Cultivation-Assisted Genome of Candidatus Fukatsuia symbiotica; the Enigmatic ”X-Type” Symbiont of Aphids
Vilas Patel, Germain Chevignon, Alejandro Manzano-Marin, Jayce W. Brandt, Michael R. Strand, Jacob A. Russell, Kerry M. Oliver

To cite this version:
Vilas Patel, Germain Chevignon, Alejandro Manzano-Marin, Jayce W. Brandt, Michael R. Strand, et al.. Cultivation-Assisted Genome of Candidatus Fukatsuia symbiotica; the Enigmatic ”X-Type” Symbiont of Aphids. Genome Biology and Evolution, Society for Molecular Biology and Evolution, 2019, 11 (12), pp.3510-3522. 10.1093/gbe/evz252. hal-02623378

HAL Id: hal-02623378
https://hal.inrae.fr/hal-02623378
Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Cultivation-Assisted Genome of Candidatus Fukatsuia symbiotica; the Enigmatic “X-Type” Symbiont of Aphids

Vilas Patel1,†, Germain Chevignon2,†, Alejandro Manzano-Marín3, Jayce W. Brandt1, Michael R. Strand1, Jacob A. Russell4, and Kerry M. Oliver 1,*

1Department of Entomology, University of Georgia
2Montpellier Cell Biology Research Center, Montpellier, France
3CBGP, INRA, CIRAD, IRD, Montpellier SupAgro, University of Montpellier, France
4Department of Biology, Drexel University

†These authors contributed equally to this work.
*Corresponding author: E-mail: kmoliver@uga.edu.

Accepted: November 13, 2019

Data deposition: This project has been deposited at GenBank under the accessions CP021659-CP021662.

Abstract

Heritable symbionts are common in terrestrial arthropods and often provide beneficial services to hosts. Unlike obligate, nutritional symbionts that largely persist under strict host control within specialized host cells, heritable facultative symbionts exhibit large variation in within-host lifestyles and services rendered with many retaining the capacity to transition among roles. One enigmatic symbiont, Candidatus Fukatsuia symbiotica, frequently infects aphids with reported roles ranging from pathogen, defensive symbiont, mutualism exploiter, and nutritional co-obligate symbiont. Here, we used an in vitro culture-assisted protocol to sequence the genome of a facultative strain of Fukatsuia from pea aphids (Acyrthosiphon pisum). Phylogenetic and genomic comparisons indicate that Fukatsuia is an aerobic heterotroph, which together with Regiella insecticola and Hamiltonella defensa form a clade of heritable facultative symbionts within the Yersiniaceae (Enterobacteriales). These three heritable facultative symbionts largely share overlapping inventories of genes associated with housekeeping functions, metabolism, and nutrient acquisition, while varying in complements of mobile DNA. One unusual feature of Fukatsuia is its strong tendency to occur as a coinfection with H. defensa. However, the overall similarity of gene inventories among aphid heritable facultative symbionts suggests that metabolic complementarity is not the basis for coinfection, unless playing out on a H. defensa strain-specific basis. We also compared the pea aphid Fukatsuia with a strain from the aphid Cinara confinis (Lachninae) where it is reported to have transitioned to co-obligate status to support decaying Buchnera function. Overall, the two genomes are very similar with no clear genomic signatures consistent with such a transition, which suggests co-obligate status in C. confinis was a recent event.

Key words: heritable symbiont, bacteria, comparative genomics, insect, evolutionary transition.

Introduction

Insects frequently harbor maternally transmitted symbionts that facilitate host resource acquisition, impact dietary breadth, or contribute defensive services (Moran et al. 2008; Oliver and Martinez 2014). Many associations are ancient (>100 Ma), and by augmenting the nutritional profiles of feeding substrates they have allowed insects to occupy and radiate on a variety of feeding niches, including plant phloem and xylem (Moran and Telang 1998; Baumann 2005; Wilson and Duncan 2015). These “obligate” symbionts are often sequestered in specific host cells called bacteriocytes, and exhibit very small (0.1–1 Mb), static genomes that contain few if any mobile genetic elements (McCutcheon and Moran 2012). More widespread are heritable facultative symbionts, which vary widely in characterized roles across a spectrum from parasitic to conditional mutualist. Heritable facultative symbionts show large variation in age of host restriction, tissue tropism, and distributions within and among host species (Duron et al. 2008; Oliver et al. 2010; Feldhaar 2011; Guo et al. 2017).
Some heritable facultative symbiont-containing lineages, including *Serratia* and *Sodalis*, also contain species of “free-living” bacteria that exist in soil, on plants or are known community members in the guts of different insects (Lo et al. 2016). Examples have also been identified where species have transitioned from facultative to obligate status by substituting for decaying functions in obligate nutritional symbionts that are present in the same host (Meseguer et al. 2017). Heritable facultative symbiont genomes are often intermediate in size (2–4 Mb) between “free-living” bacteria and obligate symbionts, but more similar to “free-living” species, often retain an abundance of mobile genetic elements and pathogenicity-associated factors (Moran et al. 2008; Lo et al. 2016).

