ELSEVIER

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm



ListiWiki: A database for the foodborne pathogen Listeria monocytogenes

Christoph Elfmann^{a,1}, Bingyao Zhu^a, Jörg Stülke^{a,**,2}, Sven Halbedel^{b,c,*,3}

^a Department of General Microbiology, Göttingen Center for Molecular Biosciences, Georg-August University Göttingen, Göttingen, Germany
^b FG11 Division of Enteropathogenic bacteria and Legionella, Robert Koch Institute, Burgstrasse 37, 38855 Wernigerode, Germany
^c Institute for Medical Microbiology and Hospital Hygiene, Otto von Guericke University Magdeburg, 39120 Magdeburg, Germany

ARTICLE INFO

Keywords: Genome annotation Genome browser Expression browser Interaction networks ListiWiki

ABSTRACT

Listeria monocytogenes is a Gram positive foodborne pathogen that regularly causes outbreaks of systemic infectious diseases. The bacterium maintains a facultative intracellular lifestyle; it thrives under a variety of environmental conditions and is able to infect human host cells. L. monocytogenes is genetically tractable and therefore has become an attractive model system to study the mechanisms employed by facultative intracellular bacteria to invade eukaryotic cells and to replicate in their cytoplasm. Besides its importance for basic research, L. monocytogenes also serves as a paradigmatic pathogen in genomic epidemiology, where the relative stability of its genome facilitates successful outbreak detection and elucidation of transmission chains in genomic pathogen surveillance systems. In both terms, it is necessary to keep the annotation of the L. monocytogenes genome up to date. Therefore, we have created the database ListiWiki (http://listiwiki.uni-goettingen.de/) which stores comprehensive information on the widely used L. monocytogenes reference strain EDG-e. ListiWiki is designed to collect information on genes, proteins and RNAs and their relevant functional characteristics, but also further information such as mutant phenotypes, available biological material, and publications. In its present form, ListiWiki combines the most recent annotation of the EDG-e genome with published data on gene essentiality, gene expression and subcellular protein localization. ListiWiki also predicts protein-protein interactions networks based on protein homology to Bacillus subtilis proteins, for which detailed interaction maps have been compiled in the sibling database SubtiWiki. Furthermore, crystallographic information of proteins is made accessible through integration of Protein Structure Database codes and AlphaFold structure predictions. ListiWiki is an easyto-use web interface that has been developed with a focus on an intuitive access to all information. Use of ListiWiki is free of charge and its content can be edited by all members of the scientific community after registration. In our labs, ListiWiki has already become an important and easy to use tool to quickly access genome annotation details that we can keep updated with advancing knowledge. It also might be useful to promote the comprehensive understanding of the physiology and virulence of an important human pathogen.

1. Introduction

Listeria monocytogenes is a dreaded human pathogen causing foodborne infections that may develop into severe systemic disease in susceptible risk groups. Case fatality of such invasive infections approaches 30–40% (Charlier et al., 2017), which is exceptionally high compared to other gastrointestinal bacterial pathogens (Werber et al., 2013). The natural reservoir of the bacterium is the environment, where it lives in the soil and is found associated with plant surfaces, but also in fresh and marine water samples, from where it is carried over into food production (Quereda et al., 2021; Vivant et al., 2013). For systemic infection, the bacterium breaches anatomical barriers such as the gut epithelium or the blood brain barrier by repeated cycles of host cell invasion, escape from phagosomes, intracellular growth in the cytoplasm of infected host cells and spread to uninfected neighbor cells (Dussurget et al., 2004; Vazquez-Boland et al., 2001).

https://doi.org/10.1016/j.ijmm.2023.151591

Received 13 October 2023; Received in revised form 9 November 2023; Accepted 27 November 2023 Available online 29 November 2023 1438-4221/© 2023 The Author(s). Published by Elsevier GmbH. This is an open

^{*} Corresponding author at: FG11 Division of Enteropathogenic bacteria and Legionella, Robert Koch Institute, Burgstrasse 37, 38855 Wernigerode, Germany. ** Corresponding author.

E-mail addresses: jstuelk@gwdg.de (J. Stülke), halbedels@rki.de (S. Halbedel).

