Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

Article

Genome features of *Asaia* sp. W12 isolated from the mosquito *Anopheles stephensi* reveal symbiotic traits

Shicheng Chen [1*], Ting Yu [2], Nicolas Terrapon [3,4], Bernard Henrissat [3,4,5], and Edward D. Walker [6]

[1] Department of Clinical and Diagnostic Sciences, School of Health Sciences, Oakland University, 433 Meadowbrook Road, Rochester, MI 48309; schen5@oakland.edu (S.C.)
[2] Agro-biological Gene Research Center, Guangdong Academy of Agricultural Sciences, Guangzhou, 510640, China; yuting@agrogene.ac.cn (T.Y.)
[3] Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique (CNRS), Aix-Marseille Université (AMU), UMR 7257, 13288 Marseille, France; nicolas.terrapon@afmb.univ-mrs.fr (N.T.); Bernard.Henrissat@afmb.univ-mrs.fr (B.T.)
[4] Institut National de la Recherche Agronomique (INRA), USC 1408 AFMB, Marseille, France
[5] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia. Bernard.Henrissat@afmb.univ-mrs.fr (B.T.)
[6] Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824 USA; walker@msu.edu (E.D.W.)
* Correspondence: schen5@oakland.edu (S.C.)

Abstract: *Asaia* bacteria commonly comprise part of the microbiome of many mosquito species in the genera *Anopheles* and *Aedes*, including important vectors of infectious agents. Their close association with multiple organs and tissues of their mosquito hosts enhances the potential for paratransgenesis for delivery of anti-malaria or anti-virus effectors. The molecular mechanisms involved in the interactions between *Asaia* and mosquito hosts, as well as *Asaia* and other bacterial members of the mosquito microbiome, remained unexplored. Here, we determined the genome sequence of the strain W12 isolated from *Anopheles stephensi* mosquitoes, compared them to other *Asaia* species associated with plants or insects, and investigated some properties of the bacteria relevant to their symbiosis with host mosquitoes. The assembled genome of strain W12 has a size of 3.94 MB, which is the largest among *Asaia* spp studied so far. At least 3,585 coding sequences were predicted. The insect-associated *Asaia* including strain W12 carried more glycoside hydrolase (GH) encoding genes (31 per genome) than those isolated from plants (22 per genome). W12 had the most predicted regulatory protein components (213) among the selected *Asaia* (ranging from 131 to 211), indicating its great capability to adapt to frequent environmental changes in the mosquito gut. Two complete operons encoding cytochrome *bo*$_3$-type ubiquinol terminal oxidases (*cyoABCD-1* and *cyoABCD-2*) were found in most of *Asaia* genomes, which possibly offer alternative terminal oxidases and allow the flexible transition of respiratory pathways. Genes involved in the production of acetoin and 2,3-butandiol have been identified in *Asaia* sp. W12.

Keywords: *Asaia*; paratransgenesis; symbiotic traits; *Anopheles stephensi*; genome features

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

1. Introduction

Bacteria of the genus *Asaia* are classified as acetic acid bacteria (AAB) in the class Alphaproteobacteria, family Acetobacteraceae; they are Gram-negative, aerobic rods (1, 2). These bacteria are frequently isolated from tropical plants such as *Bauhinia purpurea* and *Plumbago* where they metabolize plant sugars and alcohols for growth (3). Bacteria of the genus *Asaia* have been characterized as symbionts of several insect species including the orders *Diptera* (flies, including mosquitoes), *Hymenoptera* (bees and wasps), and *Hemiptera* (true bugs) which feed upon plant sugars from nectar, fruit, and sap (3, 4). Diverse *Asaia* strains have been demonstrated to be tightly associated with several species of *Anopheles* mosquitoes known to be vectors of human malaria (*e.g.*, *A. stephensi*, *A. maculipennis* and *A. gambiae*) (5). *Asaia* was one of the most predominant bacterial members found in samples of male and female mosquito midgut (5, 6). Moreover, they persist in host mosquitoes without variations due to sex, blood and sugar meals, and age (7, 8). *Asaia* species live in the mosquito midgut lumen, but also actively colonize other tissues and organs including the salivary glands and reproductive organs, indicating that if ingested in a sugar meal from a plant, *Asaia* bacteria pass through tissue body barriers such as midgut epithelium and basal lamina, reaching and infecting other tissues and organs (9). Further, infection of *Asaia* among mosquitoes occurs by horizontal transmission *per os*, through mating (venereal transmission from adult male to adult female) as well as by vertical transmission from mother to progeny via ovarian infection to eggs (10, 11).

Infection of *Asaia* bacteria in insects appears to be of a mutualistic nature, in that the bacteria contribute physiologically to their insect hosts (12). For example, they provide through certain nutrients and metabolic co-factors (such as carbon, nitrogen and vitamins) (13), positively affect mosquito growth and development as evidenced by negative effects upon their removal, and antagonize *Wolbachia* endosymbionts (14, 15). These intimate interactions between *Asaia* bacteria and mosquitoes expand to interference in development and propagation of *Plasmodium* malaria parasites in *Asaia* infected *Anopheles* mosquitoes (16), supporting potential use of these bacteria (in proof of concept) as a paratransgenesis agent to control malaria transmission (9, 16, 17). The traits favoring this idea are ease of cultivation *in vitro*, amenability for genetic manipulation, and quickly established and stably persistent infections in several mosquitoes (9, 10, 16-18). Bongio and Lampe (2015) engineered *Asaia* sp. SF2.1 to secrete anti-plasmodial proteins into the *Anopheles* midgut, resulting in inhibition of malaria parasite development (18). Further, *Asaia* spp. infection modulated mosquito innate immunity by activation of antimicrobial peptide expression, thereby repressing growth and propagation of parasites and viruses in the insect guts; and more generally appear to compete for space and nutrients with other microorganisms in insect hosts (10, 19). To resolve issues about the mutual exclusion between *Asaia* and *Wolbachia* (15), the latter unrelated bacteria infecting reproductive tissues and having anti-parasite properties supporting paratransgenesis control methods; and to make use of favorable traits from bacteria of both genera,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

Epis et al. (2020) created chimeric *Asaia* symbionts expressing a *Wolbachia* surface protein which efficiently stimulated mosquito immunity, inhibiting filarial parasite development (20).

Despite their importance for mosquito physiology and the potential for control of vector-borne disease, the mechanisms involved in establishing symbiosis between *Asaia* and their hosts, regulating mosquito host development and immunity remain largely unknown (13, 14). In this study, we isolated a new *Asaia* strain from *A. stephensi*, sequenced the genome, and performed comparative genomic studies to other *Asaia* strains from insects and plants. The goal of this study was to develop a comparative genomic analysis of the *Asaia* species that provide insight into the molecular mechanisms for transmission, colonization, and persistence in mosquitoes.

2. Materials and Methods

2.1. Bacterial strains and growth conditions

Adult *A. stephensi* mosquitoes were collected from a laboratory colony maintained at Michigan State University (21), anesthetized for five minutes at -20 °C, and surface-disinfected by soaking in 70% ethanol. Mosquito midgut tissues were dissected under sterile conditions, suspended in 200 µl of sterile saline solution (0.9% NaCl) and homogenized using a pestle. Suspensions were transferred to enrichment broth containing 2.2% D-glucose, 0.5% peptone and 0.5% yeast extract (pH 3.5). Cultures were rotated at 200 rpm and 30 °C overnight, and plated onto selective medium containing 2.2% D-glucose, 1.0% ethanol, 1.0% yeast extract, 0.7% $CaCO_3$ and 1.5% agar. Colonies showing clear zones were isolated and selected for further experiments. Trypticase soy broth (TSB) medium was then used for culture of *Asaia* isolates resulting from this procedure. One of these isolates, herein designated strain *Asaia* sp. W12, was chosen for study.

