nature biotechnology

ARTICLES

The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*

Eric Duchaud[1], Christophe Rusniok[1], Lionel Frangeul[2], Carmen Buchrieser[1], Alain Givaudan[5], Séad Taourit[1], Stéphanie Bocs[6], Caroline Boursaux-Eude[2], Michael Chandler[7], Jean-François Charles[3], Elie Dassa[4], Richard Derose[8], Sylviane Derzelle[3], Georges Freyssinet[8], Sophie Gaudriault[5], Claudine Médigue[6], Anne Lanois[5], Kerrie Powell[9], Patricia Siguier[7], Rachel Vincent[5], Vincent Wingate[9], Mohamed Zouine[1], Philippe Glaser[1], Noël Boemare[5], Antoine Danchin[3] & Frank Kunst[1]

Photorhabdus luminescens **is a symbiont of nematodes and a broad-spectrum insect pathogen. The complete genome sequence of strain TT01 is 5,688,987 base pairs (bp) long and contains 4,839 predicted protein-coding genes. Strikingly, it encodes a large number of adhesins, toxins, hemolysins, proteases and lipases, and contains a wide array of antibiotic synthesizing genes. These proteins are likely to play a role in the elimination of competitors, host colonization, invasion and bioconversion of the insect cadaver, making *P. luminescens* a promising model for the study of symbiosis and host-pathogen interactions. Comparison with the genomes of related bacteria reveals the acquisition of virulence factors by extensive horizontal transfer and provides clues about the evolution of an insect pathogen. Moreover, newly identified insecticidal proteins may be effective alternatives for the control of insect pests.**

Photorhabdus luminescens is an enterobacterium that is symbiotic with soil entomopathogenic nematodes and pathogenic to a wide range of insects. *P. luminescens* promotes its own transmission among susceptible insect populations using its nematode host as vector[1]. Its life cycle comprises a symbiotic stage in the nematode's gut and a virulent stage in the insect larvae, which it kills through toxemia and septicemia. After the nematode attacks a prey insect and *P. luminescens* is released, the bacterium produces a wide variety of virulence factors ensuring rapid insect killing. Bioconversion of the insect cadaver by exoenzymes produced by the bacteria allows the bacteria to multiply and the nematode to reproduce. During this process *P. luminescens* produces antibiotics to prevent invasion of the insect cadaver by bacterial or fungal competitors. Finally, elimination of competitors allows *P. luminescens* and the nematode to reassociate specifically before leaving the insect cadaver[2,3].

To better understand this complex life style, we determined the genome sequence of *P. luminescens* subspecies *laumondii* strain TT01[4], a symbiont of the nematode *Heterorhabditis bacteriophora* isolated on Trinidad and Tobago.

RESULTS

General features

Strain TT01 possesses a single circular chromosome of 5,688,987 bp with an average GC content of 42.8%. No plasmid replicon was found. A total of 4,839 protein-coding genes, including 157 pseudogenes, seven complete sets (23S, 5S and 16S) of ribosomal RNA operons and 85 tRNA genes, were predicted (**Fig. 1**; **Supplementary Table 1** online).

Toxins against insects

More toxin genes were predicted in the *P. luminescens* genome than in any other bacterial genome sequenced yet. A large number of these toxins may be involved in the killing of a wide variety of insects. Some may act synergistically or use redundancy for 'overkill'[5], ensuring a quick death of the host. In addition, some may kill insects by interfering with their development. In the TT01 genome, two paralogs, *plu4092* and *plu4436*, encode proteins similar to juvenile hormone esterases (JHEs) of the insect *Leptinotarsa decemlineata*[6]. Juvenile hormone maintains the insect in a larval state. Its inactivation by JHE allows metamorphosis to proceed. JHEs may be used to trigger the insect endocrine machinery at an inappropriate time and thus represents a promising approach for insect control[7]. These genes are located downstream of highly related orphan genes (*plu4093* and *plu4437*), suggesting a locus duplication.

The toxicity of the proteins encoded by these two loci was verified experimentally. Two *Escherichia coli* clones, containing the recombinant BAC1A02 and BAC8C11, were shown to be toxic toward insects. BAC1A02, which contains the locus *plu4093–plu4092*, exhibited substantial oral toxicity toward three mosquito species, *Aedes aegypti,*

[1]Laboratoire de Génomique des Microorganismes Pathogènes, [2]Génopole, Plate-Forme Intégration et Analyse Génomiques, [3]Unité de Génétique des Génomes Bactériens, [4]Unité de Programmation Moléculaire et Toxicologie Génétique, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France. [5]Laboratoire EMIP, Université Montpellier II, IFR122, Institut National de la Recherche Agronomique (UMR 1133), 34095 Montpellier Cedex 5, France. [6]Atelier de Génomique Comparative, Génoscope/CNRS-UMR 8030, 2, rue Gaston Crémieux, 91006 Evry Cedex 15, France. [7]Laboratoire de Microbiologie et de Génétique Moléculaire, CNRS, 118 Route de Narbonne, 31062 Toulouse Cedex, France. [8]Bayer CropScience, 1 rue Pierre Fontaine, 91058 Evry, France. [9]Bayer CropScience, 2 T.W. Alexander Drive, Research Triangle Park, North Carolina 27709, USA. Correspondence should be addressed to F.K. (fkunst@pasteur.fr).

