Designing Novel Antibacterials: Application of Omics Science

Yeni Antibiyotikler Tasarlamak: "Omik" Bilimin Uygulanması

Aubhishek Zaman[1], Samsad Razzaque[2]

[1]Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka, Bangladesh.
[2]Biotechnology Program, Department of Mathematics and Natural Sciences, BRAC University, Dhaka, Bangladesh

Abstract

Antibacterials are agents that act against pernicious bacterial pathogens by killing or inactivating them. At present, designing antibacterials have been assisted greatly by omics sciences- genomics, proteomics, metabolomics and interactomics. This review discusses different aspects of the omics sciences in simple to complex hierarchies. From the simplest of sequence analysis to more complex on the pyramid -structural, functional and interactional analysis- all comprises the grand ambit of omics science. Sequence comparison can reveal novel information about drug resistance in the bacteria and thus can be of momentous significance for designing improved antibacterials. On the other hand, sequence characterization of the host protein can lead to production of effective antibiotic synthetically. Nowadays, structure based molecular designing, using computational docking techniques, has become a widely accepted routine work in drug designing processes. Moreover, new high-throughput data from microarray expression, protein-protein interaction assay are opening up a new vista for detecting more and more drug targets. Extensive focus put on to understand host-pathogen interaction on systems level has greatly accelerated the process of designing effective antibacterials against tuberculosis and many such complex diseases. Summarizing, this review exemplifies various different ways how increasingly omics science is transforming the paradigm of discovering novel antibacterials; omics approach is all set to speed up the process and bring down the expenses of the antibacterials even more in time to come.
Klimik Dergisi 2013; 26(1): 2-8.

Key Words: Anti-bacterial agents, omics science, comparative genomics, structural and functional omics.

Özet

Antibiyotikler, zararlı bakteriyel patojenleri yok ederek ya da inaktif hale getirerek etki gösteren ajanlardır. Günümüzde antibiyotiklerin tasarlanmasında genom bilimi, proteom bilimi, metabolom bilimi ve interaktom bilimi gibi "om" bilimlerinden (ya da "omik" bilimlerden) büyük ölçüde yararlanılmaktadır. Bu derlemede bu bilimlerin basitten karmaşığa giden farklı yönleri tartışılmaktadır. Piramidin en altındaki basit dizi analizinden üst kısmındaki karmaşık -yapısal, fonksiyonel ve interaksiyonel- analizlerin tümü, omik bilimlerin kapsamına girmektedir. Dizi karşılaştırması, bakterilerdeki ilaç direncine ilişkin yeni bilgiler sunabilmesi nedeniyle, yeni antibiyotik geliştirilmesinde son derece önemlidir. Öte yandan, konak proteininin dizilenmesi, etkili antibiyotiklerin sentetik olarak üretilmesine olanak tanıyabilir. Günümüzde bilgisayar destekli bağlanma teknikleri kullanıldığı yapıya dayalı molekül tasarımı, ilaç tasarım süreçlerinde geniş ölçüde kabul gören rutin bir işlem haline gelmiştir. Ayrıca, mikroçip ekspresyonu, protein-protein etkileşimli essey yöntemleriyle elde edilen yeni ve yüksek verimli veriler, çok daha fazla ilaç hedefinin keşfi için yeni ufuklar açmaktadır. Konak-patojen etkileşimini sistemler düzeyinde anlamak için gösterilen yoğun çabalar, tüberküloza ve benzer pek çok karmaşık hastalığa karşı etkili antibiyotiklerin tasarlanması sürecini hızlandırmıştır. Özetle, bu derlemede omik bilimlerin, yeni antibiyotiklerin keşfi alanındaki paradigmayı, gittikçe artan bir biçimde nasıl değiştirdiğine ilişkin çeşitli örnekler sunulmaktadır. Omik yaklaşım, önümüzdeki dönemde bu süreci daha da hızlandıracak ve antibiyotiklerin maliyetini daha da düşürecektir.
Klimik Dergisi 2013; 26(1): 2-8.

Anahtar Sözcükler: Antibiyotikler, omik bilimler, karşılaştırmalı genom bilimi, yapısal ve işlevsel omik bilim.

Introduction

Antibacterials, alternatively known as antibiotics, fall into the broad group of small molecules that act against bacteria by various mechanisms (1-4). A key difference between antibiotics and vaccine is that a vaccine may work to bolster immune responses against pathogenic infection and essentially do not act directly on the microbe itself whereas in contrast antibacterials works against the pathogens directly to kill or inactivate them (5).

Address for Correspondence / Yazışma Adresi:
Aubhishek Zaman, Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka, Bangladesh
Phone/Tel.: +880 171 779 15 28 Fax/Faks: +880 2 861 55 83 E-mail/E-posta: aubhishek@gmail.com
(Received / Geliş: 9 September / Eylül 2012; Accepted / Kabul: 10 March / Mart 2013)
DOI: 10.5152/kd.2013.02

Today, antibacterial designing has become greatly assisted by omics sciences. Omics sciences include a wide spectrum of themes: genomics, proteomics, metabolomics and interactomics. Genomics deals with gene information and makes sense of it (6). Proteomics involves the study protein structure and function (7). Metabolomics investigates integrated information about metabolic pathways (8) whereas interactomics refers to the study of all the possible interactions, such as DNA-protein, protein-protein, protein-small molecule interactions within living cells (9-11). Omics sciences are helpful in elucidating the sequence, structure and functional information about both the pathogen and host genes and proteins (12-14).