2.2. Genome sequencing, assembly, and annotation

Isolation and purification of bacterial genomic DNA were performed with the Wizard Genomic DNA Purification Kit (Promega, CA, USA). A PCR amplification method was utilized to screen *Asaia* colonies by using the forward primer GCGCGTAGGCGGTTTACAC and reverse primer AGCGTCAGTAATGAGCCAGGTT (9). Next generation sequencing (NGS) libraries were prepared using the Illumina TruSeq Nano DNA Library Preparation Kit following standard procedures recommended by the manufacturer. De novo assembly was performed using CLC NGS Cell v. 10.0.1. Gene annotation was carried out by NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP 3.3). Genomic data of elected *Asaia* genomes were obtained from the NCBI genome database (https://www.ncbi.nlm.nih.gov).

2.3. Bioinformatics

Functional categorization and classification for predicted ORFs were performed by RAST server-based SEED viewer (22). For genome similarity

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

assessment, average nucleotide identity (ANI) and Digital DNA-DNA Hybridization (dDDH) were computed using GGDC (https://ggdc.dsmz.de). The pan genome, core genome, and specific genes of *Asaia* spp. were compared to representative *Asaia* genomes using EDGAR 2.0 (23). The presence of clustered regularly interspaced short palindromic repeats (CRISPRs) in the selected genomes was predicted using CRISPRFinder with default parameters (https://crispr.i2bc.paris-saclay.fr/Server/). Among them, questionable CRISPRs were omitted. Cluster analysis of orthologous groups (COGs) was carried out using Orthovenn2 (https://orthovenn2.bioinfotoolkits.net/task/create) with an e-value cutoff of 0.001 and an inflation value of 1.5. Taxonomic assignment of *Asaia* sp. W12 was conducted using GTDB-Tk (24) with the default settings, which placed this bacterium in *Asaia* genus. Genomic islands (GIs) were predicted by both IslandPick and IslandPath-DIMOB methods (25).

Eukaryotic-like proteins (ELPs) containing motifs tetratrico peptide repeats (TPRs: PF13429, PF13371, PF00515, PF13181, PF13432, PF14559, PF14561 and PF09976), ankyrin repeats (ANKs: PF12796 and PF13637), Sel1 repeats (PF08238) and fibronectin type III (PF14310) were predicted using InterProScan v5.44.79 (http://www.ebi.ac.uk/interpro/search/sequence/). Prediction of the regulatory elements was done using the p2rp program with default settings (http://www.p2rp.org). Carbohydrate-Active enZyme (CAZyme) families, including enzymes of glycan assembly (glycosyltransferases, GT) and deconstruction (glycoside hydrolases, GH, polysaccharide lyases, PL, carbohydrate esterases, CE), were semi-manually annotated using the CAZy database curation pipelines (26-28). More precisely, CAZymes were annotated based on a combination of BLASTP and HMMER searches and automatically processed when high similarity to reference CAZymes was observed and manually curated in intermediary similarity levels. Genomic context was inspected using GFF files and NCBI genome browsers (28). The identification of secondary metabolites was performed using the online server antiSMASH 5.0 with "relaxed" detection strictness (29).

2.4. Accession of the genome sequences

The data from these Whole Genome Shotgun projects have been deposited at DDBJ/ENA/GenBank under accession number PNQZ00000000.1. The BioProject designations for this project is PRJNA427835, and BioSample accession numbers is SAMN08274829.

2.5. Statistical analyses

Statistical analyses were performed using SAS (version 9.2; SAS Institute, Cary, NC).

3. Results

3.1. Genome features

GTDB-Tk analysis placed the bacterium W12 in the genus *Asaia* (24). *Asaia* sp. W12 formed a clade with *A. bogorensis* as well as insect-associated *Asaia* (i.e., *Asaia* sp. SF2.1 and *Asaia* As1742) (Figure 1); however, it departed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

from the cluster formed by *A. prunellae* JCM25354, *A. platycodi* JCM25414 and *A. astilbis* JCM15831 (Figure 1). The genome sizes (Table 1) in the insect-associated *Asaia* species (average 3.59 M, n = 4) were significantly larger (t-test, P < 0.05) than those from plant-associated ones (average 3.17 M, n = 5). *Asaia* sp. W12 had the largest genome size (3.94 M) among the selected *Asaia* (Table 1). The assemble genome of *Asaia* sp. W12 contained 229 contigs, 3,652 coding sequences (CDSs) and 6 rRNAs (Table 1). The average GC content W12 was 60.1%, consistent with those in the most of the selected *Asaia* species. It is interesting that the average GC content in *Asaia prunellae* JCM 25354 was much lower (55.8%) than that in other *Asaia* (Table 1). No plasmid sequence was retrieved from the *Asaia* sp. W12 genome, congruent with our inability to isolate plasmids from in strain W12. Genome analysis of *Asaia* sp. W12 by the RAST Server revealed at least 378 subsystems classified into 27 categories (Figure 2). Among these categories, the "amino acid and derivatives" subsystem had the largest number (294 CDSs), followed by carbohydrate metabolism (260), protein metabolism (257), and RNA metabolism (138). Moreover, the "stress response" category accounted for at least 104 CDSs. Within the "virulence, disease, and defense" subsystem (90 CDSs), 12 of them were related to invasion and intracellular resistance, while 72 were associated with resistance to antibiotics and toxic compounds.

Table 1. General features of various *Asaia* spp.

Species	Sources	Size (Mb)	GC%	Total RNA	CDS
A. bogorensis NBRC 16594	Plants	3.20	59.8	59	2,896
A. astilbis JCM 15831	Plants	3.15	58.0	49	3,482
A. prunellae JCM 25354	Plants	3.18	55.8	48	3,583
A. platycodi JCM 25414	Plants	3.15	59.3	49	3,774
A. bogorensis IPC-01	Flower	3.17	59.7	54	2,880
A. bogorensis GD-01	Mosquito	3.34	59.8	60	3,226
Asaia sp. SF2.1	Mosquito	3.52	59.8	45	3,005
Asaia sp. As-1742	Ants	3.74	59.6	54	3,308
Asaia sp. W12	Mosquito	3.94	60.1	59	3,580

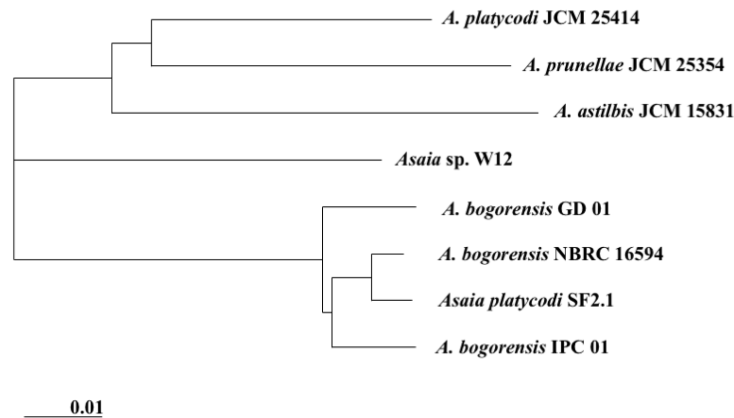


Figure 1. Phylogenetic placement of the selected *Asaia*. The tree is calculated from 829 core amino acids sequences per genome (6632 core amino acid sequences). The tree showed 100% branch support in 250 bootstrap iterations. Tree for 11 genomes, build out of a core of 471 genes per genome, 5181 in total. The selected genomes were: *Acetobacter tropicalis* strain_BDGP1 (CP022699), *A. astilbis* JCM15831 (BAJT01000001), *A. bogorensis* NBRC16594 (AP014690), *A. bogorensis* GD1 (UEGO00000000.1), *A. bogorensis* IPC01 (UBIX00000000.1), *A. platycodi* (BAKW01000001), *A. prunellae* JCM_25354 (BAJV01000001), *Asaia* sp. As1742 (VWWA00000000), *Asaia* sp. SF2.1 (AYXS00000000) and *Asaia* sp. strain W12 (PNQZ00000000.1).

