The Genomic Sequence of the Accidental Pathogen Legionella pneumophila
Minchen Chien, et al.
Science 305, 1966 (2004);
DOI: 10.1126/science.1099776

The following resources related to this article are available online at
www.sciencemag.org (this information is current as of April 17, 2009):

Updated information and services, including high-resolution figures, can be found in the online
version of this article at:
http://www.sciencemag.org/cgi/content/full/305/5692/1966

Supporting Online Material can be found at:
http://www.sciencemag.org/cgi/content/full/305/5692/1966/DC1

This article cites 20 articles, 13 of which can be accessed for free:
http://www.sciencemag.org/cgi/content/full/305/5692/1966#otherarticles

This article has been cited by 109 article(s) on the ISI Web of Science.

This article has been cited by 52 articles hosted by HighWire Press; see:
http://www.sciencemag.org/cgi/content/full/305/5692/1966#otherarticles

This article appears in the following subject collections:
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Information about obtaining reprints of this article or about obtaining permission to reproduce
this article in whole or in part can be found at:
http://www.sciencemag.org/about/permissions.dtl
The Genomic Sequence of the Accidental Pathogen Legionella pneumophila

Minchen Chien,1* Irina Morozova,1* Shundi Shi,1* Huitao Sheng,1 Jing Chen,1 Shawn M. Gomez,2 Gifty Asamani,1 Kendra Hill,1 John Nuara,1 Marc Feder,1 Justin Rineer,1 Joseph J. Greenberg,1 Valeria Steshenko,1 Samantha H. Park,1 Baohui Zhao,3 Elita Teplitzky,1 John R. Edwards,1,4 Sergey Pampou,1 Anthi Georgiou,1 I-Chun Chou,1 William Iannuccilli,1 Michael E. Ulz,1 Dae H. Kim,1 Alex Geringer-Sameth,1 Curtis Goldsberry,1 Pavel Morozov,1 Stuart G. Fischer,1 Gil Segal,5 Xiaoyan Qu,1 Andrey Rzhetsky,1 Peisen Zhang,1 Eftihia Cayanis,1 Pieter J. De Jong,3 Howard A. Shuman,1* James J. Russo 1

We present the genomic sequence of Legionella pneumophila, the bacterial agent of Legionnaires’ disease, a potentially fatal pneumonia acquired from aerosolized contaminated fresh water. The genome includes a 45-kilobase pair element that can exist in chromosomal and episomal forms, selective expansions of important gene families, genes for unexpected metabolic pathways, and previously unknown candidate virulence determinants. We highlight the genes that may account for Legionella’s ability to survive in protozoa, mammalian macrophages, and inhospitable environmental niches and that may define new therapeutic targets.

Legionella pneumophila was recognized as a human opportunistic pathogen after its isolation from patients in an outbreak of fatal pneumonia ( Legionnaires’ disease) at an American Legion convention in Philadelphia (1). L. pneumophila and other members of the genus are found within biofilms and fresh and industrial water systems worldwide, posing a significant public health concern. Inhaled Legionella spp. cause sporadic and epidemic cases of Legionnaires’ disease and the flu-like Pontiac fever (2).

L. pneumophila are aerobic Gram-negative, motile, rod-shaped bacteria of the γ-proteobacterial lineage. These intracellular pathogens use the Icm/Dot type IVB secretion system (3–4) to deliver effector proteins to the host cells (5–7) that modulate the fate of the phagocytic vacuole by preventing phagosome-lysosome fusion and vacuole acidification, and recruiting vesicles that confer on it properties of the endoplasmic reticulum (8). After proliferating within this compartment, the bacteria destroy the host cell and infect other cells.

The most intriguing aspects of L. pneumophila biology are its extraordinary ability to commandeer the organelle trafficking systems of a wide range of host cells and its ability to survive in harsh environments such as plumbing systems treated with potent biocides. Legionella species are denizens of soil and water with life-styles ranging from highly virulent (L. pneumophila) to non-pathogenic (protozoan symbionts) (2). Genomic comparisons permit the development of testable models for the molecular bases of these properties of Legionella spp. and their ability to adapt to different niches.

The genome of L. pneumophila subsp. pneumophila (serogroup 1), Philadelphia 1 strain, derived from the original 1976 isolate, consists of a single circular chromosome of 3,397,754 base pairs (bp) (GenBank accession code AE017354) with 38% G+C content (Fig. 1 and table S1) (9, 10).

As expected for an intracellular pathogen, examination of its genome indicates that L. pneumophila has undergone horizontal gene transfer. Several homologs of eukaryotic genes were identified (11), including the previously described ralF gene, encoding a guanosine triphosphatase (GTPase) modifier required for recruiting the ADP ribosylation factor (ARF) GTPase to the Legionella-containing vacuole (8). Some phage-derived and insertion sequences, constituting ~2.4% of the genome, are dispersed throughout the genome; others form nine clusters in addition to two loci containing the icm/dot genes encoding the type IVB secretion system (12, 13), the lhv/lvr cluster encoding a typical type IV secretion system (14), and a region containing F plasmid-derived tra/trb genes (15).