¹ https://orcid.org/0000-0002-7872-8232

² https://orcid.org/0000-0001-5881-5390

³ https://orcid.org/0000-0002-5575-8973

^{1438-4221/© 2023} The Author(s). Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

L. monocytogenes belongs to the *Bacillota* phylum (formerly known as firmicutes) and is the best-established model organism to study the mechanisms of bacterial growth inside mammalian host cells within this phylogenetic group (Hamon et al., 2006). The bacterium can switch between the contrasting lifestyles of a peaceful environmental soil dweller and a harmful intracellular pathogen. This ability has made *L. monocytogenes* a popular model system to study adaptation to stresses experienced in the environment, in the gastrointestinal tract or during confrontation with the immune system (Freitag et al., 2009; Leseigneur et al., 2020).

The reconstruction of the genome sequence of the widely used reference strain EGD-e in 2001 provided first insights into the genetic repertoire of the pathogen (Glaser et al., 2001). Global analyses using omics technologies further improved functional annotation of the L. monocytogenes genome. These analyses include the definition of compartment specific proteomes (Desvaux et al., 2010; Dumas et al., 2009; Ramnath et al., 2003), genome wide in silico prediction of subcellular protein localization (Renier et al., 2012) and the definition of transcriptomes under various conditions and in different strains (Behrens et al., 2014; Chatterjee et al., 2006; Hain et al., 2012). In 2014, few single nucleotide polymorphisms in the EGD-e genome were found in a genome resequencing project (Becavin et al., 2014), possibly reflecting ongoing microevolution in the lab. However, genomic analysis of L. monocytogenes isolates collected from patients, environmental and food sources in pathogen surveillance programs revealed an even higher diversity within the species that separates into at least 4 phylogenetic lineages, which even differ in their virulence gene content (Moura et al., 2016; Yin et al., 2019). While L. monocytogenes strain EGD-e is a representative of phylogenetic lineage II (Becavin et al., 2014), the majority of human infections are caused by lineage I strains (Halbedel et al., 2018; Maury et al., 2016). However, there are no reference strains of lineage I established that are used as frequently as EGD-e.

Beyond the use of genome-wide approaches, the annotation of the *L. monocytogenes* genome has been continuously improved by the identification of the function of previously uncharacterized genes. While the function of about 1000 genes remained unclear in the original annotation (Glaser et al., 2001), only about 350 genes are currently without any further functional annotation and this number will further decrease.

To combine our knowledge of the genetic elements present in the *L. monocytogenes* EGD-e genome in a single database and to keep this annotation updated on a regular basis, we have designed *Listi*Wiki, an user-editable database that integrates the available information on all protein and RNA encoding genes, their transcriptional organization, functional assignments and the mutual interactions of their gene products based on the published body of knowledge. The *Listi*Wiki dataset should not only provide a comprehensive collection of genomic annotation details of the EGD-e reference strain, it also may offer an up-to-date reference for automated genome annotations of new *L. monocytogenes* isolates.

2. Description of the database

ListiWiki (listiwiki.uni-goettingen.de) uses the same framework as a series of other model organism databases developed by our group (*SubtiWiki, SynWiki, and MycoWiki*) (Elfmann et al., 2022; Pedreira et al., 2022a; Pedreira et al., 2022b). Following the same general layout of the websites and the organization of data, the core of *ListiWiki's* structure is made up by the *Gene pages* for each individual gene. These pages are dedicated to the description of genes and the corresponding proteins. While most of the information relating to the gene in question can be found directly on the page, some aspects of the gene and the encoded product can be explored further using several interactive *browsers*, such as the *Genome, Interaction* and *Expression browsers*. Additional pages provide information about groups of genes (such as the *Categories page*) or offer more functionality to the user (such as the *Export page*).

2.1. Gene identifiers

On ListiWiki, different identifiers are used to designate genes. The gene name is usually a three-to-four letter mnemonic, using the standard nomenclature for bacterial genes (such as gpsB). The gene designations are derived from the most recent annotation for *L. monocytogenes* strain EGD-e (RefSeq accession: NC 003210.1) (Toledo-Arana et al., 2009). Moreover, we assigned 328 gene names based on similarity to B. subtilis genes. As the names should be ideally unique, manual curation was necessary in some cases to avoid duplicates. For example, the gene designation prfA is used twice in the RefSeq entry, once for the major transcriptional regulator of virulence genes in L. monocytogenes, and once for the conserved peptide chain release factor 1, which also exists in L. monocytogenes. Thus, to avoid confusion, we renamed the latter to prf1 in agreement with the UniProt database. Another identifier is the locus tag, which is a unique label assigned to each genomic feature during genome annotation (Glaser et al., 2001; Toledo-Arana et al., 2009). For example, *lmo1888* is the locus tag of the *gpsB* gene. While the names of genes can change (e.g., when a new name is assigned to a previously uncharacterized gene), the locus tag will remain constant. For some genes, only the locus tag is provided, which is then used as the primary identifier of the gene and is displayed as the title of the corresponding Gene page.

