Free-Living Enterobacterium *Pragia fontium* 24613: Complete Genome Sequence and Metabolic Profiling

Kateřina Snopková[1], Karel Sedlář[2], Juraj Bosák[1], Eva Chaloupková[1], Ivo Sedláček[3], Ivo Provazník[2] and David Šmajs[1]

[1]Department of Biology, Masaryk University, Brno, Czech Republic. [2]Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic. [3]Czech Collection of Microorganisms, Department of Experimental Biology, Masaryk University, Brno, Czech Republic.

Evolutionary Bioinformatics
Volume 13: 1–7

DOI: 10.1177/1176934317700863
SAGE

ABSTRACT: *Pragia fontium* is one of the few species that belongs to the group of atypical hydrogen sulfide-producing enterobacteria. Unlike other members of this closely related group, *P. fontium* is not associated with any known host and has been reported as a free-living bacterium. Whole genome sequencing and metabolic fingerprinting confirmed the phylogenetic position of *P. fontium* inside the group of atypical H_2S producers. Genomic data have revealed that *P. fontium* 24613 has limited pathogenic potential, although there are signs of genome decay. Although the lack of specific virulence factors and no association with a host species suggest a free-living style, the signs of genome decay suggest a process of adaptation to an as-yet-unknown host.

KEYWORDS: *Pragia fontium*, *Enterobacteriaceae*, whole genome sequence, phylogeny, free-living bacteria

RECEIVED: November 21, 2016. **ACCEPTED:** February 12, 2017.

PEER REVIEW: Seven peer reviewers contributed to the peer review report. Reviewers' reports totaled 1906 words, excluding any confidential comments to the academic editor.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Grant Agency of the Czech Republic (16-21649S) and also by internal grant MUNI/11/InGA04/2014.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: David Šmajs, Department of Biology, Masaryk University, Kamenice 753/5, 625 00 Brno, Czech Republic. Email: dsmajs@med.muni.cz

Introduction

To date, the *Enterobacteriaceae* family contains 55 genera and 248 species (www.bacterio.net, September 1, 2016). Most of the enterobacteria live in the vertebrate intestine, whereas several other enterobacterial genera/species represent plant pathogens or invertebrate endosymbionts.[1] Other enterobacteria are believed to live only in the environment, eg, *Pragia*, *Saccharobacter*, *Obesumbacterium*, *Shimwellia*,[1] *Mangrovibacter*,[2] and *Biostraticola*.[3] However, it is possible that their pathogenic/symbiotic potential will be revealed in the future, as it was for *Budvicia*.[4,5]

Pragia fontium is a gram-negative, mesophilic, rod-shaped, motile bacterium. The genus *Pragia* contains only 1 species, *P. fontium*, which was described in 1988.[6] A total of 18 strains were isolated in Czechoslovakia between 1982 and 1986. All strains, except 1, were isolated from water wells and water pipes, whereas 1 strain was obtained from the stool of a healthy woman. Another set of *Pragia* strains was isolated in Ukraine between 1996 and 1997.[7] They were mostly isolated from water (9 strains) and other environmental material (5 strains), although 2 strains came from human clinical material; their relatedness to the Czechoslovakia strains varied from 84% to 95% (based on DNA-DNA hybridization). To date, only strains from these 2 locations have been characterized, and the exact ecological niche and pathogenic potential of *Pragia* remains unclear.

Pragia fontium, as well as *Budvicia* spp. and *Leminorella* spp., is a closely related atypical enterobacterial species. Their common feature is hydrogen sulfide production, with *Budvicia diplopodorum* being the only known exception.[5] These H_2S-producing enterobacteria share several metabolic features including reduced metabolic activity that results in utilization of a limited set of substrates. The optimal growth temperature for *Pragia* and *Budvicia* is 25°C, whereas *Leminorella* is capable of growing at temperatures up to 42°C.[6] *Pragia fontium* can be differentiated from *Budvicia* spp. based on a positive (Simmons) citrate utilization test and from *Leminorella* spp. by its motility, tartrate utilization, tyrosine clearing, and inability to grow at 42°C.[6] In addition, a whole-cell protein pattern analysis of *P. fontium*, *B. aquatica*, and *Leminorella* spp. was determined and the data supported the delineation of these genera.[8] On the DNA level, *Pragia* strains were most closely related to *Budvicia* (based on DNA-DNA hybridization, relatedness 20%-37%) but barely related to other genera, eg, relatedness to *Escherichia coli* K12 was about 3%.[9]

To date, 485 completed enterobacterial genome sequences, covering 21 genera and 47 species, have been deposited in the Genomes OnLine Database (GOLD, https://gold.jgi.doe.gov/). Attention has been focused mainly on clinically and agriculturally important bacteria (eg, *Escherichia*, *Salmonella*, *Klebsiella*, and *Yersinia*), leaving the remaining genera relatively unexplored.