Published online 5 October2003; doi:10.1038/nbt886

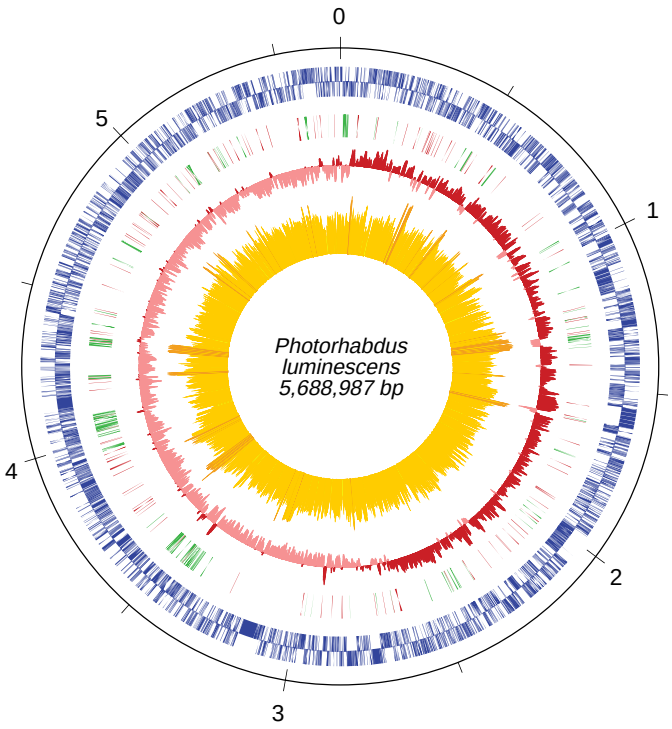

Figure 1 Circular representation of the *P. luminescens* genome. The outer scale is marked in megabases. Circle 1 and 2 (from outside to inside), genes transcribed clockwise and counterclockwise. Circle 3, transposases (red) and phage-related genes (green). Circle 4, GC bias (G – C/G + C). Circle 5, GC content with <32% G + C in light yellow, between 32% and 53.6% G + C in yellow and with >53.6% G + C in dark yellow. Shifts in GC bias correspond to positions of the predicted origin and terminus of replication or to putative prophage regions (green).

Culex pipiens and *Anopheles gambiae,* as well as toward the lepidopteran *Plutella xylostella*; BAC8C11, which contains the locus *plu4437–plu4436*, showed insecticidal activity toward *P. xylostella* (data not shown). Ingestion of undiluted recombinant *E. coli*, in which both *plu4093* and *plu4092* were expressed under the control of the P_{lac} promoter (pDIA700), led to killing of 96% of *P. xylostella* and 100% of *C. pipiens* larvae after 48 h[8] (**Table 1**). Partial deletion of *plu4092* (pDIA701) abolished toxicity for both insects, indicating that the JHE-like product of *plu4092* is required for toxicity. This is the first time that a protein lethal for mosquitoes has been found in *P. luminescens*.

Mcf, a gene previously identified in the *P. luminescens* strain W14, encodes a cytotoxin that is able to kill the insect *Manduca sexta*[5]. Strain TT01 possesses this gene as well as the paralog *plu3128*.

Four toxin-complex loci, *tca*, *tcb*, *tcc* and *tcd,* have been identified in the *P. luminescens* strain W14 (ref. 9). The *tca* locus encodes complexes with high oral toxicity for the insect *Manduca sexta*[10,11]. Strain TT01 contains the *tcc* and *tcd* loci, an incomplete *tca* locus and five newly identified *tc* loci (**Fig. 2**; **Supplementary Table 2** online). The physical linkage of the *tcdAB* and *tccC* genes was found in three other insect-associated bacteria, *Xenorhabdus nematophila, Serratia entomophila* and *Yersinia pestis*. This organization may be due to gene transfer events between bacteria sharing the same niche, namely insects.

The putative *tcc* and *rhs* gene products belong to one protein superfamily[9]. They contain repeated Tyr-Asp motifs, previously found in teneurins, transmembrane proteins of invertebrates and vertebrates[12]. An *in silico* search for the presence of Arg-Tyr × Tyr-Asp motifs (at least three, no more than one mismatch) in the TTO1 proteome identified 14 putative proteins of a superfamily including the 7 TccC paralogs, putative nematicidal proteins (Plu2222 and Plu2442) and hypothetical *rhs*-like elements as well as *Mcf* and its paralog (see **Supplementary Fig. 1** and **Supplementary Table 3** online). It has been suggested that Tyr-Asp-repeat proteins bind carbohydrates[12]. One could speculate that proteins belonging to this superfamily are exposed at the surface of the bacteria where they are important for evasion of defenses from a broad spectrum of insects.

P. luminescens contains eight genes predicted to encode hemolysin- or hemagglutinin-related proteins secreted by the two-partner secretion (TPS) pathway. These comprise the secreted PhlA protein and the PhlB protein, which allows secretion and activation of PhlA[13]. We identified six other TPS gene pairs and ten incomplete loci (**Supplementary Table 3** online).

Repeats-in-toxin (RTX) proteins constitute another family of toxins, including cytolytic toxins, metalloproteases and lipases. We identified eight genes whose deduced protein sequence is homologous to the *Vibrio cholerae* RtxA toxin (4,558 amino acids), which causes cross-linking and depolymerization of actin stress fibers in an *in vivo* model[14]. The *rtxA* homologs are clustered in two chromosomal regions and are tandemly organized. Four are complete genes and four are disrupted by frameshifts or insertion sequences (ISs; **Supplementary Table 2** online). Furthermore, the organization of the genes predicted to encode the RTX secretion system is identical to that in *V. cholerae* (**Fig. 3**). Plu4117 and Plu3668 belong to another subfamily of RTX proteins. This large set of putative RTX toxins and related proteins probably contributes to the insect pathogenicity of *P. luminescens*. In addition to the RTX family we identified a new family of putative proteins containing Gly-Asp repeats (**Supplementary Fig. 2** online). Genes encoding proteins similar to toxins from other bacteria were also predicted, including Plu0840, similar to the heat-stable cytotonic enterotoxin from *Aeromonas hydrophila*[15], and Plu1537, similar to a component of the *Bacillus thuringiensis* δ-endotoxin[16].

Toxins against competitors

One of the most important questions about the ecology of this bacterium is how *P. luminescens* defends the insect cadaver against different microbial competitors. The TT01 genome contains 33 genes, clustered in 20 loci, encoding proteins similar to polyketide and nonribosomal peptide synthases that may be part of the biosynthetic pathway of antibiotics known to be produced by *P. luminescens* (**Supplementary Table 4** online). As observed in strain W14 (ref. 17), ten genes of TT01 are similar to genes for the biosynthesis of syringomycin by *Pseudomonas syringae*, which acts as a pore-forming host cell cytolysin[18]. One gene, *plu2670*, encodes a 16,367-amino-acid putative peptide synthase that is,

Table 1 Oral larvicidal activities of *E. coli* clones on second-instar insect larvae

E. coli clones	Mortality of *P. xylostella* (%)[a]		Mortality of *C. pipiens* (%)	
	48 h	72 h	24 h	48 h
pDIA 700 (*plu4093* + *plu4092*)	96	100	30 ± 8[b]	100
pDIA 701 (*plu4093* + truncated *plu4092*)	<15	<15	0	0
Control (water)	ND	ND	0	0

[a]Percentages of larval mortality were corrected with the negative control using the Abbott equation. [b]Standard error of the mean. ND, not done.

