

Database

Open Access

## MyBASE: a database for genome polymorphism and gene function studies of *Mycobacterium*

Xinxing Zhu<sup>†1,2</sup>, Suhua Chang<sup>†2</sup>, Kechi Fang<sup>2</sup>, Sijia Cui<sup>2</sup>, Jun Liu<sup>3</sup>, Zuowei Wu<sup>4</sup>, Xuping Yu<sup>5</sup>, George F Gao<sup>4</sup>, Huanming Yang<sup>6</sup>, Baoli Zhu<sup>\*4</sup> and Jing Wang<sup>\*2</sup>

Address: <sup>1</sup>James D Watson Institute of Genome Sciences of Zhejiang University, Hangzhou 310007, PR China, <sup>2</sup>Behavioral Genetics Center, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, PR China, <sup>3</sup>Department of Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, Ontario, Canada, <sup>4</sup>Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, PR China, <sup>5</sup>Institute of Preventive Veterinary Medicine, College of Animal Sciences, Zhejiang University, Hangzhou 310029, PR China and <sup>6</sup>Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, PR China

Email: Xinxing Zhu - zhuxx@psych.ac.cn; Suhua Chang - changsh@psych.ac.cn; Kechi Fang - fangkc@psych.ac.cn; Sijia Cui - cuisj@psych.ac.cn; Jun Liu - jun.liu@utoronto.ca; Zuowei Wu - wuzw@im.ac.cn; Xuping Yu - xpyu@zju.edu.cn; George F Gao - gaof@mail.im.ac.cn; Huanming Yang - yanghm@genomics.org.cn; Baoli Zhu\* - bzhu@chori.org; Jing Wang\* - wangjing@psych.ac.cn

\* Corresponding authors †Equal contributors

Published: 20 February 2009

Received: 8 December 2008

BMC Microbiology 2009, 9:40 doi:10.1186/1471-2180-9-40

Accepted: 20 February 2009

This article is available from: <http://www.biomedcentral.com/1471-2180/9/40>

© 2009 Zhu et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Mycobacterial pathogens are a major threat to humans. With the increasing availability of functional genomic data, research on mycobacterial pathogenesis and subsequent control strategies will be greatly accelerated. It has been suggested that genome polymorphisms, namely large sequence polymorphisms, can influence the pathogenicity of different mycobacterial strains. However, there is currently no database dedicated to mycobacterial genome polymorphisms with functional interpretations.

**Description:** We have developed a **mycobacterial database** (MyBASE) housing genome polymorphism data and gene functions to provide the mycobacterial research community with a useful information resource and analysis platform. Whole genome comparison data produced by our lab and the novel genome polymorphisms identified were deposited into MyBASE. Extensive literature review of genome polymorphism data, mainly large sequence polymorphisms (LSPs), operon predictions and curated annotations of virulence and essentiality of mycobacterial genes are unique features of MyBASE. Large-scale genomic data integration from public resources makes MyBASE a comprehensive data warehouse useful for current research. All data is cross-linked and can be graphically viewed via a toolbox in MyBASE.

**Conclusion:** As an integrated platform focused on the collection of experimental data from our own lab and published literature, MyBASE will facilitate analysis of genome structure and polymorphisms, which will provide insight into genome evolution. Importantly, the database will also facilitate the comparison of virulence factors among various mycobacterial strains. MyBASE is freely accessible via <http://mybase.psych.ac.cn>.

## Background

Mycobacteria are notorious for its two species, *Mycobacterium tuberculosis* (*M. tb*) and *Mycobacterium leprae* (*M. leprae*), the causative agent of tuberculosis (TB) and leprosy, respectively. In addition to *M. tb* and *M. leprae*, a number of mycobacterial pathogens also cause human and animal diseases, including *Mycobacterium bovis* (*M. bovis*), the causative agent of classical bovine tuberculosis, and *Mycobacterium ulcerans* (*M. ulcerans*), which causes Buruli ulcers. The emergence of multi-drug resistant strains of pathogenic mycobacteria has made the development of better vaccines and new drugs and novel control strategies a top priority.

