Adaptations to Different Temperatures and Habitats

Patrick C. Y. Woo1,2,3*, Susanna K. P. Lau1,2,3*, Herman Tse1,2,3*, Jade L. L. Teng3*, Shirley O. T. Curreem3, Alan K. L. Tsang3, Rachel Y. Y. Fan3, Gilman K. M. Wong3, Yi Huang3, Nicholas J. Loman4, Lori A. S. Snyder4, James J. Cai3, Jian-Dong Huang5, William Mak3, Mark J. Pallen4, Si Lok6, Kwok-Yung Yuen1,2,3*

1 State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, Special Administrative Region, People’s Republic of China, 2 Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong, Special Administrative Region, People’s Republic of China, 3 Department of Microbiology, The University of Hong Kong, Hong Kong, Special Administrative Region, People’s Republic of China, 4 Centre for Systems Biology, University of Birmingham, Birmingham, United Kingdom, 5 Department of Biochemistry, The University of Hong Kong, Hong Kong, Special Administrative Region, People’s Republic of China, 6 Genome Research Centre, The University of Hong Kong, Hong Kong, Special Administrative Region, People’s Republic of China

Abstract

Laribacter hongkongensis is a newly discovered Gram-negative bacillus of the Neisseriaceae family associated with freshwater fish–borne gastroenteritis and traveler’s diarrhea. The complete genome sequence of L. hongkongensis HLHK9, recovered from an immunocompetent patient with severe gastroenteritis, consists of a 3,169-kb chromosome with G+C content of 62.35%. Genome analysis reveals different mechanisms potentially important for its adaptation to diverse habitats of human and freshwater fish intestines and freshwater environments. The gene contents support its phenotypic properties and suggest that amino acids and fatty acids can be used as carbon sources. The extensive variety of transporters, including multidrug efflux and heavy metal transporters as well as genes involved in chemotaxis, may enable L. hongkongensis to survive in different environmental niches. Genes encoding urease, bile salts efflux pump, adhesin, catalase, superoxide dismutase, and other putative virulence factors—such as hemolysins, RTX toxins, patatin-like proteins, phospholipase A1, and collagenases—are present. Proteomes of L. hongkongensis HLHK9 cultured at 37°C (human body temperature) and 20°C (freshwater habitat temperature) showed differential gene expression, including two homologous copies of argB, argB-20, and argB-37, which encode two isoenzymes of N-acetyl-L-glutamate kinase (NAGK)—NAGK-20 and NAGK-37—in the arginine biosynthesis pathway. NAGK-20 showed higher expression at 20°C, whereas NAGK-37 showed higher expression at 37°C. NAGK-20 also had a lower optimal temperature for enzymatic activities and was inhibited by arginine probably as negative-feedback control. Similar duplicated copies of argB are also observed in bacteria from hot springs such as Thermus thermophilus, Deinococcus geothermalis, Deinococcus radiodurans, and Roseiflexus castenholzii, suggesting that similar mechanisms for temperature adaptation may be employed by other bacteria. Genome and proteome analysis of L. hongkongensis revealed novel mechanisms for adaptations to survival at different temperatures and habitats.

Introduction

Laribacter hongkongensis is a recently discovered, Gram-negative, facultative anaerobic, motile, seagull or S-shaped, asaccharolytic, urease-positive bacillus that belongs to the Neisseriaceae family of β-proteobacteria [1]. It was first isolated from the blood and thoracic empyema of an alcoholic liver cirrhosis patient in Hong Kong [2]. In a prospective study, L. hongkongensis was shown to be associated with community acquired gastroenteritis and traveler’s diarrhea [3,4]. L. hongkongensis is likely to be globally distributed, as travel histories from patients suggested its presence in at least four continents: Asia, Europe, Africa and Central America [4–6]. L. hongkongensis has been found in up to 60% of the intestines of commonly consumed freshwater fish, such as grass carp and bighead carp [4,7,8]. It has also been isolated from drinking water reservoirs in Hong Kong [9]. Pulsed-field gel electrophoresis and multilocus sequence typing showed that the fish and patient isolates fell into separate clusters, suggesting that some clones could be more virulent or adapted to human [8,10]. These data strongly suggest that this bacterium is a potential diarrheal pathogen that warrants further investigations.

Compared to other families such as Enterobacteriaceae, Vibrionaceae, Streptococcaceae, genomes of bacteria in the Neisseriaceae family have been relatively under-studied. Within this family, Neisseria...
**Author Summary**

*Laribacter hongkongensis* is a recently discovered bacterium associated with gastroenteritis and traveler’s diarrhea. Freshwater fish is the reservoir of *L. hongkongensis*. In order to achieve a rapid understanding on the mechanisms by which the bacterium adapts to different habitats and its potential virulence factors, we sequenced the complete genome of *L. hongkongensis*, compared its gene contents with other bacteria, and compared its gene expression at 37°C (human body temperature) and 20°C (freshwater habitat temperature). We found that the gene contents of *L. hongkongensis* enable it to adapt to its diverse habitats of human and freshwater fish intestines and freshwater environments. Genes encoding proteins responsible for survival in the intestinal environments, adhesion to intestinal cells, evasion from host immune systems, and putative virulence factors similar to those observed in other pathogens are present. We also observed, in gene expression studies, that *L. hongkongensis* may be using different pathways for arginine synthesis regulated at different temperatures. Phylogenetic analysis suggested that such mechanisms for temperature adaptation may also be used in bacteria found in extreme temperatures.