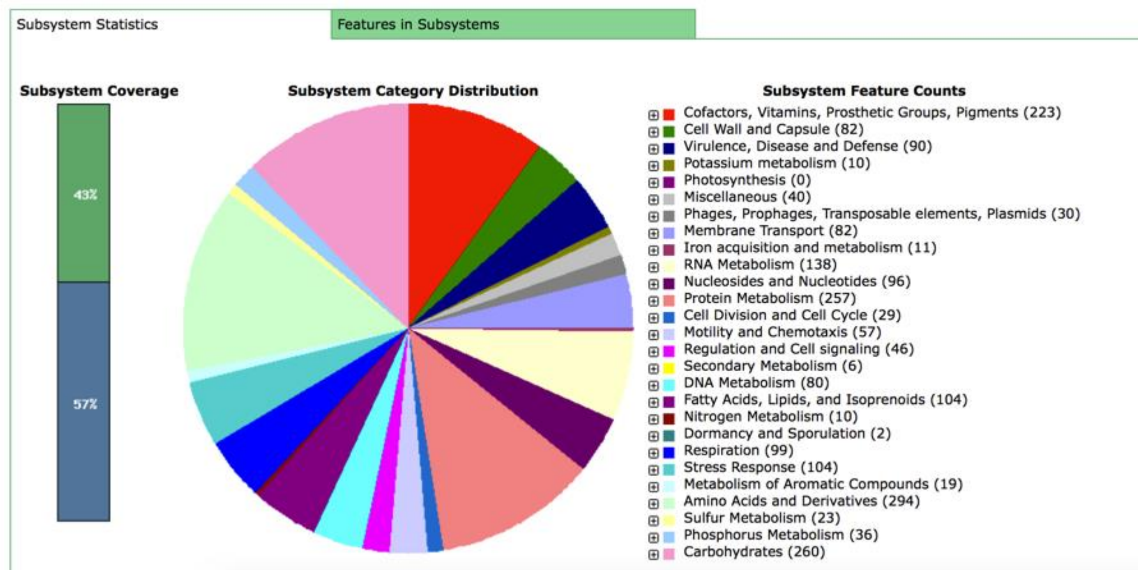
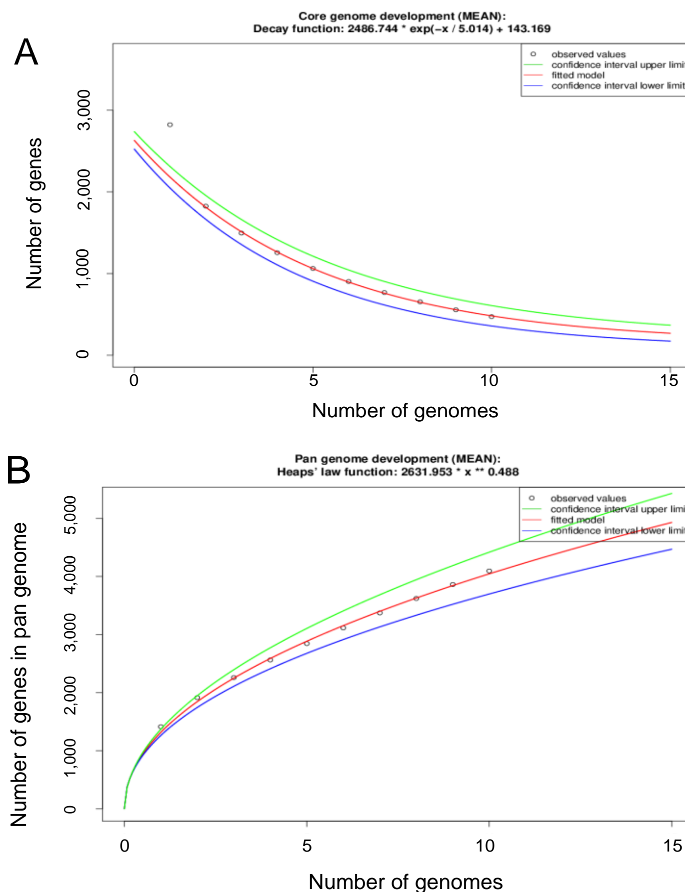


Figure 2. Subsystem category distribution of *Asaia* using SEED subsystems by RAST analysis. The pie chart represents relative abundance of each subsystem category and numbers depict subsystem feature counts.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

3.2. Gene repertoire of Asaia spp.

To investigate the pan-genome in *Asaia*, we plotted the number of annotated genes, shared genes, and unique genes as a function of the number of sequenced genomes (Figure 3). *Asaia* sp. W12 shared at least 3,649, 3,201, 2,996, 2,774, 1,971, and 1,868 CDSs, respectively, with *A. prunellae* JCM25354 (51.6% of its total encoding genes), *A. platycodi* SF2.1 (71.0%), *A. platycodi* JCM25414 (50.0%), *A. bogorensis* IPC 01 (69.6%), *A. bogorensis* GD01 (70.1%), *A. bogorensis* NBRC 16594 (71.6%) and *A. astiblis* JCM 15831 (48.5%) (Figure 4). A genome-wide comparison of orthologous clusters in different isolates provides insight into the gene structure, gene function, and molecular evolution of genomes (Figure 5). The COGs analysis of *Asaia* sp. W12 was compared with the other five genomes (Figure 5). The analysis shows that *Asaia* sp. W12 contained 2,781 COGs. Among them, 908 COGs were shared by all five strains, and 45 COGs were only present in the *Asaia* sp. W12 genome. The unique COGs existing in *Asaia* sp. W12 involved several genes functioning with transferase activity, cofactor binding, oxidoreductase activity, nucleotide binding, fatty acid elongation, and many others. However, the representative meanings of these singular genes in W12 are not clear. Further investigations to understand the features of these unique genes in W12 are warranted.



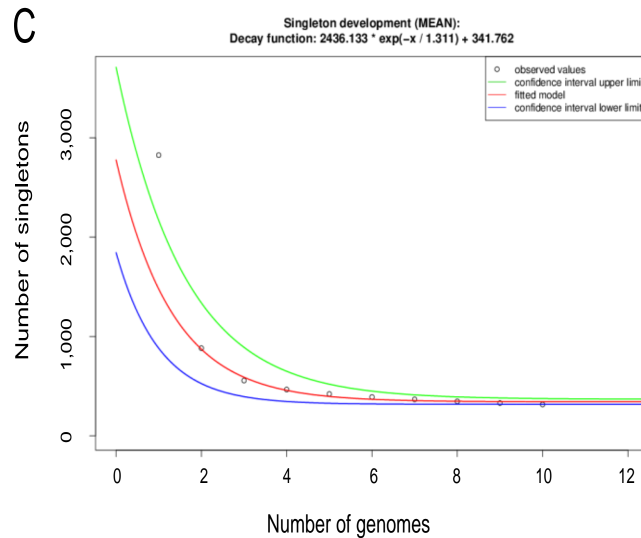


Figure 3. Core genome, pan genome and singleton genome evolution. (A) Total number of genes (pan-genome) for a given number of genomes sequentially added. (B) Number of shared genes (core genome) as a function of the number of genomes sequentially added. (C) Number of unique genes (singleton genome) as a function of the number of genomes sequentially added.

