REVIEW ARTICLE Agricultural Reviews, Volume 44 Issue 2: 264-268 (June 2023)

Metagenomics: A Novel Tool for Livestock and Poultry Improvement: A Review

Sachin S. Pawar[1], Pallavi A. Mohanapure[2], Manoj P. Brahmane[3], Mukesh P. Bhendarkar[1], Avinash V. Nirmale[1], Nitin P. Kurade[1]

10.18805/ag.R-2167

ABSTRACT

A distinct mechanism of generating metagenome is metagenomics which includes sequencing the whole DNA extracted from an environmental sample, mapping it to a reference database followed by gene annotation. Advancement in sequence-based metagenomics has significantly reduced processing costs and has acquired a rapid pace. Metagenomics has been crucial in investigating "hidden" genetic characteristics and construction biotechnological processes through the discovery of novel genes, enzymes, pathways and bioactive molecules with entirely different or improved biochemical functions. Metagenomics is a fairly recent addition to the molecular toolbox and is the simplest, impartial way of challenging the adaptive ability of microbial populations. Metagenomics is beneficial in recognising the complex consortium of bacteria, protozoa, archaea, fungi, *etc*. and association amongst them resulting in higher feed utilization and productivity of animals. It allows formulating probiotics feed materials, as well as in immunomodulation in both livestock and poultry. This review emphasizes significant recent achievements in metagenomics, offers insights into the possibilities, modern-day challenges and its utility in livestock and poultry.

Key words: Metagenomics, Metagenome, Livestock, Sequencing.

The existence of DNA molecules in a cell's different position indicates the chromosomal or extrachromosomal identity such as nuclear genome (nuclear DNA), mitochondrial genome, non-chromosomal genetic elements such as viruses, plasmids, transposable elements *etc*. PCR-RAPD is the dominant approach for studying the structure of the population without specific knowledge (Sharma *et al*., 2004); single nucleotide polymorphism is used mainly for the study of heterogeneity by genotyping by PCR-RFLP which helps with economically relevant characteristics in the association field (Sahu *et al*., 2017). The latest concept continued as metagenome which is referred to as "beyond single genome". The term metagenome applies to the collective genomes of all community members possessing a pool of microorganisms. Metagenome entails all the genetic material found in a biological study consisting of multiple genomes of individual organisms. Metagenomics has been described as a culture-independent, function-based or sequence-based collective Microbial Genome Analysis present in a certain environment (Riesenfeld *et al*., 2004). The phrase metagenomics (including the specific biological areas, known as ecogenomics, microbial community genomics or environmental genomics) to study genomes from all the microbes in an environment separated from the environment and controlled *in vitro*, as contrasted to the nucleus of one organism (Handelsman *et al*., 1998). Metagenomics relates to the analysis of collective genomes of the relevant environmental community and makes it possible to obtain information about whole species of microorganisms, such as deep oceans, soil or gut ecosystem *etc*. The poultry gut comprises millions of miscellaneous microorganisms. Most of these species remain uncharacterized and represent a huge unexplored genetic and metabolic diversity pool. Efforts have been made to record gut microbial diversity in birds. Animal nutritionists are striving for combining current knowledge of rumen activity with a potential vision on how rumen microbiology and animal nutrition could be combined with metagenomics. A better understanding of the phenomenological mechanism of manipulating amino nitrogen development and absorption can enable farm animal nutritionists to significantly improve nutritional nitrogen reformation into microbial protein. A broader understanding of the mechanistic processes that modify amino nitrogen production and uptake can help livestock nutritionists enhance the overall conversion of dietary nitrogen to microbial protein. This might relay crucial details to further develop mechanistic models that explain rumen

[1]ICAR-National Institute of Abiotic Stress Management, Baramati-413 115, Maharashtra, India.
[2]ICAR-Indian Agricultural Research Institute, Pusa, New Delhi-110 012, India.
[3]ICAR-Central Institute of Fisheries Education, Mumbai-400 061, Maharashtra, India.