The Philadelphia 1 strain possesses a plasmid-like element of 45 kb (pLP45) that exists in a circular episomal form or within the chromosome. The lhv/lvr gene region, one of the few extensive loci with an elevated G+C content (fig. S5), is found within this element, along with genes potentially involved in DNA recombination. We identified a 100-kb region (fig. S5) containing several genes encoding efflux transporters for heavy metals and other toxic substances. The presence of tRNA, phage-related genes, and transposase genes near the extremities of this region suggest that it was acquired via horizontal transfer. This “efflux island” may contribute to Legionella’s ability to flourish in plumbing sys-
tems and persist in the presence of toxic biocides.

The fully sequenced species most closely related to L. pneumophila is the obligate intracellular pathogen Coxiella burnetti, also belonging to the order Legionellales. Its proteome has a high number of basic proteins, which may explain its ability to replicate within phagolysosomes (16). L. pneumophila replicates in a vacuole resembling the endoplasmic reticulum; its proteome has a lower average isoelectric point (9). Altogether, Legionella shares ~42% of its genes with Coxiella despite differences in genome size (3.4 and 1.9 Mbp, respectively). To date the icm/dot genes have been found only in Legionella spp. and C. burnetti. Out of several identified L. pneumophila effectors delivered by the Icm/Dot system, only the recently described sid genes (6) have counterparts in C. burnetti. Very few other genes are shared exclusively by L. pneumophila and C. burnetti [the only named genes are the “enhanced entry” genes enhA and enhB (17)], implying that specific regulatory circuits or transport capabilities determine the features that distinguish them from other bacteria (table S2).

Of Legionella’s gene complement, 60% of the genes have homologs among phylogenetically diverse intracellular bacteria (Coxiella, Salmonella, Chlamydia, Rickettsia, Brucella, and Mycobacterium species), comparable to the 63% found in seven fully sequenced related γ-proteobacteria, suggesting that the species’ similar life-styles and common origin may equally affect gene complement similarity.

Paralogous gene family expansions may be associated with adaptation to specific environments, development of novel life strategies, or other species-specific adaptations. In a search for Legionella gene family expansions relative to other bacterial genomes (table S3), we identified several examples including β-lactamase, factor for inversion stimulation, and some transporters (9), even though overall Legionella does not display a high genome redundancy, as judged by its modest average gene family expansion rate. Among intracellular pathogens, Legionella groups with Brucella melitensis and Salmonella species (table S4), whereas reduced genome organisms (e.g., Rickettsia and Coxiella) exhibit low expansion values, and M. tuberculosis contains many expanded gene families including several related to lipid synthesis.

Legionella has multiple family members for genes encoding enhanced entry proteins, peptide methionine sulfoxide reductase, zinc metalloproteases, polyhydroxyalkanoic acid synthase, transposases, and effectors. The Sid effectors, identified on the basis of their ability to be transferred to another cell via the Icm/Dot type IVB system (6), form three families in L. pneumophila.

Legionella survives and replicates in axenic cultures, fresh water, soil, and in biofilms with other organisms, as well as within intracellular vacuoles of amoebae, ciliates, and human cells. To best utilize the nutrients within these diverse environments, the bacteria would need a broad range of membrane transporters. We identified more than 350 binding proteins and permeases in the Legionella genome, representing 62 separate substrate classes.

Among its many multidrug transporters, there is an expanded 11-member major facilitator family in Legionella. Nine similar proteins occur in C. burnetti, which has an unusually high overall density of multidrug transporters in its genome (18), but no clearcut representatives of this particular family exist in other intracellular pathogens or in the substantially larger genomes of Escherichia coli, Salmonella typhimurium, and Pseudomonas aeruginosa. L. pneumophila has more members of not only this particular gene family, but of multidrug transporters as a general class. It has an unusually high number of genes encoding putative effluxers for toxic compounds and heavy metals relative to other γ-proteobacteria, perhaps because the natural protozoan hosts for Legionella accumulate heavy metals, including lead and cadmium, from the environment (19).

A family of four glutamate/γ-aminobutyrate–specific amino acid antiporters is apparently represented by two members in Coxiella and only one member in 10 other species examined (GadC in E. coli). This antiporter is involved in bacterial survival within acidic environments (20); its possible functional redundancy may account for the suggested ability of L. pneumophila to survive in a putative acidic compartment late in its life cycle (16).

