Complete genome and gene expression analyses of *Asaia bogorensis* reveal unique responses to culture with mammalian cells as a potential opportunistic human pathogen

Mikihiko Kawai¹,², Norie Higashiura¹,², Kimie Hayasaki¹, Naruhei Okamoto¹, Akiko Takami¹,², Hideki Hirakawa³, Kazunobu Matsushita⁴, and Yoshinao Azuma¹,*

¹Faculty of Biology-Oriented Science and Technology, Kindai University, Kinokawa, Wakayama, Japan, ²Advanced Low Carbon Technology Research and Development Program (ALCA), Japan Science and Technology Agency (JST), Tokyo, Japan, ³Kazusa DNA Institute, Kisarazu, Chiba, Japan, and ⁴Department of Biological Chemistry, Faculty of Agriculture, Yamaguchi University, Yamaguchi, Yamaguchi, Japan

*To whom correspondence should be addressed. Tel. +81 736-77-0345. Fax. +81 736-77-4754.
E-mail: azuma@waka.kindai.ac.jp

Edited by Prof. Masahira Hattori

Received 15 June 2015; Accepted 13 August 2015

Abstract

*Asaia bogorensis*, a member of acetic acid bacteria (AAB), is an aerobic bacterium isolated from flowers and fruits, as well as an opportunistic pathogen that causes human peritonitis and bacteraemia. Here, we determined the complete genomic sequence of the *As. bogorensis* type strain NBRC 16594, and conducted comparative analyses of gene expression under different conditions of co-culture with mammalian cells and standard AAB culture. The genome of *As. bogorensis* contained 2,758 protein-coding genes within a circular chromosome of 3,198,265 bp. There were two complete operons encoding cytochrome *bo*₃-type ubiquinol terminal oxidases: *cyoABCD-1* and *cyoABCD-2*. The *cyoABCD-1* operon was phylogenetically common to AAB genomes, whereas the *cyoABCD-2* operon belonged to a lineage distinctive from the *cyoABCD-1* operon. Interestingly, *cyoABCD-1* was less expressed under co-culture conditions than under the AAB culture conditions, whereas the converse was true for *cyoABCD-2*. *Asaia bogorensis* shared pathogenesis-related genes with another pathogenic AAB, *Granulibacter bethesdensis*, including a gene coding pathogen-specific large bacterial adhesin and additional genes for the inhibition of oxidation and antibiotic resistance. Expression alteration of the respiratory chain and unique hypothetical genes may be key traits that enable the bacterium to survive under the co-culture conditions.

**Key words:** genomics, transcriptome, ubiquinol terminal oxidase, opportunistic pathogen, acetic acid bacteria

1. Introduction

Acetic acid bacteria (AAB) are obligate aerobes, and presently comprise 14 genera assigned to the family Acetobacteraceae, including *Acetobacter* (Ac.), *Gluconobacter* (Go.), *Gluconacetobacter* (Ga.), *Granulibacter* (Gr.), *Konagataeibacter* (Ko., former *Gluconacetobacter*), *Asaia* (As.), *Acidomonas* (Am.), *Kozakia*, *Swaminathania*, *Saccharibacter*, *Neoasaa*, *Tantischaroenia*, *Ameyamaea*, and *Neokomagataea*.¹² AAB incompletely oxidize a variety of sugars and alcohols, and
accumulate large amounts of the corresponding oxidation products in their environment or culture media. These traits of AAB, known as oxidative fermentation, are utilized in the industrial production of vinegar, L-sorbose, and bacterial cellulose.

Oxidative fermentation reactions occur through a variety of primary dehydrogenases in the periplasmic side of the inner cytoplasmic membrane. Electrons that are generated are transferred to ubiquinone and the reduced ubiquinol is utilized by terminal oxidases in the electron transfer chains to reduce oxygen to water and harness the energy to translocate protons across the cytoplasmic membrane. There are two types of terminal oxidases in AAB: cytochrome $b_6f_{0}$-type ubiquinol oxidase and $b_{2}$-type ubiquinol oxidase, which consist of four subunits encoded by the $cyoABCD$ operon (synonym: $cyoBACD$) and two subunits encoded by the $cydAB$ genes, respectively. Cytochrome $b_{6}$-$f_{0}$-type ubiquinol oxidase is a member of the haeme-copper oxidase superfamily, which includes cytochrome $c$ oxidase in aerobic respiration of mitochondria and aerobes.

Asaia bogorensis was first isolated from tropical flowers, and is an unusual member of AAB that oxidizes acetate and lactate to carbon dioxide and water, but not ethanol to acetic acid. A strain of As. bogorensis isolated from a flower of the orchid tree Bunga bauhinia in Bogor, Indonesia, was assigned to a type strain (≡NRBC 16594 ≡ NRIC 0311 ≡ JCM 105697). Based on DNA sequences of 16S rRNA genes, Asaia is phylogenetically located near Gluconobacter and Acetobacter, but is distant from the genera Granulibacter and Gluconacetobacter. Biochemical analysis revealed that As. bogorensis contains sugar- and sugar-alcohol-oxidizing enzymes that are specific to the respiratory chain, such as quinoprotein glyceraldehyde dehydrogenase; however, it lacks quinoprotein alcohol dehydrogenase (ADH). In addition, As. bogorensis was shown to contain a cytochrome $b_{2}$-type ubiquinol oxidase as the sole terminal oxidase in the respiratory chain.

Asaia bogorensis can also be an opportunistic pathogen in humans. Asaia bogorensis was first reported as an opportunistic human pathogen when it was isolated from the peritoneal fluids of a patient on peritoneal dialysis who showed a consistent clinical status. Asaia bogorensis is a natural but not clinical isolate. Its gene content was compared with those of other AAB genomes, and the As. bogorensis transcriptome was compared under different culture conditions.