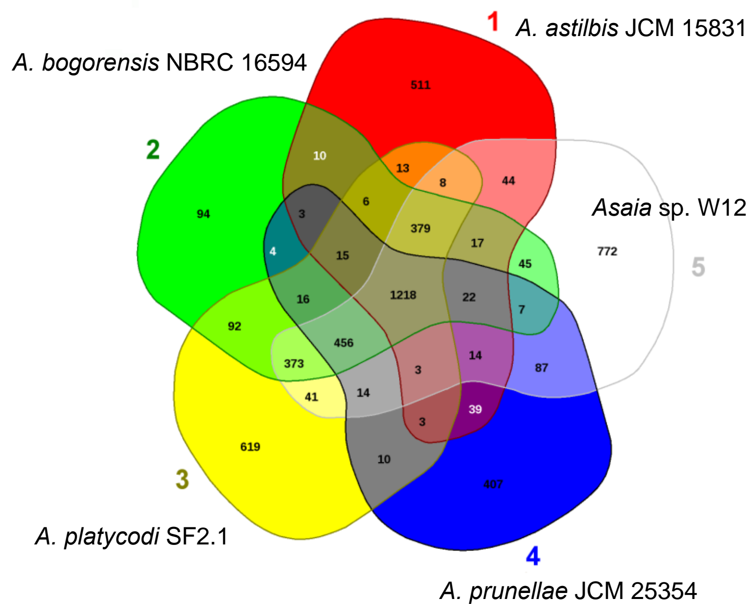


Figure 4. Venn diagram of shared and unique genes in the selected *Asaia*. The unique and shared genome among the compared genomes were determined using the BLAST score ratio approach of EDGAR 2.0 with a cutoff of 30%.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

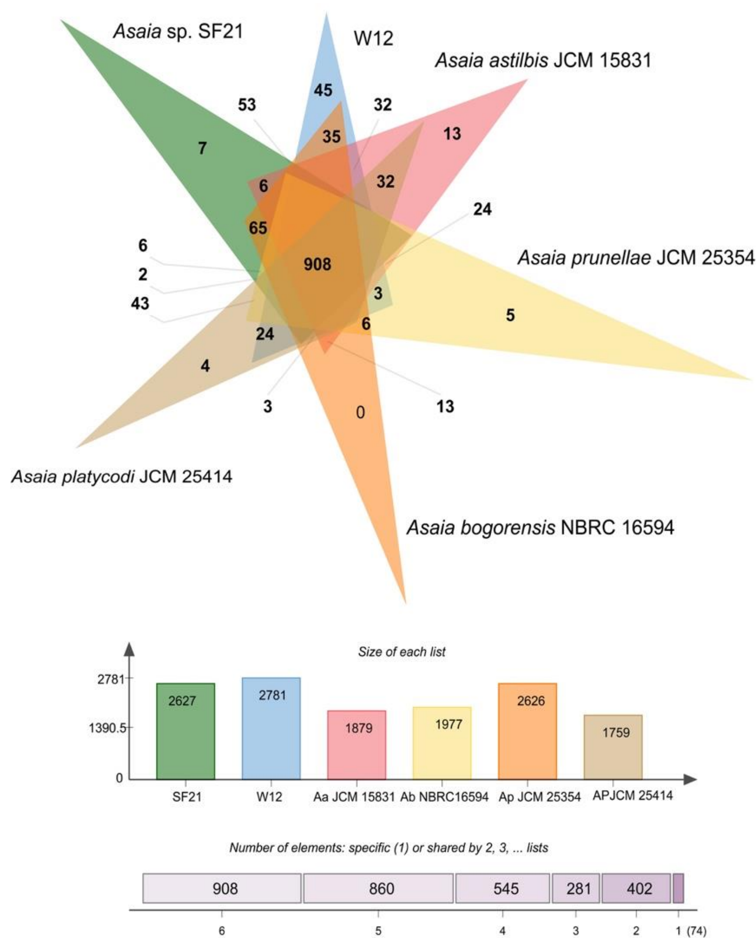


Figure 5. Proteome comparison among the selected *Asaia*. The Venn diagram and bar chart represent the numbers of unique and shared orthologous genes of each strain.

3.3. Carbohydrate-Active EnZymes (CAZymes) in Asaia spp.

The number of glycoside hydrolase (GH) genes in insect-associated *Asaia* (i.e. W12, SF 2.1, As-1742 and *A. bogorensis* GD-01) was slightly higher than those originating from plants (Table 2). Gene copies of glycosyltransferases (GTs) or carbohydrate esterases (CEs) in *A. astilbis* JCM 15831 and A. platycodi JCM 25414 were lower than those from other selected *Asaia* (Table 2). Most CAZymes in *Asaia* genomes (~30%) targeted peptidoglycan (glycoside hydrolases families GH23, GH102, GH103, GH104 and GH108), while we also observed conservation of some CAZymes dedicated to sucrose or fructose polymers (GH32 and GH68) that could, for example, allow the metabolism of plant nectar by mosquitoes. Interestingly, these genomes also encode a trehalase (GH37) commonly found in insect genomes as trehalose is their main blood sugar (Suppl. Table 1). Despite variable genome size and assembly across these genomes, two relevant CAZyme operons conserved in most species were identified. The first operon includes the bacterial glycogen operon, with alpha-glucanases (with four GH13s from distinct subfamilies for linkage specificities, three appended to carbohydrate-binding module family CBM48) and glycosyltransferases, notably from GT4 and GT5 families,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

in a single locus (Table S1). The second one included a GH8 and a glycosyltransferase of GT2 family likely involved in a cellulose-like biofilm synthesis (Table S1). The assembly of the bacterial cellulose biosynthesis operons was diverse among the *Asaia* species (Figure 5).

Table 2. The distribution of glycoside hydrolases, glycosyltransferases and carbohydrate esterases amongst different *Asaia* spp.

	Glycoside hydrolases	Glycosyltransferases	Carbohydrate esterases
A. astilbis JCM 15831	17	37	1
A. platycodi JCM 25414	15	34	1
A. prunellae JCM 25354	22	55	1
A. bogorensis NBRC 16594	28	54	3
A. bogorensis IPC-01	27	54	3
A. bogorensis GD-01	28	55	3
Asaia sp. W12	32	52	3
Asaia sp. As-1742	32	53	3
Asaia sp. SF2.1	32	53	3

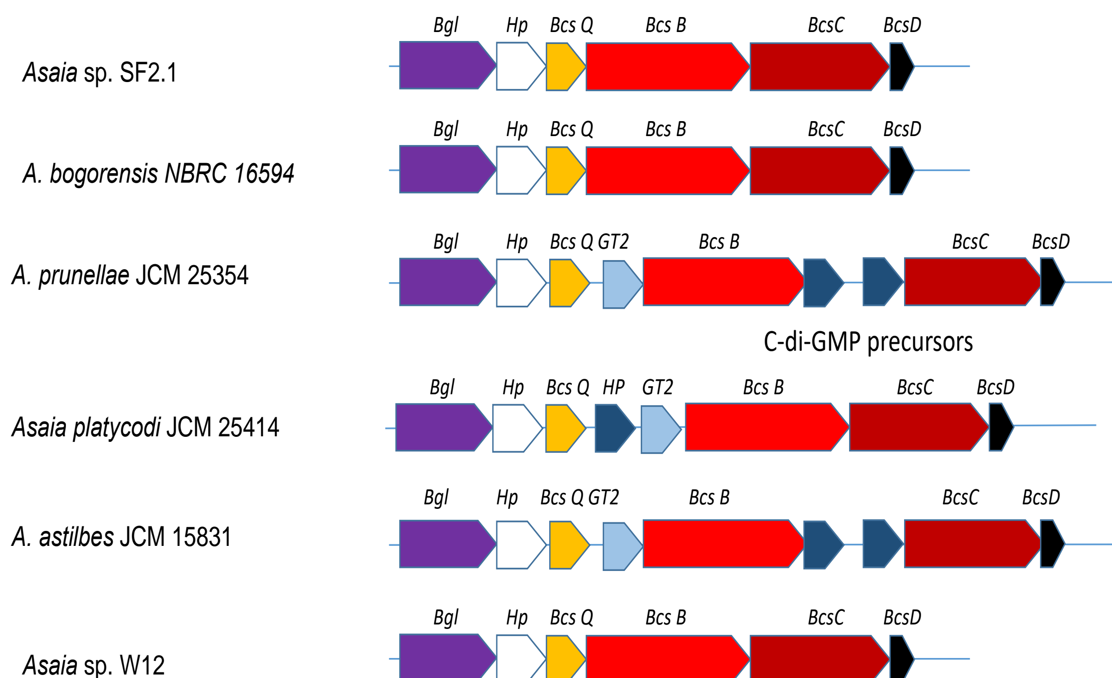


Figure 6. The operon organization of bacterial cellulose biosynthesis in selected *Asaia* strains.