Aphids are a group of phloem-feeding insects, including important pests, which are associated with diverse heritable facultative symbiont species (van Emden and Harrington 2007; Oliver et al. 2010; Guo et al. 2017). Individual aphids require infection with an obligate symbiont, usually *Buchnera aphidicola*, which supplements a N-poor phloem diet (Douglas 1998; Gunduz and Douglas 2009). Aphids may also be infected with one or more heritable facultative symbionts, which have functions that range from dietary breadth expansion to conferring protection against biotic and abiotic threats, and aiding decaying *Buchnera* function (Oliver et al. 2014; Vorburger 2014; Meseguer et al. 2017). For instance, the pea aphid, *Acyrthosiphon pisum* (Aphidinae) is variably infected with seven facultative symbionts (Ferrari et al. 2012; Russell, Weldon, et al. 2013; Rock et al. 2018) each with reported conditional benefits in defense against fungal pathogens (Scarborough et al. 2005; Łukasik et al. 2013), parasitic wasps (Oliver et al. 2003; McLean and Godfray 2015; Martinez et al. 2016) and/or thermal stress (Montllor et al. 2002; Russell and Moran 2006; Heyworth and Ferrari 2015).

Sequenced genomes are an important component to understanding the potential effects symbionts have on their hosts. One or more complete genomes are available for some pea aphid facultative symbionts, including *Hamiltonella defensa*, *Regiella insecticola*, *Rickettsiella viridis*, and *Serratia symbiotica* (Degnan et al. 2009, 2010; Burke and Moran 2011; Hansen et al. 2012; Chevignon et al. 2018; Nikoh et al. 2018), but not others, including any strain of *Candidatus* Fukatsuia symbiotica (Manzano-Marin et al. 2017), henceforth *Fukatsuia*, which have been studied under the provisional labels “X-type” or “PAXS” (Guay et al. 2009; Heyworth and Ferrari 2015, 2016; Doremus and Oliver 2017; Doremus et al. 2018). This symbiont also occurs in some Lachninae aphids where it has recently transitioned to a coinfection (Chevignon et al. 2018). Here, we used a similar approach to culture a strain of *Fukatsuia* from pea aphids. Having a culture comprised of only *Fukatsuia* enabled us to greatly increase the proportion of sequencing reads that map to this species and to fully assemble its genome using PacBio single molecule real-time (SMRT) sequencing. We then compared this *Fukatsuia* genome to other pea aphid protective symbionts and a *Fukatsuia* strain from a Lachninae aphid, *Cinara confinis*, to identify factors that potentially contribute to the unusual biology of this symbiont.

**Materials and Methods**

**In Vitro Cultivation of the *Fukatsuia* Symbiont**

Pea aphid clonal line 5D, naturally infected with *Fukatsuia* and *H. defensa*, was collected on alfalfa (*Medicago sativa*) in 2012 from Dane County, WI. A selective antibiotic cocktail was used to create a subline (i.e., same aphid genotype) carrying...
only *Fukatsuia* that was used in prior experimental studies to characterize phenotypic effects of infection (Doremus and Oliver 2017; Doremus et al. 2018). This strain of *Fukatsuia* was named Ap5D (= *A. pisum* clone 5D). We next established *Fukatsuia* strain Ap5D in culture following Brandt et al. (2017). Individual aphids were surface sterilized by washing with 5% bleach, 1% Tween 20, and 0.1% ROCCAL-D (Pfizer) in water for three minutes, followed by a 70% ethanol wash (15 s), and then rinsed in sterile, deionized water (30 s). Using sterile forceps, aphids were pierced in the abdomen to release a small amount of *Fukatsuia*-containing hemolymph into a droplet of a commercially available culture medium, TC100 (Sigma), that was supplemented by adding 10% fetal bovine serum (FBS) and TNS (BTI-TN-5B1-4) cells derived from the lepidopteran *Trichoplusia ni*. The droplet was then transferred into individual culture wells that each contained 1 ml of medium and held at 22–23°C. We conducted diagnostic PCR with primers that amplify the hpaA gene specific for *Fukatsuia* (Doremus and Oliver 2017) to verify cultivation of the expected symbiont. Cultures of *Fukatsuia* strain Ap5D were passaged weekly and ~20 times prior to sequencing.

To estimate the abundance of *Fukatsuia* and *H. defensa* cells in culture, we performed “absolute” quantitative real-time PCR of the single-copy gene *dnaK* using a standard curve with primers specific to each facultative symbiont. We collected 1 ml of medium at 24, 48, and 120 h at 22°C after passaging a culture and pelleting the bacteria present by centrifugation at 8,000×g for 10 min at 4°C. We then decanted the medium and extracted DNA from the pellet of bacteria as previously described (Weldon et al. 2013). We also estimated *Fukatsuia* abundance in fourth instar pea aphid nymphs (*N* = 6) using “whole aphid” DNA extractions and correcting for extraction efficiency using the aphid gene ef1α (as in Weldon et al. 2013). Ten microliter reactions, with three technical replicates per sample, were carried out with PerfeCta SYBR Green Master Mix (QuantaBio) using a qTower3 cycler (Analytik Jena). Primers and reaction concentrations and conditions can be found in Doremus and Oliver (2017).