**Results/Discussion**

**General Metabolism**

The *L. hongkongensis* genome lacks a complete set of enzymes for glycolysis, with orthologues of glucokinase, 6-phosphofructokinase and pyruvate kinase being absent (Table S1). This is compatible with its asaccharolytic phenotype and is consistent with other asaccharolytic bacteria, such as *Campylobacter jejuni*, *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*, in that glucokinase and 6-phosphofructokinase are also absent from their genomes [18,19]. On the other hand, the *L. hongkongensis* genome encodes the complete sets of enzymes for gluconeogenesis, the pentose phosphate pathway and the glyoxylate cycle (Table S1). Similar to *C. jejuni*, the *L. hongkongensis* genome encodes a number of extracellular proteases and amino acid transporters. These amino acids can be used as carbon source for the bacterium. The genome encodes enzymes for biosynthesis of the 21 genetically encoded amino acids and for biosynthesis and β-oxidation of saturated fatty acids (Tables S2 and S3). The *L. hongkongensis* genome encodes a variety of dehydrogenases (LHK_00527–00540, LHK_01219–01224, LHK_02418–02421, LHK_00801–00803, LHK_01861, LHK_02912–02913 and LHK_00934) that enable it to utilize a variety of substrates as electron donors, such as NADH, succinate, formate, proline, acyl-CoA and D-amino acids. The presence of three terminal cytochrome oxidases may allow *L. hongkongensis* to carry out respiration using oxygen as the electron acceptor under both aerobic conditions [type aa3 oxidase *(LHK_00169–00170, LHK_00173)*] and with reduced oxygen tension [type cb, *(LHK_00995–00996, LHK_00998)*] and type bel (LHK_02252–02253) oxidases]. The genome also encodes a number of reductases [fumarate reductase (LHK_02340–02342), nitrate reductase (LHK_02079–02083), dimethylsulfoxide (DMSO) reductase (LHK_02496–02498) and tetrathionate reductase (LHK_01476–01478), which may help carry out respiration with alternative electron acceptors to oxygen (fumarate, nitrate, DMSO and tetrathionate) under anaerobic conditions. This is supported by the enhanced growth of *L. hongkongensis* under anaerobic conditions in the presence of nitrate (data not shown). Further studies are required to confirm if the bacterium can utilize other potential electron acceptors.

**Transportsers**

There were 441 transport-related proteins (13.6% of all CDSs) in the *L. hongkongensis* genome, comprising an extensive variety of transporters, which may reflect its ability to adapt to the freshwater fish and human intestines, and freshwater environments. According to the Transporter Classification Database (TCDB) (http://www.tcdb.org/), all seven major categories of transporters are present in *L. hongkongensis*. Primary active transporters (class 3 transporters) were the most abundant class of transporters, accounting for 43.3% (191 CDSs) of all annotated CDSs related to transport, among which 104 belong to the ATP-binding cassette (ABC) transporter superfamily and 41 were oxidoreduction-driven transporters. Electrochemical potential-driven transporters (class 2 transporters) were the second most abundant class of transporters, accounting for 27.9% (123 CDSs) of all annotated CDSs related to transport, most of which (117 CDSs) are various kinds of porter including major facilitator superfamily (MFS) (19 CDSs), resistance-nodulation-cell division (RND) superfamily (22 CDSs), amino acid-polyamine-organocation family (8 CDSs), dicarboxylate/amino acid:cation symporter (DAACS) family (5 CDSs) and monovalent cation:proton antiporter-2 family (3 CDSs), and various heavy metal transporters which may be involved in detoxification and resistance against environmental hazards. Three different types of class 2 transporters, belonging to the DAAGS, tripartite ATP-independent periplasmic transporter and...
C₄-dicarboxylate uptake C family, are likely involved in the transport of malate, which can be used as the sole carbon source for \textit{L. hongkongensis} in minimal medium [unpublished data]. The remaining class 2 transporters were ion-gradient-driven energizers belonging to the TonB family (6 CDSs). The third most abundant class of transporters was the channels and pores (class 1), with 39 CDSs including 12 \(\alpha\)-type channels, 26 \(\beta\)-barrel porins. Among the 12 \(\alpha\)-type channels, four were mechanosensitive channels.
which are important for mediating resistance to mechanophysical changes. The remaining transporters belong to four other classes, namely group translocators (class 4, 9 CDSs), transport electron carriers (class 5, 16 CDSs), accessory factors involved in transport (class 6, 23 CDSs) and incompletely characterized transport system (class 9, 54 CDSs). In line with their asaccharolytic nature, the genomes of *L. hongkongensis* and *C. jejuni* do not contain genes that encode a complete phosphotransferase system. The five families of multidrug efflux transporters, including MFS (6 CDSs), RND (8 CDSs), small multidrug resistance family (2 CDSs), multidrug and toxic compound extrusion family (2 CDSs) and ABC transporter superfamily (5 CDSs), were all present in *L. hongkongensis*, which may reflect its ability to withstand toxic substances in different habitats [20]. 20 CDSs were related to iron metabolism, including hemin transporters, ABC transporters of the metal type and iron-storage proteins and the Fur protein responsible for iron uptake regulation. In contrast to *C. violaceum* which produces siderophores for iron acquisition, but similar to the genome of *C. jejuni*, which possesses two copies of such genes.

**Motility, Chemotaxis, and Signal Transduction**

Except the first strain of *L. hongkongensis* isolated from the blood and empyema pus of a patient which represented a non-motile variant, all *L. hongkongensis* strains, whether from human diarrheal stool, fish intestine or environmental water, are motile with polar flagella. The ability to sense and respond to environmental signals is important for survival in changing ecological niches. A total of 47 CDSs are related to chemotaxis, of which 27 encode methyl-accepting chemotaxis proteins (MCPs) and 20 encode chemosensory transducer proteins. While most MCPs are scattered throughout the genome, the transducer proteins are mostly arranged in three gene clusters (Figure S1). At least 38 genes in six gene clusters, are involved in the biosynthesis of flagella (Figure S2).