Corresponding Author: Sachin S. Pawar, ICAR-National Institute of Abiotic Stress Management, Baramati-413 115, Maharashtra, India. Email: sachinndri@gmail.com

How to cite this article: Pawar, S.S., Mohanapure, P.A., Brahmane, M.P., Bhendarkar, M.P., Nirmale, A.V. and Kurade, N.P. (2023). Metagenomics: A Novel Tool for Livestock and Poultry Improvement: A Review. Agricultural Reviews. 44(2): 264-268. doi: 10.18805/ag.R-2167.

Submitted: 22-01-2021 **Accepted:** 14-09-2021 **Online:** 07-10-2021

activity and analysing dietary conditions that affect the effectiveness of processing dietary nitrogen to milk protein (Firkins *et al*., 2007).

Molecular tools to study metagenomics

16S Ribosomal RNA sequencing

The occurrence of hypervariable regions within the 16S rRNA gene offers a species reliable signature sequence that is important for bacterial phase identification. The 16S rRNA sequencing is commonly used to classify species and variants in prokaryotic species and determine the phylogenetic interactions within them. The advantages of using ribosomal RNA are that many of the cells contain Ribosomes and ribosomal RNA and RNA genomes are strongly conserved in nature.

Phylogenetic oligonucleotide microarrays (PhyloChips)

PhyloChips rely entirely on standard molecules used by the rRNAs and their encoding genes for taxonomic and environmental research of microorganisms. Through this approach a sample can be analysed simultaneously with thousands of rRNA (gene) targeted samples, lending in several cases an effective diagnosis of target species. A developing application of PhyloChips is the intensely parallel study of microbial community structural feature relationships with the aid of employing *in vivo* substrate-mediated isotope labelling of rRNA (DeSantis *et al*., 2007; Loy *et al*., 2011). Such probes target all 121 demarcated prokaryotic orders and enable 8741 bacterial and archaeal taxa to be detected at almost the same time. PhyloChip technology has been used during bioterrorism monitoring, bioremediation, climate change and source detection of pathogen pollution for rapid identification of microbial biological populations (Brodie *et al*., 2007; Rastogi *et al*., 2010).

Sequencing platforms for metagenomics

Sanger sequencing or chain-terminator sequencing is one of the first innovations developed in sequencing technology. Sanger sequencing quickly became the gold standard in sequencing technology due to ease of operation and precision. However, this method has disadvantages that make it difficult to use in metagenomic sequencing. The approach of chain-terminator sequencing is biologically biased (Sorek *et al*., 2007) and the foreign DNA need to be cloned inside a bacterial vector. Sanger sequencing is an expensive and low-throughput technique. As an outcome, Sanger-based metagenomic projects are frequently confined to fosmid sequencing and bacterial artificial chromosome libraries or microbial cultures of limited variability.

Next-generation sequencing solved some of the limitations of Sanger sequencing namely, cheaper sequencing costs per base, considerably higher throughput, simpler library preparation and exclusion of cloning step. 96 sequence data (reads) per run are produced by Sanger sequencer whereas Next-Generation Sequencing (NGS) can produce between one million and billion reads per run. That is the key reason why NGS could increase the throughput dramatically compared to the traditional Sanger sequencer. With the emergence of newer sequencing technologies, the viability of metagenomics projects has also increased in recent years. These newer sequencing technologies offer cheaper, quicker and higher sequencing throughputs.

Illumina sequencing employs the method of sequence by synthesis. The sequencing process begins by ligating template DNA to an adapter chain and then to a glass flow array. Bridge amplification reveals the Template DNA, which increases each copy to nearly 1,000 copies. Using isothermal polymerase and 32 inactivated fluorescent nucleic acids, Illumina can add solitary bases to each cycle. Through base addition that reads the fluorescent tag is followed by an imaging step. Single-base integration is of considerable advantage since context-specific failures, like those induced by homopolymers, Repetitive and low-complex regions can be sequenced conveniently over and regions of repeated and low complexity are avoided. The mean error rate per sequence produced is between 1 and 2 percent. This output is at most ten times that of Sanger sequencing. Fault rates are well defined in the read, with low rates at the 5′ end, progressing to somewhat higher error rates at the 3′ end. Errors that rely on insertion/deletion are exceptionally low, with such an error margin of less than 0.01%. There are significant prejudices against G + C- or A + T-rich sequences, more likely caused by amplification processes of DNA templates (Dohm *et al*., 2008, Aird *et al*., 2011). Since Illumina's readings in prokaryotic genomes are longer than the normal perfect match repeat period (Kassai-Jáger *et al*., 2008), relatively complex metagenomes will produce completed or almost full genomes using the Illumina framework alone (Hess *et al*., 2011).

