM, Ishikawa T, et al. (2013) Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature* 496: 508-512.
27. Poole SJ, Diner EJ, Aoki SK, Braaten BA, Low DA et al. (2011). Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS Genet* 7, e1002217.
28. Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, et al. (2007) Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* 5: 883-891.
29. Stoop EJ, Bitter W, van der Sar AM (2012) Tubercle bacilli rely on a type VII army for pathogenicity. *Trends Microbiol* 20: 477-484.
30. Waksman G, Hultgren SJ (2009) Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol* 7: 765-774.
31. Busch A, Waksman G (2012) Chaperone-usher pathways: diversity and pilus assembly mechanism. *Philos Trans R Soc Lond B Biol Sci* 367: 1112-1122.
32. Lo AW, Van de Water K, Gane PJ, Chan AW, Steadman D, et al. (2014) Suppression of type 1 pilus assembly in uropathogenic *Escherichia coli* by chemical inhibition of subunit polymerization. *J Antimicrob Chemother* 69: 1017-1026.
33. Vogel H, Fischer S, Valentin K (1996) A model for the evolution of the plastid sec apparatus inferred from secY gene phylogeny. *Plant Mol Biol* 32: 685-692.
34. Simone D, Bay DC, Leach T, Turner RJ (2013) Diversity and evolution of bacterial twin arginine translocase protein, TatC, reveals a protein secretion system that is evolving to fit its environmental niche. *PLoS One* 8: e78742.
35. Stassen JH, Van den Ackerveken G (2011) How do oomycete effectors interfere with plant life? *Curr Opin Plant Biol* 14: 407-414.
36. Tyler BM (2009) Entering and breaking: virulence effector proteins of oomycete plant pathogens. *Cell Microbiol* 11: 13-20.
37. Torto-Alalibo T, Collmer CW, Lindeberg M, Bird D, Collmer A et al. (2009). Common and contrasting themes in host cell-targeted effectors from bacterial, fungal, oomycete and nematode plant symbionts described using the Gene Ontology. *BMC Microbiol* 9 Suppl 1, S3.
38. Hunter CA, Sibley LD (2012) Modulation of innate immunity by *Toxoplasma gondii* virulence effectors. *Nat Rev Microbiol* 10: 766-778.
39. Harvey KL, Gilson PR, Crabb BS (2012) A model for the progression of receptor-ligand interactions during erythrocyte invasion by *Plasmodium falciparum*. *Int J Parasitol* 42: 567-573.
40. Quentin M, Abad P, Favery B (2013) Plant parasitic nematode effectors target host defense and nuclear functions to establish feeding cells. *Front Plant Sci* 4: 53.
41. Hewezi T, Baum TJ (2013) Manipulation of plant cells by cyst and root-knot nematode effectors. *Mol Plant Microbe Interact* 26: 9-16.
42. Mitchum MG, Hussey RS, Baum TJ, Wang X, Elling AA, et al. (2013) Nematode effector proteins: an emerging paradigm of parasitism. *New Phytol* 199: 879-894.
43. Hicks SW, Galán JE (2013) Exploitation of eukaryotic subcellular targeting mechanisms by bacterial effectors. *Nat Rev Microbiol* 11: 316-326.

44. Linhartová I, Bumba L, Maňáková J, Basler M, Osička R, et al. (2010) RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol Rev* 34: 1076-1112.
45. Geissler B (2012) Bacterial toxin effector-membrane targeting: outside in, then back again. *Front Cell Infect Microbiol* 2: 75.
46. Hayashi T, Morohashi H, Hatakeyama M (2013) Bacterial EPIYA effectors--where do they come from? What are they? Where are they going? *Cell Microbiol* 15: 377-385.
47. Noto JM, Peek RM Jr (2012) The *Helicobacter pylori* cag Pathogenicity Island. *Methods Mol Biol* 921: 41-50.
48. Agbor TA, McCormick BA (2011) Salmonella effectors: important players modulating host cell function during infection. *Cell Microbiol* 13: 1858-1869.
49. Cerveny L, Straskova A, Dankova V, Hartlova A, Ceckova M, et al. (2013) Tetratricopeptide repeat motifs in the world of bacterial pathogens: role in virulence mechanisms. *Infect Immun* 81: 629-635.
50. Wang Y, Wei X, Bao H, Liu SL (2014) Prediction of bacterial type IV secreted effectors by C-terminal features. *BMC Genomics* 15: 50.