The genome of *M. tb* H37Rv was the first mycobacterial genome to be sequenced and was published in 1998 [1], which was followed by the genome of *M. leprae* in 2001 [2]. The complete sequencing of these genomes greatly contributed to the understanding of the unique physiology and pathogenesis of mycobacteria. With the development of DNA sequencing technologies in recent years, a total of 18 complete mycobacterial genomes have been available and deposited in public domains thus far. This progress offers an unprecedented opportunity to understand the virulence mechanisms of mycobacteria at the molecular level, which offers insight into the development of potential control strategies.

One of the most significant findings in mycobacterial research was from the genome-wide comparison between virulent (e.g. *M. tb* H37Rv or *M. bovis*) and avirulent strains (e.g. *M. bovis* BCG) [3]. This genomic comparison unveiled large sequence polymorphisms (LSPs), usually called regions of difference (RDs), which are believed to be the major source of genomic diversity [4,5] and probably contribute to the phenotypic differences [6]. Some of the LSPs/RDs have been shown to be important for virulence and pathogenicity. For example, RD1, which is deleted in all BCG strains but is present in virulent strains of *M. tb* or *M. bovis*, has been shown to be essential for *M. tb* virulence [7-9]. The success of systematic genetic screening of mycobacterial mutants from different environments [10-13], coupled with focused investigation into individual virulence genes, has contributed to the functional genomic data of mycobacteria [14], which has provided useful information in understanding the physiology and pathogenesis of this unique bacterial genus.

Currently, several public resources for mycobacterial research are available, including the TB database [15], which is an integrated platform of genomic data with special interest in microarray analysis; GenoList <http://genolist.pasteur.fr/>, which focuses on the gene annotation of six mycobacterial strains [16]; MycoDB from xBASE [17,18], which provides search and visualization tools for

genome comparison of mycobacteria; MycoPeronDB [19], which is a database of predicted operons in 5 mycobacterial species; MGDD [20], a mycobacterial genome divergence database derived from an anchor-based comparison approach [21]; GenoMycDB [22], a database for pair-wise comparison of six mycobacterial genomes; and MtbRegList [23], which is dedicated to the analysis of transcriptional regulation of *M. tb*. Although each of these databases provides unique and useful information, none are focused on LSPs, essential genes, and the relationship between these and virulence. MyBASE was therefore developed to meet these needs. In addition to providing a platform for analyzing all published mycobacterial genomes, MyBASE features important information on genomic polymorphisms, virulence genes, and essential genes. As such, MyBASE will help researchers to easily explore and analyze data in a user-friendly and cross-referenced fashion, thereby facilitating functional genomic studies. This will inevitably enhance our understanding on the virulence mechanisms, genome structure, and molecular evolution of mycobacteria.