Real-time (SMRT) sequencing is lately developed sequencing technique that has a significant impact on the genomics and metagenomics fields (Metzker, 2010). This technique uses real-time single-molecule (SMRT) sequencing that is similar to single-molecule DNA sequencing. SMRT sequencing uses the zero-mode waveguide (ZMW), through a single DNA polymerase enzyme is fixed as a template for a single DNA molecule to the bottom of a ZMW. The ZMW is an illuminated amount of observation small enough to allow observation of a single DNA nucleotide (also known as a base) inserted through DNA polymerase. Each of the four bases of DNA is fused to one of four different fluorescent dyes. Once a nucleotide is inserted by DNA polymerase, which diffuses from the ZMW detection region wherein the fluorescence is no longer accessible, the fluorescent tag is cleaved off. A detector identifies the fluorescent sign of nucleotide incorporation, and the base call is produced according to the dye's resulting fluorescence. Despite high read-length, this technique is restricted by a high error rate and low coverage.

Nanopore sequencing is a fourth-generation sequencing technique that employs nanopores (biological or solid-state) with advantages of label-free, ultra-long reads, high throughput and minimal material requirements

(Feng *et al*., 2015). It employs electrophoresis to move an unknown sample nanopore system also contains an electrolyte solution and when the continuous electrical field is applied, an electrical current may be detected in the system. The nanopore dimensions and structure of the DNA or RNA that the nanopore occupies governs the intensity of density in electrical current over a nanopore surface. A single DNA and RNA molecule can be sequenced with the Nanopore sequencing without PCR amplification or chemical labelling of the sample. It has the potential to provide relatively low-cost genotyping, higher test reliability and quick sample analysis with the ability to demonstrate results in real-time (Niedringhaus *et al*., 2011; Si and Aksimentiev 2017).

Next-generation sequencing (NGS) techniques are widely recognized as the most effective tools for gene sequencing, giving a deep insight into the ecology of microbial controlled activities. NGS technologies are used for a variety of purposes, including single-gene targeted sequencing to whole-genome sequencing and shotgun metagenome sequencing. NGS is also known as high-throughput sequencing and is used to Identify numerous modern environmentally sustainable RNA samples including recovery of high-quality mRNA from environmental samples, brief half-lives of mRNA species and isolation of mRNA from other RNA species. It offers affordable access to the metatranscriptome and allows a microbial population to profile the whole-genome expression. Furthermore, the transcripts can also be explicitly quantified (Carola and Daniel 2011).

Denoising

Denoising is important for 16S metagenomics data analysis which is platform-specific *i.e*., certain systems (*e.g*., Illumina) needs less denoising than others (*e.g*. pyrosequencing). Despite being computationally expensive, denoising pyrosequencing data is important due to intrinsic errors produced by pyrosequencing that can lead to erroneous operational taxonomic units (OTUs). A technique called "flowgram clustering" eliminates troublesome readings and improves taxonomic analytical accuracy. Until now, several denoising algorithms have been developed, (Reeder and Knight, 2010; Quince *et al*., 2011; Johanna *et al*., 2013; Balzer *et al*., 2013). Denoising is very effectively performed by Amplicon-Noise (Quince *et al*., 2011), a tool using the following fundamental steps:

Noise reads filtering: Reads are truncated depending on the presence of low signal intensities.

Eliminating pyrosequencing noise: The difference between the flow grams is defined and the real sequences and their frequencies are concluded by an expectation-maximization algorithm (EM).

Extracting PCR noise: The same ideas are applied to eliminate PCR errors.

Detection and elimination of chimera: On each sequence, exact pairwise alignments are performed on all sequences of equal or greater excess, which is the set of probable parents.

Whenever a large number of sequences are lost during the denoising process, high-quality sequences result are obtained, however, the level of stringency needed to obtain this high quality has been debated (Gaspar *et al*., 2012).