3.4. Regulatory systems in Asaia spp.

The *Asaia* sp. W12 genome encoded 57 two-component system proteins, 169 transcription factor proteins, and 17 other DNA-binding proteins, which possessed the most regulatory protein genes (total 243) among the selected *Asaia* (Table 3). Next to it, *Asaia* sp. As-1742 (isolated from ants) carried at least 235 predicted regulatory genes encoding 46 two-component system proteins, 174 transcription factor proteins, and 15 other DNA-binding proteins. Similar to those in the two insect isolates, the mosquito isolate SF2.1 also had abundant regulatory proteins (up to 213 regulatory protein genes and components). Remarkably, most plant-associated strains possessed relatively fewer regulatory proteins (ranging from 150 to 198) and components (Table 4). For example, the total number of transcriptional regulators (TRs, 81) and one-component systems (OCSs, ranging from 40 to 57) in these insect isolates was relatively higher than those in plant-associated ones (Table 3). However, regardless of their origin, *Asaia* bacteria have comparable numbers of RRs, phosphotransferase proteins (PPs) and sigma factors (SFs) in their genomes, indicating that these regulatory proteins play the fundamental roles in maintaining bacterial metabolism and function.

Table 3. Regulatory systems in *Asaia* spp.

	Predicted regulatory proteins							
	Two component systems			**Transcription factors**				**Other DNA-binding proteins**
	RR	**PP**	**HK**	**OCS**	**RR**	**TR**	**SF**	**ODP**
A. astilbis JCM 15831	19	1	15	27	10	65	5	8
A. prunellae JCM 25354	23	2	14	30	13	69	7	12
A. bogorensis NBRC 16594	25	0	22	42	15	62	7	11
A. bogorensis IPC-01	24	0	20	35	14	64	7	13
A. bogorensis GD-01	21	1	19	39	13	83	7	15
A. platycodi JCM 25414	16	2	16	33	13	59	8	10
Asaia SF2.1	25	1	21	40	15	88	7	16
Asaia sp. W12	30	1	26	57	18	85	9	17
Asaia sp. As-1742	24	1	21	57	14	96	7	15

Abbreviations: HK, histidine kinases; RR, response regulators; PR, phosphotransferase proteins; TR, transcriptional regulators; OCS, one-component systems; SF, sigma factors; ODP, other DNA-binding proteins.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

3.5. Genes involved in other symbiosis traits

Asaia sp. W12 has several well-known genetic traits possibly important for forming the symbiotic relationship with its host mosquitoes. For example, W12 and other selected *Asaia* (except *A. astilbis* JCM 15831) carried two complete cyoABCD operons encoding bo3-type ubiquinol oxidase genes (Table 4). All selected *Asaia* genomes carried acetoin, butanediol and inositol pathways, which possibly contribute to preventing intracellular acidification by changing the metabolism from acid production to the formation of neutral compounds. Several motifs participating in protein-protein interactions were detected in members of the genus *Asaia* (Table S2). The protein-coding genes containing the eukaryotic-like motifs/eukaryotic-like proteins (ELPs) with TPRs (7~10 per genome), ANKs (1-2 per genome), Sel1-like repeats (2~3 per genome) and Fn3 (1~3 per genome). A gene encoding a large adhesin (filamentous hemagglutinin family outer membrane protein) possibly functioning in attachment to host cells was found in all but *Asaia* sp. SF2.1. Moreover, *Asaia* had operons involved in flagella formation.

Table 4. Genes involved in symbiotic traits in *Asaia*

Species	*cyoABCD**a	Acetoin, butanediol and inositol	Flagella	Large adhesin*b
A. bogorensis NBRC 16594	2	+	+	+
A. astilbis JCM 15831	1	+	+	+
A. prunellae JCM 25354	2	+	+	+
A. platycodi JCM 25414	2	+	+	+
A. bogorensis IPC-01	2	+	+	+
A. bogorensis GD-01	2	+	+	+
Asaia sp. As-1742	2	+	+	+
Asaia sp. SF2.1	2	+	+	-
Asaia sp. W12	2	+	+	+

a) The complete operons of *cyoABCD* encode *bo3*-type ubiquinol oxidases with 4 subunits. Most bacteria carry one operon. Very few bacteria (most being symbionts of plants or animals, or are pathogens) have two or more operons; b) Filamentous hemagglutinin family outer membrane protein involved in attachment to the host.

3.6. Comparative genome plasticity

Among the selected *Asaia* spp., only *A. bogorensis* NBRC 16594, *A. astilbis* JCM 15831 and *Asaia* sp. W12 carried intact prophages (Table S3). At least 7 different prophages were predicted in the genome of *Asaia* sp. W12 including one intact (prophage IV), one questionable (prophage VI), and five incomplete ones (Table S3). The largest one (prophage IV, 30.6 Kb) contained a battery of genes encoding the phage tail, head, portal, integrase, lysin, terminase, and other component proteins. Prophage VI (23.4 Kb) carried at least 35 CDS

encoding ApaG protein, O-succinylhomoserine sulfhydrylase, DNA polymerases, primase/helicase proteins, exonuclease, lysozyme several hypothetical proteins and phage structural and assembly proteins. The other predicted prophages (size ranging from 9.0 to 19.8 Kb) seemed to be incomplete because they lacked the full set of structural or assembly elements. However, these "incomplete prophages" may utilize elements from complete ones. Some of the incomplete prophages may be involved in fitness and adaption under certain conditions. For example, the third largest one (prophage I, 19.8 Kb, incomplete) contained transcriptional regulators, siderophore interacting protein, stress response protein and several phage structure proteins (Table S3). Prophage III (19.7 Kb, incomplete) carried multiple transposases, killer proteins (HigA and HigB) and hydrolases. Moreover, the prophage IV may integrate into the host genome long ago because its GC content (61.5%) is only slightly different from the average of GC in the whole genome (60.1%). Remarkably, the GC content of two complete prophages in *A. astilbis* JCM 15831 (56.0% and 56.3% in prophage I and III, respectively) was much lower than the average GC in the genome, indicating that they are recently invaded phage elements. Similarly, a lower GC content (57.8%) of the intact prophage II predicted in *A. bogorensis* NBRC 16594 was found. It is interesting that *A. platycodi* SF2.1 associated with mosquitoes did not carry any complete prophages, which further highlights differential evolution of these *Asaia* species and strains.

Up to 17 genomic islands (GIs) were detected in *Asaia* sp. W12 genome, ranging in size from 4.05 to 72.04 Kb (Table S4). Genes encoding flagellar structural and assembly proteins, prophage components, various enzymes (e.g. lipase, proteases, lysozyme, NAD(P)H oxidoreductase, hydrolases, and amidase), DNA metabolism, transposases, regulators, modification and restriction systems and stress response systems occurred in these GIs (Table S4), indicating that *Asaia* sp. W12 possibly acquired these genes, thereby forming GIs favoring adaptation to diverse environments.