Enteric bacteria use several quorum-sensing mechanisms, including the LuxR-I, LuxS/Al-2, and Al-3/epinephrine/norepinephrine systems, to recognize the host environment and communicate across species. Unlike the genomes of *C. violaceum* and the pathogenic *Neisseria* species which encode genes involved in LuxR-I and LuxS/Al-2 systems respectively, the *L. hongkongensis* genome does not encode genes of these 2 systems. Instead, the Al-3/epinephrine/norepinephrine system, which is involved in inter-kingdom cross-signaling and regulation of virulence gene transcription and motility, best characterized in enterohemorrhagic *E. coli* [21,22], is likely the predominant quorum-sensing mechanism used by *L. hongkongensis*. Several human enteric commensals or pathogens, including *E. coli*, *Shigella*, and *Salmonella*, produce Al-3 [23]. A two-component system, QseB/C, of which QseC is the sensor kinase and QseB the response regulator, has been found to be involved in sensing AI-3 from bacteria and epinephrine/norepinephrine from host, and activation of the flagellar regulon transcription [21]. While the biosynthetic pathway of Al-3 has not been discovered, two sets of genes, LHK_00329/LHK_00328 and LHK_01812/LHK_01813, homologous to QseB/QseC were identified in the *L. hongkongensis* genome, suggesting that *L. hongkongensis* is able to utilize exogenous siderophores or host proteins for iron acquisition, which may be important for survival in different environments and hosts.

**Pathogenic Factors**

Before reaching the human intestine, *L. hongkongensis* has to pass through the highly acidic environment of the stomach. In the *L.
**Laribacter hongkongensis** genome, a cluster of genes, spanning a 12-kb region, related to acid resistance, is present. Similar to *Helicobacter pylori*, the *L. hongkongensis* genome contains a complete urease gene cluster (LHK\_01035–LHK\_01037, LHK\_01040–LHK\_01044), in line with the bacterium's urease activity. Phylogenetically, all 9 genes in the urease cassette are most closely related to the corresponding homologues in *Brucella* species (β-proteobacteria), *Yersinia* species (γ-proteobacteria) and *Photorhabdus luminescens* (γ-proteobacteria), instead of those in other members of β-proteobacteria, indicating that *L. hongkongensis* has probably acquired the genes through horizontal gene transfer after its evolution into a distinct species (Figure S3). Upstream and downstream to the urease cassette, adi (LHK\_01034) and hdeA (LHK\_01046) were found respectively. Their activities will raise the cytoplasmic pH and prevents proteins in the periplasmic space from aggregation during acid shock respectively [24,25]. In addition to the acid resistance gene cluster, the *L. hongkongensis* genome contains two arc gene clusters [arcA (LHK\_02729 and LHK\_02734), arcB (LHK\_02728 and LHK\_02733), arcC (LHK\_02727 and LHK\_02732) and arcD (LHK\_02730 and LHK\_02731)] of the arginine deiminase pathway which converts L-arginine to carbon dioxide, ATP, and ammonia. The production of ammonia increases the pH of the local environment [26,27].

Similar to other pathogenic bacteria of the gastrointestinal tract, the genome of *L. hongkongensis* encodes genes for bile resistance. These include three complete copies of arcAB (LHK\_01425–01426, LHK\_02129–02130 and LHK\_02929–02930), encoding the best studied efflux pump for bile salts, and two pairs of genes (LHK\_01373–01374 and LHK\_03132–03133) that encode putative efflux pumps homologous to that encoded by emaAB in *E. coli* [28]. Furthermore, five genes [tolQ (LHK\_00053), tolR (LHK\_03174), tolA (LHK\_03173), tolB (LHK\_03172) and pal (LHK\_03171)] that encode the Tol proteins, important in maintaining the integrity of the outer membrane and for bile resistance, are also present [29].

In the *L. hongkongensis* genome, a putative adhesin (LHK\_01901) for colonization of the intestinal mucosa, most closely related to the adhesins of diffusely adherent *E. coli* (DAEC) and enterotoxigenic *E. coli* (ETEC), encoded by aidA and tibA respectively, was observed (Figure S4) [30,31]. aidA and tibA encode proteins of the autotransporter family, type V protein secretion system of Gram-negative bacteria. All three domains (an N-terminal signal sequence, a passenger domain and a translocation domain) present in proteins of this family are found in the putative adhesin in *L. hongkongensis*. Moreover, a putative heptosyltransferase (LHK\_01902), with 52% amino acid identity to the TibC heptosyltransferase of ETEC, responsible for addition of heptose to the passenger domain, was present upstream to the putative adhesin gene (Figure S4). In addition to host cell adhesion, the passenger domains of autotransporters may also confer various virulence functions, including autogregation, invasion, biofilm formation and cytotoxicity. The *L. hongkongensis* genome encodes a putative superoxide dismutase (LHK\_01716) and catalases (LHK\_01264, LHK\_01300 and LHK\_02456), which may play a role in resistance to superoxide radicals and hydrogen peroxide generated by neutrophils.

The same set of genes that encode enzymes for synthesis of lipid A (endotoxin), the two Kdo units and the heptose units of lipopolysaccharide (LPS) are present in the genomes of *L. hongkongensis*, *C. violaceum*, *N. meningitidis*, *N. gonorrhoeae* and *E. coli*. Moreover, 9 genes [rfbD (LHK\_02995), rfbE (LHK\_02997), rfbC (LHK\_02994), rfbD (LHK\_02996), wbmF (LHK\_02799), wbmG (LHK\_02800), wbmH (LHK\_02801), wbmI (LHK\_02790) and wbmK (LHK\_02792)] that encode putative enzymes for biosynthesis of the polysaccharide side chains are present in the *L. hongkongensis* genome. In addition to genes for synthesizing LPS, a number of CDSs that encode putative cytotoxins are present, including cytotoxins that act on the cell surface [hemolysins (LHK\_00556 and LHK\_03166) and RTX toxins (LHK\_02735 and LHK\_02910)] and those that act intracellularly [patatin-like proteins (LHK\_00116, LHK\_01938, and LHK\_03115)] [32,33]. Furthermore, a number of CDSs that encode putative outer membrane phospholipase A1 (LHK\_00790) and collagenases (LHK\_00305–00306, LHK\_00451, and LHK\_02651) for possible bacterial invasion are present.