### 2. Materials and methods

#### 2.1. Genome sequencing

Asaia bogorensis strain NRBC 16594 (≡NRBC 0311 ≡ JCM 105697) was obtained from the NITE Biological Resource Center (NRBC, Japan). The genomic DNA sequence of As. bogorensis was determined by a whole-genome shotgun strategy, using a method described previously with slight modifications. Briefly, 4.2 µg of genomic DNA was extracted from As. bogorensis that was cultured aerobically in a YPGD medium (1.0% yeast extract, 1.0% polypeptide, 2.0% glycerol, and 0.5% glucose) at 30°C. After DNA fragmentation and isolation (average fragment length 189 bp), 6,345,662 reads (75 bases per read) were generated using the GAIIx sequencing system (Illumina, San Diego, CA, USA). Short DNA reads were assembled using the Velvet 1.0.3 software. Among different k-mers (minimum-overlapping length) that were tested, 49 k-mer in size provided the best result in terms of N50, resulting in 50 contigs, with a total of 3,173,055 bp (coverage: 116x). Seventy-two primers were prepared based on relatively longer 36 contigs out of the 50 ones, and DNA amplification by PCR was performed using all combinations of both ends of the contigs. DNA sequences of the fragments amplified were determined at least once in both strands by direct sequencing of PCR products using an ABI Prism 3100 Genetic Analyzer (Life Technologies). After all the gaps were filled with appropriate qualities, assembling of the genome DNA was validated by gel electrophoresis of its genomic DNA and re-mapping of the short DNA reads, resulted in no inconsistency in the assembling. No plasmids were detected in this process.

#### 2.2. Gene finding and annotation

Protein-coding gene finding was performed using three programmes: GeneMarkS, MetaGeneAnnotator, and GLIMMER3. Protein-coding genes are defined when genomic loci were predicted as protein-coding genes (i) by two or more programmes, or (ii) by only one programme but supported by TIGRFAMs hits, or by distinctive RNA expression. The start positions of the genes were determined in two ways: (i) the same position was predicted as a start position by two or more programmes, or (ii) when all three programmes differently predicted start positions, the preferred was chosen from the software in the following order: GeneMarkS, MetaGeneAnnotator, and GLIMMER3. Prediction of rRNA and tRNA genes was performed using RNAmmer and tRNAscan-SE, respectively. Gene products were annotated on the basis of similar proteins of Am. methanolica and Ko. medellinensis, or most similar proteins in databases using the BLASTP program, as well as referring to conserved protein families that were predicted using the Hmmer3 program with a noise cut-off option (–cut_tc) against TIGRFAMs and PFAM databases.

#### 2.3. Bioinformatics analyses

To analyse a taxonomic distribution of specific genes, a platform phylogenetic tree with whole prokaryotic organisms was illustrated using 42 single-copy genes conserved among As. bogorensis and the complete genome sequences of 705 prokaryotic species representing 705 different genera, including 80 archaea and 625 bacteria, using a previously described method (Supplementary Tables S1 and S2). The 42 genes were selected on the basis of the following two requirements: (i) the gene should be conserved among >95% of both archaea and bacteria and (ii) average copy numbers of the gene in both archaea...
and bacteria should be 1.10 or lower. The 42 gene products were aligned by MAFFT\textsuperscript{35} and concatenated, and phylogenetic distances of the species were then calculated using the FastTree program.\textsuperscript{36}

Phylogenetic profiles were constructed by allocation of each gene in whole genomes of six AAB, including As. bogorensis, Gr. bethesdensis, Ac. pasteurianus, Go. oxydans, Ko. medellinensis, and Ga. diazotrophicus, to a certain gene family at the amino acid level using the Hmmer\textsuperscript{3} program\textsuperscript{37} on TIGRFAMs\textsuperscript{27} and PFAM\textsuperscript{32} databases. To predict functional and evolutionary interactions among gene families, the gene families were classified into categories, such as gene families existing in only As. bogorensis but not in the other five AAB, ones in only pathogenic AAB (As. bogorensis and Gr. bethesdensis), and ones shared with AAB isolated from natural samples by a mean in which not only the presence or absence of a protein family, but also the copy number of paralogous genes in each AAB genome belonging to a protein family was considered.\textsuperscript{37} Statistical significance of this categorization was assessed using a Wilcoxon signed-rank test, which takes into account the magnitude of the observed differences.

2.4. Total RNA preparation
Asaia bogorensis was grown under three different conditions. For cells prepared under ‘co-culture conditions’, HEK293 cells (Human Embryonic Kidney 293 cells) and McCoy cells (mouse fibroblast cells) were pre-cultured in 2 ml of Dulbecco’s Modified Eagle Medium F-12 (DMEM, Sigma-Aldrich) with 5% foetal calf serum (FCS, Canon) without antibiotics in six-well plates until 60% confluence. Cells were then collected from the supernatant by centrifugation and the mammalian cells were removed from the culture. The supernatant was centrifuged at 200 × g for 5 min, and the supernatant was then grown in small samples under three different conditions. For cells prepared under ‘tissue culture conditions’, 2 ml of the pre-cultured As. bogorensis was added to a six-well plate without mammalian cells, and the cells were treated in the same way as cells prepared under ‘co-culture conditions’. As a control condition (‘AAB condition’), As. bogorensis cells were cultured in YPGD at 30°C with aeration at 200 rpm until OD\textsubscript{600} = 1.0 was achieved.