Complete genome sequences of bacterial organisms have had a revolutionary effect on the process of designing new antibiotics. The completion of nearly 30 bacterial whole-genome sequences and ongoing sequencing projects of over 100 microbial organisms will allow researchers to probe novel therapeutic targets (15). The search for new antibiotics can be extensively assisted by computational methods such as homology-based analyses, structural genomics, motif analyses, protein-protein interactions, molecular docking and experimental functional genomics (15). However the greatest obstacle of computational assays is massive volume of data from the genome sequences and making sense of it. The sequence of microbial pathogens catalogs every gene product that would be relevant for the host-parasite interaction and potential antibiotic drug target (16,17). Therefore, scientists interested in discovering antibiotics must extract useful information from genomes through comparative, functional, or structural genomics in order to simplify drug target selection.

Whole Genome Sequencing: A Gateway to Designing Better Drug

Sequence Comparison

Useful antibiotic development and the bacterial whole-genome sequences go hand in hand. However, the concurrence of the two is even more significant as increasingly resistance of commonly used antibiotics have been developed in bacteria. A growing prevalence of infections, and the emergence of new pathogenic organisms is challenging current antibiotics pool (18). When antibiotics synthesized afresh resembles that of the ones used earlier to ones already rendered ineffective, resistance is more likely to happen (19). This poses a great concern for the clinicians and world health in large, as resistant-to-antibiotic bacteria can trigger a massive global epidemic. Ideally, new antimicrobial compounds should have novel mechanisms of action.

Phylogenetic trees are an essential tool that are used to identify recognized sequence homology pattern which can be very useful to identify essential genes for the pathogens (20). One can also design a antibacterial that works against an entire phylogenetic group in order to target all of the organisms with a broad-spectrum antibiotic. Thus the concept of universal antibacterials has flourished using sequence comparison tools.

Well before whole genome sequencing for bacteria became popular, sequencing of many apparently important genes was carried out. However, with DNA sequencing becoming a household scientific exercise, sequence comparison has been an exciting technique that took the world by storm. Sequence comparison has been an exciting new technology to compare sensitive and resistant strains of the bacteria. In a widely cited paper authored by Fournier *et al.* (21) a similar study was carried out for a multidrug resistant strain. Similar studies have been carried out for *Streptococcus*, *Staphylococcus* and *Deinococcus* (22-26). Several computational tools and online databases have also gained prominence over the years (27). Some of the computational databases are enlisted in Table 1.

Sequence Characterization

Sequence characterization is another key theme associated with designing functioning antibiotic target. Often a motif or a profile is found in the genomic or peptide sequence of the antibacterial targets (73). Sequence characterization is also of great importance for analyzing actions of the peptide antibiotics (74). Thus comparison between resistant and sensitive strain requires a detail characterization of the protein or genomic sequence of the both target molecules and peptide antibiotics (75,76).

Commonly used antibiotic drugs target series-specific genes, unique enzymes and membrane transporters (77). The mechanism of action how antibiotics mediate its response is diverse; some antibiotics prevent protein synthesis and nucleic acid replication, some inhibit cell wall or membrane synthesis, some rather prevents membrane transport (78). In this regard, all the bacteria have a special set of proteins that are responsible for either causing virulence or taking hold of host machinery. Identifying those genes are of supreme importance and hence sequence characterization for bacteria for which no sequence information is deposited of yet, is the only option in that case.

To begin with, in a novel bacteria sequence characterization flow chart, scientists start characterizing all the open reading frames of bacterial sequences and make a map of all genes and gene products (79,80). Afterwards, they must pick out the genes that are essential to cell survival or growth, which are most effective as antibiotic targets. Often the line of action to detect this genes is to go for a random mutagenesis and subsequent phenotyping of the bacteria (16). However, the job today has become a lot easier as representative genome sequences from almost all the pathologically and economically important bacteria has already been done. And with this being done even the primary sequence comparison programs, like BLAST or PSI-BLAST, can determine gene functions by sequence homology.

Motif analysis is another strategy to identify potential antibiotic targets among genes with unknown functions. Many databases, including PROSITE, InterPro, BLOCKS, Pfam etc., can search for motifs in a sequence (16,81-93). The motifs may show the approximate biochemical function of the gene.