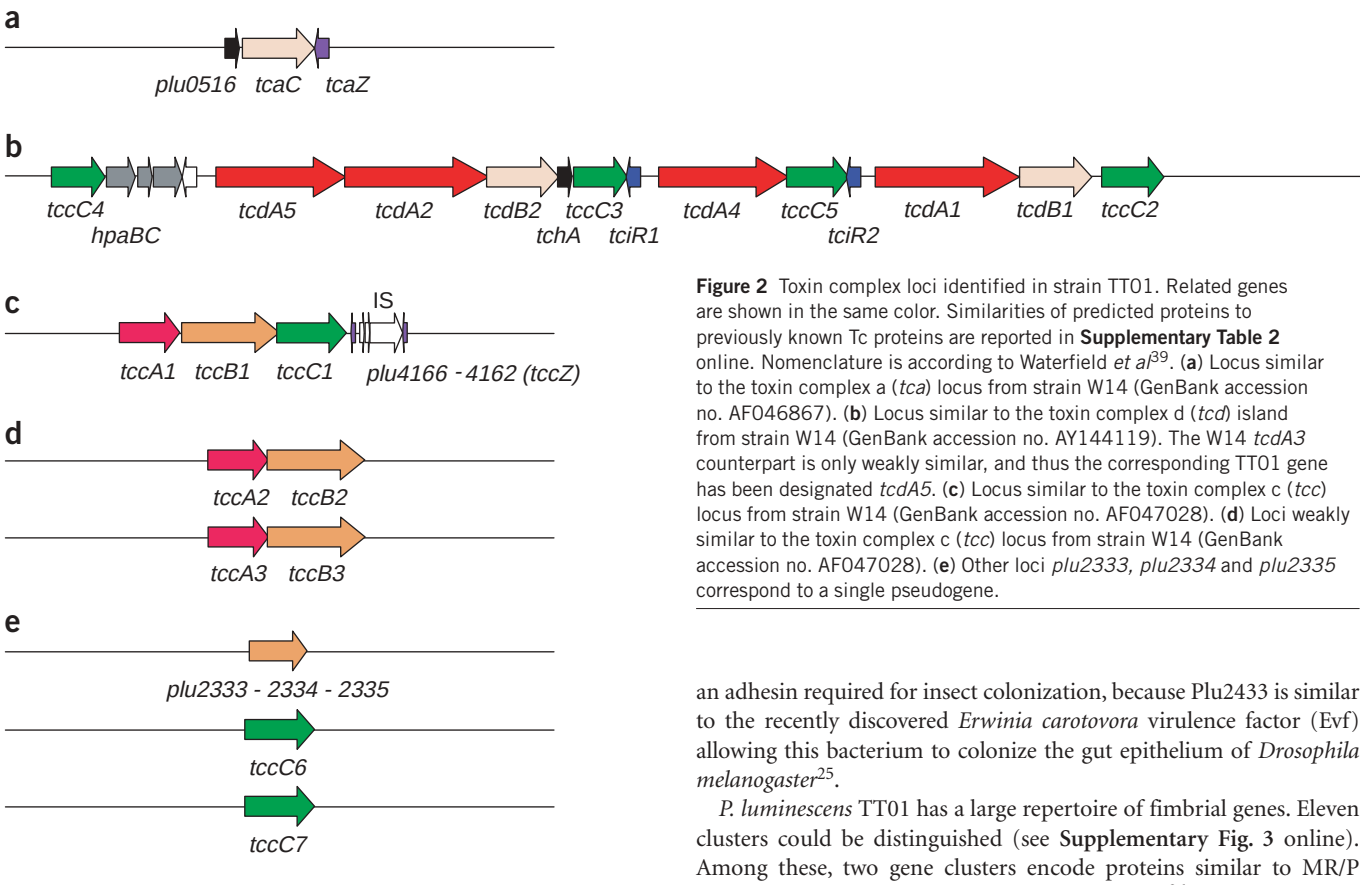

Figure 2 Toxin complex loci identified in strain TT01. Related genes are shown in the same color. Similarities of predicted proteins to previously known Tc proteins are reported in **Supplementary Table 2** online. Nomenclature is according to Waterfield *et al*[39]. (**a**) Locus similar to the toxin complex a (*tca*) locus from strain W14 (GenBank accession no. AF046867). (**b**) Locus similar to the toxin complex d (*tcd*) island from strain W14 (GenBank accession no. AY144119). The W14 *tcdA3* counterpart is only weakly similar, and thus the corresponding TT01 gene has been designated *tcdA5*. (**c**) Locus similar to the toxin complex c (*tcc*) locus from strain W14 (GenBank accession no. AF047028). (**d**) Loci weakly similar to the toxin complex c (*tcc*) locus from strain W14 (GenBank accession no. AF047028). (**e**) Other loci *plu2333, plu2334* and *plu2335* correspond to a single pseudogene.

to our knowledge, the largest protein reported for a prokaryote to date. In our work here, we frequently identified genes that were adjacent to antibiotic biosynthetic genes and that presumably encode transporters involved in antibiotic efflux. Four genes (*plu0250*, *plu2474*, *plu2475*, *plu2476*) are similar to *Burkholderia glumae* genes involved in the biosynthesis of toxoflavin. Compounds belonging to this family are active against a variety of Gram-positive bacteria and fungi[19].

Toxin-antitoxin systems are efficient in killing related bacteria, including other *P. luminescens* strains that invade the host[20]. The TT01 genome encodes three putative colicin-like factors and 17 putative immunity proteins involved in self-protection. These genes are located in three clusters, which may encode pyocin S3–like factors (*plu0884*, *plu4177*) and a colicin-like factor (*plu1894*) (**Fig. 3** and **Supplementary Table 3** online). This complex pattern of immunity-protein-coding genes may provide a selective advantage to TT01 against related bacteria producing similar toxins.