### Genome Sequencing and Assembly

To isolate *Fukatsuia* DNA for sequencing, nonadherent TNS cells were removed via centrifugation at 500×g; supernatant was then collected and centrifuged at 10,000×g for 15 m to produce a bacterial pellet. DNA extraction was performed using the DNeasy Blood and Tissue kit (Qiagen) followed by titration using a NanoDrop (Thermo Scientific) and visualization of aliquots on a 1.5% agarose gel stained with ethidium bromide to verify DNA integrity. We also Sanger sequenced the above-mentioned diagnostic PCR amplicons as well as a ~1,200 bp 16S rRNA amplicon produced by “universal” bacterial primers (primers and reactions conditions in Doremus and Oliver 2017) which were a 100% matches to prior NCBI GenBank submissions for pea aphid-associated *Fukatsuia* (accession numbers KY271023; KY271016). Samples were then sent to the Drexel University College of Medicine Genome Core Facility for SMRTBell fragment library construction using Long-Insert Genomic DNA followed by SMRT sequencing. Data were collected on a SMRTCell and the number of reads was 110,204 with mean read size of 11,342 (supplementary table S1, Supplementary Material online). *De novo* assemblies were performed with the Hierarchical Genome Assembly Process (HGAP.2) algorithm in the SMRT Portal (version 2.3.0) using default parameters.

We also reassembled and reannotated the genome of a strain of *Fukatsuia* from the Lachninae aphid *C. confinis* hereafter strain Ci for improved comparison with the pea aphid Ap5D strain. We used reads from a prior study (Meseguer et al. 2017) combined with an additional 250 bp paired-end library generated from the original source material using the Illumina HiSeq2500 platform. Preassembly treatment of reads and assembly were conducted as described in Meseguer et al. (2017). Briefly, reads were right-tail clipped (minimum quality threshold = 20) using FASTX-Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/ last accessed August 10, 2018) dropping reads shorter than 75 bp. We also used PRINSEQ v0.20.4 (Schmieder and Edwards 2011) to remove unpaired reads and those containing undefined nucleotides. The remaining reads were assembled using SPAdes v3.10.1 (Bankevich et al. 2012) using the options only-assembler and k-mer sizes of 33, 55, 77, 99, and 127 and contigs shorter than 200 bp dropped. The remaining contigs were binned using results from a BlastX (Altschul et al. 1997) search (best hit per contig) against a database consisting of the pea aphid’s proteome and a selection of aphid symbiotic bacteria proteomes (see table S4 at https://doi.org/10.1093/gbe/evy173; Manzano-Marín et al. 2018), including our newly assembled Ap5D *Fukatsuia*. Contigs assigned to *Fukatsuia* were manually confirmed using a BlastX search of the nr database. We also ran a BlastX search against the Ap5D *Fukatsuia* genome recovering the 16S and 23S rRNA-containing contigs. The resulting contigs were then used as a reference for read mapping and genome assembly using SPAdes with read error correction.

### Genome Annotation

The assembled genome for the *Fukatsuia* Ap5D strain and the reassembled *Fukatsuia* Ci strain were submitted to the National Center for Biotechnology Information (NCBI) Prokaryotic Genomes Annotation Pipeline (PGAP), Rapid Annotation using the Subsystem Technology tool kit (RASTtk) (Brettin et al. 2015). Predictions from NCBI and RAST annotation were merged and compared with Bacterial gEnome Annotation Comparis0nV (BEACON) (Kalkatawi et al. 2015) and manually curated using Geneious (www.geneious.com). Predictions for ribosomal (r), transfer (t), and transfer-messenger (tm) RNAs were retained from the PGAP annotation. ORFs that had >80% truncation or fragmented coding
sequences (CDSs) were designated pseudogenes in the final annotation. PacBio sequencing technology can generate a deletion bias in homopolymeric runs, especially in GC-rich regions (Ross et al. 2013) inflating pseudogene calls. After verifying that all homologous loci in the Ci strain were intact, we manually corrected all deletions in the Ap5D strain. Noncoding RNAs (ncRNA) were predicted using Infernal cmsearch (www.ebi.ac.uk/Tools/rna/infernal_cmscan; last accessed December 2018) and protein homology with phmmer (www.ebi.ac.uk/Tools/hmmer/search/phmmer; last accessed December 2018) and Joint Genome Institute IMG/MER (Markowitz et al. 2012). KEGG Orthology (KO) was assessed by the automatic annotation servers BlastKOALA (KEGG Orthology And Links Annotation) (Kanehisa et al. 2016). Clusters of Orthologous Groups (COGs) and Gene Ontology (GOs) were retrieved with EggNOG 4.5.1 (Huerta-Cepas et al. 2016). Transposable element predictions were made using the ISfinder (Siguier et al. 2006) and ISsaga (Insertion Sequence semiautomatic genome annotation) (Varani et al. 2011). Phage island predictions were made using PHASTER (Arndt et al. 2016) and identification and classification of secretion systems was performed with TXSScan (Abby 2016). Plasmid islands were identified manually using BLAST (Altschul et al. 1990).