Gene prediction

The process of identifying protein and RNA sequences coded on the DNA sample is gene prediction/gene calling. Gene prediction can be made on post assembly contigs on reads from the unassembled metagenome, based on the applicability and performance of the assembly (Kunin *et al*., 2008). One of the most important challenges in bioinformatics is identifying the location of protein-coding areas using computational approaches. Two classes of methods are generally adopted: similarity-based searches and ab initio prediction using gene structure as a template to detect genes (Zhou *et al*., 2004).

Applications in livestock and poultry

For millions of people around the world, the livestock industry is both an economic enterprise and a survival enterprise (Karangiya *et al*., 2016). Meat and milk produced by ruminants are important agricultural products and provide nutritionally rich food for humans. The livestock industry, however, faces major challenges due to declining natural resources and the resulting increase in the cost of production, as well as environmental effects on farming ruminants (Morgavi *et al*., 2013). The level of intake and digestibility of feed in livestock is related to methane production (Das 2018). Feed digestion and enteric methane production are essential functions that could be controlled by providing a detailed overview of the rumen microbiome. Advances in DNA sequencing methods and bioinformatics are shaping our perception of microbial diverse communities like the mammalian gastrointestinal tract. The application of these strategies to the rumen ecosystem has enabled the analysis of microbial diversity under different conditions of diet and growth. The sequencing of genomes from many bacterial and archaeal cultured rumen species provides extensive knowledge of their physiology. Microbiome research is gaining attention in livestock products, as it helps to elucidate diseases and efficiency processes. The rumen microbiota in cattle is directly associated with the digestive process of feed and accessibility of host nutrients and is considered responsible for digesting million of tons of cellulosic material worldwide to provide people with milk and meat (Hackmann and Spain, 2010). The rumen microbiota rapidly digests plant material and has recently drawn researchers interested in the cost-effective system for transforming lignocellulosic plant material into biofuel. A metagenomic study using rumen as an effective cellulose fermenter fed with switchgrass and restoration of the fiber-attached microbiome showed the rumen microbiota's ability to colonize and degrade biofuel biomass rapidly (Hess *et al*., 2011). Previous studies have related well known taxonomical groups or system composition with feed reliability or residual feed intake (Jami *et al*., 2014; Jewell *et al*., 2015; Roehe *et al*., 2016). Most of these studies used 16S rRNA sequencing as an outline of the microbiota. The gastrointestinal tract

microbial profiles of chicken and Guinea fowl was analysed using a metagenomic method to understand the microbial diversity of both avian organisms. For this analysis, DNA was collected from the chicken and Guinea fowl's gastro-intestinal environment. The region encoding hypervariable 16s rRNA was targeted to decipher the composition of microbial communities in organisms using the metagenomics approach (Palys *et al*., 1997). The efforts were made to record gut microbial diversity in birds layer White Leghorn (Gallus gallus) by evaluating datasets at Meta Genomic Rapid Annotation using Subsystem Technology (MG-RAST) (Meyer *et al*., 2008).

FUTURE PROSPECTS

Metagenomics has revolutionized research in multiple fields and has paved the way for an independent analysis of the cultivation and consumption of microbial populations existing in complex ecosystems such as livestock habitats. There are several metagenomics methods like NGS techniques now being used for the sequencing of genomes. The long-term aim of metagenomics is to rebuild the genomes of non-culturable gut microorganisms by identifying overlapping fragments in metagenomic libraries and clone-to-clone walking to assemble each chromosome (Singh *et al*., 2008). The application of metagenomics along with NGS approaches provides broad sets of data sequence and their analysis can unfold the unrevealed taxonomic and functional abundance of microbial communities in different ecosystems of livestock and poultry. Developing knowledge of the chicken gastrointestinal tract microbiome by culture-independent metagenomic evaluation has helped explain the nature and role of microbial communities in the metabolism and health of the chicken. Future work on GIT microbiota can lead to discovering the role of the microbe in mediating chicken growth under specific environmental factors including stress, nutrients availability and well-being, to determine better production quantities and efficiency. The development of metagenomic methods and bioinformatics tools would be critical for the progress of the livestock rumen and chicken gut microbiota research. Designed strategies to improve the metabolism of nutrients may be established by concentrating on functional genes in them. With advanced bioinformatics software, Strain-specific 16S rRNA gene sequencing (Eren *et al*., 2014) will track how each strain will develop under various growth and environmental conditions, then control the abundance of target stress changes by food care. It is pertinent that metagenomics will have leverage in exploring the full potential of gut microflora of ruminants and poultry.