2.2. The Front page

The Front page gives access to many of *Listi*Wiki's features (see Fig. 1). Its most important feature is the search bar, which allows to look up genes by different keywords and identifiers, but also performs full-text searches. In addition, the Front page displays various navigation elements. With the top bar, users can access the Categories page, which presents a tree-like structure of functional categories to which the genes and corresponding proteins of *L. monocytogenes* are assigned. Furthermore, the top bar gives access to a list of essential genes and contains buttons to access a random Gene page. Finally, it allows to log in or register for *Listi*Wiki.

2.3. The Gene pages

The *Gene pages* represent the core of *Listi*Wiki and its underlying data structure. As mentioned above, most of the information pertaining to a specific gene and the encoded protein can be found on the corresponding Gene page. In addition, the user can explore the role of the gene in more detail by using several browsers, which can be accessed via the top navigation bar. The latter also features a search bar for convenience, as well as log-in functionality. Fig. 2 displays the Gene page for *gpsB*.

While the data available for each gene may vary, the general structure of each Gene page is the same. The main body is introduced with the gene name as the page title, followed by a short general description.

		{ Lis	stiW	'iki	}			
		l=	de de)			
	Genes or protein	s v				Go	earch	
	Browsers							

Fig. 1. The Front page of *Listi***Wiki**. The search bar in the center of the pages allows to search for genes by identifiers and keywords. The top bar gives access to an overview of categories, a list of essential genes, and a random gene page. It also provides functionality for login and registration. Below the search bar, the interactive browsers can be directly accessed.



Fig. 2. The Gene page for *gpsB*. Each Gene page has the same general structure, but depending on the available data, individual elements might vary. (A), page header with gene name, strain and description; (B), tabular overview over basic properties of the gene and its protein; (C), Genomic Context Viewer, an interactive display showing the genomic neighborhood of the gene; (D), side bar with navigation links, the Structure Viewer and the Interaction Graph; (E), various sections on different aspects of the gene, such as assigned functional categories or the protein; (F), list of relevant publications. Some sections were omitted in the figure as indicated by the vertical ellipsis.

Next, a table summarizes basic properties of the gene, such as the locus tag, isoelectric point and molecular weight of the protein, details about the associated sequences and additional information about function and product. Utility links allow to view the associated sequences in the Genome browser, and to directly perform a BLAST search (Sayers et al., 2022) with them. The table also contains outlinks to the protein database UniProt (UniProt, 2023) the interaction database STRING (Szklarczyk et al., 2023) and PaperBLAST (Price and Arkin, 2017), a tool to find related publications based on a protein sequence. The last row displays the Cluster of Orthologous Genes (COG) (Galperin et al., 2015) the gene belongs to.

Below this tabular overview, the Genomic Context Viewer is embedded, an interactive display for the genomic neighborhood of the gene. The viewer is a simplified version of the interactive *Genome Browser*, which will be discussed below. Next, different sections of text describe the gene/protein in further detail. For example, most genes are assigned to several functional categories, which have been specifically curated for *Listi*Wiki. A list of these assigned categories can be found as one of the Gene page sections. Additional sections describe properties of the encoded protein or the regulation of the gene or operon among others, in more detail. Moreover, the Gene pages list biological materials and relevant publications.

In the side bar, several navigation links give access to other pages, such as the Categories page, or the Export page. The latter allows users to download tabular data on all entries of a certain entity, such as genes, categories or operons. The side bar also features more information on the gene or protein in the form of graphical displays. One such display is the interactive *Structure Viewer*, which displays three-dimensional protein structures (see Fig. 3 A). The user can cycle through the different structures available for the protein in question, such as experimentally derived structure predictions from the AlphaFold Protein Structure Database (Varadi et al., 2022), which we added for all proteins in the *Listi*Wiki database. The Structure Viewer features a full-screen mode and



Fig. 3. The interactive displays available in the side bar of the Gene page. (A) The Structure Viewer displays the structures available for a protein (here: GpsB). In addition to experimentally determined structures from PDB, structure predictions from the AlphaFold Protein Structure Database are presented, as well. Different rendering options and a full-screen mode are available. **(B)** Interaction Graph for GpsB showing its protein-protein interactions. Interactions of GpsB are displayed as opaque edges, while interactions between the other interaction partners are transparent.