Because L. pneumophila utilizes amino acids as carbon and energy sources and cannot ferment or oxidize carbohydrates (21), we were surprised to find an intact glycolytic chain, pyruvate dehydrogenase complex, tricarboxylic acid cycle and respiratory chain, and a glucose-6-phosphate

![Fig. 1.](image-url)
transporter. Perhaps these pathways are used when bacteria are exposed to the complex molecular content of intracellular organelles of host cells or potentially nutrient-rich biofilms. *Legionella* is reportedly autotrophic for several amino acids, so it was surprising to find potential genes for their synthetic pathways: cysteine from pyruvate or serine, methionine from cysteine, and both phenylalanine and tyrosine from phosphoenolpyruvate. Even if some of these genes are not expressed under laboratory growth conditions, their presence presumably relates to the organism’s ability to persist in diverse environments.

In addition to previously recognized virulence factors (table S5), we identified new candidates (table S6), including homologs of genes encoding virulence functions in other bacteria. Moreover, among fully sequenced organisms, this is exceeded only by the predatory *Bdellovibrio* (22). *L. pneumophila* has been proposed to utilize bacterial-induced apopotic (early) and/or necrotic pore-forming (late) events to exit infected hosts (23); its putative hydrolases may be involved in these processes.

The genome sequence of *L. pneumophila* offers the opportunity to explain its broad host range and extraordinary ability to resist eradication in water supplies. Having lists of genes unique to *Legionella* or shared with unrelated bacteria with similar life-styles, it should now be possible to determine experimentally which of them distinguish *Legionella* species displaying different host preferences or pathogenicity.

References and Notes

9. Supporting materials are available on Science Online.
11. K. Suwwan de Felipe, S. Pampou, S. Kalachikov, H. A. Shuman, unpublished data.

Nitric Oxide Represses the *Arabidopsis* Floral Transition

Yikun He, 1,2* Ru-Hang Tang, 1* Yi Hao, 1* Robert D. Stevens, 3 Charles W. Cook, 1 Sun M. Ahn, 1 Liufang Jing, 1 Zhongguang Yang, 4 Longen Chen, 4 Fangqin Guo, 5 Fabio Fiorani, 1, 7 Robert B. Jackson, 4 Nigel M. Crawford, 8 Zhen-Ming Pei 1, 9

The correct timing of flowering is essential for plants to maximize reproductive success and is controlled by environmental and endogenous signals. We report that nitric oxide (NO) represses the floral transition in *Arabidopsis thaliana*. Plants treated with NO, as well as a mutant overproducing NO (*nox1*), flowered late, whereas a mutant producing less NO (*nos1*) flowered early. NO suppressed CONSTANS and GIGANTEA gene expression and enhanced FLOWERING LOCUS C expression, which indicated that NO regulates the photoperiod and autonomous pathways. Because NO is induced by environmental stimuli and constitutively produced, it may integrate both external and internal cues into the floral decision.

The life of flowering plants is divided into two distinct phases, an initial vegetative phase during which meristems produce leaves and a subsequent reproductive phase during which meristems produce flowers. Genetic studies of the timing of flowering in *Arabidopsis* have revealed four major pathways (1). The photoperiod and vernalization pathways integrate external signals into the floral decision, whereas the autonomous and gibberellin (GA) pathways act independently of environmental cues (2).

NO plays a pivotal role in animals and has emerged as a key growth regulator in plants (3, 4). NO promotes leaf expansion, inhibits maturation and senescence, stimulates light-dependent germination, and promotes detiolation (5, 6). Excess endogenous NO reduces growth and delays development in tobacco plants (7). NO production is induced by biotic and abiotic stimuli, such as drought, salt stress, and pathogen infection (4). In addition, substantial NO is emitted from plants into the atmosphere. Conversely, atmospheric NO, a major greenhouse pollutant produced by combustion of fossil fuels, can affect plants. Thus, NO has a central role in coordinating plant growth and development with environmental conditions. However, little is known about the molecular mechanisms underlying the function of NO in plants.

Treatment of *Arabidopsis* seedlings with an NO donor, sodium nitroprusside (SNP), enhanced vegetative growth and delayed flowering (Fig. 1). SNP increased shoot growth by ~65% at low concentrations (~100 μM), although it inhibited growth at high concentrations (Fig. 1, A and B; fig S1A). The optimal SNP concentration for promoting shoot growth was ~100 μM. A similar promotive effect of

---

1Department of Biology, Duke University, Durham, NC 27708, USA.
2Department of Biology, Capital Normal University, Beijing 100037, China.
3Mass Spectrometry Laboratory, Duke University Medical Center, Research Triangle Park, NC 27709, USA.
4Orthopaedic Research Laboratory, Duke University Medical Center, Durham, NC 27710, USA.
5Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92039, USA.

*These authors contributed equally to this work.
†Present address: Department of Plant Systems Biology, VIB-Ghent University, B-9052 Ghent, Belgium.
‡To whom correspondence should be addressed. E-mail: zpei@duke.edu