Gene fusion is another computational method to infer protein interactions from genome sequences. Proteins that interact with each other tend to have homologs in other organisms. This evolutionary calling-of-function

Table 1. Databases and Online Tools Often Used for Comparative Genomics

Database/Tool	Use	Reference	Web Link
MBGD	Comparative analysis of completely sequenced microbial genomes	(28-30)	http://mbgd.genome.ad.jp/
WormBase	Information about *Caenorhabditis elegans* and related nematodes	(31)	http://www.wormbase.org/
JCVI CMR	Cross-genome analysis to identify differences and similarities between the genomes	(32)	http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi
Vista	To examine pre-computed whole-genome alignments of different species	(33)	http://genome.lbl.gov/vista/index.shtml
HOBACGEN	Comparative genome analysis using protein genes from bacteria, *Archaea*, and yeasts	(34,35)	http://pbil.univ-lyon1.fr/databases/hobacgen.html
PipMaker	Comparing two long DNA sequences to identify conserved segments and for producing informative, high-resolution displays of the resulting alignments	(36)	http://www.bx.psu.edu/miller_lab/
PLAZA	Plant comparative genomics	(37,38)	
UCSC Genome Browser	Contains the reference sequence and working draft assemblies for a large collection of genomes	(39-52)	http://genome.ucsc.edu/
Ensembl	Genome databases for vertebrates and other eukaryotic species	(53-63)	http://uswest.ensembl.org/index.html
PlantGDB	Provides genome browsers to display current gene structure models and transcript evidence from spliced alignments of EST and cDNA sequences	(64-68)	http://www.plantgdb.org/
LegumeIP	Comparative genomics and transcriptomics of model legumes	(69)	http://www.biosharing.org/biodbcore-000056
ShiBASE	Comparative genomics of *Shigella*	(70)	http://www.mgc.ac.cn/ShiBASE/
CoGemiR	Conservation of microRNAs during evolution in different animal species	(71)	http://cogemir.tigem.it/
Neisseria Base	Genome browser for *Neisseria meningitidis*	(72)	http://nbase.biology.gatech.edu

method often gives out functional information for target proteins (16).

Function Based Techniques Help Selecting Soft Targets for Designing Antibiotic against It

Assigning functions to the genes is one of the major steps involved in designing soft targets in bacteria for which drug can be designed. Identifying the genes that are essential for proper functioning of the bacteria is thus also important. Many online databases contain these information and thus can be tremendously useful for antibacterial designing (94,95).

Microarray or Fuzzy Algorithms

There are some disadvantages associated to sequence homology based methods. About 25-40% of the genes in a bacterial genome usually do not find matches with known genes (79,80). Furthermore, sequence homology is based on the assumption that similar sequences will share similar functions -an assumption that does not hold true in many cases where similar sequences are structurally and functionally diverse.

Therefore, alternatives to sequence homology techniques had to be established. To predict the function of a gene, cluster analysis of the expression profile has been extensively used. Cluster analysis uses microarray technology to analyze gene expression in order to organize genes into functional groups (96). Genes for which no annotation has been assigned can be classified into a functional group and thus can be assigned a functional annotation on the basis of microarray data instead of sequence data. Protein synthesis patterns are also useful to analyze the antimicrobial effect certain drugs would have on particular necessary or important proteins (97).

Systems Biology

Using systems biology to design drugs have been a popular approach in the post genomic era (98). Systems biology considers genes and proteins to be integrated to each other and hence takes up an integrated approach for drug designing (99,100). Protein-protein interaction (PPI) and gene regulatory networks (GRN) are one of the most recurrent themes of systems biology in designing drugs. PPI exposes novel information about whole proteome interaction status (101,102). A common strategy of the systems biology investigators has been looking for most densely interconnected protein, known as hub proteins, in the systems map; hub proteins often make a good antibiotic target (103-105). Similarly DNA-protein interaction and GRN has been an important tool to understand host pathogen interaction and host cellular mechanism (106,107).

Systems biology has been applied assaying drugs against pathogens with complex life cycle. Detecting a universal drug against these pathogens have been difficult as they drug target proteins are often poorly understood. However, today systems approach has been applied successfully to design antibacterials against tuberculosis and gastric ulcer (108-111).

Structure Based Techniques Assists Designing Molecular Medicine against Pathogens

Although assigning gene function by cluster analyses is quite useful as described in the earlier section, they are also subject to significant level of discrepancies as well. Some proteins have multiple functions and likewise, some functions require multiple proteins (8).

Therefore, structural genomics has been used as a better method of drug target selection. Function is more directly related to its structure than its sequence (96). Now even considerable number of protein 3D structures in the native tertiary form has also been deposited in structure databases. That makes possible the task of comparing different protein structures and annotating functions accordingly. Some such protein structure databases are RCSB Protein Data Bank (PDB), PDBsum, ModBase, Proteopedia, 3D Complex, SCOP etc. (112-116).

Another property of the drug target should be non-redundancy that is the target should be structurally different or nonexistent in humans. Checking for structural homology against a human genome protein structure database would determine whether the antibiotic against that drug target would also interfere with any human functions.

Another key advantage of structure based medicines is that the action of the drug is very predictable in nature. Because the drug-protein interaction involves a complementary fit to each other and not any co-expression information, they are remarkably specific in their mode of action most of the times (117-121). These very properties have made structural methods an ideal choice for selection of drug targets. However, structural databases are not complete since quality protein-crystals are difficult to form and hinders x-ray crystallography (122). However, nuclear magnetic resonance can determine 3D structure determination. Also, computational modeling is approaching accurate functional predictions based on alignment of amino acid sequences (123).