51. Ruhanen H, Hurley D, Ghosh A, O'Brien KT, Johnston CR et al (2014). Potential of known and short prokaryotic protein motifs as a basis for novel peptide-based antibacterial therapeutics: a computational survey. *Front Microbiol* 5, 4.
52. Sánchez B, Urdaci MC, Margolles A (2010) Extracellular proteins secreted by probiotic bacteria as mediators of effects that promote mucosa-bacteria interactions. *Microbiology* 156: 3232-3242.
53. von Heijne G, Abrahmsén L (1989) Species-specific variation in signal peptide design. Implications for protein secretion in foreign hosts. *FEBS Lett* 244: 439-446.
54. Willing B, Halfvarson J, Dicksved J, Rosenquist M, Jarnerot G, et al. (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm Bowel Dis* 15, 653-60.
55. Barinov A, Loux V, Hammani A, Nicolas P, Langella P, et al. (2009) Prediction of surface exposed proteins in *Streptococcus pyogenes*, with a potential application to other Gram-positive bacteria. *Proteomics* 9: 61-73.
56. Kankainen M, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J, et al. (2009) Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein. *Proc Natl Acad Sci U S A* 106, 17193-8.
57. Turrone F, Serafini F, Foroni E, Duranti S, O'Connell Motherway M, et al. (2013) Role of sortase-dependent pili of *Bifidobacterium bifidum* PRL2010 in modulating bacterium-host interactions. *Proc Natl Acad Sci U S A* 110: 11151-11156.
58. O'Connell Motherway M, Zomer A, Leahy SC, Reunanen J, Bottacini F, et al. (2011) Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proc Natl Acad Sci U S A* 108, 11217-22.
59. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* 335: 716-719.
60. Wawra S, Belmonte R, Löbach L, Saraiva M, Willems A, et al. (2012) Secretion, delivery and function of oomycete effector proteins. *Curr Opin Microbiol* 15: 685-691.
61. Kale SD, Gu B, Capelluto DG, Dou D, Feldman E, et al. (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell* 142: 284-295.
62. Tian M, Win J, Savory E, Burkhardt A, Held M, et al. (2011) 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Mol Plant Microbe Interact* 24: 543-553.
63. Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, et al. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 306: 1934-1937.
64. Bhattacharjee S, Stahelin RV, Speicher KD, Speicher DW, Haldar K (2012) Endoplasmic reticulum PI(3)P lipid binding targets malaria proteins to the host cell. *Cell* 148: 201-212.
65. Bhattacharjee S, Speicher KD, Stahelin RV, Speicher DW, Haldar K (2012) PI(3)P-independent and -dependent pathways function together in a vacuolar translocation sequence to target malarial proteins to the host erythrocyte. *Mol Biochem Parasitol* 185, 106-13.
66. Bhattacharjee S, Stahelin RV, Haldar K (2012) Host targeting of virulence determinants and phosphoinositides in blood stage malaria parasites. *Trends Parasitol* 28: 555-562.
67. Melo MB, Jensen KD, Saeij JP (2011) *Toxoplasma gondii* effectors are master regulators of the inflammatory response. *Trends Parasitol* 27: 487-495.
68. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang R H, et al. (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261-6.
69. Judelson HS (2012) Dynamics and innovations within oomycete genomes: insights into biology, pathology, and evolution. *Eukaryot Cell* 11: 1304-1312.
70. Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, et al. (2011) Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nat Commun* 2, 202.
71. Jiang RH, Tyler BM (2012) Mechanisms and evolution of virulence in oomycetes. *Annu Rev Phytopathol* 50: 295-318.
72. Wang Q, Han C, Ferreira AO, Yu X, Ye W, et al. (2011) Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell* 23, 2064-86.
73. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, et al. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-837.
74. Win J, Morgan W, Bos J, Krasileva KV, Cano LM, et al. (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* 19: 2349-2369.
75. Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, et al. (2008) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques* 45: 81-94.
76. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, et al. (2010) Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330, 1549-51.
77. Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S (2004) Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel* 17: 349-356.
78. Saunders DG, Win J, Cano LM, Szabo LJ, Kamoun S, et al. (2012) Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. *PLoS One* 7: e29847.
79. Dou D, Kale SD, Wang X, Jiang RH, Bruce NA, et al. (2008) RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *Plant Cell* 20: 1930-1947.
80. Dilks K, Rose RW, Hartmann E, Pohlschröder M (2003) Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* 185: 1478-1483.
81. Jiang RH, Tripathy S, Govers F, Tyler BM (2008) RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proc Natl Acad Sci U S A* 105: 4874-4879.
82. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16: 276-277.
83. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389-402.
84. Wheeler TJ, Eddy SR (2013) nhmmer: DNA homology search with profile HMMs. *Bioinformatics* 29: 2487-2489.
85. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through

- sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-80.
86. Hicks SW, Charron G, Hang HC, Galán JE (2011) Subcellular targeting of *Salmonella* virulence proteins by host-mediated S-palmitoylation. *Cell Host Microbe* 10: 9-20.
87. Hicks SW, Galán JE (2010) Hijacking the host ubiquitin pathway: structural strategies of bacterial E3 ubiquitin ligases. *Curr Opin Microbiol* 13: 41-46.
88. Nimchuk Z, Marois E, Kjemtrup S, Leister RT, Katagiri F, et al. (2000). Eukaryotic fatty acylation drives plasma membrane targeting and enhances function of several type III effector proteins from *Pseudomonas syringae*. *Cell* 101, 353-63.
89. Lewis JD, Abada W, Ma W, Guttman DS, Desveaux D (2008) The HopZ family of *Pseudomonas syringae* type III effectors require myristoylation for virulence and avirulence functions in *Arabidopsis thaliana*. *J Bacteriol* 190: 2880-2891.
90. Robert-Seilaniantz A, Shan L, Zhou JM, Tang X (2006) The *Pseudomonas syringae* pv. tomato DC3000 type III effector HopF2 has a putative myristoylation site required for its avirulence and virulence functions. *Mol Plant Microbe Interact* 19: 130-138.
91. Price CT, Al-Quadan T, Santic M, Jones SC, Abu Kwaik Y (2010) Exploitation of conserved eukaryotic host cell farnesylation machinery by an F-box effector of *Legionella pneumophila*. *J Exp Med* 207: 1713-1726.
92. Reinicke AT, Hutchinson JL, Magee AI, Mastroeni P, Trowsdale J, et al. (2005) A *Salmonella typhimurium* effector protein SifA is modified by host cell prenylation and S-acylation machinery. *J Biol Chem* 280: 14620-14627.
93. Ragaz C, Pietsch H, Urwyler S, Tiaden A, Weber SS, et al. (2008) The *Legionella pneumophila* phosphatidylinositol-4 phosphate-binding type IV substrate SidC recruits endoplasmic reticulum vesicles to a replication-permissive vacuole. *Cell Microbiol* 10, 2416-33.
94. Brombacher E, Urwyler S, Ragaz C, Weber SS, Kami K, et al. (2009) Rab1 guanine nucleotide exchange factor SidM is a major phosphatidylinositol 4-phosphate-binding effector protein of *Legionella pneumophila*. *J Biol Chem* 284, 4846-56.
95. Jelenska J, Yao N, Vinatzer BA, Wright CM, Brodsky JL, et al. (2007) A J domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. *Curr Biol* 17: 499-508.
96. Kenny B, Jepson M (2000) Targeting of an enteropathogenic *Escherichia coli* (EPEC) effector protein to host mitochondria. *Cell Microbiol* 2: 579-590.
97. Nougayrède JP, Donnenberg MS (2004) Enteropathogenic *Escherichia coli* EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *Cell Microbiol* 6: 1097-1111.
98. Block A, Guo M, Li G, Elowsky C, Clemente TE, et al. (2010) The *Pseudomonas syringae* type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity. *Cell Microbiol* 12: 318-330.
99. Bogdanove AJ, Schornack S, Lahaye T (2010) TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol* 13: 394-401.
100. Pitzschke A, Hirt H (2010) New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. *EMBO J* 29: 1021-1032.
101. Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., Foster, L. J. & Brinkman, F. S. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26, 1608-15.
102. McDermott JE, Corrigan A, Peterson E, Oehmen C, Niemann G, et al. (2011) Computational prediction of type III and IV secreted effectors in gram-negative bacteria. *Infect Immun* 79: 23-32.
103. Vens C, Rosso MN, Danchin EG (2011) Identifying discriminative classification-based motifs in biological sequences. *Bioinformatics* 27: 1231-1238.
104. Bendtsen JD, Kiemer L, Fausbøll A, Brunak S (2005) Non-classical protein secretion in bacteria. *BMC Microbiol* 5: 58.
105. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18: 6097-6100.
106. Meyer DF, Noroy C, Moumène A, Raffaele S, Albina E, et al. (2013) Searching algorithm for type IV secretion system effectors 1.0: a tool for predicting type IV effectors and exploring their genomic context. *Nucleic Acids Res* 41: 9218-9229.