2.5. Transcriptome analysis
Total RNA was isolated from the bacterial cells using the RNaseasy Protect Bacteria Mini Kit (QIAEN, Germantown, MD, USA), and sequenced using HiSeq2000 (Illumina) according to the manufacturer’s protocols, with the exception of a treatment to decrease a copious amount of cDNA derived from rRNA using a duplex-specific nuclease.\textsuperscript{38} After base calling and quality filtering using CASAVA (Illumina), the sequence was mapped to the As. bogorensis genomic DNA sequence and processed using the Bowtie\textsuperscript{39} and SAMtools software package.\textsuperscript{40} The number of reads that were mapped for each gene was calculated using the BedTools software.\textsuperscript{41} Reads that were mapped to 16S and 23S rRNA genes were omitted from further analyses. Genes that were differentially expressed between two sets of samples were estimated using the edgeR package\textsuperscript{42} after normalization using the TCC package.\textsuperscript{43} For cells that were treated under ‘co-culture conditions’, four samples (two samples containing HEK293 and two samples containing McCoy cells) were prepared from independent cultures. Adjustments of false discovery rates for P-values were performed for multiple comparisons using a multi-test correction. All statistical analyses were conducted using the R statistical package. After data averaging and normalization, Reads per kilobase gene per million reads (RPKM) values were used for graphical presentations.

### Table 1. Comparison of the As. bogorensis NBRC 16594 genome with other complete Acetobacteraceae genomes

| Species          | Strain         | Plasmid | Length (bp) | GC (%) | CDS\textsuperscript{5} | tRNA | rRNA operon | Transposase\textsuperscript{b} | Reference | Trait                           |
|------------------|----------------|---------|-------------|--------|------------------------|------|-------------|-------------------------------|-----------|---------------------------------|
| As. bogorensis   | NBRC 16594     | 0       | 3,198,265   | 59.8   | 2,758                  | 5    | 14 (4, 9)   | This work                     |           | Isolated from the flower of the orchid tree |
| Gr. bethesdensis | CGDNIIH1       | 0       | 2,708,335   | 59.1   | 2,437                  | 3    | 15 (4, 12)  | Greenberg et al.\textsuperscript{20} |           | Chronic granulomatous disease |
| Ac. pasteurianus | IFO 3283-01    | 6       | 3,340,249   | 53.1   | 3,050                  | 5    | 268 (52, 47)| Azuma et al.\textsuperscript{13} |           | Used for brewing traditional vinegar |
| Go. oxydans      | 621H           | 5       | 2,922,384   | 60.8   | 2,664                  | 4    | 104 (27, 24)| Prust et al.\textsuperscript{16} |           | Used for vitamin C synthesis   |
| Ko. medellinensis | NBRC 3288      | 7       | 3,308,936   | 60.6   | 3,195                  | 5    | 204 (24, 57)| Ogino et al.\textsuperscript{17} |           | Isolated from vinegar          |
| Ga. diazotrophicus | PAL 5         | 1       | 3,914,947   | 66.3   | 3,501                  | 4    | 129 (22, 26)| Giongo et al.\textsuperscript{19} |           | Endophyte of sugarcane         |

\textsuperscript{a}As: Asaia; Gr.: Gramlibacter; Ac.: Acetobacter; Go.: Gluconobacter; Ko.: Komagataeibacter; Ga.: Gluconacetobacter.
\textsuperscript{b}The number of transposase genes. Numbers in the parentheses represent the copy number of the largest family and the number of families.
\textsuperscript{c}One tRNA (Arg) is split by a sequence homologous to a group I intron.
\textsuperscript{d}Formerly known as Gluconacetobacter xylinus NBRC 3288.\textsuperscript{2}
from *Gr. bethesdensis*, suggesting that pathogenicity and the common traits of *As. bogorensis* and *Gr. bethesdensis* independently evolved during the evolutionary adaptation of the two AAB.

### 3.2. Characteristic genes in the *As. bogorensis* genome

To clarify the characteristic genes, especially potentially pathogenic genes (described as pathogenic genes below), data from the phylogenetic profiles were used (Supplementary Table S3). Several genes are likely to be associated with pathogenicity of *As. bogorensis* in the gene families of only two pathogenic AAB (*As. bogorensis* and *Gr. bethesdensis*), and in three AAB able to habitat in eukaryotic organisms (*As. bogorensis*, *Gr. bethesdensis*, and *Ga. diazotrophicus*; Table 2). For example, *As. bogorensis* and *Gr. bethesdensis* contained genes encoding a large adhesin protein and its associated haemolysin.

---

**Figure 1.** Circular representation of the *As. bogorensis* NBRC 16594 genome. From the outermost concentric circles: protein-coding sequences clockwise (red) and anti-clockwise (blue); tRNA genes (purple), rRNA genes (green), and other non-coding RNA genes (orange); transposases (blue) and the horizontally acquired region (high probability to low were shown in red to grey); genes present only in *As. bogorensis* and *Gr. bethesdensis* but absent in three acetate forming AAB (*Ac. pasteurianus*, *Ko. medellinensis*, and *Go. oxydans*) (orange); genes present in *As. bogorensis* and *Gr. bethesdensis* (blue); genes present only in *As. bogorensis* (magenta). The inner histograms show a plot of clockwise GC content and GC-skew (high and low regions in yellow green and violet, respectively). Horizontally acquired regions were predicted using Alien_hunter.