different style options, for example coloring the protein surface based on hydrophobicity. The rendering of the molecules is performed with NGL Viewer (Rose and Hildebrand, 2015). Another feature of the sidebar is the Interaction Graph, an interactive display presenting protein-protein interactions (see Fig. 3B). The user can click on the protein nodes to navigate to the corresponding Gene page or select edges to show further information on the corresponding interactions. A basic version of this graph was already implemented in *SubtiWiki* (Pedreira et al., 2022b), but we recently expanded its capabilities. Before, only the interactions between the protein in question and its direct interaction partners where shown, but now interactions *between* these partners are included, as well. The latter are displayed as additional transparent edges between

the nodes.

2.3.1. The Browsers

Using the top bar of any Gene page, the user can access several browsers, which are pages dedicated to specialized interactive tools. They allow users to explore certain aspects of the organism's biology in more detail. When accessed from the top bar of a Gene page, the data corresponding to that gene is highlighted in the selected browser. But each browser also features its own search bar, which makes it easier to access and compare information on different genes or proteins. Currently, three browsers are featured on *Listi*Wiki.

The Genome Browser (Fig. 4) allows the user to view the genomic properties of the gene in question, such as its coordinates and orientation, as well as its nucleotide and amino acid sequence. The featured interactive graphical display also shows the neighborhood of the gene, which the user can scroll through. By clicking on any gene marker in the display, the associated sequences are loaded. The user can then search the sequences for substrings and toggle the reverse complement. The tabs at the top of the page can be used to perform advanced searches, such as querying arbitrary regions of the genome or genes including their flanking regions.

In addition to the aforementioned Interaction Graph, the Interaction Browser (Fig. 5) lets the user explore protein-protein interaction networks in even greater detail. The browser shows a dynamic network of interactions with a range of different options. The user can click on the nodes representing proteins to display more information in the form of a tooltip or drag them around to rearrange them. The settings menu offers several options, such as adjusting the *radius* of the network, which corresponds to the neighborhood depth of the graph. On a low setting, only the selected protein and its direct interaction partner are included. But by increasing the radius, more and more interaction partners of the partners can be included. Another setting, the *spacing*, controls the distance between nodes in the graph. The user can also highlight a protein via a search bar and change the color scheme for nodes and edges. Rightclicking empty space in the interaction display opens a context menu which allows to export the viewed network in different formats. The info box in top left corner contains a search bar as well as the proportion of currently loaded proteins.

The last browser, the Expression Browser (Fig. 6), features a line chart representing the raw intensities of transcript levels for different experimental conditions. The data stems from an analysis of global gene expression in the mouse spleen during infection (Camejo et al., 2009). Using the search bar, the expression of multiple genes can be compared in the same chart. Upon clicking on data points, an explanation of the corresponding experimental condition appears. Data can be exported using the buttons in the top right corner. The plot is drawn using the Apache ECharts visualization library.

3. Implementation and data

*Listi*Wiki uses the same framework as its sibling platforms *Subti*Wiki, *Syn*Wiki and *Myco*Wiki. As a result, the overall design, functionality and



MTSEQFEYHL TGKEILEKEF KTGLRGYSPE DVDEFLDMVI KDYSTFTQEI EALQAENIRL VQELDNAPLR TSTQPAPTFQ AAAQPAGTTN FDILKRLSNL EKHVFGNKLD DNE

Fig. 4. The Genome Browser displaying the genomic context of *gpsB*. The graphical element in the page center allows the user to interactively explore the genome. Below, sequences corresponding to the current selection are shown. The tabs at the page top provide more complex options for searching the genome.

Genome Browser



Fig. 5. The Interaction Browser with GpsB selected. This interactive tool displays a dynamic network of protein-protein interactions. The user can rearrange the nodes, highlight a specified protein, or increase the radius to include interactions of interaction partners.



Fig. 6. The Expression Browser comparing the transcript levels of *gpsB* (red line), *prkA* (dark blue) and *ezrA* (turqoise) before and during infection. The search bar allows to load data on additional genes into the chart. When hovering over a data point, the corresponding value and description of the experimental condition are displayed in a tooltip. If the user clicks on the data point, the description is also shown below the search bar.

organization of data are generally the same. However, availability of certain data and corresponding frontend features might vary between the websites. The backend is written in PHP and hosted with a Apache HTTP server, while the frontend uses simple JavaScript without a specialized framework. The relational database used for data storage is MySQL.