The use of computational methods and expression profiling all point to the need for a non-redundant, complete database of structural and functional annotation of the proteins from known pathogenic bacteria genomes and the human genome, once it is completed. The organization, accuracy, and easy accessibility of such databases are crucial in the hunt for novel antibiotics. Perhaps a program can be specifically designed to highlight antibiotic drug targets in query sequences. This program would scan structural databases and other bacterial genomes for homology and similar folds. The program could be complemented by a central, tailored database that reorganizes data for the most efficient search of novel antibiotic targets, for example, each protein or gene that is essential to certain bacterial species. For example, the database could include the protein's phylogenetic group, 3D structure, proteins of similar structural homology, and whether any similar protein exists in humans. It could also use foreign keys to connect to other databases that catalogue which known antibiotics and inhibitors are used against similar targets.

Conclusion

Summarizing, the computational methods and omics sciences has become an integral part of designing novel antibacterials. Therefore, along with structure function annotation to ensure rapid and effective communication of the *in vitro* results has become an absolute essential in this regard. *In silico* experimentations have also added to the data wealth and thus an efficient data management is also the call of the hour.

One should not get carried away by the whole-genome sequence data available; there are still many hurdles to overcome. One of the major problems is the localization of the drug target and an efficient drug delivery that can hardly be analyzed by omics tools (124). Therefore requirement of drug response databases -databases that deposit pharmacokinetic and pharmacodynamic data- are also gaining prominence. Despite limitations there are some databanks growing for small molecules and their properties (125). Approval and strict patent laws is another administrative hurdle that makes commercialized antibacterials something extremely hard to materialize. Lastly but more importantly, an indiscriminate use of antibiotics is, in an unprecedented speed, making more and more bacterial species resistant and this is putting unmanageable pressure to the demand of antibiotic. Scaling up antibacterial production has improved to a great extent over the last few decades, yet the supply of antibacterials hardly catches up to the demand (126). All these pose severe challenges for designing antimicrobials. However, it is to be noted that never before technologies, computer aided tools, huge workforce and human intelligence worked in so unified nature. Hence, it would not be an overstatement to say that the days today for novel antibacterials designing are brighter than it has ever been before.

Conflict of Interest

No conflict of interest was declared by the authors.

References

1. Alcocer F, López E, Calva JJ, Herrera MF. Antibiotic therapy in secondary peritonitis: towards a definition of its optimal duration. *Rev Invest Clin.* 2001; 53(2): 121-5.
2. Waksman, SA. Definition of antibiotic. *Int Rec Med Gen Pract Clin.* 1956; 169(2): 87-8.
3. Saddiqe Z, Naeem I, Maimoona A. A review of the antibacterial activity of Hypericum perforatum L. *J Ethnopharmacol.* 2010; 131(3): 511-21. [CrossRef]
4. Nevalainen TJ, Graham GG, Scott KF. Antibacterial actions of secreted phospholipases A2. Review. *Biochim Biophys Acta.* 2008; 1781(1-2): 1-9. [CrossRef]
5. Di Guilmi AM, Dessen A. New approaches towards the identification of antibiotic and vaccine targets in Streptococcus pneumoniae. *EMBO Rep.* 2002; 3(8): 728-34. [CrossRef]