**Adaptability to Different Environmental Temperatures**

To better understand how *L. hongkongensis* adapts to human body and freshwater habitat temperatures at the molecular level, the types and quantities of proteins expressed in *L. hongkongensis* HLJHK9 cultured at 37°C and 20°C were compared. Since initial 2D gel electrophoresis analysis of *L. hongkongensis* HLJHK9 proteins under a broad range of pH and molecular weight conditions revealed that the majority of the proteins reside on the weakly acidic to neutral portion, with a minority on the weak basic portion, consistent with the median pI value of 6.63 calculated for all putative proteins in the genome of *L. hongkongensis* HLJHK9, we therefore focused on IPG strips of pH 4–7 and 7–10. Comparison of the 2D gel electrophoresis patterns from *L. hongkongensis* HLJHK9 cells grown at 20°C and 37°C revealed 12 differentially expressed protein spots, with 7 being more highly expressed at 20°C than at 37°C and 5 being more highly expressed at 37°C than at 20°C (Table 2, Figure 3). The identified proteins were involved in various functions (Table 2). Of note, spot 8 [N-acetyl-L-glutamate kinase (NAGK)-37, encoded by argB-37] was up-regulated at 37°C, whereas spot 1 (NAGK-20, encoded by argB-20), was up-regulated at 20°C (Figures 3, 4A and 4B). These two homologous copies of argB encode two isoenzymes of NAGK [NAGK-20 (LHK\_02829) and NAGK-37 (LHK\_02373)], which catalyze the second step of the arginine biosynthesis pathway.

The transcription levels of argB-20 and argB-37 at 20°C and 37°C were quantified by real time RT-PCR. Results showed that the mRNA level of argB-20 at 20°C was significantly higher that at 37°C and the mRNA level of argB-37 at 37°C was significantly higher that at 20°C (Figure 4C and 4D), suggesting that their expressions, similar to most other bacterial genes, were controlled at the transcription level. When argB-20 and argB-37 were cloned, expressed and the corresponding proteins NAGK-20 and NAGK-37 purified for enzyme assays, their highest enzymatic activities were observed at 37–45°C and 45–50°C respectively (Figure 4E). Moreover, NAGK-20, but not NAGK-37, was inhibited by 0.25–10 mM of arginine (Figure 4F).

*L. hongkongensis* probably regulates arginine biosynthesis at temperatures of different habitats using two pathways with two isoenzymes of NAGK. *L. hongkongensis* and wild type *E. coli* ATCC 25922, but not *E. coli* JW5553-1 (argB deletion mutant), grew in minimal medium without arginine, indicating that *L. hongkongensis* contains a functional arginine biosynthesis pathway. NAGK-20 is expressed at higher level at 20°C than 37°C, whereas NAGK-37 is expressed at higher level at 37°C than 20°C. Bacteria use either of two different pathways, linear and cyclic, for arginine biosynthesis. Similar to NAGK-20 of *L. hongkongensis*, NAGK of *Pseudomonas aeruginosa* and *Thermotoga maritima*, which employ the cyclic pathway, can be inhibited by arginine as the rate-limiting enzyme for negative feedback control [34–37]. On the other hand, similar to NAGK-37 of *L. hongkongensis*, NAGK of *E. coli*, which employs the linear pathway, is not inhibited by arginine [35,36]. We speculate that *L. hongkongensis* can use different pathways with the...
Table 2. Differential protein expression of *L. hongkongensis* at 20°C and 37°C.

| Spot no. | Gene no.     | Theoretical MW (kDa) | Theoretical p/ | Proteins                                      | Functional categories (COG)                              | Sequence coverage (%) | Peptide matched | MOWSE score | 20°C/37°C ratio | p value |
|----------|--------------|----------------------|----------------|-----------------------------------------------|--------------------------------------------------------|-----------------------|----------------|--------------|----------------|---------|
| Higher expression at 20°C |              |                      |                |                                               |                                                        |                       |                |              |                |         |
| 1        | LHK_02829    | 30.0                 | 5.03           | Acetylglutamate kinase, NAGK-20              | Amino acid transport and metabolism                     | 24                    | 6               | 9.49e+003   | +2.50          | <0.005  |
| 2        | LHK_01819    | 39.0                 | 6.30           | Probable porin protein                       | Cell wall/membrane/envelope biogenesis                  | 36                    | 9               | 3.86e+005   | +2.41          | <0.005  |
| 3        | LHK_00236    | 49.5                 | 7.06           | Peptidase M16 domain protein precursor       | —                                                      | 33                    | 14              | 2.47e+005   | +2.04          | <0.05   |
| Only at 20°C |              |                      |                |                                               |                                                        |                       |                |              |                |         |
| 4        | LHK_00011    | 16.9                 | 4.99           | Protein-L-isospartate (D-aspartate) O-methyltransferase | Post-translational modification, protein turnover, chaperones | 41                    | 9               | 6.28e+03   | —              | <0.01   |
| 5        | LHK_01152    | 49.0                 | 5.44           | Probable phage sheath protein                | —                                                      | 28                    | 10              | 1.11e+005   | —              | <0.05   |
| 6        | LHK_01472    | 23.4                 | 4.70           | Ribonuclease activity regulator protein RaA  | General function prediction only                       | 36                    | 7               | 4.13e+03   | —              | <0.01   |
| 7        | LHK_01018    | 30.3                 | 6.66           | Methylene tetrahydro-folate dehydrogenase/cyclohydrolase | Coenzyme transport and metabolism                      | 45                    | 14              | 6.67e+07   | —              | <0.05   |
| Higher expression at 37°C |              |                      |                |                                               |                                                        |                       |                |              |                |         |
| 8        | LHK_02337    | 31.5                 | 5.12           | Acetylglutamate kinase, NAGK-37              | Amino acid transport and metabolism                     | 38                    | 13              | 3.48e+004   | −2.09          | <0.005  |
| 9        | LHK_02119    | 23.7                 | 9.22           | Hypothetical protein                         | —                                                      | 41                    | 8               | 1.24e+004   | −3.06          | <0.005  |
| Only at 37°C |              |                      |                |                                               |                                                        |                       |                |              |                |         |
| 10       | LHK_02507    | 44.6                 | 9.11           | Probable N-acetylmuramoyl-L-alanine amidase  | Cell wall/membrane/envelope biogenesis                  | 58                    | 20              | 6.01e+009   | —              | <0.001  |
| 11       | LHK_02507    | 44.6                 | 9.11           | Probable N-acetylmuramoyl-L-alanine amidase  | Cell wall/membrane/envelope biogenesis                  | 68                    | 23              | 7.66e+010   | —              | <0.05   |
| 12       | LHK_03194    | 48.9                 | 7.94           | Survival protein SurA precursor             | Post-translational modification, protein turnover, chaperones | 33                    | 10              | 3.79e+006   | —              | <0.005  |