Host interactions

Another intriguing question is how *P. luminescens* interacts with the nematode midgut and the insect hemocytes. We identified a large number of genes that could be involved in adhesion. *plu2096* encodes a protein similar to the *Pseudomonas aeruginosa* lectin PA-I, which recognizes specific carbohydrates exposed on the host cells and functions as an adhesin and a cytotoxin that plays a role in the initiation of infection[21]. The putative *plu1561* gene product is similar to a Ca^{2+}-dependent adhesion molecule of the cadherin family in *Dictyostelium discoideum*[22]. We have further identified one gene, *plu2057*, encoding a protein 45% similar to *inv* and three *ail*-like paralogs (*plu2480*, *plu2481* and *plu1967*). In *Yersinia* species, Inv and Ail play a role in surface adhesion or invasion[23,24]. Another *P. luminescens* gene, *plu2433*, may encode an adhesin required for insect colonization, because Plu2433 is similar to the recently discovered *Erwinia carotovora* virulence factor (Evf) allowing this bacterium to colonize the gut epithelium of *Drosophila melanogaster*[25].

P. luminescens TT01 has a large repertoire of fimbrial genes. Eleven clusters could be distinguished (see **Supplementary Fig. 3** online). Among these, two gene clusters encode proteins similar to MR/P (mannose-resistant) fimbriae of *Proteus mirabilis*[26] as well as type IV pili of pathogenic *E. coli* and *Salmonella enterica*[27]. This large repertoire of pili may help *P. luminescens* colonize the nematode gut and invade the different insect compartments, as the *ngrA* gene (an *ngrA* homolog has previously been identified in *Photorhabdus luminescens* strain W14; see ref. 28), belonging to cluster VIII, is important for bacterial-nematode interactions[28].

Secreted proteins

P. luminescens secretes many enzymes that contribute to insect death and result in bioconversion of the insect cadaver[29]. In *X. nematophila* a zinc metalloprotease, PrtA, is involved in the immunosuppression of the insect[30]. In *P. luminescens,* the protein is encoded by the operon *prtA-inh-prtBCD* and is secreted by a type I secretory system[31]. Other putative proteases were identified, such as Plu2820, similar to a serine protease of the subtilisin family, Plu1382, similar to an extracellular metalloproteinase, and Plu2455, similar to calpain.

Lipases are another class of secreted proteins. TT01 encodes ten triacylglycerol lipase-, phospholipase A- and D-like proteins. For example, the proteins predicted to be encoded by genes *plu3370* and *plu3369* are similar to phospholipase A and its accessory protein, respectively. Gene *plu0830* encodes a similar protein that is more distantly related to phospholipase A. Both contain the lipase-specific consensus sequence G-x-S-x-G-G in their amino-terminal moiety. *Yersinia enterocolitica* phospholipase A contributes to pathogenesis in a mouse model[32], suggesting a role in virulence for the TT01 homolog. Plu1971 is remarkably similar to the *Y. pestis* plasmid (pMT1)-encoded *Yersinia* murine toxin (Ymt). It contains the two phospholipase D motifs H-x-K-x_4-D-x_6-G-G. Genes *plu1971* and *ymt* are 56% identical, with the same GC content (38%), substantially different from the 50% GC content of the pMT1 plasmid; this suggests that *ymt* has been acquired