6. Hrmova M, Fincher GB. Functional genomics and structural biology in the definition of gene function. *Methods Mol Biol.* 2009; 513: 199-227. [CrossRef]
7. Lan N, Montelione GT, Gerstein M. Ontologies for proteomics: towards a systematic definition of structure and function that scales to the genome level. *Curr Opin Chem Biol.* 2003; 7(1): 44-54. [CrossRef]
8. Braaksma M, Bijlsma S, Coulier L, Punt PJ, van der Werf MJ. Metabolomics as a tool for target identification in strain improvement: the influence of phenotype definition. *Microbiology*. 2011; 157(Pt 1): 147-59. [CrossRef]
9. Ito T. Trends in interactomics. *Tanpakushitsu Kakusan Koso*. 2005; 50(Suppl. 16): 2263-8.
10. Cesareni G, Ceol A, Gavrila C, Palazzi LM, Persico M, Schneider MV. Comparative interactomics. *FEBS Lett.* 2005; 579(8): 1828-33. [CrossRef]
11. Collura V, Boissy G. From protein-protein complexes to interactomics. *Subcell Biochem*. 2007; 43: 135-83. [CrossRef]
12. Wild CP. OMICS technologies: an opportunity for "two-way" translation from basic science to both clinical and population-based research. *Occup Environ Med*. 2010; 67(2): 75-6. [CrossRef]
13. Garcia I, Tabak LA. Beyond the "omics": translating science into improved health. *J Am Dent Assoc.* 2008; 139(4): 392-5. [CrossRef]
14. Holmes C, McDonald F, Jones M, Ozdemir V, Graham JE. Standardization and omics science: technical and social dimensions are inseparable and demand symmetrical study. *OMICS*. 2010; 14(3): 327-32. [CrossRef]
15. Loferer H. Mining bacterial genomes for antimicrobial targets. *Mol Med Today.* 2000; 6(12): 470-4. [CrossRef]
16. Hood DW. The utility of complete genome sequences in the study of pathogenic bacteria. *Parasitology*. 1999; 118(Suppl.): 3-9. [CrossRef]
17. Finken M, Kirschner P, Meier A, Wrede A, Böttger EC. Molecular basis of streptomycin resistance in Mycobacterium tuberculosis: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Mol Microbiol*. 1993; 9(6): 1239-46. [CrossRef]
18. Rosamond J, Allsop A. Harnessing the power of the genome in the search for new antibiotics. *Science*. 2000; 287(5460): 1973-6. [CrossRef]
19. Hughes D. Exploiting genomics, genetics and chemistry to combat antibiotic resistance. *Nat Rev Genet*. 2003; 4(6): 432-41. [CrossRef]
20. Gerstein M. Integrative database analysis in structural genomics. *Nat Struct Biol*. 2000; 7 (Suppl.): 960-3. [CrossRef]
21. Fournier PE, Vallenet D, Barbe V, *et al.* Comparative genomics of multidrug resistance in Acinetobacter baumannii. *PLoS Genet.* 2006; 2(1): 7. [CrossRef]
22. Yamamoto T, Takano T, Higuchi W, *et al.* Comparative genomics and drug resistance of a geographic variant of ST239 methicillin-resistant Staphylococcus aureus emerged in Russia. *PLoS One*. 2012; 7(1): e29187. [CrossRef]
23. Hu P, Yang M, Zhang A, *et al.* Comparative genomics study of multi-drug-resistance mechanisms in the antibiotic-resistant Streptococcus suis R61 strain. *PLoS One.* 2011; 6(9): 24988. [CrossRef]
24. Christianson S, Golding GR, Campbell J; Canadian Nosocomial Infection Surveillance Program, Mulvey MR. Comparative genomics of Canadian epidemic lineages of methicillin-resistant Staphylococcus aureus. *J Clin Microbiol*. 2007; 45(6): 1904-11. [CrossRef]
25. Hua Y, Gao G. Comparative genomics of genes contributed to DNA repair in the radiation-resistant Deinococcus radiodurans. *Wei Sheng Wu Xue Bao*. 2003; 43(1): 120-6.
26. Makarova KS, Aravind L, Wolf YI, *et al.* Genome of the extremely radiation-resistant bacterium Deinococcus radiodurans viewed from the perspective of comparative genomics. *Microbiol Mol Biol Rev.* 2001; 65(1): 44-79. [CrossRef]
27. Pennacchio LA, Rubin EM. Comparative genomic tools and databases: providing insights into the human genome. *J Clin Invest.* 2003; 111(8): 1099-106. [CrossRef]
28. Uchiyama I. MBGD: microbial genome database for comparative analysis. *Nucleic Acids Res.* 2003; 31(1): 58-62. [CrossRef]
29. Uchiyama I. MBGD: a platform for microbial comparative genomics based on the automated construction of orthologous groups. *Nucleic Acids Res.* 2007; 35(Suppl. 1): D343-6. [CrossRef]
30. Uchiyama I, Higuchi T, Kawai M. MBGD update 2010: toward a comprehensive resource for exploring microbial genome diversity. *Nucleic Acids Res.* 2010; 38(Suppl. 1): D361-5. [CrossRef]
31. Harris TW, Lee R, Schwarz E, *et al.* WormBase: a cross-species database for comparative genomics. *Nucleic Acids Res.* 2003; 31(1): 133-7. [CrossRef]
32. Peterson JD, Umayam LA, Dickinson T, Hickey EK, White O. The comprehensive microbial resource. *Nucleic Acids Res.* 2001; 29(1): 123-5. [CrossRef]
33. Lukashin I, Novichkov P, Boffelli D, *et al.* VISTA Region Viewer (RViewer)--a computational system for prioritizing genomic intervals for biomedical studies. *Bioinformatics*. 2011 Sep 15;27(18):2595-7.
34. Larimer F. HOBACGEN: homologous bacterial genes database. *Brief Bioinform*. 2000; 1(4): 415-6. [CrossRef]
35. Perrière G, Duret L, Gouy M. HOBACGEN: database system for comparative genomics in bacteria. *Genome Res.* 2000; 10(3): 379-85. [CrossRef]
36. Schwartz S, Zhang Z, Frazer KA, *et al.* PipMaker--a web server for aligning two genomic DNA sequences. *Genome Res.* 2000; 10(4): 577-86. [CrossRef]
37. Van Bel M, Proost S, Wischnitzki E, *et al.* Dissecting plant genomes with the PLAZA comparative genomics platform. *Plant Physiol*. 2012; 158(2): 590-600. [CrossRef]
38. Proost S, Van Bel M, Sterck L, *et al.* PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *Plant Cell*. 2009; 21(12): 3718-31. [CrossRef]
39. Kent WJ, Sugnet CW, Furey TS, *et al.* The human genome browser at UCSC. *Genome Res.* 2002; 12(6): 996-1006. [CrossRef]
40. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinform*. 2013; 14(2): 144-61. [CrossRef]
41. Chan PP, Holmes AD, Smith AM, Tran D, Lowe TM. The UCSC Archaeal Genome Browser: 2012 update. *Nucleic Acids Res.* 2012; 40(Suppl. 1): D646-52. [CrossRef]
42. Karolchik D, Hinrichs AS, Kent WJ. The UCSC genome browser. *Curr Protoc Hum Genet.* 2011; Chapter 18: Unit18.6.
43. Fujita PA, Rhead B, Zweig AS, *et al.* The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.* 2011; 39(Suppl. 1): D876-82. [CrossRef]
44. Rhead B, Karolchik D, Kuhn RM, *et al.* The UCSC Genome Browser database: update 2010. *Nucleic Acids Res.* 2010; 38(Suppl. 1): D613-9. [CrossRef]
45. Kuhn RM, Karolchik D, Zweig AS, *et al.* The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res.* 2009; 37(Suppl. 1): D755-61. [CrossRef]
46. Zweig AS, Karolchik D, Kuhn RM, Haussler D, Kent WJ. UCSC genome browser tutorial. *Genomics*. 2008; 92(2): 75-84. [CrossRef]
47. Karolchik D, Hinrichs AS, Kent WJ. The UCSC Genome Browser. *Curr Protoc Bioinformatics.* 2007; Chapter 1: Unit 1.4.
48. Karolchik D, Kuhn RM, Baertsch R, *et al.* The UCSC Genome