*a* A negative number indicates the inverse ratio of that indicated in the column heading.
doi:10.1371/journal.pgen.1000416.t002
two NAGK isoenzymes with differential importance at different temperatures of different habitats.

Phylogenetic analysis of NAGK-20 and NAGK-37 showed that they were more closely related to each other than to homologues in other bacteria (Figure 5). The topology of the phylogenetic tree constructed using NAGK was similar to that constructed using 16S rRNA gene sequences (data not shown). This suggested that the evolution of \textit{argB} genes in general paralleled the evolution of the corresponding bacteria, and \textit{argB} gene duplication has probably occurred after the evolution of \textit{L. hongkongensis} into a separate species. The requirement to adapt to different temperatures and habitats may have provided the driving force for subsequent evolution to 2 homologous proteins that serve in different environments. Notably, among all 465 bacterial species with complete genome sequences available, only \textit{Thermus thermophilus}, \textit{Deinococcus geothermalis}, \textit{Deinococcus radiodurans}, \textit{Roseiflexus castenholzi} and \textit{Roseiflexus} sp. RS-1 possessed two copies of \textit{argB}, whereas \textit{Anaeromyxobacter} sp. Fw109-5 and \textit{Anaeromyxobacter dehalo-
Figure 4. Proteomics, transcriptional and biochemical analysis of N-acetyl-L-glutamate kinase (NAGK)-20 and NAGK-37. Differential expressions of NAGK-20 (pink arrow) and NAGK-37 (blue arrow) in L. hongkongensis cultured at 20°C (A) and 37°C (B). Normalized mRNA levels of argB-20 (C) and argB-37 (D) in L. hongkongensis cells cultured at 20°C and 37°C. Data were analyzed by unpaired Student’s t-test. The data represent...
genus 2CP-C possessed one copy of argB and another fused with argJ (Figure 5). The clustering of argB in two separate groups in these bacteria suggests that argB gene duplication has probably occurred in their ancestor, before the divergence into separate species. The prevalence of *T. thermophilus*, *Deinococcus* species and *Roseiflexus* species in hot springs suggested that this novel mechanism of temperature adaptation may also be important for survival at different temperatures in other bacteria. Further experiments on differential expression of the two isoenzymes at different temperatures in these bacteria will verify our speculations.

Conclusion

Traditionally, complete genomes of bacteria with medical, biological, phylogenetic or industrial interests were sequenced only after profound phenotypic and genotypic characterization of the bacteria had been performed. With the advance in technology and bioinformatics tools, complete genome sequences of bacteria can be obtained with greater ease. In this study, we sequenced and analyzed the complete genome of *L. hongkongensis*, a newly discovered bacterium of emerging medical and phylogenetic interest, and performed differential proteomics and downstream characterization of important pathways. In addition, putative virulence factors and a putative novel mechanism of arginine biosynthesis regulation at different temperatures were discovered, further characterization of which will lead to better understanding of their contributions to the survival and virulence of *L. hongkongensis*, the *Neisseriaceae* family and other bacteria. A similar “reverse genomics” approach can be used for the study of other newly discovered important bacteria.

Materials and Methods

Sequencing and Assembly

The genome sequence of *L. hongkongensis* HLHK9 was determined with the whole-genome shotgun method. Three shotgun libraries were generated: one small-insert (2–4 kb) library and one medium-insert (5–6 kb) library in pcDNA2.1, and a large-insert (35–45 kb) fosmid library in pCC2FOS. DNA sequencing was performed using dye-terminator chemistries on ABI3700.

The means of three independent experiments. Error bars represent standard deviations. Significant changes are represented by asterisks (*, $P<0.05$).