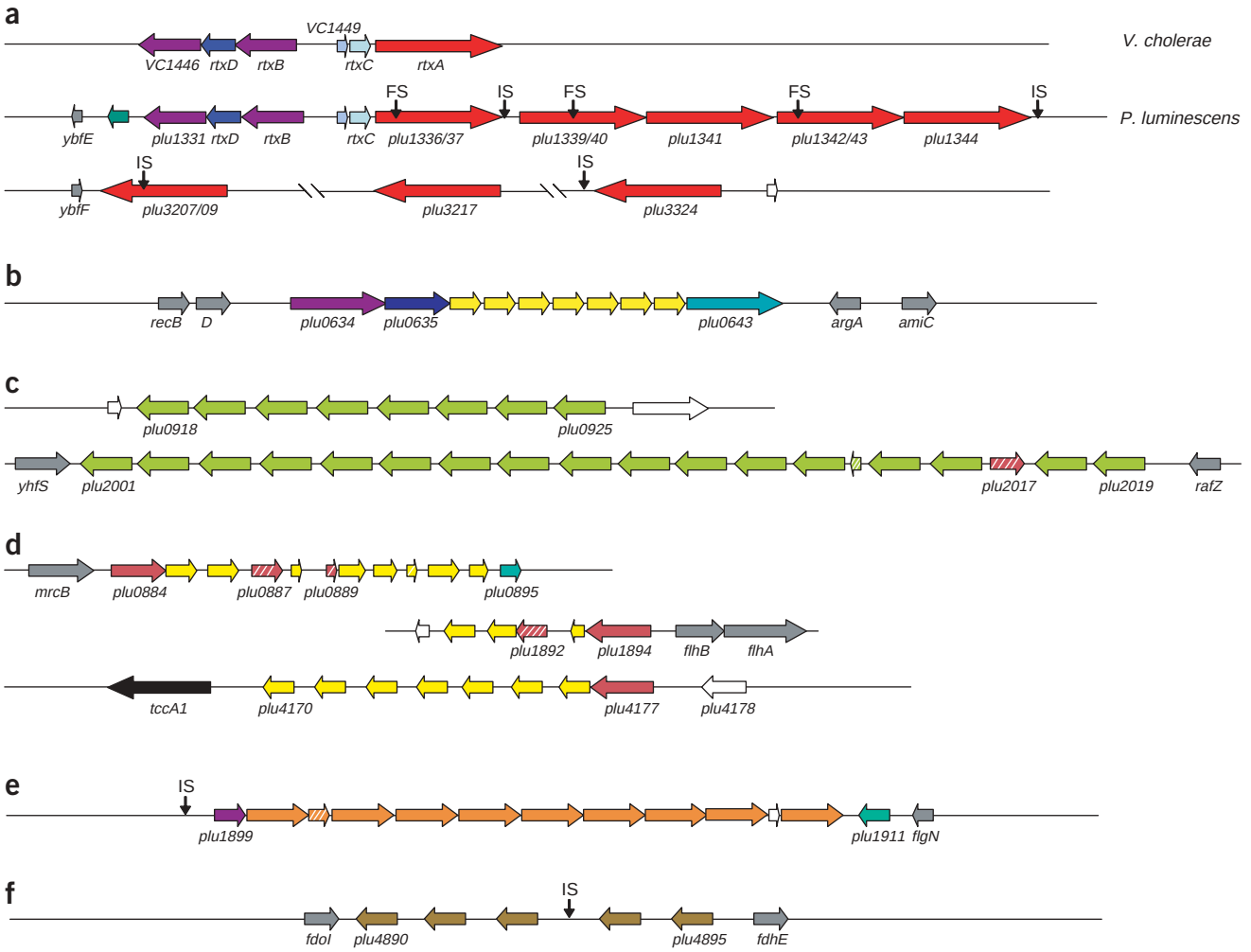

Figure 3 Schematic representation of tandemly repeated genes. Related genes are color coded: blue-purple, genes encoding putative secretion systems; gray, genes belonging to the core genome; white, genes of unknown function; dashed, pseudogenes. ↓ IS indicates insertion-sequence element. (**a**) Red, tandemly repeated *rtxA* genes and genes encoding the RTX secretory machinery as compared to *V. cholerae*. (**b**) Yellow, genes encoding unknown proteins, probably secreted by the Type I secretory system encoded by *plu0634* and *plu0635*. (**c**) Green, *luxR* gene clusters. (**d**) Red, clusters encoding putative pyocin- and colicin-like factors; yellow, immunity proteins. (**e**) Orange, highly related genes of unknown function. (**f**) Brown, putative *O*-methyltransferase-encoding genes.

by *Y. pestis* from *P. luminescens* or a close relative. Because Ymt is essential for flea colonization by *Y. pestis*[33], *P. luminescens* and *Y. pestis* may use similar genes for insect colonization.

Besides the *lip-1* gene, encoding a triacylglycerol lipase[34], we predicted four triacylglycerol lipases unrelated to Lip-1, encoded by the tandemly repeated paralogs *plu1519*, *plu1518*, *plu1517* and *plu1516*. Another unrelated lipase, Plu2266, is similar to LipP (carboxylesterase) of *Pseudomonas* sp. strain B11-1[35], and Plu2313 is highly similar to the extracellular lipase, LipA, from *Serratia marcescens*[36]. Plu2313 lacks an N-terminal signal peptide but contains instead, in its C terminus, four almost exact repeats of the glycine-rich motif L-x-G-G-x-G-x-D of RTX toxins, suggestive of a similar secretion pathway[37]. We predicted a chitin-binding-like protein (Plu2352) and chitinase-like proteins (Plu2235, Plu2458 and Plu2461) as expected for an insect pathogen.

Metabolism

Within an insect, *P. luminescens* needs to get access to the available nutrients and has to deal with low-iron conditions. The overall intermediary metabolism of *P. luminescens* is similar to that found in most *Enterobacteriaceae*, but it comprises many degradation pathways that are absent from the laboratory strain *E. coli* K12, including urease, proline or histidine degradation pathways that may help the bacteria to multiply in the hemocele of the insect larvae, while providing the bacteria and the nematode host with building blocks for their growth.

The TT01 genome contains many genes encoding proteins similar to monooxygenases, dioxygenases and hydroxylases that could be implicated in rapid elimination of insect polyphenols, such as *plu4258*, adjacent to glutathione transferase (*plu4259*), or operons such as the *hca* operon or *plu1434–plu1437*. Many gene products may be involved in the detoxification of reactive oxygen species generated by the invaded host; the bacterium may take advantage of the reactivity of dioxygen to metabolize insect products. We also identified counterparts of a putative steroid monooxygenase (Plu4232), glycine oxidase (Plu2242) and aromatic monooxygenase (Plu1313) as well as five tandemly organized paralogs (*plu4890–plu4895*), possibly encoding proteins similar to *O*-methyltransferase that may play a similar role (**Fig. 3**).

Iron acquisition seems to be of particular importance for the life cycle of *P. luminescens*, as this bacterium has the largest known set of iron, heme, hemin and siderophore transporters. Five loci have counterparts in *Y. pestis*, one of which is similar to the genes involved in biosynthesis of the siderophore yersiniabactin (**Supplementary Table 3** and **Supplementary Fig. 4** online). A sixth locus (*fecI–fecE*) is similar to the *E. coli* iron(III)-dicitrate transport system. Two putative binding proteins and six outer membrane receptors of the ISVH (iron-siderophores, vitamin B_{12} and hemin) family may be involved in iron, heme, hemin or siderophore transport[38] (**Supplementary Table 3** online).

P. luminescens encodes four complete type I, one type II and one type III secretion systems[39]. The latter one is highly similar to the plasmid-encoded type III system of *Y. pestis* and is probably involved in the secretion of virulence proteins or in immunomodulation of the insect response. The phage elements present in the genome may participate in the release of molecules involved in pathogenesis by partial lysis of the bacteria (S.G., unpublished data).

Regulation

To adapt to invertebrate environments such as the nematode gut, the insect hemolymph and the insect cadaver, *P. luminescens* needs to sense changes in such conditions as nutrient and cation availability, osmolarity and bacterial density. Strain TT01 encodes homologs of the five *E. coli* sigma factors RpoD, RpoS, RpoH, FliA and RpoN. In contrast to *E. coli*, which has only one ECF sigma factor, TT01 possesses five such factors. We predicted 192 transcriptional regulators (**Supplementary Table 5** online) from the sequence. Interestingly, the LuxR family of regulators is over-represented (32 genes) and the majority of these genes are tandemly located in two regions (**Fig. 3**). This tandem organization of the *luxR* genes may play a role in the response of TT01 to acylhomoserine lactone (AHL) signals produced by TT01 or by closely related bacteria. LuxR regulators may then, in association with AHL-responsive systems, affect diverse biological functions including bioluminescence. Although TT01 contains a complete *lux* operon responsible for the production of bioluminescence, we could identify only one gene, *plu2238*, that might be part of an AHL-responsive system. It encodes a protein highly similar (76% identity) to an *Agrobacterium tumefaciens* N-AHL-lactonase[40].

Two other highly represented classes of regulators are phage gene repressors (37) and Ner-like regulators (15), which is consistent with the large number of prophages or prophage remnants present in the genome. Bacteriophages may have affected bacterial-host interactions. It has been shown that overexpression of a Ner-like regulator switches the primary bacterial variant, which supports nematode growth, to the secondary variant, which does not[41]. Nineteen two-component regulators and 20 LysR–type regulators were identified. A LysR-type regulator, also called HexA or LrhA, was reported to control symbiosis negatively and pathogenicity positively, suggesting that it is involved in the transition from symbiosis to pathogenicity[42]. However, the relatively low number of two-component regulators identified in the *P. luminescens* genome (compared to the 89 systems of *P. aeruginosa*) might be related to the quite stable environment encountered by *P. luminescens*, obviating the need for metabolic adaptation to a wide range of conditions.