Browser Database: 2008 update. *Nucleic Acids Res.* 2008; 36(Suppl. 1): D773-9. [CrossRef]
49. Kuhn RM, Karolchik D, Zweig AS, *et al.* The UCSC genome browser database: update 2007. *Nucleic Acids Res.* 2007; 35(Suppl. 1): D668-73. [CrossRef]
50. Hinrichs AS, Karolchik D, Baertsch R, *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* 2006; 34(Suppl. 1): D590-8. [CrossRef]
51. Schneider KL, Pollard KS, Baertsch R, Pohl A, Lowe TM. The UCSC Archaeal Genome Browser. *Nucleic Acids Res.* 2006; 34(Suppl. 1): D407-10. [CrossRef]
52. Karolchik D, Baertsch R, Diekhans M, *et al.* The UCSC Genome Browser Database. *Nucleic Acids Res.* 2003; 31(1): 51-4. [CrossRef]
53. Hubbard T, Barker D, Birney E, *et al.* The Ensembl genome database project. *Nucleic Acids Res.* 2002; 30(1): 38-41. [CrossRef]
54. Birney E, Andrews D, Bevan P, *et al.* Ensembl 2004. *Nucleic Acids Res.* 2004; 32(Suppl. 1): D468-70. [CrossRef]
55. Birney E, Andrews TD, Bevan P, *et al.* An overview of Ensembl. *Genome Res.* 2004; 14(5): 925-8. [CrossRef]
56. Hammond MP, Birney E. Genome information resources - developments at Ensembl. *Trends Genet.* 2004; 20(6): 268-72. [CrossRef]
57. Birney E; Ensembl Team. Ensembl: a genome infrastructure. *Cold Spring Harb Symp Quant Biol.* 2003; 68: 213-5. [CrossRef]
58. Hubbard T, Andrews D, Caccamo M, *et al.* Ensembl 2005. *Nucleic Acids Res.* 2005; 33(Suppl. 1): D447-53. [CrossRef]
59. Hubbard TJ, Aken BL, Beal K, *et al.* Ensembl 2007. *Nucleic Acids Res.* 2007; 35(Suppl. 1): D610-7. [CrossRef]
60. Flicek P, Aken BL, Beal K,*et al.* Ensembl 2008. *Nucleic Acids Res.* 2008; 36(Suppl. 1): D707-14. [CrossRef]
61. Hubbard TJ, Aken BL, Ayling S, *et al.* Ensembl 2009. *Nucleic Acids Res.* 2009; 37(Suppl. 1): D690-7. [CrossRef]
62. Flicek P, Amode MR, Barrell D, *et al.* Ensembl 2011. *Nucleic Acids Res.* 2011; 39(Suppl. 1): D800-6. [CrossRef]
63. Flicek P, Amode MR, Barrell D, *et al.* Ensembl 2012. *Nucleic Acids Res.* 2012; 40(Suppl. 1): D84-90. [CrossRef]
64. Brendel V. Gene structure annotation at PlantGDB. *Methods Mol Biol.* 2007; 406: 521-33.
65. Duvick J, Fu A, Muppirala U, *et al.* PlantGDB: a resource for comparative plant genomics. *Nucleic Acids Res.* 2008; 36(Suppl. 1): D959-65. [CrossRef]
66. Dong Q, Lawrence CJ, Schlueter SD, *et al.* Comparative plant genomics resources at PlantGDB. *Plant Physiol.* 2005; 139(2): 610-8. [CrossRef]
67. Dong Q, Schlueter SD, Brendel V.PlantGDB, plant genome database and analysis tools. *Nucleic Acids Res.* 2004; 32(Suppl. 1): D354-9. [CrossRef]
68. Schlueter SD, Dong Q, Brendel V. GeneSeqer@PlantGDB: Gene structure prediction in plant genomes. *Nucleic Acids Res.* 2003; 31(13): 3597-600. [CrossRef]
69. Li J, Dai X, Liu T, Zhao PX. LegumeIP: an integrative database for comparative genomics and transcriptomics of model legumes. *Nucleic Acids Res.* 2012; 40(Suppl. 1): D1221-9. [CrossRef]
70. Yang J, Chen L, Yu J, Sun L, Jin Q. ShiBASE: an integrated database for comparative genomics of Shigella. *Nucleic Acids Res.* 2006; 34(Suppl. 1): D398-401. [CrossRef]
71. Maselli V, Di Bernardo D, Banfi S. CoGemiR: a comparative genomics microRNA database. *BMC Genomics.* 2008; 9: 457. [CrossRef]
72. Katz LS, Humphrey JC, Conley AB, *et al.* Neisseria Base: a comparative genomics database for Neisseria meningitidis. *Database (Oxford).* 2011; 2011: bar035.