(E) Specific kinase activities of purified NAGK-20 (pink) and NAGK-37 (blue) measured at different temperatures (25–60 °C). The data represent the means of three independent experiments. Error bars represent standard deviations. (F) Effects of arginine on kinase activities of purified NAGK-20 (pink) and NAGK-37 (blue) measured in arginine at different concentrations (0.25–20 mM). The data represent the means of three independent experiments. Error bars represent standard deviations.

doi:10.1371/journal.pgen.1000416.g004
sequences. Shotgun sequences were assembled with Phrap. Fosmid end sequences were mapped onto the assembly using BACCardI [30] for validation and support of gap closing. Sequences of all large repeat elements (rRNA operons and prophages) were confirmed by primer walking of fosmid clones. The nucleotide sequence for the complete genome sequence of *L. hongkongensis* HLHK9 was submitted to Genbank under accession number CP001154.

**Genome Annotation**

Gene prediction was performed by Glimmer [39] version 3.02, and results post-processed using TICO [40] for improving predictions of translation initiation sites. Automated annotation of the finished sequence was performed by a modified version of AutoFACT [41], supplemented by analysis by InterProScan [42]. Manual curation of annotation results was done with support from the software tool GenDB [43]. In addition, annotation of membrane transport proteins was done by performing BLAST search of all predicted genes against the curated TCDB [44]. Ribosomal RNA genes were annotated using the online RNAmer service [45]. Putative prophage sequences were identified using Prophage Finder [46]. Frameshift errors were predicted using ProFED [47]. CRISPRs (Clustered Regularly Interspaced Short Palindromic Repeats) were searched by using PILER-CR [48], CRISPRFinder [49] and CRT (CRISPR recognition tool) [50].

**Proteomic Analysis by 2D Gel Electrophoresis**

Single colony of *L. hongkongensis* HLHK9 was inoculated into brain heart infusion (BHI) medium for 16 h. The bacterial cultures were diluted 1:100 in BHI medium and growth was continued at 20°C for 20 h and 37°C for 6 h, respectively, with shaking to OD600 of 0.6. After centrifugation at 6,500 × g for 15 min, cells were lysed in a sample buffer containing 7 M urea, 2 M thiourea and 4% CHAPS. The crude cell homogenate was sonicated and centrifuged at 16,000 × g for 20 min. Immobilized pH gradient (IPG) strips (Bio-Rad Laboratories) (17 cm) with pH 4–7 and 7–11 (IPG strip of pH 7–10) (GE Healthcare) were then thermal-cycled in 40 cycles of 95°C for 15 s and 60°C for 60 s. puts UV light to KB-1000 (GE Healthcare). ImageMaster 2D Platinum 6.0 (GE Healthcare) and ImageMaster 2D Platinum 6.0 (GE Healthcare) were used for image analysis. For MALDI-TOF analysis, protein spots were manually excised from gels and subjected to protein identification by MALDI-TOF MS analysis. For MALDI-TOF MS analysis, protein spots were manually excised from gels and subjected to protein identification by MALDI-TOF MS analysis. Three independent experiments for each growth condition were performed.

**Essentiality of Arginine for Growth of *L. hongkongensis* HLHK9**

*L. hongkongensis* HLHK9 cells were grown in minimal medium M63 [51] supplemented with 20 mM L-malate as carbon source and 19 mM potassium nitrate as nitrogen source, and 1 mM each of vitamin B1 and vitamin B12. The pH of all media was adjusted to 7.0 with KOH. Essentiality of arginine for growth of *L. hongkongensis* HLHK9 was determined by transferring the bacterial cells to the modified M63 medium with or without 100 mM of L-arginine. *Esherichia coli* ATCC 25922 and JW5534-1 (argB deletion mutant) [52] were used as positive and negative controls respectively. All cultures were incubated at 37°C with shaking for 5 days. Growth in each medium was determined by measuring absorbance spectrophotometrically at OD600. The experiment was performed in duplicate.

**Real-Time RT-PCR**

mRNA levels of *argB-20* and *argB-37* in *L. hongkongensis* HLHK9 cells grown in 20°C and 37°C were compared. Total RNA was extracted from culture of *L. hongkongensis* HLHK9 (OD600 of 0.6) grown in conditions described in proteomic analysis by using RNeasy kit (Qiagen) in combination with RNPackt Bacteria Reagent (Qiagen) as described by the manufacturer. Genomic DNA was removed by DNase digestion using RNase-free DNase I (Roche). The total nucleic acid concentration and purity were estimated using a A260/A280 values measured by NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Bacteria were harvested from three independent replicate cultures. cDNA was synthesized by RT using random hexamers and SuperScript III kit (Invitrogen) as described previously [53,54]. cDNA was amplified by TaqMan PCR Core Reagent kit (Applied Biosystems) in an ABI Prism 7000 Sequence Detection System (Applied Biosystems). Briefly, 2 μl of cDNA was amplified in a 25 μl reaction containing 2.5 μl of TaqMan buffer A, 5.3 μM of MgCl2, 0.2 μM of each deoxynucleoside triphosphates (dNTPs), 0.8 μM of each primer, 0.8 μM of gene-specific TaqMan probe with a 5’-[6-carboxyfluorescein (FAM)] reporter dye and a 3’- [6-carboxytetramethylrhodamine (TAMRA)] quencher dye, 2.5 U of AmpErase Uracil N-glycosylase (UNG) and 0.625 U AmpliTaq Gold polymerase (Applied Biosystems). Primers and TaqMan probes were designed using Primer Express software, version 2.0 (Applied Biosystems) (Table S4). Reactions were first incubated at 50°C for 10 min in duplicate wells. Reactions were then thermal-cycled in 40 cycles of 95°C for 15 s and 60°C for 1 min. Absolute standard curve method was used for determination of transcript level for each gene. Standard curves were made by using serial dilutions from plasmids containing the target sequences with known quantities. Housekeeping gene RNA polymerase beta subunit, *rpoB*, was used as an internal control. Triplicate assays using RNAs extracted in three independent experiments confirmed that transcript levels of *rpoB* were not significantly different (*P*>0.05) at 20°C compared with 37°C (data not shown). The transcript levels of *argB-20* and *argB-37* were then normalized to that of *rpoB*. Triplicate assays using RNAs extracted in three independent experiments were performed for each target gene.