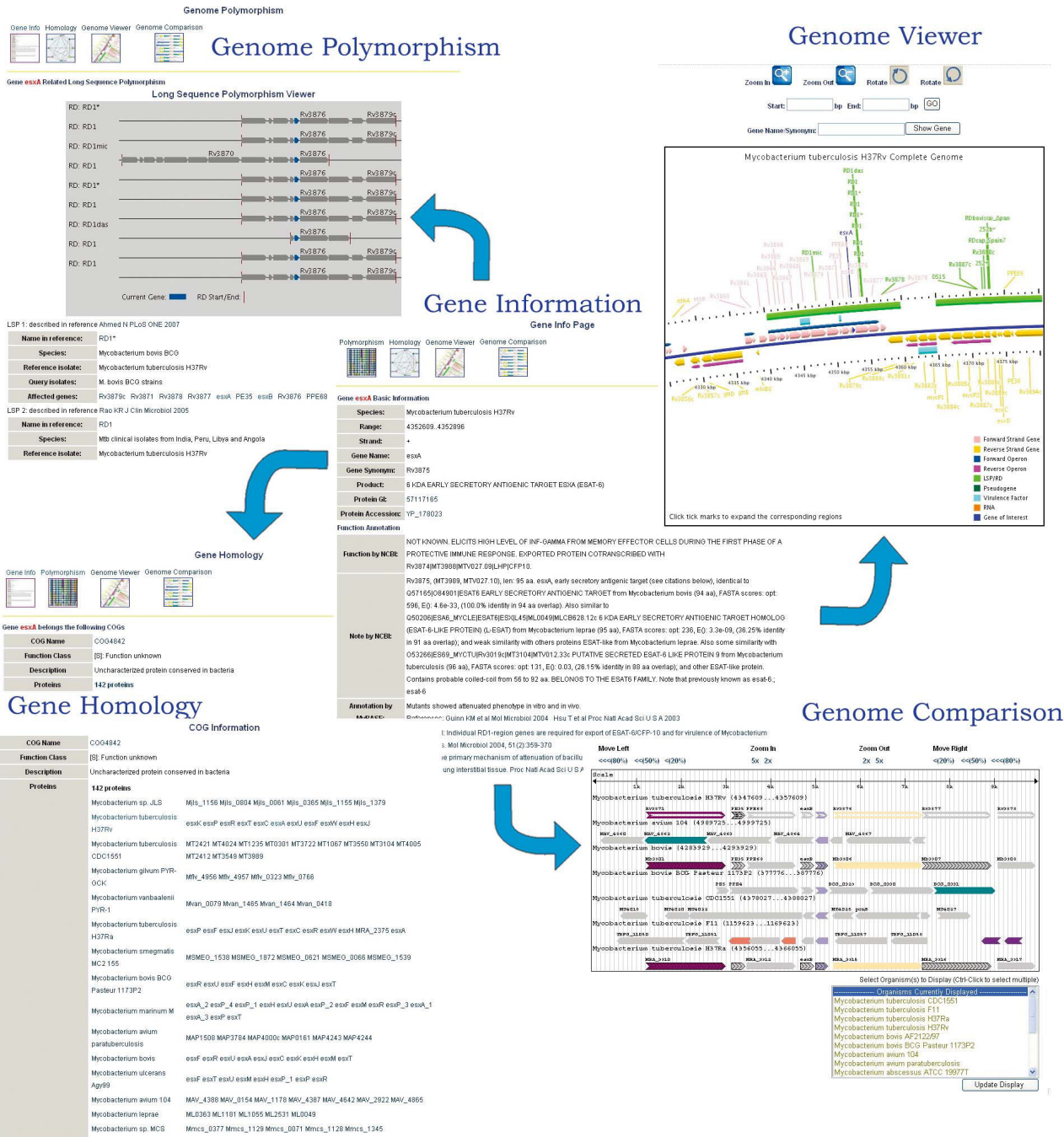
## Construction and content

### Data sources and curation

MyBASE contains data from both our own experiments and public resources. There are four main types of data: 1) genome sequences with curated annotations, 2) genome polymorphism data, particularly LSPs identified among different mycobacterial genomes, 3) functional gene annotations with a specific focus on virulence genes and essential genes, and 4) predicted operons.

All complete genome sequences and original annotation files were downloaded from NCBI <ftp://ftp.ncbi.nih.gov/genomes/Bacteria>. Curations were made to clarify inconsistencies resulting from different annotations provided by the original sequence providers. For Clusters of Orthologous Groups (COGs) that were inconsistently designated [24], we refined the COGs using the algorithm previously described [25].

We have recently used the NimbleGen tiling microarray method for whole-genome comparison of 13 BCG strains with subsequent confirmation by DNA re-sequencing [26]. A total of 42 deletions were identified, four of which are novel [26]. These novel deletions are believed to have an impact on virulence or immunogenicity of the corresponding BCG strains [26]. All data and analytical results were incorporated into MyBASE. In addition to our self-generated data, other polymorphism datasets, particularly LSPs/RDs that were included in MyBASE were extracted from public literatures. After the first usage of microarray to study genome polymorphism in 1999 [3], a growing trend emerged to generate systematic genome polymorphism data [27-29]. We performed an extensive literature



**Figure 1** Schematic representation of the data repository and the interrelation of functional modules in MyBASE. After the gene of interest is found, users can check whether it is in a genomic polymorphic region, compare the selected genome with MCV, explore the details of its genome segment with Genome Viewer or view its homolog distributions.

review to extract information about each LSP/RD from original experiments. We found inconsistencies between the nomenclature of LSPs (RDs) used by different groups and so to avoid further confusion, we have kept the original nomenclature from each group. However, we have provided the reference information and a hyperlink to the PubMed entry for each LSP/RD dataset.