The whole genome sequence and the pilot assembly of *P. fontium* 24613 were published in 2015.[10] In this study, we characterized *P. fontium* based on genomic data, including the relationship of *Pragia* to other genera, and compared metabolic pathways with the results of phenotypic metabolic fingerprinting.

Materials and Methods

Bacterial strains and cultivation conditions

The strains used in this study came from the collection of the Department of Biology, Masaryk University, Brno, Czech Republic (*P. fontium* 24613, originally stored at the National Institute of Public Health, Prague, Czech Republic); from the Czech National Collection of Type Cultures, Prague, Czech Republic (*Budvicia aquatica* CNCTC 6285^T); and from the Czech Collection of Microorganisms, Brno, Czech Republic (*Leminorella grimontii* CCM 4003^T). *Pragia fontium* 24613 came from the same set of strains as *P. fontium* DSM 5563^{T6}. Strains were cultivated in TY medium (8 g casein, 5 g yeast extract, 5 g sodium chloride, pH 7.5; HiMedia, Mumbai, India) at 30°C for 24 hours.

Pragia fontium 24613 genome sequencing and annotation

In our previous study, protocols for DNA extraction, whole genome sequencing, and annotation of *P. fontium* 24613 were described in detail.[10] For additional gene mining and genome comparisons, annotation was manually curated based on results of a RAST (Rapid Annotation using Subsystem Technology) pipeline[11] and DOE-JGI (US Department of Energy-Joint Genome Institute) Microbial Genome Annotation Pipeline.[12] Detected proteins were assigned to Clusters of Orthologous Group (COG) categories based on DOE-JGI results. Methylome was characterized using PacBio single-molecule real-time sequencing (1× SMRT cell) of kinetic data collected during the genome sequencing process.[13] SMRT analysis version 2.3, using the "RS_Modification_and_Motif_Analysis.1" protocol, was used for genome-wide base modification and detection of the affected motifs. Regarding sequencing coverage, a default quality score value of 30 (corresponding to a *P* value of .001) was used for motif determination. The detected motifs were uploaded and further analyzed using the REBASE database.[14] The complete genome was also scanned for homologues of restriction-modification system genes (using a Basic Local Alignment Search Tool [BLAST] search, with the BLASTX algorithm) against the REBASE and GenBank databases.

Phylogenetic position of P. fontium

The genome sequence of *P. fontium* 24613 was compared with other enterobacterial genera on a genome-wide level. Whole genome sequences were downloaded from the GOLD (https://gold.jgi.doe.gov/); their accession numbers are listed in Table S1. Each genus was represented by 1 sequence (except for *Pragia* where both the type strain DSM 5563^T and strain 24613 were used). If available, the sequence of the type strain was used. For genera *Biostraticola*, *Cosenzaea*, *Gibbsiella*, *Mangrovibacter*, *Obesumbacterium*, *Saccharobacter*, and *Samsonia*, no sequences were available. A whole genome phylogenetic analysis was built using PhyloPhlAn 0.99,[15] which compared more than 400 selected protein sequences conserved across bacterial domains. The genes were identified using an internal PhyloPhlAn database by translated mapping with USEARCH 8.1.[16] The topology was computed using the neighbor-joining algorithm in conjunction with the Jukes-Cantor evolution model. Moreover, the CAT model, with gamma correction, was used to optimize and rescale the tree. The final tree was reconstructed, using FastTree 2.1,[17] from protein subsequences of the genes concatenating their most informative amino acid positions, and each was aligned using MUSCLE 3.8.[18] The tree was visualized in MEGA 6.06.[19] Dot plot diagrams between genomes were constructed using the Integrated Microbial Genome platform.[12] The core genome of *P. fontium*, *B. aquatica*, and *L. grimontii* was determined based on orthologous clusters produced by OrthoVenn[20] using a modified OrthoMCL heuristic approach. Default parameters (E-value 1e−5 and inflation value 1.5) were used. Metabolic pathway analysis of *P. fontium* 24613, *Wigglesworthia glossinidia* (acc. no. CP003315), and *Buchnera aphidicola* G002 (acc. no. CP002701) was performed using the KEGG PATHWAY database,[21] which is part of KEGG Web services (http://www.genome.jp/kegg/).

Analyses of metagenomics data

Data from the Human Microbiome Project database (http://hmpdacc.org) and EBI Metagenomics database (https://www.ebi.ac.uk/metagenomics/) were searched with BLASTN 2.2.22[22] using a consensus sequence of 7 16S ribosomal RNA (rRNA) genes of *P. fontium* 24613. The first database contained a complete set of human microbiome data (associating data from several human sites), and the latter database covered data from different environmental sources.

