Defensive Bacteriome Symbiont with a Drastically Reduced Genome

Atsushi Nakabachi,1,2* Reiko Ueoka,3,4 Kenshiro Oshima,6 Roberta Teta,7 Alfonso Mangoni,7 Mihaela Gurgui,5 Neil J. Oldham,8 Gerhild van Echten-Deckert,9 Keiko Okamura,1 Kohei Yamamoto,1 Hiromitsu Inoue,9 Moriya Ohkuma,2 Yuichi Hongoh,10 Shin-ya Miyagishima,11 Masahira Hattori,6 Jörn Piel,3,4 and Takeka Fukatsu12

1Electronics-Inspired Interdisciplinary Research Institute (EIIRIS), Toyohashi University of Technology, Toyohashi 441-8580, Japan
2Japan Collection of Microorganisms, RIKEN BioResource Center, Tsukuba 305-0074, Japan
3Institute of Microbiology, Eidgenössische Technische Hochschule (ETH) Zurich, 8093 Zurich, Switzerland
4Kekulé Institute of Organic Chemistry and Biochemistry, Graduate School of Frontier Sciences, University of Tokyo, Tokyo 152-8550, Japan
5Kekulé Institute of Organic Chemistry and Biochemistry, Graduate School of Frontier Sciences, University of Tokyo, Tokyo 152-8550, Japan
6School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK
7Citrus Research Division, National Institute of Fruit Tree Science, Kuchinozou 859-2501, Japan
8Graduate School of Frontier Sciences, University of Tokyo, Kashiwa 277-8561, Japan
9School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK
10Institute of Microbiology, Eidgenössische Technische Hochschule (ETH) Zurich, 8093 Zurich, Switzerland
11Center for Frontier Research, National Institute of Genetics, Mishima 411-8540, Japan
12Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8566, Japan

Summary

Diverse insect species harbor symbiotic bacteria, which play important roles such as provisioning nutrients and providing defense against natural enemies [1–6]. Whereas nutritional symbioses are often indispensable for both partners, defensive symbioses tend to be of a facultative nature [1–12]. The Asian citrus psyllid Diaphorina citri is a notorious agricultural pest that transmits Liberibacter spp. (Alphaproteobacteria), causing the devastating citrus greening disease or Huanglongbing [13, 14]. In a symbiotic organ called the bacteriome, D. citri harbors two distinct intracellular symbionts: a putative nutrition provider, Carsonella DC (Gammaproteobacteria), and an unnamed betaproteobacterium with unknown function [15], for which we propose the name “Candidatus Profftella armatura.” Here we report that Profftella is a defensive symbiont presumably of an obligate nature with an extremely streamlined genome. The genomes of Profftella and Carsonella DC were drastically reduced to 464,857 bp and 174,014 bp, respectively, suggesting their ancient and mutually indispensable association with the host. Strikingly, 15% of the small Profftella genome encoded horizontally acquired genes for synthesizing a novel polyketide toxin. The toxin was extracted, pharmacologically and structurally characterized, and designated diaphorin. The presence of Profftella and its diaphorin-biosynthetic genes was perfectly conserved in the world’s D. citri populations.

Results and Discussion

Proposal of a Candidate Name

On account of the distinct genomic and microbiological traits reported in this study, we propose the designation “Candidatus Profftella armutura” for the betaproteobacterial symbiont in the syncytial bacteriome of D. citri. The generic name refers to the German scientist Joachim Profft, who provided the first comprehensive histological description of psyllid-microbe symbiotic associations [16]. The specific name indicates the defensive property of this bacterium (armatura is armament, the feminine form).

Genome Sequencing of the Bacteriome Symbionts of D. citri, Carsonella DC, and Profftella