**Phylogenetic Characterization**

The phylogenetic relationships among NAGK-20 and NAGK-57 of *L. hongkongensis* HLHK9 and their homologues in other bacteria with complete genomes available were analyzed. Phylogenetic tree was constructed by the neighbor-joining method using Kimura’s two-parameter correction with ClustalX 1.83. Three hundred and eleven positions were included in the analysis.
Cloning and Purification of (His)_6-Tagged Recombinant NAGK Proteins of *L. hongkongensis* HHK9

Cloning and purification of (His)_6-tagged recombinant NAGK proteins of *L. hongkongensis* HHK9 was performed according to our previous publications [33,35]. To produce plasmids for protein purification, primers 5’- GGAATTC-TATGTCGTGGCTCAGACGCCC -3’ and 5’- GGAATTC-TATGTCCAAGGCTGCCGCGATCAT -3’ for argB-20 and 5’-GGAATTC-ATATGGTTATTCAATCTGAAGT -3’ and 5’-GGAATTC-ATATGTCAGAGCGTGGTACAGAT -3’ for argB-37 were used to amplify the genes encoding NAGK-20 and NAGK-37, respectively, by PCR. The sequence coding for amino acid residues of the complete NAGK-20 and NAGK-37 was amplified and cloned, respectively, into the NdeI site of expression vector PET-28b(+) (Novagen) in frame and downstream of the series of six histidine residues. The two recombinant NAGK proteins were expressed and purified using the Ni^{2+}-loaded HiTrap Chelating System according to the manufacturer’s instructions (GE Healthcare).

**Enzyme Assays**

Purified NAGK-20 and NAGK-37 were assayed for N-acetyl-L-glutamate kinase activity using Haas and Leisinger’s method [56], with modifications. The reaction mixtures contained 400 mM NH_4OH·HCl, 400 mM Tris·HCl, 40 mM N-acetyl-L-glutamate, 20 mM MgCl_2, 10 mM ATP and 2 μg of enzyme in a final volume of 1.0 ml at pH 7.0. After incubation at 25°C, 30°C, 37°C, 45°C, 50°C or 60°C for 30 min, the reaction was terminated by adding 1.0 ml of a stop solution containing 5% (w/v) FeCl_3·6H_2O, 8% (w/v) trichloroacetic acid and 0.3 M HCl. The absorbance of the hydroxamate·Fe^{3+} complex was measured with a spectrophotometer at A_{460} [57]. Inhibition of the kinase activities of NAGK-20 and NAGK-37 were examined with and without 0.25, 0.5, 0.75, 1, 2.5, 5, 10, and 20 mM of L-arginine and incubated at 37°C for 30 min. One unit of N-acetyl-L-glutamate kinase is defined as the amount of enzyme required to catalyze the formation of 1 μmol of product per min under the assay conditions used. Each assay was performed in duplicate. Results were presented as means and standard deviations of three independent experiments.

**Supporting Information**

**Figure S1**  Physical map of the chemotaxis-related genes in *L. hongkongensis*. While the three gene clusters contain the transducer proteins and some of the methyl-accepting proteins (MCPs), most MCPs are scattered outside the clusters. Genes in orange are coding for chemotaxis transducer proteins; genes in green are coding for MCPs; genes in grey are coding for hypothetical proteins. The numbers refer to the coding sequences in the *L. hongkongensis* genome. Found at: doi:10.1371/journal.pgen.1000416.s001 (1.62 MB TIF)

**Figure S2**  Physical map of six gene clusters of flagellar genes of *L. hongkongensis*. The numbers refer to the coding sequences in the *L. hongkongensis* genome. Genes in pink are regulatory genes for flagellar gene expression; genes in light blue are coding for export apparatus proteins; genes in red are coding for proteins of motor complex/basal body; genes in yellow are coding for hook proteins; genes in green are coding for filament proteins; genes in grey are coding for hypothetical proteins or proteins with other functions; the gene in orange is coding for chemotaxis-related protein. Found at: doi:10.1371/journal.pgen.1000416.s002 (1.77 MB TIF)

**Figure S3**  Phylogenetic analysis of the structural (UreA, UreB and UreC) and accessory proteins (UreE, UreF, UreG, UreD and UreI) in the urease cassette of *L. hongkongensis*. The trees were constructed by the neighbor-joining method and bootstrap values calculated from 1,000 trees. One hundred, 131, 572, 190, 231, 211, 317, and 330 amino acid positions in UreA, UreB, UreC, UreE, UreF, UreG, UreD and UreI, respectively, were included in the analysis. The corresponding amino acid sequences of *S. annus* were used as outgroups. The scale bar indicates the estimated number of substitutions per 20 or 50 amino acids as indicated. All names and accession numbers are given as cited in the GenBank database. Found at: doi:10.1371/journal.pgen.1000416.s003 (0.19 MB PDF)

**Figure S4**  Schematic diagram of the putative heptosyltransferase and adhesin in *L. hongkongensis* and the corresponding homologues in enterotoxigenic E. coli (ETEC). The three functional domains of the putative adhesin are depicted [SS = N-terminal signal sequence (amino acid residues 1–36), a passenger or α-domain (amino acid residues 37–756), translocation or β-domain (amino acid residues 757–1023)]. Alignment of amino acid sequences of the passenger domain of the putative adhesin in *L. hongkongensis* and that of T1fA adhesin in ETEC. Residues that match the putative acceptor sites for the heptosyltransferase are boxed. The shaded boxes represent the consensus motifs [Y/V/1/F/W-X-(F/W) at the last three residues of the translocation domains. Found at: doi:10.1371/journal.pgen.1000416.s004 (0.11 MB TIF)