Substrate diversity studies

The Biolog GN2 MicroPlate analysis platform (Biolog, Inc., Hayward, CA, USA) was used for determination of the biochemical profiles of *P. fontium* 24613, *B. aquatica* CNCTC 6285^T, and *L. grimontii* CCM 4003^T cultivated on Biolog Universal Growth (BUG) agar at 30°C for 24 hours. Utilization of 95 carbon sources was tested[23] (Table S2). Media and all reagents were supplied by Biolog and used according to the manufacturer's protocol. Plates were incubated in parallel under aerobic and anaerobic conditions and tests were read after 24 hours of incubation.

Results

Genome analyses of P. fontium 24613

Complete genome sequence of P. fontium 24613. A complete genome sequence for *P. fontium* 24613 represents a single circular chromosome with a length of 4 094 629 bp.[10] The *P. fontium* 24613 genome was compared with 3 draft genomes

Table 1. Genome features of *Pragia fontium* 24613 in comparison with the draft genome of the type strain and the draft genomes of closely related hydrogen sulfide producers.

FEATURE	*P. FONTIUM* 24613	*P. FONTIUM* DSM 5563^T	*B. AQUATICA* DSM 5075^T	*L. GRIMONTII* DSM 5078^T
Genome status	Complete	Draft	Draft	Draft
Genome size	4 094 629 bp	3 950 845 bp	5 670 930 bp	4 222 128 bp
GC content	45.38%	45.23%	45.68%	53.86%
No. of CDS	3579	3464	5130	3878
No. of rRNA genes	22 (8–7–7)	10 (2–5–3)	7 (5–2–0)	16 (8–6–2)
No. of tRNA genes	72	58	57	57
No. of pseudogenes	146 (4.1%)	NA	144 (2.8%)	62 (1.6%)
No. of genes with predicted function	2809 (78.49%)	2862 (82.62%)	3896 (75.95%)	3083 (79.50%)
No. of genes assigned to COG	2601 (72.67%)	2613 (75.43%)	2601 (72.67%)	2804 (72.31%)
No. of genes assigned to KEGG pathways	1160 (32.41%)	1172 (33.83%)	1419 (27.66%)	1217 (31.38%)

Abbreviations: CDS, coding sequences; COG, Clusters of Orthologous Group; KEGG, Kyoto Encyclopedia of Genes and Genomes; rRNA, ribosomal RNA; tRNA, transfer RNA.
Accession numbers of the whole genome sequences of the type strains are listed in Table S1. Order of the rRNA genes in parentheses: 5S-16S-23S. NA—data not available in the GenBank and Genomes OnLine databases.

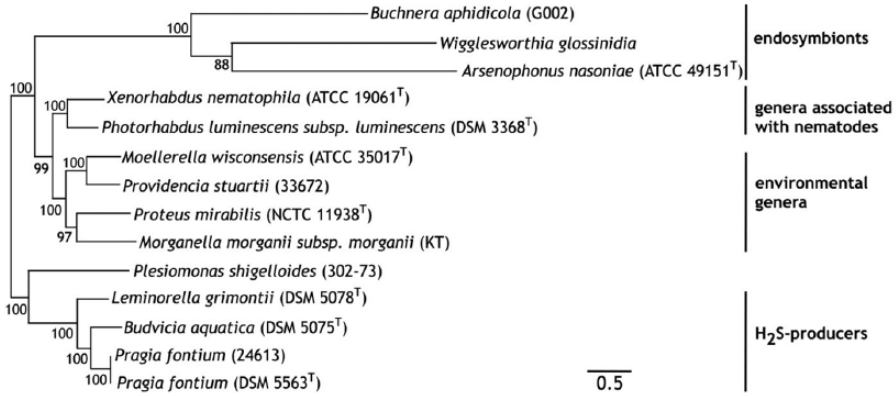

Figure 1. Phylogenetic position of the genus *Pragia* based on a whole genome sequence tree. Only the relevant part of the *Enterobacteriaceae* tree is shown (the whole tree is depicted in Figure S2). All branches are supported with high bootstrap values. The tree was drawn to scale; the scale bar represents the estimated number of amino acid changes per site per unit of branch length.

of related bacteria, including the draft genome of *P. fontium* DSM 5563^T (Table 1). Both the *B. aquatica* DSM 5075^T genome and the *L. grimontii* DSM 5078^{T}genome were larger in size and gene count compared with the complete genome sequence of *P. fontium* 24613. Moreover, the proportion of pseudogenes was larger in *P. fontium* (4.1%) than in the draft genomes of other H_2S-producing enterobacteria (ie, 2.8% and 1.6% for *B. aquatica* and *L. grimontii*, respectively), suggesting genome decay in *P. fontium*. In addition, a clearly higher GC content was found in the *L. grimontii* DSM 5078^T (*L. grimontii*) genome. The draft status was likely responsible for the lower number of predicted rRNA and transfer RNA genes in the *P. fontium* DSM 5563^T, *B. aquatica*, and *L. grimontii* genomes.