4. Discussion

The mosquito gut harbors a dynamic and diverse microbial community structure because hosts live in both aquatic and territory environments (depending on their developmental stages) and require various diets (microorganisms and detritus as larval food; sugar and blood as adults) (16). Bacteria of the genus *Asaia* are one of the most predominant and stably-associated bacterial members in *A. stephensi*. They carry several genetic traits facilitate formation of symbiotic relations with insects (13, 19, 30). For example, the genome of *Asaia sp. W12* had more signaling components and glycoside hydrolase genes, compared to many other selected *Asaia* genomes. Many well-known eukaryotic-like motifs/eukaryotic-like proteins (ELPs) and large adhesins involved in protein-protein interactions for a range of cellular processes were discovered in its genome. Remarkably, most *Asaia* species carried 2 copies of *cyoABCD*, encoding *bo3*-type ubiquinol oxidase genes (encoding

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

the terminal respiratory chain protein), which was a unique characteristic for them to be symbionts in both plants and insects (31). Metabolic genes including the acetoin and butanediol synthesis possibly play the important roles in preventing intracellular acidification.

Abundant genes encoding regulatory proteins such as HKs (>20), OCSs (>40) and TRs (>81) were found in *Asaia* genomes (Table 4). Differences in the repertoires of these regulatory proteins are likely to facilitate the adaptation of *Asaia* to different hosts and/or could be responsible for the different symbiosis or disease characteristics induced (3, 6, 18). Strain W12 carried the most regulatory protein genes among the selected genomes, which may indicate a high degree of adaptability of this organism to both plant and insect environments (31). *Asaia* sp. F2.1 was reported to locate in multiple sites in mosquitoes including slavery gland, midguts, and reproduction organs (16). These niches obviously represent very different microenvironments physiologically (3, 30). One possible explanation is that these signaling molecules trigger the expression of genes responsible for stress and regulate development of the symbiosis relationship or intercellular communication (13, 15, 32). Several signaling proteins are present in insect-associated *Asaia* while commonly absent in the plant-associated *Asaia* (Table 4). However, the challenge remains to associate these differences in TCS proteins to specific traits of *Asaia*.

Most of the CAZymes in *Asaia* species are involved in metabolizing simple sugars rather than complex plant polysaccharides (33), which is consistent with their living conditions (8). The increases number of glycoside hydrolases in insect-associated *Asaia* strains were only limited to a few peptidoglycan lyases or chitinase-like hydrolases, showing that the adaptation to a plant or insect host was not marked by the acquisition of specific CAZymes (3, 33). It has been well documented that *Asaia* and other acetic acid bacteria produce a bacterial cellulose (BC), which contributes to cell adherence to plant surface or insect epithelium, or other tissues, as one of the biofilm components (34) (35). The physiological role of BC may be important for symbiotic relationship persistence between *Asaia* and mosquitoes by 1) hindering flagellar rotation, 2) limiting cell motility and 3) promoting biofilm formation (34). The mechanisms involved in bacterial cellulose formation and its regulation in *Asaia* are not clear; our findings suggest that bacterial cellulose biosynthesis is diverse among *Asaia* strains, implying that the regulatory mechanisms of BC-related biofilm are different (Figure 5). It is well known that the cellulose matrix extracted from *Asaia* has thinner fibrils, highlighting the biological origination differences from those well-known bacteria (36). The cellulose-producing bacterium *Gluconacetobacter xylinus* has endo-1,4-β-glucanase and β-glucosidase genes (locating adjacent to the cellulose synthesis operon) which played an important role in regulation of the cellulose biosynthesis (37). However, β-glucosidase genes in *A. bogorensis* were outside of the cellulose synthesis, which produced the different type of β-glucosidases (38). The different cellulose productivities between *A. bogorensis* and *G. xylinus* may partly be linked to the different β-1,4-glucanases (38). *bscA* was not present in this operon as typical ones discovered in *E. coli* and others (39). The second gene in the operon (*bcsB*) encodes the catalytic subunit of cellulose synthase

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

(40). Bacterial strains mutated in the *bcsA* locus were found to be deficient in cellulose synthesis due to the lack of cellulose synthase and diguanylate cyclase activities (40).

Investigation of eukaryote-like proteins (ELPs) in *Asaia* genomes may lead to the discovery of novel mechanisms underlying host-symbiont interactions (41). These domains may be acquired through horizontal gene transfer (HGT) from the eukaryote hosts or through convergent evolution (42). ELPs have only been reported in a few symbiotic bacteria in mosquitoes (43) (44, 45). However, they are likely to be more widespread. *Asaia* species carried several genes encoding ELPs (10~13 per genome) including TPR, Ank, Sel1 repeats and Fn3. ELPs in The mosquito symbiont *Wolbachia pipientis* *w*Mel, carried at least 23 ANK-containing genes (44, 45), indicating that the contents of ELPs were beyond the strict pathogens. Further, Klasson et al. (2008) reported that *W. pipientis w*Pip had up to 60 ANK proteins with some of unique presence in its own genome (46). *Cardinium hertigii* (another inherited bacterial symbionts causing CI) was reported to have many ELPs (such as ANK-containing proteins) (47). Large adhesion was discovered in most of the *Asaia* genomes except *Asaia* sp. SF 2.1.

Even though bacterial niches in mosquito tissues and organs are not strictly anaerobic, some tissues may have limited oxygen concentrations. Coon *et al.* (2017) showed that bacteria (wild-type *E. coli*) reduced midgut oxygen concentrations below 5% in both nonsterile or gnotobiotic larvae. However, *E. coli* mutants lacking cytochrome *bd* oxidase genes did not. In the same study, they further demonstrated that hypoxia mediated by bacterial respiratory functioned as an important signal of larval development and ecdysone-induced molting. Together, their findings elucidated the importance of aerobic respiration by gut bacteria in mosquito development. Here, our comparative genome analysis showed that, regardless the species' origination, most *Asaia* genomes carried two distinct *cyo* operons (except *A. astilbis* JCM 15831). Previously, Chouaia *et al*. (2014) found only eight AAB species (animal and plant pathogens or symbionts including *A. bogorensis* NBRC 16594) among the 705 bacterial genomes contained two complete *cyo* operons. The expression level of two *cyo* operons was differently regulated, dependent on their culture conditions. One can assume that the additional copy of fully functional modules of *bo3* and *bd* may allow bacteria *Asaia* to handle with oxidative stress conditions in mosquito. They may actively create hypoxia (a signal for molting) condition in mosquito gut using their enhanced respiratory capability. It will be necessary to test which *cyo* operon(s) directly participate in the above process, which may serve to create tools for interrupting growth of larval mosquitoes. In many microbes, pyruvate can be channeled via alpha-acetolactate and acetoin into 2,3-butanediol, the production of which is induced if oxygen is limited and the pH is lowered (48, 49).

Horizontal gene transfer (HGT), transposons, and prophages often lead to genome plasticity. Among them, prophages are very important for genetic diversification by delivering the functional genes among the different strains.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

For example, some transcriptional regulators, iron uptake proteins, stress response proteins, killer proteins (HigA and HigB) and hydrolase were found in the prophages of *Asaia*. The size of genome, the numbers and size of GIs in W12 were more than those in other selected *Asaia*, implying that W12 had more flexible response to environment change. The flagellar structural and assembly protein genes, various enzymes (e.g. lipase, proteases, lysozyme, NAD(P)H oxidoreductase, hydrolases, and amidase), DNA metabolism, transposases, modification and restriction systems and stress response systems were found in these GIs. *Asaia* bacteria possibly employ a complex set of chemosensory pathways to swim towards different insect organs and adapt to insect physiological demands. However, the detailed mechanisms need to be further investigated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2021 doi:10.20944/preprints202101.0526.v1

Supplementary Materials: Table S1: The CAZy prediction amongst different *Asaia* spp., Table S2: Eukaryotic-like proteins (ELPs) containing motifs predicted in various *Asaia*, Table S3: The putative prophages predicted in various *Asaia*, Table S4: The total genomic islands predicted in various *Asaia* spp.