Comparative Genomics
The 77 contigs identified in the original sequencing of the Fukatsuia Ci strain (Manzano-Marin et al. 2017) along with the 381 contigs identified in the reassembly of this genome during the current study were aligned to the complete genome generated for the Fukatsuia Ap5D strain using nucmer from the MUMmer package (Kurtz et al. 2004). Alignments were plotted using Circos (Krzywinski et al. 2009). We then used the tool “annotation from” in Geneious v10.2 (www.geneious.com) to map our annotation for the Fukatsuia Ap5D strain onto the Fukatsuia Ci strain followed by extensive manual curation. This allowed us to identify Fukatsuia Ap5D CDSs not present in Fukatsuia Ci and fragmented CDSs in Fukatsuia Ci which were usually positioned at contig ends. Finally, we conducted pairwise alignment of the 2,340 CDSs shared by both Fukatsuia strains using MAFFT (Katoh and Standley 2016). Whole genome alignments were conducted using both the ProgressiveMauve and Mauve Contig Mover algorithms in Geneious Prime 2019.2.1. (Rissman et al. 2009; Darling et al. 2010).

Phylogenetic Reconstruction
Fragments of seven single-copy orthologous genes (accD, dnaA, gyrB, murE, ptsI, recJ, and rpoS) were used for reconstructing phylogenetic relationships of Fukatsuia and other enterobacterial relatives. Partial sequences for these loci from 26 species were downloaded from NCBI (supplementary table S2, Supplementary Material online). Protein alignments were conducted using the Geneious aligner v10.2 with Blosum-62 cost matrix and aligned orthologs were manually trimmed and concatenated into a supermatrix (Henikoff and Henikoff 1992; Kease et al. 2012). We used ModelFinder (Kalyaanamoorthy et al. 2017) to identify the best models of sequence evolution for each locus separately (supplementary Table S2, Supplementary Material online). We then used IQ-TREE v1.6 (Nguyen et al. 2015) to estimate evolutionary relationships using maximum likelihood with partitioned models (Chernomor et al. 2016) and branch support approximated using UFBoot2 with 1,000 replicates (Hoang et al. 2018). The phylogenetic tree was edited with TreeGraph2 v2.14.0-771beta (Stower and Muller 2010).

Base Modification
Base modification analysis and motif detection were performed using the RS Modification_and_Motif_Analysis.1 algorithm in the SMRT Portal with standard settings (minimum-modification QV of 30). To reduce false positive motif detection due to high level base coverage, we reanalyzed the data output with specific scripts provided in Base Modification Tools accessible on the Pacific Biosciences GitHub pages (https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/BaseModification-Tools) using a minimum-modification QV of 100 for base modification and motif prediction. Motifs were then manually curated per PacBio guidelines.

Results and Discussion
Candidatus Fukatsuia symbiotica Can Be Cultured In Vitro to Produce an Isolated Template for Genomics
The in vitro culture of Fukatsuia strain Ap5D allowed us to generate DNA template for genome sequencing that was free of contaminating aphid, Buchnera or other bacterial DNA with the large majority (85%) of PacBio SMRT sequencing reads (N50 read length = 27,056) mapping to Fukatsuia and providing 291x coverage. See supplementary table S1, Supplementary Material online, for a summary of sequencing metrics. Cultures of Fukatsuia Ap5D were routinely passaged each week and maintained at an abundance per ml that was within an order of magnitude across examined time points (2.1, 1.9, and 1.6x10^8 per ml at 24, 48, and 120 h after passage, respectively) of that observed in fourth instar (6 day old) aphid nymphs (≈ 6.8x10^8 per aphid). We also separately cultured Fukatsuia and H. defensa at the same time to confirm that both facultative symbionts could be cultured under identical conditions. Although H. defensa abundance and growth rates were higher in culture (1.9, 2.5, and
3.3 × 10^9 per ml for 24, 28, and 120 h, respectively) compared with *Fukatsuia*, they were also substantially higher in aphids (Brandt et al. 2017; Doremus and Oliver 2017). The ability to maintain *Fukatsuia* and *H. defensa* at relatively high titers under similar conditions provides opportunities to study function when biology observed in vivo can be recapitulated in vitro.