CONCLUSION

Presently, metagenomics is evolving as a powerful and effective method for sequencing the metagenome obtained from environmental samples. The prospects for use of metagenomics in investigating microbial composition in unknown ecosystems will advance with the development and use of newer approaches and techniques and the reduction in the cost of sequencing. Analysing gut microbial diversity would help in classifying the strains that are responsible for the animal's adaptability to natural conditions. Studying gut microbial diversity in different species of livestock and poultry can help in recognizing the probiotics communities with both growth-promoting and immune response-boosting properties.

REFERENCES

Aird, D., Ross, M.G., Chen, W.S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C. and Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome Biology. 12: R18.

Balzer, S., Malde, K., Grohme, M.A. and Jonassen, I. (2013). Filtering duplicate reads from 454 pyrosequencing data. Bioinformatics. 29: 830-836.

Brodie, E.L., DeSantis, T.Z., Parker, J.P.M., Zubietta, I.X., Piceno, Y.M. and Andersen, G.L. (2007). Urban aerosols harbor diverse and dynamic bacterial populations. Proceedings of the National Academy of Sciences U.S.A. 104: 299-304.

Carola Simon, C. and Daniel, R. (2011). Metagenomic analyses: past and future trends. Applied and Environmental Microbiology. 77: 1153-1161.

Das, S.K. (2018). Impact of climate change (heat stress) on livestock: adaptation and mitigation strategies for sustainable production. Agricultural Reviews. 39: 130-136.

DeSantis, T.Z., Brodie, E.L., Moberg, J.P., Zubieta, I.X., Piceno, Y.M. and Andersen, G.L. (2007). High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microbial Ecology. 53: 371-383.

Dohm, J.C., Lottaz, C., Borodina, T. and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Research. 36: e105.

Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G. and Sogin, M.L. (2014). Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. Methods in Ecology and Evolution. 4: 1111-1119.

Feng, Y., Zhang, Y., Ying, C., Wang, D. and Du, C. (2015). Nanopore-based fourth-generation DNA sequencing technology. Genomics Proteomics Bioinformatics. 13: 4-16.

Firkins, J.L., Yu, Z. and Morrison, M. (2007). Ruminal nitrogen metabolism: Perspectives for integration of microbiology and nutrition for dairy. Journal of Dairy Science. 90: E1-E16.

Gaspar, J.M. and Thomas, W. (2012). The consequences of denoising marker-based metagenomic data. BMC Proceedings. 6: p11.

Hackmann, T.J. and Spain, J.N. (2010). Invited review: Ruminant ecology and evolution: Perspectives useful to ruminant livestock research and production. Journal of Dairy Science. 93: 1320-1334.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J. and Goodman, R.M. (1998). Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. Chemistry and Biology. 5: R245-R249.

Hess, M., Sczyrba, A., Egan, R., Kim, T.W., Chokhawala, H., *et al.* (2011). Metagenomic discovery of biomass degrading genes and genomes from cow rumen. Science. 331: 463-467.

Jami, E., White, B.A. and Mizrahi, I. (2014). Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. PLoS One. 9: e85423.

Jewell, K.A., McCormick, C.A., Odt, C.L., Weimer, P.J. and Suen, G. (2014). Ruminal bacterial community composition in dairy cows is dynamic over the course of two lactations and correlates with feed efficiency. Applied and Environmental Microbiology. 81: 4697-4710.

Johanna Brodin, J., Mild, M., Hedskog, C., Sherwood, E., Leitner, T., Andersson, B. and Albert, J. (2013). PCR-induced transitions are the major source of error in cleaned ultra-deep pyrosequencing data. PLoS One. 8: e70388-e70388.