The data featured on *Listi*Wiki is a mixture of manually curated gene specific information and bulk imports from other sources. The foundation of data is laid by the genome annotation of *L. monocytogenes* strain EGD-e (RefSeq accession: NC_003210.1) (Toledo-Arana et al., 2009). Gene essentiality data were obtained from a Tn-Seq study (Fischer et al., 2022) complemented by individual reports on unsuccessful gene deletion attempts. Expression data were taken from a study that compared global gene expression during growth in BHI broth at 37 °C with gene

expression in infected mice at one, two and three days post infection (Camejo et al., 2009). Using specific PubMed queries, 1285 publications were added as relevant references to the Gene pages. Based on homologies to *B. subtilis* genes, the database was populated using *Subti*Wiki as a rich source of information. Besides the aforementioned 328 newly assigned gene names, 1636 protein-protein interactions were adopted from *Subti*Wiki for orthologous proteins.

Similar to its sibling platforms, *Listi*Wiki received a specifically curated set of categories to appropriately group and describe genes in a standardized manner. The categories are structured like a tree, with six top-level categories which are further subdivided for a more finegrained characterization of genes. For example, genes are grouped by their roles in cellular processes and metabolism, genomic properties, localization, and essentiality, among others. Table 1 gives an overview

Table 1 –

List of top-level categories and their subcategories used in *Listi*Wiki, and the number of genes assigned to each of them.

Top-level category	Subcategory	Number of genes
Cellular processes	Cell envelope and cell division	126
centuar processes	Transporters	318
	Homeostasis	21
Metabolism	Flectron transport and ATP	24
Metubolishi	synthesis	21
	Carbon metabolism	86
	Amino acid/ nitrogen	106
	metabolism	
	Lipid metabolism	37
	Nucleotide metabolism	62
	Additional metabolic pathways	181
Information processing	Genetics	38
	RNA synthesis and degradation	32
	Protein synthesis, modification and degradation	279
	Regulation of gene expression	70
Lifestyles	Exponential and early post-	40
	exponential lifestyles	
	Lifestyles/ miscellaneous	107
	Virulence	79
Prophages, mobile genetic	Prophages	80
elements, genomic islands	Mobile genetic elements	8
	Genomic islands	5
Groups of genes	Essential genes	399
	Membrane proteins	672
	GTP-binding proteins	11
	Phosphoproteins	126
	Universally conserved proteins	28
	Poorly characterized/ putative	42
	enzymes	
	Proteins of unknown function	298
	Short peptides	2
	NcRNA	64
	Pseudogenes	6
	Efp-dependent proteins	1
	Secreted proteins	76
	LPXTG proteins	41
	GW proteins	9
	LysM proteins	5
	Lipoproteins	65

over the top-level categories and their immediate subcategories.

4. Curation and community

As keeping up with the newest scientific findings is a continuous effort, we invite the community of *Listeria* researchers to contribute to *Listi*Wiki. It is possible to request a registration using the "Register" button, which can be found on the Front page and on any Gene page. Although exploring the data on *Listi*Wiki does not require any user account, a login is needed for editing the available data and contributing new information. Changes to the database are regularly reviewed to ensure the quality of provided information. We welcome addition of any kind of data that helps to better describe the function of the elements annotated in the genome, such as mutant phenotypes (including observations made with mutants in other *Listeria* species), interaction partners, characteristic expression patterns, assignments to regulons and gene categories or even posttranslational modifications.

The presentation of up-to-date information and the possibility for the scientific community to contribute to the effort are important advantages of *Listi*Wiki over other genome databases available for *L. monocytogenes*. *Listi*List as the first of them was published in 2001, along with the announcement of the *L. monocytogenes* EGD-e genome, to facilitate genome handling through a web interface called GenoList (Glaser et al., 2001; Moszer et al., 1995). *Listi*List has been a useful reference for many years and even is still accessible while the service of several other GenoList databases has been terminated. However, the

annotation used in ListiList has never been updated since its release and also cannot be edited. As a result, the presented information is outdated. Active participation of the scientific community in database curation also discriminates ListiWiki from Listeriomics, a multi-omics platform which focusses on the integration of multiple transcriptomic, proteomic and genomic data sets (among others) (Becavin et al., 2017). A link to the relevant Listeriomics pages is included in the ListiWiki record for each gene. Moreover, ListiWiki has been optimized for proper display on smartphone screens, which is another advantage compared to Listeriomics. Another genome browser has been developed based on the BioCyc platform for the second widely used L. monocytogenes reference strain 10403 S and a few other L. monocytogenes strains. The 10403 S BioCyc browser provides additional tools for the visualization of biochemical pathways and transcriptional networks (Orsi et al., 2015). However, the use of the BioCyc database requires a paid subscription, whereas ListiWiki is available to all users without any subscription or charges.