Author Contributions: Conceptualization, S.C., N.T., B.H., and T.Y.; methodology, S.C.; validation, S.C.; formal analysis, S.C.; investigation, S.C. E.D.W.; resources, S.C. and E.D.W.; data curation, S.C.; writing—original draft preparation, S.C. and N.D.W.; writing—review and editing, S.C., T.Y., N.T., B.H., N.D.W.; supervision, S.C. and N.D.W..; project administration, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by NIH grant R37AI21884

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analyzed during this study are available from the corresponding author on reasonable request.

Acknowledgments: We thank Dr. Jochen Blom at Justus-Liebig-University helped with EDGAR analysis.

Conflicts of Interest: None declared.

References

1. **Favia G, Ricci I, Damiani C, Raddadi N, Crotti E, Marzorati M, Rizzi A, Urso R, Brusetti L, Borin S, Mora D, Scuppa P, Pasqualini L, Clementi E, Genchi M, Corona S, Negri I, Grandi G, Alma A, Kramer L, Esposito F, Bandi C, Sacchi L, Daffonchio D.** 2007. Bacteria of the genus Asaia stably associate with Anopheles stephensi, an Asian malarial mosquito vector. Proceedings of the National Academy of Sciences of the United States of America **104:**9047-9051.
2. **Yamada Y, Katsura K, Kawasaki H, Widyastuti Y, Saono S, Seki T, Uchimura T, Komagata K.** 2000. Asaia bogorensis gen. nov., sp. nov., an unusual acetic acid bacterium in the alpha-Proteobacteria. International Journal of Systematic and Evolutionary Microbiology **50:**823-829.
3. **Bassene H, Niang EHA, Fenollar F, Doucoure S, Faye O, Raoult D, Sokhna C, Mediannikov O.** 2020. Role of plants in the transmission of Asaia sp., which potentially inhibit the Plasmodium sporogenic cycle in Anopheles mosquitoes. Scientific Reports **10:**7144.
4. **Crotti E, Damiani C, Pajoro M, Gonella E, Rizzi A, Ricci I, Negri I, Scuppa P, Rossi P, Ballarini P, Raddadi N, Marzorati M, Sacchi L, Clementi E, Genchi M, Mandrioli M, Bandi C, Favia G, Alma A, Daffonchio D.** 2009. Asaia, a versatile acetic acid bacterial symbiont, capable of cross-colonizing insects of phylogenetically distant genera and orders. Environmental Microbiology **11:**3252-3264.
5. **Chouaia B, Rossi P, Montagna M, Ricci I, Crotti E, Damiani C, Epis S, Faye I, Sagnon N, Alma A, Favia G, Daffonchio D, Bandi C.** 2010. Molecular evidence for multiple infections as revealed by typing of Asaia bacterial symbionts of four mosquito species. Appl Environ Microbiol **76:**7444 - 7450.
6. **Minard G, Tran F, Raharimalala F, Hellard E, Ravelonandro P, Mavingui P, Valiente M.** 2013. Prevalence, genomic and metabolic profiles of Acinetobacter and Asaia associated with field-caught Aedes albopictus from Madagascar. FEMS Microbiol Ecol **83:**63 - 73.
7. **Chen S, Zhao J, Joshi D, Xi Z, Norman B, Walker ED.** 2016. Persistent infection by *Wolbachia w*AlbB has no effect on composition of the gut microbiota in adult female *Anopheles stephensi*. Front Microbiol **7**.
8. **Minard G, Tran FH, Raharimalala FN, Hellard E, Ravelonandro P, Mavingui P, Valiente Moro C.** 2013. Prevalence, genomic and metabolic profiles of Acinetobacter and Asaia associated with field-caught Aedes albopictus from Madagascar. FEMS Microbiology Ecology **83:**63-73.

9. **Favia G, Ricci I, Damiani C, Raddadi N, Crotti E, Marzorati M, Rizzi A, Urso R, Brusetti L, Borin S, Mora D, Scuppa P, Pasqualini L, Clementi E, Genchi M, Corona S, Negri I, Grandi G, Alma A, Kramer L, Esposito F, Bandi C, Sacchi L, Daffonchio D.** 2007. Bacteria of the genus *Asaia* stably associate with *Anopheles stephensi*, an Asian malarial mosquito vector. Proc Natl Acad Sci USA **104:**9047 - 9051.
10. **Damiani C, Ricci I, Crotti E, Rossi P, Rizzi A, Scuppa P, Capone A, Ulissi U, Epis S, Genchi M, Sagnon N, Faye I, Kang A, Chouaia B, Whitehorn C, Moussa G, Mandrioli M, Esposito F, Sacchi L, Bandi C, Daffonchio D, Favia G.** 2010. Mosquito-bacteria symbiosis: the case of *Anopheles gambiae* and *Asaia*. Microb Ecol **60:**644 - 654.
11. **Damiani C, Ricci I, Crotti E, Rossi P, Rizzi A, Scuppa P.** 2008. Paternal transmission of symbiotic bacteria in malaria vectors. Curr Biol **18**.
12. **Capone A, Ricci I, Damiani C, Mosca M, Rossi P, Scuppa P.** 2013. Interactions between *Asaia*, *Plasmodium* and *Anopheles*: new insights into mosquito symbiosis and implications in malaria symbiotic control. Parasite Vector **6**.
13. **Ricci I, Damiani C, Rossi P, Capone A, Scuppa P, Cappelli A, Ulissi U, Mosca M, Valzano M, Epis S, Crotti E, Daffonchio D, Alma A, Sacchi L, Mandrioli M, Bandi C, Favia G.** 2011. Mosquito symbioses: from basic research to the paratransgenic control of mosquito-borne diseases. J Appl Entomol **135:**487-493.
14. **Mancini MVD, Claudia; Short, Sarah M.; Cappelli, Alessia; Ulissi, Ulisse; Capone, Aida; Serrao, Aurelio; Rossi, Paolo; Amici, Augusto; Kalogris, Cristina; Dimopoulos, George; Ricci, Irene; Favia, Guido. .** 2020. Inhibition of Asaia in Adult Mosquitoes Causes Male-Specific Mortality and Diverse Transcriptome Changes. Pathogens **9(5):380.**
15. **Rossi P, Ricci I, Cappelli A, Damiani C, Ulissi U, Mancini MV, Valzano M, Capone A, Epis S, Crotti E, Chouaia B, Scuppa P, Joshi D, Xi Z, Mandrioli M, Sacchi L, O'Neill SL, Favia G.** 2015. Mutual exclusion of *Asaia* and *Wolbachia* in the reproductive organs of mosquito vectors. Parasite Vector **8:**1-10.
16. **Favia G, Ricci I, Marzorati M, Negri I, Alma A, Sacchi L, Bandi C, Daffonchio D.** 2008. Bacteria of the genus *Asaia*: a potential paratransgenic weapon against malaria, p 49-59. *In* Aksoy S (ed), Transgenesis and the management of vector-borne disease doi:10.1007/978-0-387-78225-6_4. Springer New York, New York, NY.
17. **Favia G, Ricci I, Marzorati M, Negri I, Alma A, Sacchi L.** 2008. Bacteria of the genus *Asaia*: A potential weapon against malaria. Adv Exp Med Biol **627**.
18. **Bongio NJ, Lampe DJ.** 2015. Inhibition of *Plasmodium berghei* development in mosquitoes by effector proteins secreted from *Asaia* sp. Bacteria using a novel native secretion signal. PLoS ONE **10:**e0143541.
19. **Crotti E, Rizzi A, Chouaia B, Ricci I, Favia G, Alma A, Sacchi L, Bourtzis K, Mandrioli M, Cherif A, Bandi C, Daffonchio D.** 2010. Acetic Acid Bacteria, Newly Emerging Symbionts of Insects. Applied and Environmental Microbiology **76:**6963-6970.
20. **Epis S, Varotto-Boccazzi I, Crotti E, Damiani C, Giovati L, Mandrioli M, Biggiogera M, Gabrieli P, Genchi M, Polonelli L, Daffonchio D, Favia G, Bandi C.** 2020. Chimeric symbionts expressing a Wolbachia protein stimulate mosquito immunity and inhibit filarial parasite development. Communications Biology **3:**105.
21. **Chen S, Bagdasarian M, Walker ED.** 2015. *Elizabethkingia anophelis*: molecular manipulation and interactions with mosquito hosts. Appl Environ Microbiol doi:10.1128/aem.03733-14.
22. **Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R.** 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res **42:**D206-D214.
23. **Blom J, Kreis J, Spänig S, Juhre T, Bertelli C, Ernst C, Goesmann A.** 2016. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. Nucleic Acids Res doi:10.1093/nar/gkw255**:**W22–W28.
24. **Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH.** 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics **36:**1925-1927.