We dissected the bacteriomies from an isofemale strain of D. citri (Figure 1A), constructed shotgun libraries, and determined the complete symbiont genomes. Of 33,024 Sanger reads in total, 2,054 and 16,799 were assembled into a 5,458 bp circular plasmid (7.6-fold coverage) and 459,399 bp (26-fold coverage) circular bacterial chromosomes, respectively (Figure 2 and Table S1 available online). Judging from the 16S ribosomal RNA (rRNA) gene sequences encoded on the chromosomes, the former represented Carsonella DC, while the latter was attributed to Profftella. In situ hybridization targeting 16S rRNAs confirmed the previously reported in vivo localization of the symbionts: Carsonella DC is found in uninucleate bacteriocytes on the surface of the bacteriome, while Profftella is located in syncytial cytoplasm at the center of the bacteriome (Figure 1B) [15]. In addition, 59 reads were assembled into a 5,458 bp circular plasmid (7.6-fold coverage) (Figure 2), whose in situ localization agreed with Profftella (Figure 1C). Transmission electron microscopy identified Carsonella DC and Profftella as large and pleomorphic bacterial cells with distinct morphological traits (Figure 1D). The gene repertoire of Carsonella DC demonstrated its nutritional role, as reported in other Carsonella lineages [17, 18]. More-detailed features of the genome of Carsonella DC are provided in the Supplemental Results and Table S1 and Table S2.

Drastic Genome Reduction in Profftella Comparable to Obligate Nutritional Symbionts

In general, obligate nutritional symbionts are characterized by features such as perfect infection in host populations, specific localization to the host symbiotic organ, host-symbiont coexistence reflecting strictly vertical symbiont transmission over evolutionary time, and drastic genome reduction down to less than 1 Mb in size [1, 19, 20]. On the other hand, facultative symbionts are typically characterized by imperfect infection frequencies in host populations, systemic infection in various...
cosymbiont genomes of other psyllids, Ctenarytaina eucalypti and Heteropsylla cubana [18], the Profftella genome contained no genes for synthesis of the essential amino acids tryptophan and histidine, which are present in the Carsonella_DC genome but deficient in the other Carsonella genomes [17, 18] (see the Supplemental Results and Figure S1B). These results highlight a complementary aspect in the Profftella-Carsonella_DC cosymbiosis and also illustrate a relatively limited nutritional capacity of Profftella.

Large Portions of the Profftella Genome Encode Genes for Polyketide Biosynthesis

Notably, as much as 15.0% (69,678 bp/464,857 bp) of the highly reduced genome of Profftella was devoted to 20 genes constituting the polyketide synthase (PKS) biosynthetic gene clusters (Figure 2 and Table S2). The PKS system of Profftella is separated into two loci on the genome (Figure 2) and exhibited remarkable similarities to PKS clusters responsible for the biosynthesis of pederin, onnaminde, and psymberin (Figure 3), which are cytotoxic metabolites produced by symbiotic bacteria of beetles and sponges [7, 26–28]. Since an especially close resemblance was found with the pederin PKS (ped), where Ped proteins were 42%–78% identical to their orthologs in Profftella, we designated the PKS loci of Profftella as dip clusters (after Diaphorina pederin-like polyketide). Pederin is a defensive polyketide accumulated in the body fluid of Paederus rove beetles for deterring predators [7, 26, 29]. This toxin is produced by an uncultured Pseudomonas symbiont that resides in up to 90% of females but not in males [26, 29, 30]. The dip and ped systems differed in a divergent architecture of the multidomain PKS gene dipO, which resembled its ortholog pedH only in the upstream half (Figures 3A and 3B), and by the absence of orthologs of the methyltransferase genes pedA and pedO [31] in the Profftella genome (Figure 3A and Table S2).

Prediction of the Polyketide Synthetic Pathway of Profftella

The dip PKS was classified into the trans-acyltransferase (trans-AT) subgroup of PKSs [32, 33]. These PKSs are large multimodular proteins that assemble polyketide chains from starter units and several elongation units. A module contains the core chain-elongating ketosynthase (KS) domain and an acyl carrier protein (ACP) domain that serves as an anchor for the growing polyketide. In addition, various optional domains can be present to catalyze a wide range of redox and other reactions. The phylogeny of the KS domains and the overall domain architecture typically mirror the structure of the synthesized polyketide and can therefore be used to predict natural products from PKS sequences [32, 33]. On the grounds that DipP and DipT are virtually identical to Pedl (42% identical to DipP) and PedF (44% identical to DipT), respectively (Figures 3A and 3B and Table S2), it was expected that the dip biosynthetic product should contain at least an entire pederin core structure (Figures 3B and 3C). In the case of pederin, the polyketide portion attributable to PedH is not found in the natural product purified from the insect, presumably due to a chain cleavage catalyzed by the oxygenase PedG [26]. Since the dip system also encodes a close ortholog of this oxygenase, DipN (78% identical to PedG) (Figures 3A and 3B and Table S2), we expected that the dip biosynthetic product also lacks the corresponding polyketide portion and thus the PKS DipO (45% identical to the N-terminal half of PedH) does not contribute to the metabolic prediction (Figure 3B). In addition, considering the lack of two O-methyltransferases...
orthologous to PedA and PedO in the dip system, we hypothesized that the dip biosynthetic product of Profftella may exhibit a structure very similar to pederin but lacking at least two O-methyl groups (Figure 3B).