Genome comparison

The *P. luminescens* life cycle and a number of its phenotypic traits resemble those of *Y. pestis*. Both transit and colonize insects and both are pathogenic bacteria. However, *P. luminescens* has been used for biological control of insects without causing any harm to humans. Several common phenotypic characteristics are reflected in the genome. We identified 2,107 genes in *P. luminescens* that have an ortholog in *Y. pestis*. In addition, 77% (1,621 out of 2,107) of the orthologous genes of TT01 and *Y. pestis* CO92 (ref. 43) are syntenic. Interestingly, *Y. pestis* and *P. luminescens* share not only the chromosomal backbone of *Enterobacteriaceae*, which both share also with *E. coli,* but also many putative mobile regions encoding toxins, virulence factors and proteins of unknown function (**Supplementary Fig. 5** online). Orthologs of genes present on the *Y. pestis* plasmids (pCD1 and pMT1) were also identified in the genome of *P. luminescens.*

The impressive number of mobile genetic elements or their remnants suggest that the TT01 genome is subject to continuously ongoing gene transfer (**Supplementary Table 6** online). We identified phage remnants representing 4% of the genome, 195 ISs or IS fragments, putative transposons, and 711 inverted repeats of the enterobacterial repetitive intergenic consensus (ERIC) sequences[44], in contrast to only 21 ERIC sequences present in the *E. coli* K12 chromosome. In addition, several gene classes are over-represented (**Supplementary Table 3** online), suggesting frequent rearrangements (duplication, recombination) and a high degree of genome plasticity. This redundancy may contribute not only to the impressive arsenal of toxins but also to the generation of variability, which is an advantage for a pathogen subject to strong selective constraints. *P. luminescens* kills a wide variety of insects and thus has to circumvent a multitude of host defenses.

DISCUSSION

The availability and the functional analysis of the genome sequence of *P. luminescens* TT01 should lead to several useful applications, such as the development of new entomotoxins for crop protection and the genetic engineering of the bacterium-nematode pair for use as biological control agents. The identification of new antibiotic biosynthetic genes, which could be manipulated to generate new biological activities by domain shuffling, provides a promising resource for fighting microbial infections in the future. Furthermore, the availability of the first genome sequence of a symbiotic as well as entomopathogenic bacterium is an important advance in helping to decipher the relation between pathogenesis and symbiosis.

METHODS

Cloning, sequencing, assembly and annotation. Genome sequencing was performed using the whole-genome shotgun strategy[45], as described previously[46,47]. Two libraries (1–2 kb and 2–3 kb inserts) were generated by random mechanical shearing of genomic DNA and cloning into pcDNA-2.1 (Invitrogen) and a medium-size insert library (5–10 kb) was generated in the low copy number vector pSYX34[48]. The BAC library was constructed as previously described[49]. Assembly and annotation are detailed in **Supplementary Methods** online.

BAC1A02 encompassing the JHE-like toxin locus *plu4093–plu4092* extends from position 4,729,455 to position 4,789,585 of the *P. luminescens* genome; BAC8C11 encompassing the JHE-like toxins locus *plu4437–plu4436* extends from position 5,179,747 to position 5,205,606.

To construct the pDIA700 plasmid, a DNA fragment containing the two genes *plu4093* and *plu4092*, was generated by PCR with primers JHE2 (5′-AACTGCAGCATTGAAGCAGAGCGTTGACAT-3′) and JHE3 (5′-CGGGATCCGACGTCGGCAAGTGCATCAAAT-3′). The amplified DNA fragment of 2,060 bp was purified and cloned into the pBluescript SK vector (Stratagene) at the *Eco*RV site.

To inactivate *plu4092*, the pDIA700 was digested with *Eco*RI, leading to a deletion of the DNA region located between an *Eco*RI site located at the 270th codon of *plu4092* and the pBluescript *Eco*RI site, and self-ligated to yield pDIA701, which contains a 3′ truncated *plu4092* gene.

Insecticidal assays.

P. xylostella *leaf bioassay.* The recombinant *E. coli* XL1-Blue strain containing pDIA700 or pDIA701 and *E. coli* DH10B containing BAC1A02 or BAC8C11

were grown for 20 h at 28 °C in 50 ml of LB broth supplemented with 100 μg/ml ampicillin or 12.5 μg/ml chloramphenicol, respectively. For each clone, six cabbage leaves (3 cm diameter) were incubated in undiluted bacterial cultures for 1 h with 0.05% Tween 20 (Sigma). Treated leaves were put onto 6-well plates with agar beds containing 15 g/l of agar (Difco) and 30 mg/l of the fungicide nipagine (Sigma). Five second-instar larvae were placed into each well. All the assays were done at 28 °C with a photoperiod of 11 h:13 h (night/day). After 2 d the treated leaves were replaced by untreated leaves. The larval mortality was recorded at day 2 and day 3.

Recombinant *E. coli* strains containing the pBluescript SK vector without insert and *P. luminescens* TT01 cultures were used as negative and positive controls, respectively. The percentage of larval mortality was corrected with respect to the negative control by using the Abbott equation[50]. The toxicity of each clone was tested on 30 larvae, and each experiment was done in triplicate.

Mosquitocidal activity. Derivatives of *E. coli* TG1 containing plasmids pDIA700 and pDIA701, or *E. coli* DH10B containing BAC1A02 or BAC8C11, were grown at 30 °C in 50 ml LB broth supplemented with 100 μg/ml ampicillin or 12.5 μg/ml chloramphenicol, respectively, until the optical density at 600 nm (OD_{600}) reached a value of 2, then harvested by centrifugation and resuspended at the same optical density. Ten second-instar *C. pipiens* larvae were placed in Petri dishes (2.5 cm diameter) containing 5 ml bacterial suspensions at OD_{600} = 2 (or water as a control). Yeast cells were given as a food source in all dishes (to avoid mortality in the control), and larval mortality was recorded at days 1 and 2. Two independent experiments were conducted at 24 ± 2 °C, each one in duplicate. The same procedure was used for toxicity assays involving *Aedes aegypti* and *Anopheles gambiae*.