Virulence, essentiality and other functional annotations of mycobacterial genes were extracted and corrected through data mining of public resources [10-14,30]. Virulence of mycobacterial genes was evaluated by phenotypic outcomes observed from animal and cellular models of *M. tb* infections (e.g., mouse, guinea pig, macrophages, etc.) for the corresponding mutants [14]. Recently, with the success of genetic manipulation of mycobacterial genes, a number of new virulence factors have been uncovered [31-35]. Since the role of some of these genes in pathogenesis are still in dispute [36,37], the annotations of experimental evidence for virulence have been provided to facilitate further investigation.

Operons of mycobacterial genomes were predicted using methods described by Alm *et al.* [38]. This method combines a comparative genomic approach with genome-specific distance models, and has shown some improvements in operon prediction [39].

### **System design and implementation**

MyBASE was developed using our established pipeline for biological databases [40-44]. It consists of three hardware components: a World Wide Web server, a database server, and a server for sequence analysis. The system is based on a MySQL relational database and the front end consists of a set of JSP scripts running on a Tomcat web server. Hibernate, a high-performance object/relational persistence and query service for Java, was used for system development. The search engine, Multi-genome Comparison Viewer, was developed using Java. Genome Viewer was implemented using CGView [45].

### **Utility and discussion**

#### **Database usage and the toolbox**

All the data in MyBASE can be easily explored using the toolbox. The keyword-based search engine enables a multiple keyword (e.g. gene name, COG number, etc.) search across MyBASE, while the BLAST-based sequence search engine allows user to quickly find similar genes to the query.

LSP/RD data is a distinct feature of MyBASE. The Polymorphism-LSP/RD module was developed to explore and mine the LSP/RD datasets. Users can search for a genomic polymorphism region by its name (e.g. RD1), the name of reference strain and query strain in the experiment, start/end positions within its genome, or by literature informa-

tion. Users can also visualize the distributions of selected RDs in the genome by using LSP/RD Viewer. RDs in the same dataset are present in one solid line according to its position along the genome (upper-left in Figure 1). Experimental information can be seen when users mouse over the LSP/RD region. To keep the data content in MyBASE most up-to-date, the "LSP/RD upload" module was designed for the user to upload their own LSP/RD data to MyBASE.

The Multi-genome Comparison Viewer (MCV) allows users to rapidly align and compare mycobacterial genome synteny by selecting an anchor gene of interest. This module is helpful for genome structure and evolutionary analysis of mycobacteria. Users can select any number of genomes, zoom in or out and move upstream or downstream along the genome in the viewer. Genes in MCV with the same color-coding are predicted homologs via COG designation, while grey indicates that no homolog was detected. More importantly, MCV also displays various featured annotations in MyBASE with different legends. Virulence factors, pseudogenes, and genes in an operon or polymorphic region are all presented in this graphic way. By clicking the gene, users can either re-anchor the viewer with this gene or navigate to the detailed gene information page.

Genome Viewer allows users to explore individual genomes with customized featured annotations, which include operons, LSPs/RDs, pseudogenes, and virulence factors. In addition, users can visualize a particular segment of a genome by zooming in/out, rotating or defining the start and end positions.

All data and tools in MyBASE are cross-linked. Users can start from searching a particular gene, for example, *esxA*, which is a virulence determinant that encodes a secretory protein [6,46,47], and then search each functional module, including polymorphisms (LSPs/RDs) for related LSP information. Furthermore, MCV and Genome Viewer can be used to compare the genome structure among selected genomes and to check other genomic features within the corresponding segment, respectively. Using these tools, we can see that *esxA* is located in RD1 and that its functional properties are represented by different legends. Users may also begin from a polymorphism search and then navigate to a gene page, MCV or Genome Viewer. Overall, MyBASE forms a highly-integrated and inter-correlated platform for efficient utilization and exploration of functional and comparative genomic data (Figure 1).

#### **Future developments**

The goal of MyBASE is to provide the mycobacterial research community with a useful resource and analysis platform for the functional and evolutionary investigation of mycobacteria. Newly generated genomic data and func-

tional annotations by the research community will be added to MyBASE periodically to keep the database up-to-date. The functionality of the LSP search and viewer will be enriched and enhanced. In addition, new tools, such as software packages for phylogenomic study will be integrated. Finally, MyBASE also provides an opportunity for the mycobacterial research community to standardize nomenclature, data formats of gene, and polymorphism annotations.

## Conclusion

MyBASE is a unique data warehouse and analysis platform for the mycobacterial research community, which features a collection and curation of a large amount of LSP and functional genomic data. By developing various tools, MyBASE can help researchers to easily explore and investigate genome deletions, virulence factors, essential genes, and operon structure of mycobacteria.