Phylogenetic position of P. fontium. A whole genome phylogenetic approach was used to compare the genome sequence of *P. fontium* 24613 with genome sequences of other enterobacterial genera. The relevant part of the *Enterobacteriaceae* tree is shown in Figure 1. Strong support was found for a close relationship among *Pragia* and other atypical H_2S producers, including *Budvicia* and *Leminorella*. The high similarity among genomes was also supported by a dot plot analysis of H_2S producer genomes (Figure S1). Another related genus was *Plesiomonas*, an oxidase-positive genus recently reclassified into the *Enterobacteriaceae* family.[24] A sister clade contains a cluster of genera occurring frequently in the (1) environment (*Providencia*, *Moellerella*, *Proteus*, and *Morganella*), (2) genera associated with nematodes (*Xenorhabdus*, *Photorhabdus*), and (3) endosymbionts (*Arsenophonus*, *Buchnera*, and *Wigglesworthia*). Except for the delineation of *Proteus* vs *Morganella* and endosymbionts *Buchnera* vs *Wigglesworthia*, all other branches were supported by bootstrap values higher than 99%.

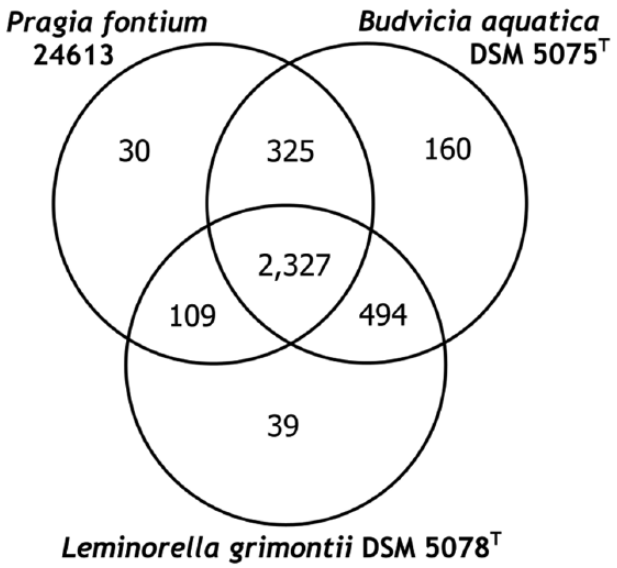

Figure 2. The Venn diagram represents the core genome and pangenome of *Pragia* and the closely related atypical H_2S producers. The numbers represent the gene clusters shared by corresponding group of genera. The diagram shows the close relationships among those inside the group of atypical H_2S producers.

The core genome of enterobacterial hydrogen sulfide producers. The core genomes of *P. fontium*, *B. aquatica*, and *L. grimontii* contain 2327 gene clusters (ie, at least 1 gene from each cluster was found in each genome; Figure 2). The number of gene clusters exclusively shared by 2 genomes was higher for the *P. fontium* and *B. aquatica* genomes (325) compared with the *P. fontium* and *L. grimontii* genomes (109), whereas there were 494 clusters shared by the *L. grimontii* and *B. aquatica* genomes. These data indicate a higher degree of relatedness between *P. fontium* and *B. aquatica* compared with *P. fontium* and *L. grimontii*. A set of 30 gene clusters was unique for the *P. fontium* genome; these clusters encoded homologues to fimbrial genes found in *Serratia* spp. and *Proteus* spp. and also homologues to pyocin S3 and its immunity protein–encoding genes. In total, 10 clusters encoded genes for hypothetical proteins.

Genome-based metabolic and virulence analyses

Analysis of metabolic pathways in the P. fontium genome. Based on the genomic data analysis from KEGG PATHWAY and DOE-JGI, aerobic and facultative anaerobic metabolism of *P. fontium* 24613 was predicted. Oxidized nitrogen and sulfur compounds were capable of serving as alternative terminal electron acceptors under anaerobic conditions. Identification of thiosulfate reductase, responsible for H_2S production, corresponded to previously detected enzyme activity.[6] The genes involved in glycolysis, citrate cycle, and pentose phosphate pathway could also be found in the *P. fontium* genome in addition to genes responsible for amino acid, fatty acid synthesis, lipid, and nucleotide metabolism. *Pragia* was found to be auxotrophic for L-tryptophan, L-histidine, and L-leucine and deficient in biotin synthesis. Compared with *Budvicia* and *Leminorella*, *Pragia* was able to synthetize L-arginine but lacked the genes for fatty acid degradation. In addition, the *P. fontium* genome contained fewer genes involved in carbohydrate metabolism compared with the *L. grimontii* and *B. aquatica* genomes (Table S3).