**Fukatsuia Forms a Clade with the Protective Aphid Symbionts *H. defensa* and *R. insecticola***

The order Enterobacterales (class Gammaproteobacteria) traditionally contained the single large family Enterobacteriaceae. However, recent analyses propose seven families, including the Yersiniaceae, which contains the well-known genera *Yersinia* and *Serratia* (Adeolu et al. 2016). Our maximum likelihood phylogeny indicates that *Fukatsuia* resides in the Yersiniaceae and is the sister group to *Regiella* with 100% bootstrap support (fig. 1). *Fukatsuia*, *R. insecticola*, and *H. defensa* further form a highly supported clade (99% bootstrap support) of insect heritable facultative symbionts that are primarily associated with aphids although *H. defensa* is also found in other sternorrhynchan Hemiptera, including whiteflies, and possibly a coccinellid predator (Clark et al. 1992; Russell et al. 2003; Majerus and Majerus 2010). The sister-relationship between *Regiella* and *Fukatsuia* contrasts with previously reported phylogenies using fewer loci that instead showed *Fukatsuia* as basal to a clade comprised of *H. defensa* and *R. insecticola* (Doremus and Oliver 2017; Manzano-Marin et al. 2017).

**Overview of the Fukatsuia Genome Compared with Related Facultative Symbionts**

Long-read sequencing of the *Fukatsuia* Ap5D strain yielded a fully assembled genome that consisted of a circularized main chromosome of 2,824,275 bp and three plasmids named pFSSD.1 (148,330 bp), pFSSD.2 (91,928 bp), and pFSSD.3 (67,451 bp) (supplementary table S1, Supplementary Material online). As is the case for other aphid facultative symbionts, genome-based inferences of central metabolism indicated that *Fukatsuia* is a host-dependent, aerobic heterotroph with a total genome size (3.1 Mb) intermediate between the obligate nutritional symbiont, *Buchnera aphidicola* (640 kb) and the related free-living pathogen *Yersinia pestis* (4.7 Mb) (fig. 2 and table 1). GC content (43.5%) was nested within the range of other aphid facultative symbionts (40–52%) while a total of 2,607 CDSs were predicted for synthesis of some essential amino acids (EAAs) (supplementary table S3, Supplementary Material online, for specific gene comparisons). The *Fukatsuia* Ap5D genome further contained a relatively larger repertoire of genes with predicted functions in replication, recombination and repair, cell motility, inorganic ion transport and metabolism, secondary metabolism, and signal transduction. The proportion of the genome containing mobile genetic elements was also similar to that found in *H. defensa* and *R. insecticola* (see supplementary table S4, Supplementary Material online, for complete annotation).

Annotating the main chromosome for the *Fukatsuia* Ap5D strain indicated that the pathways required for biosynthesis of nonessential AAs, cofactors (coenzyme A, isoprenoids, ubiquinone), and B vitamins (B1, 2, 3, 6, 7, and 9) are intact, which is similar to related symbionts (*H. defensa* and *R. insecticola*) that reside in the same clade (supplementary table S3, Supplementary Material online). In addition, similar to related symbionts, Ap5D lacked genes in the pathways required for biosynthesis of some essential amino acids (EAAs) (supplementary table S3, Supplementary Material online). However, the presence of EAA transporter genes, suggests *Fukatsuia* Ap5D likely acquires these products from the aphid host and *Buchnera* and persists, similar to other facultative symbionts, as a nutritional parasite. In contrast, *Fukatsuia* Ap5D differed from all known strains of *H. defensa* or *R. insecticola* in that it encodes all genes required for synthesis of branched-chain amino acids (BCAAs).
such as valine, leucine, and isoleucine. Although the nutritional symbiont *Buchnera* lacks the terminal genes in the BCAA pathways, the presence and expression of these genes in aphids enables coordinated BCAA biosynthesis (Shigenobu et al. 2000; Wilson et al. 2010; Russell, Bouvaine, et al. 2013). Whether an intact pathway for BCAA synthesis by *Fukatsuia* Ap5D affects the biology of aphids, symbiont phenotypes, or interactions with other bacterial symbionts is currently unclear.

Annotation of the three plasmids in the Ap5D strain indicated that pFSSD.1 contains 134 CDSs including 17 pseudogenes, 35 TEs, partial components of T1SS (ToIC, HylD) and
T4SS (tra and pil genes), and toxin–antitoxin systems. Ten genes comprising 34,136 bp (spanning two loci) encode a nonribosomal peptide synthase (NRPS) module with homology to one found in betaproteobacterial Chromobacterium species. NRPS are ribosome-independent units often involved in the production of bio-active molecules (Izore and Cryle 2018), which together with its presence on a plasmid, warrants further investigation. pFS5D.2 contained 121 CDSs.