**Figure 2.** Phylogenetic tree based on 42 prokaryotic universal genes. Amino acid sequences of 42 conserved prokaryotic genes were concatenated and aligned. Numbers on branching points are confidence values of the Shimodaira-Hasegawa test (SH-test). Scale bar indicates substitutions per amino acid residue (change/amino acid site).
activator protein, peroxidase genes, catalase genes, and antibiotic-resistant genes (beta-lactamases and penicillin amidase), as well as a type II secretion system. The type II secretion system and the large adhesin family genes are prevalent among a wide range of animal and plant pathogens.\textsuperscript{46,47} Interestingly, the \textit{As. bogorensis} genome contained two complete operons for the \textit{bo}$_{1}$-type ubiquinol oxidase, which is described below (see Section 3.4). The \textit{As. bogorensis} genome also uniquely contained an operon with two genes that were homologous to rhodopsin and beta-carotene \textit{15,15}]-monooxygenase, which may function as a light-driven ion pump and produce a retinal cofactor of rhodopsin, respectively. \textit{Asaia} \textit{bogorensis} and \textit{Ko. medellinensis} are known to produce cellulose fibres,\textsuperscript{17,48} and both contained a cellulose synthase \textit{bcs} operon (synonyms: \textit{cel} and \textit{ces}) composed of \textit{bcsAB}, \textit{bcsC}, and \textit{bcsD} genes (Asbog\textunderscore 01433–Asbog\textunderscore 01431). Those operons were classified into the gene families conserved in only \textit{As. bogorensis} and \textit{Ko. medellinensis} (Supplementary Table S3A).

### 3.3. Transcriptome analysis

\textit{Asaia} \textit{bogorensis} and \textit{Am. methanolica} have been reported as opportunistic human pathogens.\textsuperscript{8–10} \textit{Asaia bogorensis} NBRC 16594 and \textit{Am. methanolica} NBRC104435 were shown to grow at 37°C under ‘tissue culture conditions’ (Supplementary Fig. S1).\textsuperscript{23} To identify \textit{Asaia} \textit{bogorensis} genes that contribute to its growth in human bodies, comparative transcriptome analysis by RNA-seq was performed using \textit{Asaia} \textit{bogorensis} cultured under three different conditions: ‘co-culture conditions’, involving tissue culture at 37°C with mammalian cells, ‘tissue culture conditions’, involving the same culture conditions at 37°C but without mammalian cells, and in a YPGD medium at 30°C (‘AAB condition’). The results for ‘co-culture conditions’ and ‘tissue culture conditions’ showed high reproducibility for RNA-seq experiments (Supplementary Fig. S2). For the 2,833 genes of \textit{Asaia} \textit{bogorensis}, the expression of 101 and 78 genes was significantly higher and lower under ‘co-culture conditions’ than under the ‘AAB condition’, respectively (Fig. 3 and Supplementary Table S4). Individual results for genes that encode terminal oxidases, primary dehydrogenases, genes that are unique to \textit{Asaia} \textit{bogorensis}, and genes related to stress responses are described below.

### 3.4. Genes for electron transport chain

Two complete operons of \textit{cyoABCD} encoding \textit{bo}$_{1}$-type ubiquinol oxidases were identified in the \textit{As. bogorensis} genome and assigned as \textit{cyoABCD}-1 (Asbog\textunderscore 01765–Asbog\textunderscore 01762) and \textit{cyoABCD}-2 (Asbog\textunderscore 00581–Asbog\textunderscore 00578) (Table 2 and Fig. 4A). Within 705 prokaryotic species that represent 705 different genera (Supplementary Tables S1 and S2), 98 species contained \textit{cyoABCD} genes, but only eight, besides \textit{Asaia} \textit{bogorensis}, contained two or more complete \textit{cyo} operons. Most of the eight species are animal and plant pathogens or symbionts, including \textit{Bordetella bronchiseptica} RB50, \textit{Burkholderia pseudomallei} K96243, \textit{Stenotrophomonas maltophilia} K279a, and \textit{Ralstonia solanacearum} GMI100. Other AAB contained only one operon for the oxidase in each genome. The \textit{cyoABCD}-1 operon belonged to a monophyletic clade formed by \textit{cyoABCD} genes from the other AAB (Fig. 4B and C), whereas the \textit{cyoABCD}-2 operon belonged to a clade composed of various \textit{Proteobacteria} (Fig. 4B and D). Genes encoding haeme O synthase, \textit{cyoE}, and assembly factor \textit{surf1} were not in the flanking regions for both operons.

The structural and mutagenic studies performed on the \textit{bo}$_{1}$-type ubiquinol oxidase of \textit{Escherichia coli} strongly suggest that the ubiquinol-binding site is in the transmembrane helices I and II, which contain a cluster of conserved polar residues in the subunit I, \textit{CyoB}.\textsuperscript{36} Most of the residues conserved among ubiquinol oxidase \textit{CyoB} were also conserved in both \textit{CyoB}-1 and \textit{CyoB}-2, except for three residues, 75E, 78V, and 102L, in \textit{CyoB}-2 (Fig. 4E). Therefore, both \textit{CyoB}-1 and \textit{CyoB}-2 may be functional if those genes are expressed properly. Comparative transcriptome analysis indicated that \textit{Asaia} \textit{bogorensis} expressed lower levels of \textit{cyoABCD}-1 under ‘co-culture

### Table 2. Summary of characteristic gene content of \textit{As. bogorensis} NBRC 16594 among six Acetobacteraceae genomes

| Species          | Strain        | Large adhesin | Peroxidase | Catalase | Penicillin amidase | Beta-lactamase$^a$ | Type II secretion | \textit{cyoABCD} |
|-----------------|---------------|---------------|------------|----------|-------------------|-------------------|------------------|-----------------|
| \textit{As. bogorensis} | NBRC 16594   | 1             | 1          | 2        | 1                 | 3                 | 1                | 2               |
| \textit{Gr. bethesdensis} | CGDNIH1      | 1             | 1          | 3        | 1                 | 3                 | 1                | 1$^b$           |
| \textit{Ga. diazotrophicus} | PAI 5        | 0             | 0          | 0        | 0                 | 2                 | 0                | 1$^b$           |
| \textit{Ac. pasteurianus} | IPO 3283-01  | 0             | 0          | 1        | 0                 | 2                 | 0                | 0               |
| \textit{Go. oxydans} | 621H         | 0             | 0          | 1        | 0                 | 2                 | 0                | 0               |
| \textit{Ko. medellinensis} | NBRC 3288    | 0             | 0          | 1        | 0                 | 2                 | 0                | 1               |

$^a$Beta-lactamase family (PFAM accession PF13354.1).