Genome methylation pattern. Analysis of PacBio sequencing data revealed 24 814 methylated positions of the m6A type, but only a single sequence motif (GATC) was found in all these modifications. More than 80% (21 735 of 26 606) of the GATC positions in the genome were methylated. Methylation type m4C was not found. Kinetic signatures of m5C were subtler than signatures of m6A and m4C and harder to detect using PacBio SMRT sequencing[25]; therefore, they were not assessed. The results of *P. fontium* genome methylation were deposited in the REBASE PacBio database (http://rebase.neb.com/cgi-bin/pacbiolist).[14] In total, 8 different putative restriction-modification systems, all of them type II, were predicted in the genome (Table S4). Seven of them consisted of only methyltransferases, whereas the last one modifying m5C consisted of methyltransferase, mismatch repair endonuclease, and restriction endonuclease.

Virulence and antimicrobial genes in the P. fontium genome. In silico analysis of virulence determinants of the *P. fontium* genome revealed genes involved in iron acquisition (encoding Fe^{2+} and Fe^{3+} transport systems), adhesion (encoding P pili and type I pili), secretion systems (T1SS and T6SS), and antibiotic resistance (encoding AmpC β-lactamase and several efflux pump) (see Table S5).

Production of tailocins, ie, R-type and F-type bacteriocins resembling phage tails, was previously detected in several *Pragia* strains.[26] Gene clusters similar to the phage genes were described as being responsible for production of these antimicrobial compounds.[27] A total of 6 clusters homologous to phage genes were predicted in the *P. fontium* genome, and one of them was likely responsible for tailocin production (see Table S5). The genome search also detected a gene encoding a colicin-like bacteriocin, a homologue of pyocin S3.

Metabolic profiling of P. fontium 24613

The carbohydrate utilization pattern resulting from the testing of various saccharides, carboxylic acids, alcohols, amino acids, aromatic compounds, and their derivatives was determined for *P. fontium* 24613, *B. aquatica* CNCTC 6285T, and *L. grimontii* CCM 4003T. In general, the data obtained from the Biolog assay revealed low levels of metabolic activity in all tested strains. Substrate utilization profiles differed for the 3 tested H_2S producers in 17 substrates (Table S2). *Pragia fontium* 24613 was able to utilize 15 substrates (out of 95; 16%) under aerobic conditions and 22 (out of 95; 23%) under anaerobic conditions. *Pragia* utilized monosaccharides and their derivatives (α-D-glucose, α-D-glucose-1-phosphate, D-glucose-6-phosphate, *N*-acetyl-D-glucosamine, and β-methyl-D-glucoside), monocarboxylic acids (D,L-lactic acid, and D-gluconic acid), dicarboxylic acids (α-keto-glutaric acid, and L-glutamic acid), alcohols and their derivatives (glycerol, D,L-α-glycerol phosphate, myo-inositol, and xylitol), amino acids (D-serine), and aromatic compounds (uridine and thymidine). In addition to substrates

utilized under aerobic conditions, anaerobically cultivated *Pragia* utilized L-arabinose, pyruvic acid methyl ester, D-glucuronic acid, bromosuccinic acid, L-aspartic acid, glycyl-L-aspartic acid, and L-serine. Although *Budvicia* utilized 16 substrates (17%) aerobically and 24 (25%) anaerobically, *Leminorella* utilized only 13 substrates (14%) aerobically and 18 (19%) anaerobically. *Budvicia* and *Leminorella* were able to metabolize several amino acids and their derivatives (L-asparagine, L-aspartic acid, and glycyl-L-aspartic acid) as well as derivatives of organic acids from Krebs cycle (pyruvic acid methyl ester, bromosuccinic acid), which were not utilized by *Pragia*. The complete results of this assay are shown in Table S2. In most of the substrate tests, which differed among H_2S producers, the genes encoding corresponding catabolic enzymes or enzymes possibly involved in metabolism of these compounds were found (Table S6). The only exception was the *B. aquatica* genome, where some of the genes responsible for catabolism of uridine were not found.

Discussion

Pragia belongs to a relatively small group of H_2S-producing enterobacteria containing *P. fontium*, *Budvicia* spp., and *Leminorella* spp. Although all members of this small group are closely related and have a relatively similar biochemical profile, they occupy quite different ecological niches. Although *Budvicia* was originally isolated from freshwater,[28] several other isolates have been described from the intestinal microflora of insects,[26,29] *Diplopoda*,[5] and salmonids.[30] A possible clinical relevance for *B. aquatica* was reported by Corbin et al[4] when this bacterium was isolated from a human clinical sample. *Leminorella* spp. have been exclusively isolated from human clinical specimens and no environmental sources have been reported. Although its clinical significance is unclear,[1] *Leminorella* spp. appear to be associated with urinary tract infections and other human nosocomial infections.[31] In contrast to *Budvicia* and *Leminorella*, *Pragia* has been isolated almost exclusively from environmental sources. Only 3 isolates originated from human clinical samples; there is no information on the role of these strains in infection or disease.[6,7] Because the prevalent habitat of other *Pragia* strains is drinking water, these cases likely reflect accidental isolations. Inspection of metagenomics data revealed the absence of *Pragia* 16S ribosomal DNA (rDNA) in both environmental and host-associated data sets (data not shown). From all the available data, *Pragia* appears to be the only H_2S producer occupying environmental niches with no association with humans or other hosts.