**Profftella Produces Diaphorin, an Analog of Pederin**

To search for the predicted compound, we extracted over 1,000 adult individuals of *D. citri* and analyzed them by liquid chromatography electrospray ionization mass spectrometry using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) (Figures S2 A and S2B). A major signal at m/z 484.2514 was detected, consistent with an ion of formula [C_{22}H_{39}O_{9}N+Na]^+ (theoretical m/z 484.2517, Δm = 0.62 ppm), which corresponds to a tridesmethylpe-derin. Comparison of the electron-induced dissociation tandem mass spectrometry spectra of the *D. citri* polyketide and pederin revealed similar fragmentation patterns, corroborating their common structural backbone (Figures S2C–S2J). We designated this pederin analog as “diaphorin,” after the generic name of the psyllid *Diaphorina*.

**Structural Characterization of Diaphorin**

For unequivocal structural characterization, diaphorin was purified by reversed-phase high performance liquid chromatography (HPLC), yielding 1.3 ± 0.3 mg from the entire extract. The planar structure (Figure 3D) was solved through the interpretation of one-dimensional and two-dimensional nuclear magnetic resonance (see the Supplemental Results, Figures S2K and S2L, and Table S4), which confirmed diaphorin as a novel pederin congener lacking three O-methyl groups. The relative configuration of the two tetrahydropyran rings was determined by rotating Overhauser enhancement and exchange spectroscopy (ROESY) analysis and was shown to be the same as in pederin. Thus, diaphorin exhibits remarkable structural similarity to pederin and represents the first member of this compound family with a nonalkylated hemiaminal moiety.

**Horizontally Acquired PKS Genes in Profftella**

Molecular phylogenetic analyses revealed that many, if not all, dip genes of *Profftella* are monophyletic with corresponding ped genes of the *Paederus* symbiont with high levels of statistical support (Figure S3). The disparate phylogenetic affiliations of PKS-encoding bacteria (betaproteobacterial *Profftella*, gammaproteobacterial *Paederus* symbiont, and various distantly related free-living bacteria; see Figure S3) strongly suggest that the PKS genes have been horizontally transferred across the phylogenetically and ecologically divergent bacterial lineages. The horizontal transmission route and the evolutionary relationship between dip genes and ped genes are elusive, but on the basis that *Paederus* rove beetles often feed on hemipteran insects [34], it is conceivable, although speculative, that prey-predator relationship might be involved in the transfer process. The timing of horizontal gene transfer (HGT) is also currently uncertain. The GC skew values on the *Profftella* genome appeared to shift at the boundaries franking the dip cluster regions (Figure 2), which could hint at a relatively recent acquisition [35]. However, neither G + C content [36] (23.0% average G + C in dip genes versus 25.8% average G + C in other protein-coding genes; Figure 2) nor codon adaptation index (CAI) [37] of dip genes (average = 0.792) showed statistically significant differences from the rest of the genome (p > 0.05, Kolmogorov-Smirnov test).

**Diaphorin Has Significant Cytotoxicity**

Purified diaphorin was examined for its cytotoxicity against rat neuroblastoma B104 cells and human HeLa cells (Figure 4). Cell viability assays with a range of diaphorin concentrations revealed IC_{50} values of ~1 μM for B104 cells and ~2 μM for HeLa cells, respectively. In contrast to *Paederus* beetles, wherein only up to 90% of female insects and no male insects...
Figure 3. The dip Genes, Proposed Biosynthetic Pathway, and Relationship to Other Pederin-Type Metabolites

(A) Map of the dip PKS genes with that of ped genes for comparison. The dots denote regions that separate PKS loci.

(B) Architecture of the Dip and Ped PKS proteins and predicted biosynthetic pathway for diaphorin.