**Table S1**  Comparison of metabolic pathways for carbohydrate metabolism deduced from the genomes of *L. hongkongensis*, *C. violaceum*, *N. gonorrhoeae* and *N. meningitidis*. Found at: doi:10.1371/journal.pgen.1000416.s005 (0.03 MB DOC)

**Table S2**  Comparison of metabolic pathways for amino acid metabolism deduced from the genomes of *L. hongkongensis*, *C. violaceum*, *N. gonorrhoeae* and *N. meningitidis*. Found at: doi:10.1371/journal.pgen.1000416.s006 (0.04 MB DOC)

**Table S3**  Comparison of metabolic pathways for fatty acid metabolism deduced from the genomes of *L. hongkongensis*, *C. violaceum*, *N. gonorrhoeae* and *N. meningitidis*. Found at: doi:10.1371/journal.pgen.1000416.s007 (0.03 MB DOC)

**Table S4**  Primers and probes for quantitative RT-PCR. Found at: doi:10.1371/journal.pgen.1000416.s008 (0.03 MB DOC)

**Acknowledgments**

We are grateful to Professor Lap-Chee Tsui’s advice on sequencing strategies, the support of Professor Paul Tam and the Genome Research Centre, The University of Hong Kong, on the genomic sequencing platform, and Crystal Lai, Ian Melhado, Angel Ma, Wing Tong and Carol Lau for technical support.

**Author Contributions**

Conceived and designed the experiments: PCYW SKPL KYY. Performed the experiments: JLLT SOTC AKLT RYYF. Contributed reagents/materials/analysis tools: PCYW SKPL HT JLLT. Performed genome assembly, bioinformatics analysis, and genome annotation: HT JLLT SOTC AKLT RYYF. GKMW YH NJL LASS JJC.
References

1. Woo PC, Lau SK, Teng JL, Yuen KY (2005) Current status and future directions of Laribacter hongkongensis, a novel bacterium associated with gastroenteritis and traveler's diarrhea. Curr Opin Infect Dis 18: 413–419.

2. Yuen KY, Lau PC, Teng JL, Leung KW, Wong MK, et al. (2001) Laribacter hongkongensis gen. nov., sp. nov., a novel Gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. J Clin Microbiol 39: 4227–4233.

3. Lau SK, Woo PC, Fan RW, Lee RG, Teng JL, et al. (2007) Seasonal and tissue distribution of Laribacter hongkongensis, a novel bacterium associated with gastroenteritis, in retail freshwater fish in Hong Kong. Int J Food Microbiol 113: 62–66.

4. Teng JL, Woo PC, Ma SS, Sit TH, Ng LT, et al. (2000) Characterization of a chromosomal class C inducible expression shuttle vector for Listeria monocytogenes. J Bacteriol 182: 507–513.

5. Lau SK, Woo PC, Fan RW, Lee RG, Teng JL, et al. (2007) Antimicrob Agents Chemother 49: 1957–1964.

6. Antimicrob Agents Chemother 49: 1957–1964.

7. Lau SK, Woo PC, Li MW, Teng JL, et al. (2003) Use of cefoperazone MacConkey agar for selective isolation of Laribacter hongkongensis. J Clin Microbiol 41: 4839–4841.

8. Pseudomonas aeruginosa. Catalytic and regulatory properties. Eur J Biochem 139: 297–299.

9. Steinberg CM, McNicholl PJ, Fuggle KR, et al. (2005) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.

10. Dal Canto M, Tschumy M (2003) N-acetylglutamate and its changing role through evolution. Biochem J 372: 259–270.

11. Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, et al. (2000) The Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science 287: 1809–1813.

12. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine deiminase of Escherichia coli. J Bacteriol 168: 1160–1163.

13. Laribacter hongkongensis: a novel bacterium associated with gastroenteritis, from drinking water reservoirs in Hong Kong. J Appl Microbiol 103: 507–515.

14. Lau SK, Woo PC, Fan RW, Lee RG, Teng JL, et al. (2007) Laribacter hongkongensis isolated from a community-acquired gastroenteritis in Hangzhou City. J Clin Microbiol 45: 253–256.

15. Laribacter hongkongensis: a novel bacterium associated with gastroenteritis, traveler's diarrhea. BMC Microbiol 7: 356–361.

16. Barnerji S, Flieger A (2004) Patulin-like proteins: a new family of lipolytic enzymes present in bacteria? Microbiology 150: 322–325.

17. Tettelin H, Sands BJ, Heidelberg JF, et al. (2000) Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella bronchiseptica, and Bordetella parapertussis. Nat Genet 31: 32–40.

18. Lau SK, Yuen KY, Lee RG, Teng JL, et al. (2003) Use of cefoperazone MacConkey agar for selective isolation of Laribacter hongkongensis. J Clin Microbiol 41: 4839–4841.

19. Parkhill J, Sebaihia M, Preston A, Murphy LD, Thomson N, et al. (2003) The Complete genome sequence of Staphylococcus aureus strain N315. Nature 423: 162–170.

20. Nikaido H (1996) Multidrug efflux pumps of gram-negative bacteria. J Bacteriol 178: 1497–1504.

21. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.

22. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.

23. Dawson KL, Gray MW, Lang BF, Burger G (2005) AutoFACT: a web tool for improving predictions of prokaryotic translation initiation sites. Bioinformatics 21: 3568–3569.

24. Foster JW (2004) Classification Database for membrane transport protein analyses and information. Nucleic Acids Res 32: D136–D141.

25. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.

26. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.

27. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.

28. Marquis RE, Bender GR, Murray DR, Wong A (1987) Arginine biosynthesis in Thermotoga maritima: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J Bacteriol 186: 6411–6419.