A possible interaction between *Pragia* and a host species was examined by identification and analysis of genes encoding virulence factors. Several common virulence factors shared by most enterobacterial species (even saprophytic ones) were detected. Genes for adhesion, antibiotic resistance, iron uptake, and 2 secretion systems were found. Adhesion and the ability to acquire iron are key factors required for colonization and survival in a host (animal or plant).[32-34] These findings indirectly support an association between *Pragia* and an as-yet-unknown host. We can speculate that if a host organism exists, it will likely be similar to those of the closely related genus *Budvicia*, ie, nonvertebrate hosts such as insects or nematodes. Although the presence of *Pragia* has been detected in the intestines of freshwater salmon,[35] the much more frequent isolation from deepwater wells[6] tends to support a free-living lifestyle of *Pragia*. Both detected secretion systems, T1SS and T6SS, are widely distributed in gram-negative bacteria[36,37] and could mediate interaction with a host or with another bacterium.[38] Although the contribution of T6SS to pathogenesis has been described for several bacteria, eg, *Pseudomonas*[39] and *E coli*,[40] T6SS has also been found in saprophytic bacteria, where it was involved in interactions across the microbial community.[38] Several bacteriocin types have been suggested as putative virulence factors, whereas the importance of others was demonstrated in interactions across microbial community.[41,42] Although the function of *P. fontium* bacteriocins remains unknown, both tailocins and colicin-like homologues were found in the *Pragia* genome. The GATC methylation motif was found in the *P. fontium* genome, and because the corresponding gene for the restriction enzyme recognizing this motif was not found, methylation appears to be more connected to gene expression regulation[43] and not to degradation of foreign nucleic acid molecules.

Metabolic profiling revealed a metabolic pattern for *Pragia*, *Budvicia*, and *Leminorella*, which was quite distinct from other enterobacteria,[44] supporting the distinctness of enterobacterial H_2S producers and also the close relationship of these bacteria within this group. Despite their overall similarity, H_2S-producing enterobacteria revealed several differences in their ability to utilize substrates. Analyses of genomic data supported the metabolic findings, with only 1 case in which some of the genes encoding expected enzymatic activity were not found. This is likely a result of an incomplete genomic sequence in *Budvicia*. Surprisingly, all species were able to degrade multiple substrates under anaerobic conditions suggesting that alternative electron acceptors (nitrate, reduced sulfur compounds) could be used under anaerobic conditions. Nitrogen oxidation could be carried out using the "nitrite reduction to ammonium pathway" for which the corresponding genes were found in the *P. fontium* genome. This pathway is preferred for respiration under anaerobic conditions, and it is common across *Enterobacteriaceae* and in other facultatively anaerobic bacteria.[45]

A comparative genomics approach revealed that almost 80% of the gene clusters were shared by H_2S-producing enterobacteria, whereas only 49% were shared when *E coli* K12 was added to the analysis. Analysis of the complete genome sequence of *Pragia* revealed that the genome contains genes involved in essential metabolic pathways, in nutrient metabolism, and also in the synthesis of most of the amino acids.

However, the "fatty acid degradation pathway" is missing from the *P. fontium* 24613 genome. This pathway is present in most enterobacterial genomes but not in invertebrate endosymbionts with a reduced genome, such as *Wigglesworthia* and *Buchnera*. Nevertheless, when compared with these endosymbionts, the *P. fontium* genome is relatively large and also contains an additional set of genes, eg, those responsible for degradation of more complex polysaccharides. However, *P. fontium* 24613 has a relatively small genome in comparison with other enterobacteria, even in comparison with the genus *Budvicia*. In addition, the proportion of pseudogenes was larger in *Pragia* compared with other closely related bacteria (despite their draft status, which is prone to assembly errors). Larger proportions of pseudogenes have also been observed in bacteria that were associated with or dependent on eukaryotic hosts.[46] Nevertheless, this analysis comes from a limited number of genome sequences per species and it is known that the prevalence of pseudogenes is quite variable even among closely related strains.[47] A reduction in genome size and an increased number of pseudogenes are common signs of bacterial adaptation to a eukaryotic host. In addition, the *P. fontium* genome contains fewer genes involved in carbohydrate utilization compared with other H_2S producers; a large battery of degradation enzymes is important mainly for free-living bacteria. The traces of genome decay (ie, small genome, absence of fatty acid degradation pathways, the small number of genes associated with carbohydrate utilization, and a larger proportion of pseudogenes) suggest an ongoing process of adaptation to a particular host organism. Although no such host has been identified for *P. fontium*, the recent progress in metagenome studies could help to answer this question in the near future.