The annotation currently stored in *Listi*Wiki has been compiled from global and gene-specific data sets available in the scientific literature, either by automated or manual import. This initiative needs to be continued and therefore we invite the *L. monocytogenes* scientific community to keep this annotation up to date in a joint effort. *Listi*Wiki is a dynamic database that constantly expands along with the progress of research and which is curated by active scientists in the field. Thus, it may become a useful resource for the easy and barrier-free access of current *L. monocytogenes* functional genome annotation.

5. Future perspectives

With ListiWiki we have released a new and self-explanatory tool for the intuitive exploration of the genome of this pathogen more than two decades after the release of ListiList, the first genome browser that was available for L. monocytogenes. By consideration of the gene designations that have been proposed since 2001, ListiWiki allows translating the results from over 20 years of L. monocytogenes research into an updated genome annotation. Moreover, the database features an expanded toolbox with modern functionalities to facilitate the generation of hypotheses on the function and the mutual interactions of genes and their gene products. ListiWiki is a dynamic database and can be complemented and expanded when further data become accessible. Currently, global analyses on protein-protein interactions between L. monocytogenes proteins or even on the interactions of L. monocytogenes proteins with the host proteome are not available. Likewise, the determination of essential gene sets under varying conditions has just started recently (Fischer et al., 2022; Wu et al., 2022). When information on gene essentiality in different mutant backgrounds becomes available one day, implementation and visualization of such synthetic lethal interactions could possibly be a useful extension. As further steps, genomes of other reference strains or of relevant genes only found outside these strains could be implemented. Moreover, useful database extensions can be initiated through users by proposing them to the ListiWiki admins. However, we would recommend the Listeriomics database for exploration of complex transcriptomic data sets (Becavin et al., 2017). Besides all these aspects, ListiWiki represents the fourth member of a family of similar databases that includes MycoWiki, SubtiWiki, and SynWiki. Thus, ListiWiki will profit from the introduction of novel features that will be developed for these other three databases.

Acknowledgements

We are grateful to Marie-Theres Thieme, Fabian Commichau, Richard Lüdtke and Johannes Gibhardt for their help with the initial work on *Listi*Wiki. JS and SH would like to thank the German Research Foundation for the longstanding financial support of the research in their laboratories.

C. Elfmann et al.

References

Becavin, C., Bouchier, C., Lechat, P., Archambaud, C., Creno, S., Gouin, E., Wu, Z., Kuhbacher, A., Brisse, S., Pucciarelli, M.G., Garcia-del Portillo, F., Hain, T., Portnoy, D.A., Chakraborty, T., Lecuit, M., Pizarro-Cerda, J., Moszer, I., Bierne, H., Cossart, P., 2014. Comparison of widely used *Listeria monocytogenes* strains EGD, 10403S, and EGD-e highlights genomic variations underlying differences in pathogenicity. mBio 5, e00969–00914.

Becavin, C., Koutero, M., Tchitchek, N., Cerutti, F., Lechat, P., Maillet, N., Hoede, C., Chiapello, H., Gaspin, C., Cossart, P., 2017. Listeriomics: an Interactive Web Platform for Systems Biology of *Listeria*. mSystems 2.

Behrens, S., Widder, S., Mannala, G.K., Qing, X., Madhugiri, R., Kefer, N., Abu Mraheil, M., Rattei, T., Hain, T., 2014. Ultra deep sequencing of *Listeria monocytogenes* sRNA transcriptome revealed new antisense RNAs. PLoS One 9, e83979.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res 28, 235–242.

Camejo, A., Buchrieser, C., Couve, E., Carvalho, F., Reis, O., Ferreira, P., Sousa, S., Cossart, P., Cabanes, D., 2009. In vivo transcriptional profiling of *Listeria monocytogenes* and mutagenesis identify new virulence factors involved in infection. PLoS Pathog. 5, e1000449.