Karangiya, V.K., Savsani, H.H. and Ribadiya, N.K. (2016). Use of densified complete feed blocks as ruminant feed for sustainable livestock production: A review. Agricultural Reviews. 37: 141-147.

Kassai-Jáger, E., Ortutay, C., Tóth, G., Vellai, T. and Gáspári, Z. (2008). Distribution and evolution of short tandem repeats in closely related bacterial genomes. Gene. 410: 18-25.

Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K. and Hugenholtz, P. (2008). A bioinformatician's guide to metagenomics. Microbiology and Molecular Biology Reviews. 72: 557-578.

Loy, A., Pester, M. and Steger, D. (2011). Phylogenetic microarrays for cultivation-independent identification and metabolic characterization of microorganisms in complex samples. Methods Molecular Biology. 688: 187-206.

Metzker, M.L. (2010). Sequencing technologies- The next generation. Nature Reviews Genetics. 11: 31-46.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass,. EM., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. and Edwards, R.A. (2008). The metagenomics RAST server- a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics. 9: 386.

Morgavi, D.P., Kelly, W.J., Janssen, P.H. and Attwood, G.T. (2013). Rumen microbial (meta) genomics and its application to ruminant production. Animal. 1: 184-201.

Niedringhaus, T.P., Milanova, D., Kerby, M.B., Snyder, M.P. and Barron, A.E. (2011). Landscape of next-generation sequencing technologies. Analytical Chemistry. 83: 4327-4341.

Palys, T., Nakamura, L.K. and Cohan, F.M. (1997). Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. International Journal of Systematic and Evolutionary Microbiology. 47: 1145-1156.

Quince, C., Lanzen, A., Davenport, R.J. and Turnbaugh, P.J. (2011). Removing noise from pyrosequenced amplicons. BMC Bioinformatics. 12: 38.

Rastogi, G., Osman, S., Vaishampayan, P.A., Andersen, G.L., Stetler, L.D. and Sani, R.K. (2010). Microbial diversity in uranium mining-impacted soils as revealed by high-density 16S microarray and clone library. Microbial Ecology. 59: 94-108.

Reeder, J. and Knight, R. (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. Nature Methods. 7: 668-669.

Riesenfeld, C.S., Schloss, P.D. and Handelsman, J. (2004). Metagenomics: genomic analysis of microbial communities. Annual Review Genetics. 38: 525552.

Roehe, R., Dewhurst, R.J., Duthie, C.A., Rooke, J.A., McKain, N., Ross, D.W., Hyslop, J.J., Waterhouse, A., Freeman, T.C., Watson, M. and Wallace, R.J. (2016). Bovine host genetic variation infuences rumen microbial methane production with best selection criterion for low methane emitting and efciently feed converting hosts based on metagenomic gene abundance. PLoS Genetics. 12: e1005846.

Sahu, A.R., Jeichitra, V., Rajendran, R. and Raja, A. (2017). Polymorphism in exon 3 of myostatin (MSTN) gene and its association with growth traits in Indian sheep breeds. Small Ruminant Research. 149: 81-84.

Sharma, A.K., Bhushan, B., Kumar, S., Kumar, P., Sharma, A. and Kumar, S. (2004). Molecular characterization of Rathi and Tharparkar indigenous cattle (*Bos indicus*) breeds by RAPD-PCR. Asian-Australasian Journal of Animal Sciences. 17: 1204-1209.

Si, W. and Aksimentiev, A. (2017). Nanopore sensing of protein folding. ACS Nano. 11: 7091-7100.

Singh, B., Bhat, T.K., Kurade, N.P. and Sharma, O.P. (2008). Metagenomics in animal gastrointestinal ecosystem: A microbiological and biotechnological perspective. Indian Journal of Microbiology. 48: 216-227.

Sorek, R., Zhu, Y., Creevey, C.J., Francino, M.P., Bork, P., Rubin, E.M. (2007). Genome-wide experimental determination of barriers to horizontal gene transfer. Science. 318: 1449-1452.

Zhuo, W., Chen, Y. and Li, Y. (2004). A brief review of computational gene prediction methods. Genomics Proteomics Bioinformatics. 2: 216-221.