## Availability and requirements

The database is freely available on <http://mybase.psych.ac.cn>.

## Authors' contributions

XZ designed the database, collected, curated the data and wrote the manuscript. SC analyzed the data and developed the database. KF and SC developed the database and did the programming work. JL, ZW, and XY performed the microarray experiments and analyzed the data. GFG, and HY revised the manuscript. BZ and JW supervised the work, manage the team and wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

This work was sponsored by the National Natural Science Foundation of China (NSFC, Grant No. 30700441, 30221004) and Beijing Municipal Science and Technology Commission (Grant No: Z0005190043521).

## References

- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd, et al.: **Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.** *Nature* 1998, **393(6685)**:537-544.
- Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, et al.: **Massive gene decay in the leprosy bacillus.** *Nature* 2001, **409(6823)**:1007-1011.
- Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM: **Comparative genomics of BCG vaccines by whole-genome DNA microarray.** *Science* 1999, **284(5419)**:1520-1523.
- Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, et al.: **A new evolutionary scenario for the Mycobacterium tuberculosis complex.** *Proc Natl Acad Sci USA* 2002, **99(6)**:3684-3689.
- Garnier T, Eiglmeier K, Camus JC, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, et al.: **The complete genome sequence of Mycobacterium bovis.** *Proc Natl Acad Sci USA* 2003, **100(13)**:7877-7882.
- Behr MA, Sherman DR: **Mycobacterial virulence and specialized secretion: same story, different ending.** *Nat Med* 2007, **13(3)**:286-287.
- Majlessi L, Brodin P, Brosch R, Rojas MJ, Khun H, Huerre M, Cole ST, Leclerc C: **Influence of ESAT-6 secretion system I (RDI) of Mycobacterium tuberculosis on the interaction between mycobacteria and the host immune system.** *J Immunol* 2005, **174(6)**:3570-3579.
- Junqueira-Kipnis AP, Basaraba RJ, Gruppo V, Palanisamy G, Turner OC, Hsu T, Jacobs WR Jr, Fulton SA, Reba SM, Boom WH, et al.: **Mycobacteria lacking the RDI region do not induce necrosis in the lungs of mice lacking interferon-gamma.** *Immunology* 2006, **119(2)**:224-231.
- Billeskov R, Vingsbo-Lundberg C, Andersen P, Dietrich J: **Induction of CD8 T cells against a novel epitope in TB10.4: correlation with mycobacterial virulence and the presence of a functional region of difference-1.** *J Immunol* 2007, **179(6)**:3973-3981.
- Sassetti CM, Boyd DH, Rubin EJ: **Comprehensive identification of conditionally essential genes in mycobacteria.** *Proc Natl Acad Sci USA* 2001, **98(22)**:12712-12717.
- Sassetti CM, Boyd DH, Rubin EJ: **Genes required for mycobacterial growth defined by high density mutagenesis.** *Mol Microbiol* 2003, **48(1)**:77-84.
- Sassetti CM, Rubin EJ: **Genetic requirements for mycobacterial survival during infection.** *Proc Natl Acad Sci USA* 2003, **100(22)**:12989-12994.