$^b$Contains another terminal oxidase, a cytochrome \textit{c} oxidase, as previously reported.\textsuperscript{81}

**Figure 3.** Comparison of whole RNA-seq results of different culture conditions. RPKM values from RNA-seq analyses under the ‘AAB’ and ‘co-culture conditions’ were plotted as X-Y axes for each gene. Coloured dots indicate genes showing significantly different expression ($P=0.001$, listed in Supplementary Table S4), and categories are shown inside of the graph.
conditions' than under the 'AAB condition', whereas the opposite was true for cyoABCD-2 (Figs 3 and 5A). The results indicate that cyoABCD-1 and cyoABCD-2 could contribute to growth in the tissue culture condition by altering the expressions.

*Escherichia coli* contains cyo and cyd operons that encode cytochrome bo3 oxidase and cytochrome d oxidase, respectively. Expression of *E. coli* cyd genes is upregulated under an oxygen-limited condition, so that the organism can survive under oxygen starvation conditions.50 *Asaia bogorensis* also has cydAB of cytochrome bd oxidase that belongs to the subfamily of cyanide-insensitive terminal quinol oxidases. In contrast, expression levels of *As. bogorensis* cyd genes were relatively consistent under the different conditions (Fig. 5A).

*Asaia bogorensis* and three other AAB, such as *Ac. pasteurianus*, *Ko. medellinensis*, and *Gr. bethesdensis*, contain one nuo operon for type I NADH dehydrogenases (Complex I), whereas *Ga. diazotrophicus* contains two, and *Go. oxydans* does not contain any. The phylogenetic analysis based on concatenated amino acid sequences encoded by six-well-conserved nuo genes (nuoHIJKLN) indicated that nuo operon of *As. bogorensis* and one operon of *Ga. diazotrophicus* formed a clade that was distinctive from one composed by nuo operons of *Ac. pasteurianus*, *Ko. medellinensis*, *Gr. bethesdensis*, and the other operon of *Ga. diazotrophicus* (Supplementary Fig. S3). Taken together with the phylogenetic analysis presented in Fig. 2, it appears that *As. bogorensis* and *Ga. diazotrophicus* gained an extra nuo gene,

---

**Figure 4.** Cytochrome bo3-type ubiquinol oxidases. (A) Genetic organization of cyoABCD-1 and cyoABCD-2 operons of *Asaia bogorensis*. (B) The phylogenetic analysis based on an alignment of a concatenated amino acid sequence of four genes, cyoABCD. Clades (i) and (ii) as indicated by broken lines include cyoABCD-1 and cyoABCD-2, respectively. (C and D) enlarged views of the clades (i) and (ii), respectively. (E) Putative ubiquinol-binding site. Residues in bold denote the conserved residues for the ubiquinol-binding site.49 Asterisks denote alteration of residues in *Asaia bogorensis* cyoB-2. The numbers indicate amino acid positions of *Pseudomonas denitrificans* cytochrome aa3 (PDB 1AR1).
possibly by the horizontal gene transfer, and *As. bogorensis* and *Go. oxydans* lost one *nuo* operon that is common among AAB.

The *As. bogorensis* genome included three more operons for the electron transport: *ndh* (type II NADH-dehydrogenase [NDH-2]), *sdh* (membrane-associated FAD-dependent succinate dehydrogenase [complex II]), and *atp* (an F1F0-type ATP synthase [complex V]) operons. Expression of the *nuo*, *sdh*, and *atp* operons were relatively lower under ‘co-culture conditions’ than under the ‘AAB condition’, whereas *ndh* expression was higher under ‘co-culture conditions’ than the ‘AAB condition’ (Supplementary Table S4 and Fig. S4A). These results regarding *sdh* operon differed from a previous report that showed no activities were detected. It suggests that production of complex II is too low to detect biochemical activity, or that genetically it lost its function.

### 3.5. Membrane-associated primary dehydrogenases

AAB genomes contain several genes that encode membrane-associated quinone reductases coupling substrate oxidation with quinone reduction (primary dehydrogenases). *Asaia bogorensis* genome included genes for six membrane-associated primary dehydrogenases assigned based on subunit structures and amino acid sequences, i.e. SLDH1, GADH, GDH, GLDH, IDH (PQQ1), and LDH (Table 3). However, genes that encode membrane-associated ADH and aldehyde dehydrogenase (ALDH) were not found in the *As. bogorensis* genome. This result was consistent with reports that *As. bogorensis* does not oxidize ethanol to acetic acid.