URLs. The genome sequence has been submitted to EMBL under the accession number BX470251. The complete data set of DNA and protein sequences, linked to the relevant annotations and functional assignments, is available online at http://genolist.pasteur.fr/PhotoList.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This work received financial support from the Institut Pasteur, the Centre National de la Recherche Scientifique, the Institut National de la Recherche Agronomique and the Ministère de l'Industrie et des Finances (Après séquençage des Génomes). We wish to thank Tatiana Vallaeys and Evelyne Krin for helpful discussions, Ivan Moszer and Eduardo Rocha for help with bioinformatics, Elisabeth Couvé, Christina Nielsen-Le Roux and Hafed Nedjari for expert technical assistance and Tim Stinear for critical reading of the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 7 May; accepted 18 August 2003
Published online at http://www.nature.com/naturebiotechnology/

1. Boemare, N.E., Akhurst, R.J. & Mourant, R.G. DNA relatedness between *Xenorhabdus* spp. (*Enterobacteriaceae*), symbiotic bacteria of entomopathogenic nematodes, and a proposal to transfer *Xenorhabdus* luminescens to a new genus, *Photorhabdus* gen. nov. *Int. J. Syst. Bacteriol.* **43**, 249–255 (1993).
2. Forst, S., Dowds, B., Boemare, N. & Stackebrandt, E. *Xenorhabdus* and *Photorhabdus* spp.: bugs that kill bugs. *Annu. Rev. Microbiol.* **51**, 47–72 (1997).
3. ffrench-Constant, R. *et al. Photorhabdus*: towards a functional genomic analysis of a symbiont and pathogen. *FEMS Microbiol. Rev.* **26**, 433–456 (2003).
4. Fischer-Le Saux, M., Viallard, V., Brunel, B., Normand, P. & Boemare, N.E. Polyphasic classification of the genus *Photorhabdus* and proposal of new taxa: *P. luminescens* subsp. *luminescens* subsp. nov., *P. luminescens* subsp. *akhurstii* subsp. nov., *P. luminescens* subsp. *laumondii* subsp. nov., *P. temperata* sp. nov., *P. temperata* subsp. *temperata* subsp. nov. and *P. asymbiotica* sp. nov. *Int. J. Syst. Bacteriol.* **49**, 1645–1656 (1999).
5. Daborn, P.J. *et al.* A single *Photorhabdus* gene, makes caterpillars floppy (*mcf*), allows *Escherichia coli* to persist within and kill insects. *Proc. Natl. Acad. Sci. USA* **99**, 10742–10747 (2002).
6. Vermunt, A.M., Koopmanschap, A.B., Vlak, J.M. & de Kort, C.A. Evidence for two juvenile hormone esterase-related genes in the Colorado potato beetle. *Insect Mol. Biol.* **7**, 327–336 (1998).
7. Bonning, B.C. & Hammock, B.D. Development of recombinant baculoviruses for insect control. *Annu. Rev. Entomol.* **41**, 191–210 (1996).
8. patent PCT-FR/0301238 Duchaud, E. *et al.* Insecticidal proteins of *Photorhabdus luminescens.* International patent PCT-FR03/01238. April 17, 2002.
9. Waterfield, N.R., Bowen, D.J., Fetherston, J.D., Perry, R.D. & ffrench-Constant, R.H. The *tc* genes of *Photorhabdus:* a growing family. *Trends Microbiol.* **9**, 185–191 (2001).
10. Bowen, D. *et al.* Insecticidal toxins from the bacterium *Photorhabdus luminescens. Science* **280**, 2129–2132 (1998).
11. Marokhazi, J. *et al.* Using a DNA microarray to investigate the distribution of insect virulence factors in strains of *Photorhabdus. J. Bacteriol.* **185**, 4648–4656 (2003).
12. Minet, A.D. & Chiquet-Ehrismann, R. Phylogenetic analysis of teneurin genes and comparison to the rearrangement hot spot elements of *E. coli. Gene* **257**, 87–97 (2000).
13. Brillard, J., Duchaud, E., Boemare, N., Kunst, F. & Givaudan, A. The PhlA hemolysin from the entomopathogenic bacterium *Photorhabdus luminescens* belongs to the two-partner secretion family of hemolysins. *J. Bacteriol.* **184**, 3871–3878 (2002).
14. Fullner, K.J. & Mekalanos, J.J. *In vivo* covalent cross-linking of cellular actin by the *Vibrio cholerae* RTX toxin. *EMBO J.* **19**, 5315–5323 (2000).
15. Sha, J., Kozlova, E.V. & Chopra, A.K. Role of various enterotoxins in *Aeromonas hydrophila*-induced gastroenteritis: generation of enterotoxin gene-deficient mutants and evaluation of their enterotoxic activity. *Infect. Immun.* **70**, 1924–1935 (2002).
16. Moellenbeck, D.J. *et al.* Insecticidal proteins from *Bacillus thuringiensis* protect corn from corn rootworms. *Nat. Biotechnol.* **19**, 668–672 (2001).
17. ffrench-Constant, R.H. *et al.* A genomic sample sequence of the entomopathogenic bacterium *Photorhabdus luminescens* W14: potential implications for virulence. *Appl. Environ. Microbiol.* **66**, 3310–3329 (2000).
18. Bender, C.L., Alarcon-Chaidez, F. & Gross, D.C. *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol. Mol. Biol. Rev.* **63**, 266–292 (1999).
19. Nagamatsu, T. *et al.* Syntheses of 3-substituted 1-methyl-6-phenylpyrimido[5,4-e]-1,2,4-triazine-5,7(1H,6H)-diones (6-phenyl analogs of toxoflavin) and their 4-oxides, and evaluation of antimicrobial activity of toxoflavins and their analogs. *Chem. Pharm. Bull. (Tokyo)* **41**, 362–368 (1993).
20. Sharma, S. *et al.* The lumicins: novel bacteriocins from *Photorhabdus luminescens* with similarity to the uropathogenic-specific protein (USP) from uropathogenic *Escherichia coli. FEMS Microbiol. Lett.* **214**, 241 (2002).
21. Avichezer, D., Katcoff, D.J., Garber, N.C. & Gilboa-Garber, N. Analysis of the amino acid sequence of the *Pseudomonas aeruginosa* galactophilic PA-I lectin. *J. Biol. Chem.* **267**, 23023–23027 (1992).
22. Wong, E.F., Brar, S.K., Sesaki, H., Yang, C. & Siu, C.H. Molecular cloning and characterization of DdCAD-1, a Ca^{2+}-dependent cell-cell adhesion molecule, in *Dictyostelium discoideum. J. Biol. Chem.* **271**, 16399–16408 (1996).
23. Gustavsson, A. *et al.* Role of the β1-integrin cytoplasmic tail in mediating invasin-promoted internalization of *Yersinia. J. Cell. Sci.* **115**, 2669–2678 (2002).
24. Miller, V.L., Beer, K.B., Heusipp, G., Young, B.M. & Wachtel, M.R. Identification of regions of Ail required for the invasion and serum resistance phenotypes. *Mol. Microbiol.* **41**, 1053–1062 (2001).
25. Basset, A., Tzou, P., Lemaitre, B. & Boccard, F. A single gene that promotes interaction of a phytopathogenic bacterium with its insect vector, *Drosophila melanogaster. EMBO Rep.* **4**, 205–209 (2003).
26. Zhao, H., Li, X., Johnson, D.E., Blomfield, I. & Mobley, H.L. *In vivo* phase variation of MR/P fimbrial gene expression in *Proteus mirabilis* infecting the urinary tract. *Mol. Microbiol.* **23**, 1009–1019 (1997).
27. Srimanote, P., Paton, A.W. & Paton, J.C. Characterization of a novel type IV pilus locus encoded on the large plasmid of locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect. Immun.* **70**, 3094–3100 (2002).
28. Ciche, T.A., Bintrim, S.B., Horswill, A.R. & Ensign, J.C. A Phosphopantetheinyl transferase homolog is essential for *Photorhabdus luminescens* to support growth and reproduction of the entomopathogenic nematode *Heterorhabditis bacteriophora. J. Bacteriol.* **183**, 3117–3126 (2001).
29. Bowen, D., Blackburn, M., Rocheleau, T., Grutzmacher, C. & ffrench-Constant, R.H. Secreted proteases from *Photorhabdus luminescens:* separation of the extracellular proteases from the insecticidal Tc toxin complexes. *Insect. Biochem. Mol. Biol.* **30**, 69–74 (2000).
30. Caldas, C., Cherqui, A., Pereira, A. & Simoes, N. Purification and characterization of an extracellular protease from *Xenorhabdus nematophila* involved in insect immuno-suppression. *Appl. Environ. Microbiol.* **68**, 1297–1304 (2002).
31. Valens, M., Broutelle, A.C., Lefebvre, M. & Blight, M.A. A zinc metalloprotease inhibitor, Inh, from the insect pathogen *Photorhabdus luminescens. Microbiology* **148**, 2427–2437 (2002).
32. Schmiel, D.H., Wagar, E., Karamanou, L., Weeks, D. & Miller, V.L. Phospholipase A of *Yersinia enterocolitica* contributes to pathogenesis in a mouse model. *Infect. Immun.* **66**, 3941–3951 (1998).
33. Hinnebusch, B.J. *et al.* Role of *Yersinia* murine toxin in survival of *Yersinia pestis* in the midgut of the flea vector. *Science,* **296**, 733–735 (2002).
34. Thaler, J.O., Duvic, B., Givaudan, A. & Boemare, N. Isolation and entomotoxic properties of the *Xenorhabdus nematophilus* F1 lecithinase. *Appl. Environ. Microbiol.* **64**, 2367–2373 (1998).
35. Choo, D.W., Kurihara, T., Suzuki, T., Soda, K. & Esaki, N. A cold-adapted lipase of an Alaskan psychrotroph, *Pseudomonas* sp. strain B11-1: gene cloning and enzyme purification and characterization. *Appl. Environ. Microbiol.* **64**, 486–491 (1998).
36. Akatsuka, H. *et al.* The *lipA* gene of *Serratia marcescens* which encodes an extracellular lipase having no N-terminal signal peptide. *J. Bacteriol.* **176**, 1949–1956 (1994).
37. Li, X., Tetling, S., Winkler, U.K., Jaeger, K.E. & Benedik, M.J. Gene cloning, sequence analysis, purification, and secretion by *Escherichia coli* of an extracellular