Charlier, C., Perrodeau, E., Leclercq, A., Cazenave, B., Pilmis, B., Henry, B., Lopes, A., Maury, M.M., Moura, A., Goffinet, F., Dieye, H.B., Thouvenot, P., Ungeheuer, M.N., Tourdjman, M., Goulet, V., de Valk, H., Lortholary, O., Ravaud, P., Lecuit, M., group, Ms, 2017. Clinical features and prognostic factors of listeriosis: the MONALISA national prospective cohort study. Lancet Infect. Dis. 17, 510–519.

Chatterjee, S.S., Hossain, H., Otten, S., Kuenne, C., Kuchmina, K., Machata, S., Domann, E., Chakraborty, T., Hain, T., 2006. Intracellular gene expression profile of *Listeria monocytogenes*. Infect. Immun. 74, 1323–1338.

Desvaux, M., Dumas, E., Chafsey, I., Chambon, C., Hebraud, M., 2010. Comprehensive appraisal of the extracellular proteins from a monoderm bacterium: theoretical and empirical exoproteomes of *Listeria monocytogenes* EGD-e by secretomics. J. Proteome Res 9, 5076–5092.

Dumas, E., Desvaux, M., Chambon, C., Hebraud, M., 2009. Insight into the core and variant exoproteomes of *Listeria monocytogenes* species by comparative subproteomic analysis. Proteomics 9, 3136–3155.

Dussurget, O., Pizarro-Cerda, J., Cossart, P., 2004. Molecular determinants of *Listeria* monocytogenes virulence. Annu. Rev. Microbiol. 58, 587–610.

Elfmann, C., Zhu, B., Pedreira, T., Hossbach, B., Lluch-Senar, M., Serrano, L., Stulke, J., 2022. MycoWiki: Functional annotation of the minimal model organism *Mycoplasma pneumoniae*. Front. Microbiol. 13, 935066.

Fischer, M.A., Engelgeh, T., Rothe, P., Fuchs, S., Thurmer, A., Halbedel, S., 2022. *Listeria monocytogenes* genes supporting growth under standard laboratory cultivation conditions and during macrophage infection. Genome Res 32, 1711–1726.

Freitag, N.E., Port, G.C., Miner, M.D., 2009. Listeria monocytogenes - from saprophyte to intracellular pathogen. Nat. Rev. Microbiol 7, 623–628.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., Koonin, E.V., 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43, D261–D269.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Dartuvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K.D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J.C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehland, J., Cossart, P., 2001. Comparative genomics of Listeria species. Science 294, 849–852.

Hain, T., Ghai, R., Billion, A., Kuenne, C.T., Steinweg, C., Izar, B., Mohamed, W., Mraheil, M.A., Domann, E., Schaffrath, S., Karst, U., Goesmann, A., Oehm, S., Puhler, A., Merkl, R., Vorwerk, S., Glaser, P., Garrido, P., Rusniok, C., Buchrieser, C., Goebel, W., Chakraborty, T., 2012. Comparative genomics and transcriptomics of lineages I, II, and III strains of *Listeria monocytogenes*. BMC Genom. 13, 144.

Halbedel, S., Prager, R., Fuchs, S., Trost, E., Werner, G., Flieger, A., 2018. Whole-genome sequencing of recent *Listeria monocytogenes* isolates from germany reveals population structure and disease clusters. J. Clin. Microbiol 56.

Hamon, M., Bierne, H., Cossart, P., 2006. Listeria monocytogenes: a multifaceted model. Nat. Rev. Microbiol 4, 423–434.

Leseigneur, C., Le-Bury, P., Pizarro-Cerda, J., Dussurget, O., 2020. Emerging evasion mechanisms of macrophage defenses by pathogenic bacteria. Front. Cell. Infect. Microbiol. 10, 577559.

Maury, M.M., Tsai, Y.H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E.P., Brisse, S., Lecuit, M., 2016. Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. Nat. Genet 48, 308–313.

Moszer, I., Glaser, P., Danchin, A., 1995. SubtiList: a relational database for the Bacillus subtilis genome. Microbiol. (Read.) 141 (Pt 2), 261–268.

Moura, A., Criscuolo, A., Pouseele, H., Maury, M.M., Leclercq, A., Tarr, C., Bjorkman, J. T., Dallman, T., Reimer, A., Enouf, V., Larsonneur, E., Carleton, H., Bracq-Dieye, H., Katz, L.S., Jones, L., Touchon, M., Tourdjman, M., Walker, M., Stroika, S.,

Cantinelli, T., Chenal-Francisque, V., Kucerova, Z., Rocha, E.P., Nadon, C., Grant, K., Nielsen, E.M., Pot, B., Gerner-Smidt, P., Lecuit, M., Brisse, S., 2016. Whole genomebased population biology and epidemiological surveillance of *Listeria monocytogenes*. Nat. Microbiol 2, 16185.