Conclusions

Analysis of the complete genome sequence of *P. fontium* 24613 and metabolic profiling confirmed the close relatedness of this bacterium to other H_2S-producing enterobacteria, *Budvicia* spp. and *Leminorella* spp., although for each genus a different environmental niche has been described. Virulence gene mining and the absence of *Pragia* 16S rDNA sequences in the human metagenomics data suggest limited pathogenic potential for *Pragia*, consistent with the previously described free-living lifestyle of this bacterium. On the contrary, reduced genome size, limited number of encoded enzymes for carbohydrate and fatty acid degradation, and frequent presence of pseudogenes suggest a process of adaptation to an as-yet-unknown host.

Acknowledgements

The authors thank Jan Šmarda for providing the *P. fontium* 24613 strain.

Author Contributions

KSn and DS conceived, designed, and performed the experiments (genome sequencing). IS performed phenotypical characterization. KSn, KSe, and IP analyzed the data. KSn wrote the first draft of the manuscript. DS and JB contributed to the writing of the manuscript. All authors reviewed and approved the final manuscript.

Disclosures and Ethics

As a requirement of publication, author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including, but not limited to, the following: authorship and contributorship, conflicts of interest, privacy and confidentiality, and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

REFERENCES

1. Octavia S, Lan R. The family *Enterobacteriaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, eds. *The Prokaryotes*. 4th ed. Berlin, Germany: Springer-Verlag; 2014:225–275.
2. Rameshkumar N, Lang E, Nair S. *Mangrovibacter plantisponsor* gen. nov., sp. nov., a nitrogen-fixing bacterium isolated from a mangrove-associated wild rice (*Porteresia coarctata* Tateoka). *Int J Syst Evol Microbiol*. 2010;60:179–186.
3. Verbarg S, Frühling A, Cousin S, et al. *Biostraticola tofi* gen. nov., spec. nov., a novel member of the family *Enterobacteriaceae*. *Curr Microbiol*. 2008;56:603–608.
4. Corbin A, Delatte C, Besson S, et al. *Budvicia aquatica* sepsis in an immunocompromised patient following exposure to the aftermath of Hurricane Katrina. *J Med Microbiol*. 2007;56:1124–1125.
5. Lang E, Schumann P, Knapp BA, Kumar R, Spröer C, Insam H. *Budvicia diplopodorum* sp. nov. and emended description of the genus *Budvicia*. *Int J Syst Evol Microbiol*. 2013;63:260–267.
6. Aldová E, Hausner O, Brenner DJ, et al. *Pragia fontium* gen. nov., sp. nov. of the family *Enterobacteviaceae*, isolated from water. *Int J Syst Evol Microbiol*. 1988;38:183–189.
7. Pokhyl SI. The properties of *Pragia fontium* bacteria isolated on the territories of Ukraine and the Czech Republic. *Mikrobiol Z*. 2000;62:3–10.
8. Schindler J, Potužníková B, Aldová E. Classification of strains of *Pragia fontium, Budvicia aquatica* and of *Leminorella* by whole-cell protein pattern. *J Hyg Epidemiol Microbiol Immunol*. 1992;36:207–216.
9. Hickman-Brenner FW, Vohra MP, Huntley-Carter GP, et al. *Leminorella*, a new genus of *Enterobacteriaceae*: identification of *Leminorella grimontii* sp. nov. and *Leminorella richardii* sp. nov. found in clinical specimens. *J Clin Microbiol*. 1985;21:234–239.
10. Snopková K, Sedlář K, Bosák J, Chaloupková E, Provazník I, Šmajs D. Complete genome sequence of *Pragia fontium* 24613, an environmental bacterium from the family *Enterobacteriaceae*. *Genome Announc*. 2015;3:e00740–15.
11. Aziz RK, Bartels D, Best AA, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. 2008;9:75.
12. Markowitz VM, Chen I-MA, Palaniappan K, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res*. 2014;42:D560–D567.
13. Flusberg BA, Webster DR, Lee JH, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods*. 2010;7:461–465.
14. Roberts RJ, Belfort M, Bestor T, et al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res*. 2003;31:1805–1812.
15. Segata N, Börnigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun*. 2013;4:2304.
16. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinforma Oxf Engl*. 2010;26:2460–2461.
17. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE*. 2010;5:e9490.

18. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 2004;5:113.
19. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–2729.
20. Wang Y, Coleman-Derr D, Chen G, Gu YQ. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res*. 2015;43:W78–W84.
21. Zhou T. Computational reconstruction of metabolic networks from KEGG. *Methods Mol Biol Clifton NJ*. 2013;930:235–249.
22. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–3402.
23. Garland JL, Mills AL. Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. *Appl Environ Microbiol*. 1991;57:2351–2359.
24. Janda JM. Genus XXVII. *Plesiomonas*, Habs and Schubert 1962, 324AL. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, eds. *Bergey's Manual of Systematic Bacteriology*. Vol 2. 2nd ed, New York, NY: Springer; 2005;740–744.
25. Davis BM, Chao MC, Waldor MK. Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Curr Opin Microbiol*. 2013;16:192–198.
26. Tagliavia M, Messina E, Manachini B, Cappello S, Quatrini P. The gut microbiota of larvae of *Rhynchophorus ferrugineus* Oliver (Coleoptera: Curculionidae). *BMC Microbiol*. 2014;14:136.
27. Ghequire MGK, De Mot R. Ribosomally encoded antibacterial proteins and peptides from *Pseudomonas*. *FEMS Microbiol Rev*. 2014;38:523–568.
28. Aldová E, Hausner O, Gabrhelová M, Schindler J, Petráš P, Braná H. A hydrogen sulphide producing Gram-negative rod from water. *Zentralblatt Für Bakteriol Mikrobiol Hyg A*. 1983;254:95–108.
29. Gupta AK, Rastogi G, Nayduch D, Sawant SS, Bhonde RR, Shouche YS. Molecular phylogenetic profiling of gut-associated bacteria in larvae and adults of flesh flies. *Med Vet Entomol*. 2014;28:345–354.
30. Skrodenyte-Arbaciauskiene V, Sruoga A, Butkauskas D. Assessment of microbial diversity in the river trout *Salmo trutta fario* L. intestinal tract identified by partial 16S rRNA gene sequence analysis. *Fish Sci*. 2006;72:597–602.
31. Blekher L, Siegman-Igra Y, Schwartz D, Berger SA, Carmeli Y. Clinical significance and antibiotic resistance patterns of *Leminorella* spp., an emerging nosocomial pathogen. *J Clin Microbiol*. 2000;38:3036–3038.
32. Kline KA, Dodson KW, Caparon MG, Hultgren SJ. A tale of two pili: assembly and function of pili in bacteria. *Trends Microbiol*. 2010;18:224–232.
33. Rossez Y, Holmes A, Wolfson EB, et al. Flagella interact with ionic plant lipids to mediate adherence of pathogenic *Escherichia coli* to fresh produce plants. *Environ Microbiol*. 2014;16:2181–2195.
34. Boyer E, Bergevin I, Malo D, Gros P, Cellier MF. Acquisition of Mn(II) in addition to Fe(II) is required for full virulence of *Salmonella enterica* serovar Typhimurium. *Infect Immun*. 2002;70:6032–6042.
35. Skrodenyte-Arbaciauskiene V, Sruoga A, Butkauskas D, Skrupskelis K. Phylogenetic analysis of intestinal bacteria of freshwater salmon *Salmo salar* and sea trout *Salmo trutta trutta* and diet. *Fisheries Sci*. 2008;74:1307–1314.
36. Costa TRD, Felisberto-Rodrigues C, Meir A, et al. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol*. 2015;13:343–359.
37. Thomas S, Holland IB, Schmitt L. The Type 1 secretion pathway—the hemolysin system and beyond. *Biochim Biophys Acta*. 2014;1843:1629–1641.
38. Basler M. Type VI secretion system: secretion by a contractile nanomachine. *Philos Trans R Soc Lond B Biol Sci*. 2015;370:20150021.
39. Mougous JD, Cuff ME, Raunser S, et al. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science*. 2006;312:1526–1530.
40. Zhou Y, Tao J, Yu H, et al. Hcp family proteins secreted via the type VI secretion system coordinately regulate *Escherichia coli* K1 interaction with human brain microvascular endothelial cells. *Infect Immun*. 2012;80:1243–1251.
41. Riley MA, Gordon DM. The ecological role of bacteriocins in bacterial competition. *Trends Microbiol*. 1999;7:129–133.
42. Šmajs D, Micenková L, Šmarda J, et al. Bacteriocin synthesis in uropathogenic and commensal *Escherichia coli*: colicin E1 is a potential virulence factor. *BMC Microbiol*. 2010;10:288.
43. Sánchez-Romero MA, Cota I, Casadesús J. DNA methylation in bacteria: from the methyl group to the methylome. *Curr Opin Microbiol*. 2015;25:9–16.
44. Janda JM. New members of the family *Enterobacteriaceae*. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, eds. *The Prokaryotes: Volume 6: Proteobacteria: Gamma Subclass*. 3rd ed. New York, NY: Springer; 2006:5–40.
45. Tiedje JM. Ecology of denitrification and dissimilatory nitrate reduction to ammonium. In: Zehnder AJB, ed. *Environmental Microbiology of Anaerobes*. New York, NY: John Wiley & Sons; 1988:179–244.
46. Lerat E, Ochman H. Psi-Phi: exploring the outer limits of bacterial pseudogenes. *Genome Res*. 2004;14:2273–2278.
47. Kuo C-H, Ochman H. The extinction dynamics of bacterial pseudogenes. *PLoS Genet*. 2010;6:e1001050.