lipase from *Serratia marcescens*. *Appl. Environ. Microbiol.* **61**, 2674–2680 (1995).
38. Dassa, E. & Bouige, P. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Res. Microbiol.* **152**, 211–229 (2001).
39. Waterfield, N.R., Daborn, P.J. & ffrench-Constant, R.H. Genomic islands in *Photorhabdus*. *Trends Microbiol.* **10**, 541–545 (2002).
40. Zhang, H.B., Wang, L.H. & Zhang, L.H. Genetic control of quorum-sensing signal turnover in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. USA* **99**, 4638–4643 (2002).
41. O'Neill, K.H., Roche, D.M., Clarke, D.J. & Dowds, B.C. The *ner* gene of *Photorhabdus:* effects on primary-form-specific phenotypes and outer membrane protein composition. *J. Bacteriol.* **184**, 3096–3105 (2002).
42. Joyce, S.A. & Clarke, D.J. A *hexA* homologue from *Photorhabdus* regulates pathogenicity, symbiosis and phenotypic variation. *Mol. Microbiol.* **47**, 1445–1457 (2003).
43. Parkhill, J. *et al*. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* **413**, 523–527 (2001).
44. Hulton, C.S.J., Higgins, C.F. & Sharp, P.M. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli, Salmonella typhimurium* and other enterobacteria. *Mol. Microbiol.* **5**, 825–834 (1991).
45. Fleischmann, R.D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
46. Frangeul, L. *et al*. Cloning and assembly strategies in microbial genome projects. *Microbiology* **145**, 2625–2634 (1999).
47. Frangeul, L. *et al*. CAAT-Box, Contigs-Assembly and Annotation tool-box for genome sequencing projects. *Bioinformatics*, in the press.
48. Xu, S.Y. & Fomenkov, A. Construction of pSC101 derivatives with Cam^r and Tet^r for selection or LacZ' for blue/white screening. *Biotechniques* **17**, 57 (1994).
49. Buchrieser, C. *et al*. The 102-kilobase pgm locus of *Yersinia pestis*: sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect. Immun.* **67**, 4851–4861 (1999).
50. Abbott, W.S. A method for computing the effectiveness of an insecticide. *J. Econ. Entomol.* **18**, 265–267 (1925).