- Rengarajan J, Bloom BR, Rubin EJ: **Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages.** *Proc Natl Acad Sci USA* 2005, **102(23)**:8327-8332.
- Smith I: **Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence.** *Clin Microbiol Rev* 2003, **16(3)**:463-496.
- Reddy TB, Riley R, Wymore F, Montgomery P, Decaprio D, Engels R, Gellesch M, Hubble J, Jen D, Jin H, et al.: **TB database: an integrated platform for tuberculosis research.** *Nucleic Acids Res* 2008, **37(Database issue)**:D499-D508.
- Lechat P, Hummel L, Rousseau S, Moszer I: **GenoList: an integrated environment for comparative analysis of microbial genomes.** *Nucleic Acids Res* 2008:D469-474.
- Chaudhuri RR, Pallen MJ: **xBASE, a collection of online databases for bacterial comparative genomics.** *Nucleic Acids Res* 2006:D335-337.
- Chaudhuri RR, Loman NJ, Snyder LA, Bailey CM, Stekel DJ, Pallen MJ: **xBASE2: a comprehensive resource for comparative bacterial genomics.** *Nucleic Acids Res* 2008:D543-546.
- Ranjan S, Gundu RK, Ranjan A: **MycoperoDB: a database of computationally identified operons and transcriptional units in Mycobacteria.** *BMC Bioinformatics* 2006, **7(Suppl 5)**:S9.
- Vishnoi A, Srivastava A, Roy R, Bhattacharya A: **MGDD: Mycobacterium tuberculosis genome divergence database.** *BMC Genomics* 2008, **9**:373.
- Vishnoi A, Roy R, Bhattacharya A: **Comparative analysis of bacterial genomes: identification of divergent regions in mycobacterial strains using an anchor-based approach.** *Nucleic Acids Res* 2007, **35(11)**:3654-3667.
- Catanho M, Mascarenhas D, Degraeve W, Miranda AB: **GenoMycDB: a database for comparative analysis of mycobacterial genes and genomes.** *Genet Mol Res* 2006, **5(1)**:115-126.
- Jacques PE, Gervais AL, Cantin M, Lucier JF, Dallaire G, Drouin G, Gaudreau L, Goulet J, Brzezinski R: **MtbRegList, a database dedicated to the analysis of transcriptional regulation in Mycobacterium tuberculosis.** *Bioinformatics* 2005, **21(10)**:2563-2565.
- Tatusov RL, Koonin EV, Lipman DJ: **A genomic perspective on protein families.** *Science* 1997, **278(5338)**:631-637.
- Tatusov RL, Galperin MY, Natale DA, Koonin EV: **The COG database: a tool for genome-scale analysis of protein functions and evolution.** *Nucleic Acids Res* 2000, **28(1)**:33-36.
- Leung AS, Tran V, Wu Z, Yu X, Alexander DC, Gao GF, Zhu B, Liu J: **Novel genome polymorphisms in BCG vaccine strains and impact on efficacy.** *BMC Genomics* 2008, **9**:413.
- Kato-Maeda M, Rhee JT, Gingeras TR, Salamon H, Drenkow J, Smittipat N, Small PM: **Comparing genomes within the species Mycobacterium tuberculosis.** *Genome Res* 2001, **11(4)**:547-554.
- Semret M, Zhai G, Mostowy S, Cleto C, Alexander D, Cangelosi G, Cousins D, Collins DM, van Soolingen D, Behr MA: **Extensive genomic polymorphism within Mycobacterium avium.** *J Bacteriol* 2004, **186(18)**:6332-6334.
- Tsolaki AG, Hirsh AE, DeRiemer K, Enciso JA, Wong MZ, Hannan M, Goguet de la Salmoniere YO, Aman K, Kato-Maeda M, Small PM:

- Functional and evolutionary genomics of Mycobacterium tuberculosis: insights from genomic deletions in 100 strains.** *Proc Natl Acad Sci USA* 2004, **101(14)**:4865-4870.
30. Yang J, Chen L, Sun L, Yu J, Jin Q: **VFDB 2008 release: an enhanced web-based resource for comparative pathogenomics.** *Nucleic Acids Res* 2008, **36(Database issue)**:D539-D542.
  31. Velmurugan K, Chen B, Miller JL, Azogue S, Gurses S, Hsu T, Glickman M, Jacobs WR Jr, Porcelli SA, Briken V: **Mycobacterium tuberculosis nuoG is a virulence gene that inhibits apoptosis of infected host cells.** *PLoS Pathog* 2007, **3(7)**:e110.
  32. Kana BD, Gordhan BG, Downing KJ, Sung N, Vostroktunova G, Machowski EE, Tsenova L, Young M, Kaprelyants A, Kaplan G, et al.: **The resuscitation-promoting factors of Mycobacterium tuberculosis are required for virulence and resuscitation from dormancy but are collectively dispensable for growth in vitro.** *Mol Microbiol* 2008, **67(3)**:672-684.
  33. Bhatt A, Fujiwara N, Bhatt K, Gurcha SS, Kremer L, Chen B, Chan J, Porcelli SA, Kobayashi K, Besra GS, et al.: **Deletion of kasB in Mycobacterium tuberculosis causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice.** *Proc Natl Acad Sci USA* 2007, **104(12)**:5157-5162.
  34. Alteri CJ, Xicohtencatl-Cortes J, Hess S, Caballero-Olin G, Giron JA, Friedman RL: **Mycobacterium tuberculosis produces pili during human infection.** *Proc Natl Acad Sci USA* 2007, **104(12)**:5145-5150.
  35. Calamita H, Ko C, Tyagi S, Yoshimatsu T, Morrison NE, Bishai WR: **The Mycobacterium tuberculosis SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence.** *Cell Microbiol* 2005, **7(2)**:233-244.
  36. Malhotra V, Sharma D, Ramanathan VD, Shakila H, Saini DK, Chakravorty S, Das TK, Li Q, Silver RF, Narayanan PR, et al.: **Disruption of response regulator gene, devR, leads to attenuation in virulence of Mycobacterium tuberculosis.** *FEMS Microbiol Lett* 2004, **231(2)**:237-245.
  37. Parish T, Smith DA, Kendall S, Casali N, Bancroft GJ, Stoker NG: **Deletion of two-component regulatory systems increases the virulence of Mycobacterium tuberculosis.** *Infect Immun* 2003, **71(3)**:1134-1140.
  38. Alm EJ, Huang KH, Price MN, Koche RP, Keller K, Dubchak IL, Arkin AP: **The MicrobesOnline Web site for comparative genomics.** *Genome Res* 2005, **15(7)**:1015-1022.
  39. Price MN, Huang KH, Alm EJ, Arkin AP: **A novel method for accurate operon predictions in all sequenced prokaryotes.** *Nucleic Acids Res* 2005, **33(3)**:880-892.
  40. He X, Chang S, Zhang J, Zhao Q, Xiang H, Kusonmano K, Yang L, Sun ZS, Yang H, Wang J: **MethyCancer: the database of human DNA methylation and cancer.** *Nucleic Acids Res* 2008:D836-841.
  41. Chang S, Zhang J, Liao X, Zhu X, Wang D, Zhu J, Feng T, Zhu B, Gao GF, Wang J, et al.: **Influenza Virus Database (IVDB): an integrated information resource and analysis platform for influenza virus research.** *Nucleic Acids Res* 2007, **35**:D376-380.
  42. Wang J, He X, Ruan J, Dai M, Chen J, Zhang Y, Hu Y, Ye C, Li S, Cong L, et al.: **ChickVD: a sequence variation database for the chicken genome.** *Nucleic Acids Res* 2005:D438-441.
  43. Wang J, Xia Q, He X, Dai M, Ruan J, Chen J, Yu G, Yuan H, Hu Y, Li R, et al.: **SilkDB: a knowledgebase for silkworm biology and genomics.** *Nucleic Acids Res* 2005:D399-402.
  44. Zhao W, Wang J, He X, Huang X, Jiao Y, Dai M, Wei S, Fu J, Chen Y, Ren X, et al.: **BGI-RIS: an integrated information resource and comparative analysis workbench for rice genomics.** *Nucleic Acids Res* 2004:D377-382.
  45. Stothard P, Wishart DS: **Circular genome visualization and exploration using CGView.** *Bioinformatics* 2005, **21(4)**:537-539.
  46. Guinn KM, Hickey MJ, Mathur SK, Zakel KL, Grotzke JE, Lewinson DM, Smith S, Sherman DR: **Individual RDI-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis.** *Mol Microbiol* 2004, **51(2)**:359-370.
  47. Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M, et al.: **The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue.** *Proc Natl Acad Sci USA* 2003, **100(21)**:12420-12425.
  48. Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C, et al.: **Genome plasticity of BCG and impact on vaccine efficacy.** *Proc Natl Acad Sci USA* 2007, **104(13)**:5596-5601.
  49. Rehren G, Walters S, Fontan P, Smith I, Zarraga AM: **Differential gene expression between Mycobacterium bovis and Mycobacterium tuberculosis.** *Tuberculosis (Edinb)* 2007, **87(4)**:347-359.
  50. Golby P, Hatch KA, Bacon J, Cooney R, Riley P, Allnutt J, Hinds J, Nunez J, Marsh PD, Hewinson RG, et al.: **Comparative transcriptomics reveals key gene expression differences between the human and bovine pathogens of the Mycobacterium tuberculosis complex.** *Microbiology* 2007, **153(Pt 10)**:3323-3336.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