73. Spratt BG. Resistance to antibiotics mediated by target alterations. *Science.* 1994; 264(5157): 388-93. [CrossRef]
74. Tam JP, Lu YA, Yang JL, Chiu KW. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc Natl Acad Sci USA.* 1999; 96(16): 8913-8. [CrossRef]
75. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology.* 2009; 155(Pt 3): 733-40. [CrossRef]
76. Deng Y, Lu Z, Bi H, Lu F, Zhang C, Bie X. Isolation and characterization of peptide antibiotics LI-F04 and polymyxin B6 produced by Paenibacillus polymyxa strain JSa-9. *Peptides.* 2011; 32(9): 1917-23. [CrossRef]
77. Galperin MY, Koonin EV. Searching for drug targets in microbial genomes. *Curr Opin Biotechnol.* 1999; 10(6): 571-8. [CrossRef]
78. Navarro Velázquez PA. Mechanism of antibiotic action. *Rev Fac Odontol Tucuman.* 1978; (12): 53-7.
79. Opperman T, Ling LL, Moir DT. Microbial pathogen genomes - new strategies for identifying therapeutic and vaccine targets. *Expert Opin Ther Targets.* 2003; 7(4): 469-73. [CrossRef]
80. Smith DR. Microbial pathogen genomes--new strategies for identifying therapeutics and vaccine targets. *Trends Biotechnol.* 1996; 14(8): 290-3. [CrossRef]
81. Sigrist CJ, Cerutti L, de Castro E, *et al.* PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res.* 2010; 38(Suppl. 1): D161-6. [CrossRef]
82. Hulo N, Bairoch A, Bulliard V, *et al.* The PROSITE database. *Nucleic Acids Res.* 2006; 34(Suppl. 1): D227-30. [CrossRef]
83. Pietrokovski S, Henikoff JG, Henikoff S. The Blocks database-a system for protein classification. *Nucleic Acids Res.* 1996; 24(1): 197-200. [CrossRef]
84. Punta M, Coggill PC, Eberhardt RY, *et al.* The Pfam protein families database. *Nucleic Acids Res.* 2012; 40(Suppl. 1): D290-301. [CrossRef]
85. Finn RD, Mistry J, Tate J, *et al.* The Pfam protein families database. *Nucleic Acids Res.* 2010; 38(Suppl. 1): D211-22. [CrossRef]
86. Bateman A, Birney E, Cerruti L, *et al.* The Pfam protein families database. *Nucleic Acids Res.* 2002; 30(1): 276-80. [CrossRef]
87. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. *Nucleic Acids Res.* 2000; 28(1): 263-6. [CrossRef]
88. Hunter S, Jones P, Mitchell A, *et al.* InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res.* 2012; 40(Suppl. 1): D306-12. [CrossRef]
89. Hunter S, Apweiler R, Attwood TK, *et al.* InterPro: the integrative protein signature database. *Nucleic Acids Res.* 2009; 37(Suppl. 1): D211-5. [CrossRef]
90. Mulder NJ, Kersey P, Pruess M, Apweiler R. In silico characterization of proteins: UniProt, InterPro and Integr8. *Mol Biotechnol.* 2008; 38(2): 165-77. [CrossRef]
91. Mulder NJ, Apweiler R, Attwood TK, *et al.* New developments in the InterPro database. *Nucleic Acids Res.* 2007; 35(Suppl. 1): D224-8. [CrossRef]
92. Mulder NJ, Apweiler R, Attwood TK, *et al.* InterPro, progress and status in 2005. *Nucleic Acids Res.* 2005; 33(Suppl. 1): D201-5. [CrossRef]
93. Apweiler R, Attwood TK, Bairoch A, *et al.* InterPro--an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics.* 2000; 16(12): 1145-50. [CrossRef]
94. Zhang CT, Zhang R. Gene essentiality analysis based on DEG, a database of essential genes. *Methods Mol Biol.* 2008; 416: 391-400. [CrossRef]