---

**Table 1**

<table>
<thead>
<tr>
<th></th>
<th>Buchnera aphidicola APS</th>
<th>Hamiltonella defensa SAT</th>
<th>Regiella insecticola LSR1</th>
<th>Serratia symbiotica Tucson</th>
<th>Fukatsuia symbiotica Ap5D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome (bp)</td>
<td>640,681</td>
<td>2,110,331</td>
<td>2,035,106</td>
<td>2,789,218</td>
<td>2,824,275</td>
</tr>
<tr>
<td>Extrachromosomal elements</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>unknown</td>
<td>3</td>
</tr>
<tr>
<td>Total G+C (%)</td>
<td>26.2</td>
<td>40.3</td>
<td>42.4</td>
<td>52.0</td>
<td>43.5</td>
</tr>
<tr>
<td>Total predicted CDSs</td>
<td>571</td>
<td>2,243</td>
<td>1,761</td>
<td>2,098</td>
<td>2,929</td>
</tr>
<tr>
<td>Total pseudogenes (including plasmids)</td>
<td>13</td>
<td>203</td>
<td>214</td>
<td>550</td>
<td>189</td>
</tr>
<tr>
<td>tRNAs</td>
<td>32</td>
<td>42</td>
<td>36</td>
<td>44</td>
<td>48</td>
</tr>
<tr>
<td>Lifestyle</td>
<td>Obligate</td>
<td>Facultative</td>
<td>Facultative</td>
<td>Facultative</td>
<td>Facultative</td>
</tr>
</tbody>
</table>

---

**Fig. 3.**—Metabolic map of *Fukatsuia symbiotica* strain Ap5D constructed using BLASTKOALA (Kanehisa et al. 2016) and IMG-ER (Markowitz et al. 2012). Amino acid pathways are shown in boxes; vitamins as hexagons. Solid black dots represent intact genes, solid red circles depict absent genes and open red circles indicate pseudogenes.
(21 repeats, 9 pseudogenes), including genes associated with T4SS (virB), toxin–antitoxin genes (higA/higB and vapBC), and transcription regulator genes, while pFS5D.3 contained 70 CDSs (23 repeats, 4 pseudogenes) and was also very similar in gene content to the plasmid p5DHAT in *H. defensa*.

Overall, the ability to culture *Fukatsuia* Ap5D using conditions identical to those used for *H. defensa* is consistent with the high similarity in genome content among this related clade of facultative symbionts, and suggests *R. insecticola* may be similarly cultivated.

**Diverging Paths for *Fukatsuia***?

Our rereassembly and -annotation of previously generated short-read sequence data yielded a genome for the *Fukatsuia* Ci strain that was comprised of 381 contigs (up from 77) and had an estimated size of 2.79 Mb (up from 2.03) (supplementary table S1, Supplementary Material online). We also identified two of the three plasmids (pFS5D.1 and pFS5D.3) present in the Ap5D strain and numerous fragmented CDSs. As earlier noted, the *Fukatsuia* Ci strain was previously hypothesized to support the loss of biosynthetic capability in *C. continis*-associated *Buchnera* for riboflavin (B2) and biotin (B7), and thus has potentially transitioned from facultative to co-obligate status (Meseguer et al. 2017). These pathways are intact in the *Fukatsuia* Ci strain, but are also present in the Ap5D strain as well as other common heritable facultative symbionts (supplementary table S3, Supplementary Material online).

One expectation for a facultative symbiont transitioning toward co-obligacy would be a further reduction in genome size compared with facultative strains. The estimated size of the genome for the Ci strain is 337,000 bp smaller than for Ap5D (total CDS of main chromosome + plasmid = 2,800 vs. 2,932, respectively). However, we note that a large proportion of this difference is due to the apparent absence of pFS5D.2 in the Ci strain, while most other “missing” genes in the Ci strain are associated with mobile genetic element islands rather than single copy orthologs. Thus, while the Ci strain is a subset of the Ap5D strain (fig. 4 and supplementary table S3, Supplementary Material online), which is consistent with genome reductions following a transition to co-obligacy, the smaller estimated genome size for the Ci strain more likely reflects an incomplete assembly owing to the short read illumina data generated to sequence this strain.

A recent transition to co-obligate status might also lead to differences in the number of pseudogenes present between strains. Proliferation of mobile elements and pseudogenes is often observed after transition to a host-restricted lifestyle (Moran and Plague 2004). We identified more pseudogenes in the Ap5D (*N* = 189) than Ci strain (*N* = 126) (supplementary table S1, Supplementary Material online) but also noted that about half (46) of the pseudogenes found only in the Ap5D strain occur as fragments at contig ends in the Ci strain, which prevented us from classifying them as intact. A total of 23 pseudogenes in the Ap5D strain had no homolog (intact or pseudogenized) in the Ci strain, whereas no pseudogenes in the Ci strain were missing homologs in the Ap5D strain. This pattern could reflect intensified genome degradation associated with transition to co-obligacy by the Ci strain, since a similar pattern was observed for a range of *S. symbiotica* strains that vary from having a facultative to obligate association with their hosts (Manzano-Manín and Latorre 2016). However, it is more likely the differences observed between the Ci and Ap5D strains reflect the differential quality of the genomes we were able to assemble. When considering only pairwise comparisons (i.e., genes present in both genomes; fig. 4), 117 pseudogenes are shared by both strains, 8 pseudogenes in the Ci strain are intact in Ap5D, and 10 pseudogenes in the Ap5D strain are intact in Ci. Most of these pseudogenes are also in mobile regions of the genomes.