(C) The structures of psymberin, pederin, and onnamide A for comparison.
also subjected to diagnostic PCR of Profftella’s PKS genes, dipP, dipO, and dipT. All of the PKS genes exhibited 100% detection rates in all the populations examined (Table S5). No sequence variation was identified in any gene amplicons derived from nine populations. Hence, not only is Profftella infection highly conserved across the world’s populations of D. citri, but so is its genetic capacity for synthesizing diaphorin. These findings suggest crucial roles of the symbionts and their secondary metabolite for the insect and present a sharp contrast to previously known defensive symbionts of other insects that exhibit imperfect infection frequencies in host populations [7–12, 26, 29, 30], while some marine invertebrates were reported to consistently harbor defensive symbionts that are specific to the host species or sibling species [38–40].

**Profftella as a Defensive Symbiont with a Drastically Reduced Genome**

In summary, (1) the genome size of Profftella is only 460 kb, which is comparable to that of obligate nutritional symbionts, (2) as much as 15% of the reduced Profftella genome is dedicated to defense-related genes that constitute the PKS system for synthesizing the bioactive polyketide compound, diaphorin, (3) the novel pederin-like polyketide, diaphorin, is present in D. citri, (4) diaphorin has cytotoxic activity, (5) the titer of diaphorin in D. citri is sufficient to exert a significant cytotoxicity, and (6) the presence of Profftella and its PKS genes is highly conserved across the world’s populations of D. citri. These results indicate that Profftella is a previously unknown type of defensive symbiont.

**Evolutionary History of Profftella**

The 640 kb genome of Profftella metabolically complements the ancient nutritional cosymbiont and also encodes a defensive toxin. How such a compact dual-functioning symbiont genome has evolved is of interest. One plausible scenario is that an ancestral lineage of bacteriome nutritional symbiont leading to Profftella acquired the PKS genes via HGT and began to play a defensive role in addition to its pre-existing metabolic function. Previous studies on bacteriome symbionts have highlighted the evolutionary trends of massive and irreversible genome reduction, with the lack of novel gene acquisitions through HGT, which is a reflection of the fact that they are confined in the host symbiotic organ and sequestered from external microbial populations [1, 19, 20]. Thus, the Profftella genome may exemplify an unprecedented case in which a bacteriome symbiont acquired foreign functional genes via HGT. Another plausible scenario is that ancestral Profftella had acquired the dip genes before or just after it started the symbiotic association with the host. In this context, a recently reported symbiont of the marine tunicate Lissoclinum patella is notable [41]. This alphaproteobacterial symbiont, Candidatus Endolithosclinion faulkneri, has a 1.48 Mb genome that contains a 86 kb (5.8% of the genome) region encoding toxin-producing trans-AT PKS pathway genes presumably acquired through ancient HGT. While the Endolithosclinion genome is considerably larger than the Profftella genome, its coding
density is as low as 57%, suggesting that *Endolissocinulum* is at an early stage of reductive genome evolution. On the other hand, *Profftella* is likely at a more advanced stage of reductive genome evolution, in which degenerative DNA regions have been purged whereas the toxin-producing PKS genes have been conserved. Genomic analyses of currently unexplored symbionts closely related to *Profftella* would provide a clue to understanding the evolutionary process leading to this unique degenerative defensive symbiont genome.

**Accession Numbers**

The sequences reported in this paper have been deposited in the GenBank database under accession numbers CP003467 (Carsonella DC chromosome), CP003468 (*Profftella* chromosome), and CP003469 (*Profftella* plasmid).

**Supplemental Information**

Supplemental Information includes Supplemental Results, Supplemental Experimental Procedures, three figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.06.027.

**Acknowledgments**

We thank K. Furuya, C. Shindo, H. Inaba, E. Iioka, M. Takayama, E. Omori, M. Kiuchi, and Y. Hattori for technical support. We thank Wayne Hunter, Siti Subandiyah, Rie Ukuda, Chuan-Chan Wang, and Chun-Lin Li for providing supplemental information. The authors acknowledge funding from the global COE project entitled "Genome Information Big Bang" to M.H. and K.O. We also acknowledge funding from the global COE project entitled "Genome Information Big Bang" to M.H. and K.O.

Received: April 12, 2013

Revised: May 31, 2013

Accepted: June 12, 2013

Published: July 11, 2013

**References**