95. Lehner KR, Stone MM, Farber RA, Petes TD. Ninety-six haploid yeast strains with individual disruptions of open reading frames between YOR097C and YOR192C, constructed for the Saccharomyces genome deletion project, have an additional mutation in the mismatch repair gene MSH3. *Genetics*. 2007; 177(3): 1951-3. [CrossRef]
96. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA*. 1998; 95(25): 14863-8. [CrossRef]
97. Frosch M, Reidl J, Vogel U.Genomics in infectious diseases: approaching the pathogens. *Trends Microbiol*. 1998; 6(9): 346-9.
98. Rokem JS, Lantz AE, Nielsen J.Systems biology of antibiotic production by microorganisms. *Nat Prod Rep*. 2007; 24(6): 1262-87. [CrossRef]
99. Sigurdsson G, Fleming RM, Heinken A, Thiele I. A systems biology approach to drug targets in Pseudomonas aeruginosa biofilm. *PLoS One*. 2012; 7(4): 34337. [CrossRef]
100. Loeb JA. A human systems biology approach to discover new drug targets in epilepsy. *Epilepsia*. 2010; 51(Suppl. 3): 171-7. [CrossRef]
101. Huang S. Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery. *J Mol Med (Berl)*. 1999; 77(6): 469-80. [CrossRef]
102. Rush TS 3rd, Grant JA, Mosyak L, Nicholls A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J Med Chem*. 2005; 48(5): 1489-95. [CrossRef]
103. Ray MK, Connerton IF, Griffiths DE. DNA sequence analysis of the Olir2-76 and Ossr1-92 alleles of the Oli-2 region of the yeast Saccharomyces cerevisiae. Analysis of related amino-acid substitutions and protein-antibiotic interaction. *Biochim Biophys Acta*. 1988; 951(1): 213-9. [CrossRef]
104. Otani T. Conformation studies on and assessment by spectral analysis of the protein-chromophore interaction of the macromolecular antitumor antibiotic C-1027. *J Antibiot (Tokyo)*. 1993; 46(5): 791-802. [CrossRef]
105. Zapp ML, Stern S, Green MR. Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell*. 1993; 74(6): 969-78. [CrossRef]
106. Reinert KE, Stutter E, Schweiss H. Aspects of specific protein-DNA interaction; multi-mode binding of the oligopeptide antibiotic netropsin to (A.T)-rich DNA segments. *Nucleic Acids Res*. 1979; 7(5): 1375-92. [CrossRef]
107. Reinert KE. Aspects of specific DNA-protein interaction; local bending of DNA molecules by in-register binding of the oligopeptide antibiotic distamycin. *Biophys Chem*. 1981; 13(1): 1-14. [CrossRef]
108. Beste DJ, McFadden J. Systems biology of the metabolism of Mycobacterium tuberculosis. *Biochem Soc Trans*. 2010; 38(5): 1286-9. [CrossRef]
109. Young D, Stark J, Kirschner D. Systems biology of persistent infection: tuberculosis as a case study. *Nat Rev Microbiol*. 2008; 6(7): 520-8. [CrossRef]
110. Chandra N, Kumar D, Rao K. Systems biology of tuberculosis. *Tuberculosis (Edinb)*. 2011; 91(5): 487-96. [CrossRef]
111. Franke R, Müller M, Wundrack N, *et al*. Host-pathogen systems biology: logical modelling of hepatocyte growth factor and Helicobacter pylori induced c-Met signal transduction. *BMC Syst Biol*. 2008; 2: 4. [CrossRef]
112. Lo Conte L, Ailey B, Hubbard TJ, Brenner SE, Murzin AG, Chothia C. SCOP: a structural classification of proteins database. *Nucleic Acids Res*. 2000; 28(1): 257-9. [CrossRef]
113. Lo Conte L, Brenner SE, Hubbard TJ, Chothia C, Murzin AG. SCOP database in 2002: refinements accommodate structural genomics. *Nucleic Acids Res*. 2002; 30(1): 264-7. [CrossRef]
114. Laskowski RA. PDBsum new things. *Nucleic Acids Res*. 2009; 37(Suppl. 1): D355-9. [CrossRef]
115. Pieper U, Eswar N, Webb BM, *et al*. MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res*. 2009; 37(Suppl. 1): D347-54. [CrossRef]
116. Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D complex: a structural classification of protein complexes. *PLoS Comput Biol*. 2006; 2(11): 155. [CrossRef]
117. Walkinshaw MD. Protein targets for structure-based drug design. *Med Res Rev*. 1992; 12(4): 317-72. [CrossRef]
118. Verlinde CL, Hol WG. Structure-based drug design: progress, results and challenges. *Structure*. 1994; 2(7): 577-87. [CrossRef]
119. Blundell TL. Structure-based drug design. *Nature*. 1996; 384(Suppl. 6604): 23-6.
120. Marrone TJ, Briggs JM, McCammon JA. Structure-based drug design: computational advances. *Annu Rev Pharmacol Toxicol*. 1997; 37: 71-90. [CrossRef]
121. Shanmugam A, Natarajan J. Homology modeling and docking analyses of M. leprae Mur ligases reveals the common binding residues for structure based drug designing to eradicate leprosy. *J Mol Model*. 2012; 18(6): 2659-72. [CrossRef]
122. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. *J Mol Biol*. 1993; 233(1): 123-38. [CrossRef]
123. Grigoriev IV, Kim SH. Detection of protein fold similarity based on correlation of amino acid properties. *Proc Natl Acad Sci USA*. 1999; 96(25): 14318-23. [CrossRef]
124. Kotra LP, Vakulenko S, Mobashery S. From genes to sequences to antibiotics: prospects for future developments from microbial genomics. *Microbes Infect*. 2000; 2(6): 651-8. [CrossRef]
125. Knox C, Law V, Jewison T, *et al*. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res*. 2011; 39(Suppl. 1): D1035-41. [CrossRef]
126. Essack SY, Schellack N, Pople T, *et al*. Part III. Antibiotic supply chain and management in human health. *S Afr Med J*. 2011; 101(8 Pt 2): 562-6.