The pyruvate oxidase operon (*Asbog_02184–Asbog_02187*) was unique to *As. bogorensis* and showed features conserved in membrane-associated primary dehydrogenases. *Asbog_02184* encoded thiamin- and FAD-binding dehydrogenase similar to a membrane-associated cytoplasmic protein, PoxB (Supplementary Fig. S5A and B). *Asbog_02185* encoded a protein with seven transmembrane helices. *Asbog_02186* encoded a protein with a domain of gluconate 2-dehydrogenase subunit 3 similar to FAD-binding small subunits of primary dehydrogenases, such as *Asbog_00352* for *sldh* and *Asbog_02524* for *gadh*. *Asbog_02187* encoded a glucose-methanol-choline family oxidoreductase similar to FAD-binding large subunits of primary dehydrogenases, such as *Asbog_00353* for *sldh* and *Asbog_02525* for *gadh*. The pyruvate oxidase seems a membrane-associated primary dehydrogenase locating in periplasmic space. But, the possibility has not been excluded that the pyruvate oxidase localizes on the cytoplasmic side of the cytoplasmic membrane, because it is still very difficult to predict topological locations of proteins by bioinformatics approaches.

---

**Figure 5.** Gene expression in different conditions. (A) Ten genes that encode subunits of terminal oxidases. (B) Five heat shock genes. In every panel, lanes 1–5 indicate various conditions (lanes 1 and 2, co-culture with HEK293; lanes 3 and 4, co-culture with McCoy; lane 5, ‘tissue culture condition’; lane 6, ‘AAB condition’), and vertical axis shows values of RPKM.
Interestingly, the pyruvate oxidase operon was highly expressed under ‘co-culture conditions’ comparing with other two conditions, whereas gene expression levels of the other primary dehydrogenase were not altered under the three culture conditions (Supplementary Table S4 and Figs S4B and S5C). The expression profile of pyruvate oxidase operon was similar to one of cyo-2 under the three conditions (Fig. S5A). Alterations in gene expression for the pyruvate oxidase may be related to culturing statuses, such as nutrition, oxygen dissolved in media, and temperature, as well as interaction with mammalian cells; however, the functional implications of the alterations are unknown.

The uncharacterized PQQ-containing oxidoreductase PQQ2 subfamily (e.g. GOX1969 and Asbog_01692) partially shares its structure with BamB, which is a lipoprotein component of the BAM (beta-barrel assembly machinery) complex and catalyses the essential process of assembling outer membrane proteins. Thus, proteins in the PQQ2 subfamily were omitted from a list of primary dehydrogenases in this work. Similarly, quinone reductases involved in processes other than oxidative fermentation were omitted from the data (Table 3), such as FAD-dependent succinate dehydrogenase (Asbog_00368–Asbog_00371), FAD-dependent malate: quinone oxidoreductase (Asbog_01178), and proline dehydrogenase (Asbog_02561 and Asbog_00494).

Table 3. Primary dehydrogenases in As. bogorensis NBRC 16594

| Family | Cofactor | Subfamily | Genes |
|--------|----------|-----------|-------|
| Sorbitol dehydrogenase | FAD | SLDH | Asbog_00352–Asbog_00354 |
| Gluconate 2-dehydrogenase | FAD | GADH | Asbog_02524–Asbog_02526 |
| Glucose dehydrogenase | PQQ | GDH | Asbog_00903 |
| Glycerol dehydrogenase | PQQ | GLDH | Asbog_02636 and Asbog_02637 |
| PQQ-dependent family | PQQ | IDH | Asbog_00451 |
| δ-lactate dehydrogenase | FAD | LDH | Asbog_00865 |
| Pyruvate oxidase | TPP and FAD | POX | Asbog_02184–Asbog_02187 |

*Quinone reductases involved in oxidative fermentation are presented. Other quinone reductases that function in different processes are omitted from this table, such as FAD-dependent succinate dehydrogenase (Asbog_00368–Asbog_00371), FAD-dependent malate: quinone oxidoreductase (Asbog_01178), and proline dehydrogenase (Asbog_02561 and Asbog_00494).

4. Discussion

Asaia bogorensis is a unique member of AAB and was reported as an opportunistic pathogen that causes human peritonitis and bacteremia. Here, we determined the complete genome sequence of an As. bogorensis type strain NBRC 16594, which was isolated from flowers and fruits. The comparative genome analyses with other complete genomes showed that the bacterium genome contained putative pathogenic genes, including genes encoding type II secretion system regulated proteins and large cell adhesin, and genes responsible for elimination of reactive oxygen species and antibiotics. The genome data may provide basic information for further characterization of As. bogorensis pathogenicity using clinical isolates.

RNA-seq analyses using various culture conditions illuminated that expression of many short hypothetical genes, most of which are unique to As. bogorensis, were significantly higher under the ‘co-culture conditions’ with two types of mammal cells than under the ‘AAB condition’. It was previously reported that small proteins were involved in the adaptation to stressful environments, and thus short hypothetical genes may response to stressors under the ‘co-culture conditions’. Expression levels of certain stress-responsive genes were also significantly high under the ‘co-culture conditions’, including genes encoding RpoH, CsbD, Dps, heavy-metal resistance domain proteins, thioredoxin peroxidase, and cold shock transcriptional regulator. Most of the AAB are mesophilic and can grow only at 35°C or below. But certain AAB species, including As. bogorensis and Ac. pasteurianus, could be adapted to higher temperature environments managing those stress-responsive genes. Asaia bogorensis increased gene expression of RpoH under the ‘co-culture conditions’. However, gene expression of the assumable regulon of RpoH, including typical heat shock protein (HSP) genes, groEL, groES, and dnaJ/K, were significantly lower under ‘co-culture conditions’. Many reports showed that bacterial HSPs stimulate immune responses of their hosts and host cells produce proinflammatory cytokines leading to inflammation. The low expression of HSPs under the ‘co-culture conditions’ may provide a clue to understand how As. bogorensis avoids immunosurveillance and establishes chronic infection in human bodies.