Altogether, the most striking feature in our comparison of the Ci and Ap5D genomes is just how similar they are overall: 92% of complete pairwise CDSs share 100% nucleotide identity and 98% share ≥99% similarity, with what little variation exists occurring in genes associated with mobile genetic elements. Although we cannot order the 381 contigs from the Ci strain and hence determine long range rearrangements, synteny within larger contigs is similar to the Ap5D strain (supplementary fig. S1, Supplementary Material online). Thus, despite an abundance of mobile genetic elements (see below), that are often associated with genome rearrangements (Wu et al. 2004; Degnan et al. 2010; Chevignon et al. 2018), gene order to the extent we can compare it, appears conserved between the Ap5D and Ci strains. This similarity becomes more impressive given the very different impacts posited for the Ci and Ap5D strains in their respective aphid hosts (Doremus and Oliver 2017; Meseguer et al. 2017), which belong to distantly related subfamilies (Ortiz-Rivas and Martínez-Torres 2010). Such similarity may result from *Fukatsuia* persisting as a low genetic diversity heritable symbiont with a widespread distribution or this similarity might reflect a relatively recent horizontal transfer event between these distantly related aphid hosts. The overall similarity between the two strains also confirms that the transition to co-obligate status in Lachninae aphids was likely a recent event as proposed by Meseguer et al. (2017). In Lachninae aphids several other facultative symbionts have likewise transitioned to co-obligate status, and exhibit repeated turnover, which may prevent the accumulation of genomic changes seen with more established co-obligate symbionts (Manzano-Manín and Latorre 2016).

**Why is *Fukatsuia* Frequently Associated with *H. defensa***?

Many groups of sap-feeding insects harbor multiple obligate nutritional symbionts in their bacteriomes that exhibit metabolic complementation between species (Wu et al. 2006;}

---

Metabolic interdependence is also the basis for the transition of facultative symbionts into co-obligate status when the obligate symbiont falters due to genome degradation (Lamelas et al. 2011). Prior screening indicated that *Fukatsuia* has a strong tendency to co-occur with specific *H. defensa* strains (Doremus and Oliver 2017). Metabolic complementation is obviously one possible explanation for the frequent co-occurrence of *Fukatsuia* and *H. defensa* in pea aphids. However, results from this study indicate the nonmobiliome components of the *Fukatsuia* genome are very similar to the *H. defensa* genome as well as the genomes of other aphid facultative symbionts. There are a few genes missing or pseudogenized in *Ap5D* that are present in *H. defensa*, but these genes are either present or also missing in other aphid facultative symbiont species (supplementary table S3, Supplementary Material online), and so does not

![Diagram comparing the classification of CDSs between Fukatsuia strains](https://academic.oup.com/gbe/article-abstract/11/12/3510/5625627)
explain Fukatsuia preferentially co-occurring with H. defensa. Thus, if metabolic complementation is the basis for coinfection, it must play out on an H. defensa strain-specific basis. Alternative explanations for Fukatsuia and H. defensa coinfections include improved competitive abilities against third party heritable facultative symbionts (Rock et al. 2018), or tamed Fukatsuia virulence toward aphid hosts under coinfection. Interestingly, Fukatsuia encodes siderophores which are molecules involved in iron-chelation and uptake and are often important pathogenicity factors (Buckling et al. 2007). Since these are secreted molecules that can be utilized by confecting bacterial symbionts, this potentially selects for reduced siderophore production under coinfection resulting in lower virulence toward the insect host (Vorburger and Perlman 2018). Consistent with this hypothesis, we have observed reductions in infection costs when Fukatsuia coinhabits aphids with H. defensa relative to infection with Fukatsuia alone (Doremus and Oliver 2017).

Mobile DNA in Fukatsuia

Similar to other aphid facultative symbionts (Plague et al. 2008; Degnan et al. 2009, 2010; Burke and Moran 2011; Manzano-Marín and Latore 2016), a substantial portion of the Fukatsuia main chromosome is comprised of mobile elements, including prophage and prophage-like elements, plasmid elements, and transposable elements (TEs). Of the 2,607 total CDSs of the main chromosome, 833 CDS (32%) occurred in mobile elements (fig. 2 and supplementary tables S5 and S6, Supplementary Material online). We identified TEs from 20 families, with IS630 being the most abundant (supplementary table S7, Supplementary Material online). We identified 33 prophage elements that varied in size from 291 bp (1 CDS) to 43 kb (59 CDSs) that shared homology to domains. However, prophage island 31 (32.6 kb) showed prophages because they lacked one or more essential genes. Nearly all of these phage islands, however, were not intact since in Fukatsuia coinhabits aphids with H. defensa relative to infection with Fukatsuia alone (Doremus and Oliver 2017).

Putative Toxins and Virulence Factors Associated with Fukatsuia

The Ap5D strain of Fukatsuia is costly to aphid fitness when occurring as a single infection, while other strains of Fukatsuia are reported to have diverse protective effects if additional facultative symbiont species are present (Heyworth and Ferrari 2015, 2016; Doremus and Oliver 2017). Fukatsuia Ap5D carries numerous genes with homology to Type 1–4 bacterial secretion systems (supplementary table S8, Supplementary Material online) that are dispersed across the main chromosome and a plasmid. The T2SS and T3SS (SPI-1 and 2) Fukatsuia Ap5D were further noted to have a twin arginine translocation (Tat) pathway, which transports unfolded proteins and is involved in bacterial pathogenesis. All sequenced strains of R. insecticola but not H. defensa, also have a similar twin Tat pathway.