Orsi, R.H., Bergholz, T.M., Wiedmann, M., Boor, K.J., 2015. The *Listeria monocytogenes* strain 10403S BioCyc database. Database (Oxf.) 2015.

Pedreira, T., Elfmann, C., Singh, N., Stulke, J., 2022a. SynWiki: Functional annotation of the first artificial organism Mycoplasma mycoides JCVI-syn3A. Protein Sci.: a Publ. Protein Soc. 31, 54–62.

Pedreira, T., Elfmann, C., Stulke, J., 2022b. The current state of SubtiWiki, the database for the model organism *Bacillus subtilis*. Nucleic Acids Res 50, D875–D882.

Price, M.N., Arkin, A.P., 2017. PaperBLAST: text mining papers for information about homologs. mSystems 2.

Quereda, J.J., Moron-Garcia, A., Palacios-Gorba, C., Dessaux, C., Garcia-Del Portillo, F., Pucciarelli, M.G., Ortega, A.D., 2021. Pathogenicity and virulence of *Listeria monocytogenes*: A trip from environmental to medical microbiology. Virulence 12, 2509–2545.

Ramnath, M., Rechinger, K.B., Jansch, L., Hastings, J.W., Knochel, S., Gravesen, A., 2003. Development of a *Listeria monocytogenes* EGDe partial proteome reference map and comparison with the protein profiles of food isolates. Appl. Environ. Microbiol 69, 3368–3376.

Renier, S., Micheau, P., Talon, R., Hebraud, M., Desvaux, M., 2012. Subcellular localization of extracytoplasmic proteins in monoderm bacteria: rational secretomics-based strategy for genomic and proteomic analyses. PLoS One 7, e42982.

Rose, A.S., Hildebrand, P.W., 2015. NGL Viewer: a web application for molecular visualization. Nucleic Acids Res 43, W576–W579.

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B.W., Pruitt, K.D., Sherry, S.T., 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50, D20–D26.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N.T., Pyysalo, S., Bork, P., Jensen, L.J., von Mering, C., 2023. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 51, D638–D646.

Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., Loh, E., Gripenland, J., Tiensuu, T., Vaitkevicius, K., Barthelemy, M., Vergassola, M., Nahori, M.A., Soubigou, G., Regnault, B., Coppee, J.Y., Lecuit, M., Johansson, J., Cossart, P., 2009. The *Listeria* transcriptional landscape from saprophytism to virulence. Nature 459, 950–956.

UniProt, C., 2023. UniProt: the universal protein knowledgebase in 2023. Nucleic Acids Res. 51, D523–D531.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zidek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., Velankar, S., 2022. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 50, D439–D444.

Vazquez-Boland, J.A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J., Kreft, J., 2001. *Listeria* pathogenesis and molecular virulence determinants. Clin. Microbiol. Rev. 14, 584–640.

Vivant, A.L., Garmyn, D., Piveteau, P., 2013. *Listeria monocytogenes*, a down-to-earth pathogen. Front. Cell. Infect. Microbiol. 3, 87.

Werber, D., Hille, K., Frank, C., Dehnert, M., Altmann, D., Müller-Nordhorn, J., Koch, J., Stark, K., 2013. Years of potential life lost for six major enteric pathogens, Germany, 2004-2008. Epidemiol. Infect. 141, 961–968.

Wu, Y., Pang, X., Liu, X., Wu, Y., Zhang, X., 2022. Functional genomics identified novel genes involved in growth at low temperatures in *Listeria monocytogenes*. Microbiol Spectr. 10, e0071022.

Yin, Y., Yao, H., Doijad, S., Kong, S., Shen, Y., Cai, X., Tan, W., Wang, Y., Feng, Y., Ling, Z., Wang, G., Hu, Y., Lian, K., Sun, X., Liu, Y., Wang, C., Jiao, K., Liu, G., Song, R., Chen, X., Pan, Z., Loessner, M.J., Chakraborty, T., Jiao, X., 2019. A hybrid sub-lineage of *Listeria monocytogenes* comprising hypervirulent isolates. Nat. Commun. 10, 4283.