Two cyo operons were identified in the As. bogorensis genome and showed opposite expression profiles under different conditions. There are five draft genomes of Asaia species in the public database,
including *Asaia astilbis* JCM 15831 (GCF_000613845.1), *As. sp. SF2.1* (NCBI assembly ID GCF_000505765.1), *Asaia prunellae* JCM 25354 (GCF_000613855.1), *As. platycodi* JCM 25414 (GCF_000614545.1), and *As. platycodi* SF2.1 (GCF_000724025.1). All *Asaia* species, except *As. astilbis*, have two cyo operons that correspond to cyoABCD-I and cyoABCD-2. *Asaia astilbis* genome contained mutated cyoABCD-I but no cyoABCD-2 operon. It may be because the draft genome DNA sequence was highly erroneous or because its mutation rate was high. *Asaia* species were identified in the gut and salivary gland of *Anopheles stephensi*, an Asian malarial mosquito vector. It is possible that generation of the two cyo operons in an *Asaia* ancestor was a key evolutionary event that led to the survival in a variety of environments, including high temperature conditions and inside of multicellular organisms. Besides *As. bogorensis*, only 8 species among 705 prokaryotic species contained two or more complete cyo operons, and most of the species are animal and plant pathogens or symbionts. Further analyses are required to clarify the evolutionary relationship among those species and the meaning of the gene expression alteration.

In conclusion, complete genome sequencing, genome comparison with other bacteria, and transcriptome analysis of *As. bogorensis* identified genes that may confer the ability to survive in various and changing environments.

**Accession numbers**

The accession number of the genome sequence and the BioProject ID of *As. bogorensis* NBRC 16594 reported in this paper is DDBJ: AP014690.1 and PRJDB519, respectively.

**Supplementary data**

Supplementary data are available at www.dnaresearch.oxfordjournals.org

**Funding**

This work was financially supported by the Advanced Low Carbon Technology Research and Development Program (ALCA), and a grant-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (KAKEN-HI: 22510222). Funding to pay the Open Access publication charges for this article was provided by the Advanced Low Carbon Technology Research and Development Program (ALCA).

**References**

1. Saichana, N., Matsushita, K., Adachi, O., Frébort, I. and Frebortova, J. 2014, Acetic acid bacteria: a group of bacteria with versatile biotechnological applications, *Biotechnol. Adv.*, doi:10.1016/j.biotechadv.2014.10.001.

2. Yamada, Y. 2014, Transfer of *Gluconacetobacter kacaitae*, *Gluconacetobacter medellinensis* and *Gluconacetobacter malaicei* to the genus *Komagataeibacter* as *Komagataeibacter kacaitae* comb. nov., *Komagataeibacter medellinensis* comb. nov. and *Komagataeibacter malaicei* comb. nov., *Int. J. Syst. Evol. Microbiol.*, 64, 1670–2.

3. Adachi, O., Moonmangmee, D., Toyama, H., Yamada, M., Shinagawa, E. and Matsushita, K. 2003, New developments in oxidative fermentation, *Appl. Microbiol. Biotechnol.*, 60, 643–53.

4. Sousa, F.L., Alves, R.J., Ribeiro, M.A., Pereira-Leal, J.B., Teixeira, M. and Pereira, M.M. 2012, The superfamily of heme-copper oxygen reductases: types and evolutionary considerations, *Biochim. Biophys. Acta*, 1817, 629–37.

5. Yamada, Y., Katsura, K., Kawasaki, H., et al. 2000, *Asaia bogorensis* gen. nov., sp. nov., an unusual acetic acid bacterium in the alpha-Proteobacteria, *Int. J. Syst. Evol. Microbiol.*, 50 (Pt 2), 823–9.

6. Greenberg, D.E., Porcella, S.F., Stock, F., et al. 2006, *Granulibacter bethesdensis* gen. nov., sp. nov., a distinctive pathogenic acetic acid bacterium in the family Acetobacteraceae, *Int. J. Syst. Evol. Microbiol.*, 56, 2609–16.

7. Ano, Y., Toyama, H., Adachi, O. and Matsushita, K. 2008, Energy metabolism of a unique acetic acid bacterium, *Asaia bogorensis*, that lacks ethanol oxidation activity, *Biocoi. Biotechnol. Biochem.*, 72, 989–97.

8. Snyder, R.W., Rude, J., Koeborn, S., et al. 2004, *Asaia bogorensis* peritonitis identified by 16S ribosomal RNA sequence analysis in a patient receiving peritoneal dialysis, *Am. J. Kidney Dis.*, 44, e15–7.

9. Tuimmen, T., Hemsämski, T. and Kertmala, T. 2006, First report of bacteremia by *Asaia bogorensis*, in a patient with a history of intravenous-drug abuse, *J. Clin. Microbiol.*, 44, 3048–50.

10. Tuimmen, T., Roggenkamp, A. and Vuopio-Varkila, J. 2007, Comparison of two bacteremic *Asaia bogorensis* isolates from Europe, *Eur. J. Clin. Microbiol. Infect. Dis.*, 26, 523–4.

11. Alauzet, C., Teyssier, C., Jumais-Bilak, E. et al. 2010, *Gluconobacter* as well as *Asaia* species, newly emerging opportunistic human pathogens among acetic acid bacteria, *J. Clin. Microbiol.*, 48, 3935–42.

12. Chase, J.M., Holland, S.M., Greenberg, D.E., Marshall-Battky, K., Zelazny, A.M. and Church, J.A. 2012, *Acidomonas methanolaica*-associated necrotizing lymphadenitis in a patient with chronic granulomatous disease, *J. Clin. Immunol.*, 32, 1193–6.