25. **Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group, Lau BY, Hoad G, Winsor GL, Brinkman FS.** 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Research **45:**W30-W35.
26. **Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B.** 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res **42:**D490-D495.
27. **Terrapon N, Lombard V, Drula É, Lapébie P, Al-Masaudi S, Gilbert HJ, Henrissat B.** 2018. PULDB: the expanded database of polysaccharide utilization loci. Nucleic Acids Res **46:**D677-D683.
28. **Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B.** 2008. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Research **37:**D233-D238.
29. **Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T.** 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Research **47:**W81-W87.
30. **Alonso DP, Mancini MV, Damiani C, Cappelli A, Ricci I, Alvarez MVN, Bandi C, Ribolla PEM, Favia G.** 2018. Genome Reduction in the Mosquito Symbiont Asaia. Genome Biology and Evolution **11:**1-10.
31. **Chouaia B, Gaiarsa S, Crotti E, Comandatore F, Degli Esposti M, Ricci I, Alma A, Favia G, Bandi C, Daffonchio D.** 2014. Acetic Acid Bacteria Genomes Reveal Functional Traits for Adaptation to Life in Insect Guts. Genome Biology and Evolution **6:**912-920.
32. **Shane JL, Bongio NJ, Favia G, Lampe DJ.** 2014. Draft Genome Sequence of Asaia sp. Strain SF2.1, an Important Member of the Microbiome of Anopheles Mosquitoes. Genome Announcements **2:**e01202-01213.
33. **Lacombe-Harvey M-È, Brzezinski R, Beaulieu C.** 2018. Chitinolytic functions in actinobacteria: ecology, enzymes, and evolution. Applied Microbiology and Biotechnology **102:**7219-7230.
34. **Augimeri RV, Varley AJ, Strap JL.** 2015. Establishing a Role for Bacterial Cellulose in Environmental Interactions: Lessons Learned from Diverse Biofilm-Producing Proteobacteria. Frontiers in Microbiology **6**.
35. **Kregiel D, Otlewska A, Antolak H.** 2014. Attachment of <i>Asaia bogorensis</i> Originating in Fruit-Flavored Water to Packaging Materials. BioMed Research International **2014:**514190.
36. **Penttilä PA, Imai T, Capron M, Mizuno M, Amano Y, Schweins R, Sugiyama J.** 2018. Multimethod approach to understand the assembly of cellulose fibrils in the biosynthesis of bacterial cellulose. Cellulose **25:**2771-2783.
37. **Liu M, Liu L, Jia S, Li S, Zou Y, Zhong C.** 2018. Complete genome analysis of Gluconacetobacter xylinus CGMCC 2955 for elucidating bacterial cellulose biosynthesis and metabolic regulation. Scientific Reports **8:**6266.
38. **Kumagai A, Mizuno M, Nozaki K, M. Saxena I, Amano Y.** 2011. Comparative Study on the Ability to Produce Gentiobiose in Cellulose-producing Bacteria <i>Asaia bogorensis</i> and <i>Gluconacetobacter xylinus</i>. Journal of Applied Glycoscience **58:**147-150.
39. **Krasteva PV, Bernal-Bayard J, Travier L, Martin FA, Kaminski P-A, Karimova G, Fronzes R, Ghigo J-M.** 2017. Insights into the structure and assembly of a bacterial cellulose secretion system. Nature Communications **8:**2065.
40. **Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, Amikam D, Benziman M, Gelfand DH, Meade JH, Emerick AW.** 1990. Genetic organization of the cellulose synthase operon in Acetobacter xylinum. Proceedings of the National Academy of Sciences **87:**8130.
41. **Díez-Vives C, Moitinho-Silva L, Nielsen S, Reynolds D, Thomas T.** 2017. Expression of eukaryotic-like protein in the microbiome of sponges. Molecular Ecology **26:**1432-1451.
42. **Wybouw N, Pauchet Y, Heckel DG, Van Leeuwen T.** 2016. Horizontal Gene Transfer Contributes to the Evolution of Arthropod Herbivory. Genome Biology and Evolution **8:**1785-1801.
43. **Siozios S, Ioannidis P, Klasson L, Andersson SGE, Braig HR, Bourtzis K.** 2013. The Diversity and Evolution of Wolbachia Ankyrin Repeat Domain Genes. PLOS ONE **8:**e55390.

44. **Walker T, Klasson L, Sebaihia M, Sanders MJ, Thomson NR, Parkhill J, Sinkins SP.** 2007. Ankyrin repeat domain-encoding genes in the wPip strain of Wolbachia from the Culex pipiens group. BMC Biology **5**:39.
45. **Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, Brownlie JC, McGraw EA, Martin W, Esser C, Ahmadinejad N, Wiegand C, Madupu R, Beanan MJ, Brinkac LM, Daugherty SC, Durkin AS, Kolonay JF, Nelson WC, Mohamoud Y, Lee P, Berry K, Young MB, Utterback T, Weidman J, Nierman WC, Paulsen IT, Nelson KE, Tettelin H, O'Neill SL, Eisen JA.** 2004. Phylogenomics of the Reproductive Parasite Wolbachia pipientis wMel: A Streamlined Genome Overrun by Mobile Genetic Elements. PLOS Biology **2:**e69.
46. **Klasson L, Walker T, Sebaihia M, Sanders MJ, Quail MA, Lord A, Sanders S, Earl J, O'Neill SL, Thomson N, Sinkins SP, Parkhill J.** 2008. Genome Evolution of Wolbachia Strain wPip from the Culex pipiens Group. Molecular Biology and Evolution **25:**1877-1887.
47. **Penz T, Schmitz-Esser S, Kelly SE, Cass BN, Müller A, Woyke T, Malfatti SA, Hunter MS, Horn M.** 2012. Comparative Genomics Suggests an Independent Origin of Cytoplasmic Incompatibility in Cardinium hertigii. PLOS Genetics **8:**e1003012.
48. **Zhang L, Sun J, Hao Y, Zhu J, Chu J, Wei D, Shen Y.** 2010. Microbial production of 2, 3-butanediol by a surfactant (serrawettin)-deficient mutant of Serratia marcescens H30. J Ind Microbiol Biotechnol **37**.
49. **Köpke M, Mihalcea C, Liew F, Tizard JH, Ali MS, Conolly JJ, Al-Sinawi B, Simpson SD.** 2011. 2,3-butanediol production by acetogenic bacteria, an alternative route to chemical synthesis, using industrial waste gas. Applied and environmental microbiology **77:**5467-5475.