The most common toxins found in Fukatsuia are those from the repeats-in toxin RTX superfamily, including virulence factors with homology to hemolysin (hlyA) and MARTX (rtxA). RTX toxins occur in diverse gram negative bacterial species and possess aspartic acid/glycine repeats with a variety of proteins typically secreted by T1SS (Linhartova et al. 2010). All of the 30+ RTX variants present in Fukatsuia also occur in R. insecticola and/or H. defensa, suggesting these may have diversified in the ancestor to these symbionts with subsequent inactivation of specific copies in different lineages. This same pattern of similar inventory but variable inactivation also occurs between the two Fukatsuia strains (supplementary table S9, Supplementary Material online), suggesting these toxins may change rapidly when transferred to new hosts. Other putative insecticidal toxins are present, including tcc and tcd toxins typically associated with the nematode symbionts Xenorhabdus nematophila and Photorhabdus luminescens, and pathogens including Y. pestis and Senatia entomophila, but not reported in H. defensa or R. insecticola. We also identified nine mcf (make caterpillars floppy) homologs which are associated with insecticidal activity and also occur in Photorhabdus and R. insecticola (ffrench-Constant et al. 2007; Hansen et al. 2012). Other putative virulence factors were identified, such as cytotoxal distending toxin subunit B (cdtB), which is located adjacent to a plasmid island. A cdtB gene is present in two APSE haplotypes (APSE2 and APSE8) that persistently infect particular strains of H. defensa, which have antiparasitoid functions in aphids (Moran et al. 2005; Degnan and Moran 2008; Martinez et al. 2014), but as noted above APSEs are absent in Fukatsuia. Furthermore, the cdtB encoded by Fukatsuia is more similar to cdtB genes present in certain enteric pathogens than APSEs, arguing against recent exchange among common aphid endosymbionts. Helogenes of additional pathogeneity factors can be found in supplementary table S10, Supplementary Material online.
Restriction Modification System

Bacteria commonly harbor genes coding for DNA methyltransferases (MTases) and restriction endonucleases (REases). Together these form restriction modification (RM) systems involved in protection against foreign DNA, which also function in the regulation of various cellular processes (Vasu and Nagaraja 2013). Through PacBio SMRT sequencing, we were able to identify two putative MTase/REase recognition motifs in Fukatsuiia strain Ap5D: 1) $G^{m6}ATC$ motifs present in 26,436 copies and for which 95% were detected as modified, 2) $TGCC^{m6}A$ motifs present in 1,186 copies and for which 90% were detected as modified (supplementary table S11, Supplementary Material online). Furthermore, submission of the F. symbiotica strain Ap5D genome to the restriction enzyme database (REBASE: http://rebase.neb.com/rebase/rebase.html) identified two putative genes involved in the methylation pattern observed in F. symbiotica strain Ap5D. The gene CCS41_11020 codes for an orphan DAM dependent N6-adenine DNA MTas, which recognizes the Ap5D. The gene CCS41_07300 codes for a fused type IIG REase/MTase recognizing TGGCCA motif. Whereas the orphan MTase modifying the GATC motif is probably involved in cell cycle regulation (Murphy et al. 2013), the fused REase/MTase encodes the two enzymatic activities and is therefore very likely involved in protection against foreign DNA infection.

Conclusions

This study shows that a culture-assisted protocol used to generate quality DNA template for whole genome sequencing of the heritable symbiont *H. defensa* (Brandt et al. 2017; Chevignon et al. 2018) can be used for other aphid facultative symbionts. These tools, along with other recent successes in culturing arthropod-associated bacteria (Masson et al. 2018), not only provide techniques for improved genomics, but also new avenues for functional studies in historically intractable systems. Our results also show that a facultative strain of *Fukatsuiia* from the pea aphid is strikingly similar to a strain from Lachninae aphids supporting the conclusion this transition to a new role occurred in the recent past. Along with the closely related *H. defensa* and *R. insecticola*, a substantial portion of the *Fukatsuiia* genome is comprised of mobile DNA encoding diverse toxins and pathogenicity factors. Together these properties equip this clade of symbionts with a diverse arsenal of factors allowing for heritable persistence via a variety of mechanisms in diverse hemipteran hosts.

Supplementary Material

Supplementary data are available at Genome Biology and Evolution online.

Acknowledgments

This work was supported by the US National Science Foundation awards to K.M.O. and M.R.S. (IOS 1256974) and K.M.O. and J.A.R. (IOS 1754302). Also, the Marie-Curie AgreenSkills+ fellowship program cofunded by the European 5 Union’s Seventh and a Fondation pour la Recherche Médicale award (ARF201809006914) to G.C. Framework Program (FP7-609938) supported A.M.M.

Literature Cited