13. Aruma, Y., Hosoyama, A., Matsutani, M., et al. 2009, Whole-genome analyses reveal genetic instability of *Acetobacter pasteurianus*, *Nucleic Acids Res.*, 37, 5768–83.

14. Illegems, K., De Vuyst, L. and Weckx, S. 2013, Complete genome sequence and comparative analysis of *Acetobacter pasteurianus* 3868, a strain well-adapted to the cocoa bean fermentation ecosystem, *BMC Genomics*, 14, 526.

15. Ge, X., Zhao, Y., Hou, W. et al. 2013, Complete genome sequence of the industrial strain *Gluconobacter oxydans* H24, *Genome Announc.*, 1, e00003–13.

16. Prust, C., Hoffmeister, M., Liesegang, H. et al. 2005, Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*, *Nat. Biotechnol.*, 23, 195–200.

17. Ogino, H., Aruma, Y., Hosoyama, A., et al. 2011, Complete genome sequence of NBRC 3288, a unique cellulase-nonproducing strain of *Gluconacetobacter xylinus* isolated from vinegar, *J. Bacteriol.*, 193, 6997–8.

18. Bertalan, M., Albano, R., de Padua, V. et al. 2009, Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconactobacter diazotrophicus* PA1, *BMC Genomics*, 10, 450.

19. Giorgi, A., Tyler, H.L., Zipperer, U.N. and Triplett, E.W. 2010, Two genome sequences of the same bacterial strain, *Gluconacetobacter diazotrophicus* PA1 suggest a new standard in genome sequence submission, *Stand Genomic Sci.*, 2, 309–17.

20. Greenberg, D.E., Porcella, S.F., Zelazny, A.M. et al. 2007, Genome sequence analysis of the emerging human pathogenic acetic acid bacterium *Granulibacter bethesdensis*, *J. Bacteriol.*, 189, 8727–36.

21. Shin, S.C., Kim, S.H., You, H. et al. 2011, *Drosophila* microbiome modulates host developmental and metabolic homeostasis via insulin signaling, *Science*, 334, 670–4.

22. Chouaia, B., Gaia, S., Crotti, E. et al. 2014, Acetic acid bacteria genomes reveal functional traits for adaptation to life in insect guts, *Genome Biol. Evol.*, 6, 912–20.

23. Hijashizu, N., Hadano, H., Hirakawa, H. et al. 2014, Draft genomic DNA sequence of the facultatively methylo trophic bacterium *Acidomo nas methanolicus* type strain MB58, *FEMS Microbiol. Lett.*, 351, 9–13.

24. Besemer, J., Lomsadze, A. and Borodovsky, M. 2001, GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions, *Nucleic Acids Res.*, 29, 2607–18.

25. Noguchi, H., Taniguchi, T. and Itoh, T. 2008, MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene
prediction in anonymous prokaryotic and phage genomes, DNA Res., 15, 387–96.
26. Delcher, A.L., Beatle, K.A., Powers, E.C. and Salzberg, S.L. 2007, Identifying bacterial genes and endosymbiont DNA with Glimmer, Bioinformatics, 23, 673–9.
27. Haft, D.H., Selengut, J.D. and White, O. 2003, The TIGRFAMs database of protein families, Nucleic Acids Res., 31, 371–3.
28. Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.H., Rogne, T. and Ussery, D.W. 2007, RNAmer: consistent and rapid annotation of ribosomal RNA genes, Nucleic Acids Res., 35, 3100–8.
29. Lowe, T.M. and Eddy, S.R. 1997, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Res., 25, 955–64.
30. Altschul, S.F., Madden, T.L., Schäffer, A.A., et al. 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res., 25, 3389–402.
31. Eddy, S.R. 2011, Accelerated profile HMM searches, PLoS Comput. Biol., 7, e1002195.
32. Finn, R.D., Mistry, J., Tate, J., et al. 2010, The Pfam protein families database, Nucleic Acids Res., 38, D211–22.
33. Uchiyama, I. 2006, Hierarchical clustering algorithm for comprehensive orthologous-domain classification in multiple genomes, Nucleic Acids Res., 34, 647–58.
34. Kawai, M., Furuta, Y., Yahara, K., et al. 2011, Evolution in an oncogenic bacterial species with extreme genome plasticity: maximum-likelihood trees for large alignments, Proc. Natl. Acad. Sci. USA, 99, 13142–47.
35. Homma, M.R., Paul, B.J., Miranda-Rios, J., Zhang, A., Soltanzad, N. and Hemm, M.R. 2008, Protein secretion and outer membrane assembly in Alphaproteobacteria, FEMS Microbiol. Rev., 32, 995–1009.
36. Matsutani, M., Fukushima, M., Kato, N., et al. 2011, Ultrafine cellulose fibers produced by Asaia bogorensis, an acetic acid bacterium, Biomacromolecules, 12, 2813–21.
37. Matsutani, M., Fukushima, M., Kato, N., et al. 2014, Replacement of a terminal cytochrome c oxidase by ubiquinol oxidase during the evolution of acetic acid bacteria, Biochim. Biophys. Acta, 1837, 1810–20.
38. Douzi, B., Filhoux, A. and Voulhoux, R. 2012, On the path to uncover the bacterial type II secretion system, Philos. Trans. R. Soc. Lond. B. Biol. Sci., 367, 1059–72.
39. Kawai, M., Ham, J.H., Deng, W.L., Doyle, J.J. and Collmer, A. 2002, HEmA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of Erwinia chrysanthemi EC16 on Nicotiana clevelandii seedlings, Proc. Natl. Acad. Sci. USA, 99, 13142–47.