THINK INDIA JOURNAL

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

In-Silico Gene Expression Analysis of Waterborne Pathogens- A Metagenomics Study

Amir Dar, Joginder Singh and Vikas Kaushik*

Department of Bioinformatics, School of BioEngineering and BioSciences, Lovely Professional University, Punjab, India

*Corresponding Author E-mail: vikas.14664@lpu.co.in

Abstract

Water is essential for the survival, the supply for the user should be safe to use and contamination-free, because contaminated water carries a number of pathogenic bacteria's which further lead to millions of deaths every year. Water used for different purposes should be supplied as safe and pure as we can. Severe microbial infections and disease occur due to ingestion of contaminated water due to which people are getting infected. Water born diseases are not only affecting the developing countries but also affecting the developed countries according to the WHO report. In the USA, it has been reported and analyzed that each year 560,000 people suffer severe water-borne disease and 7.1M from mild and moderate infections which lead to the death of 12,000 every year. There is a number of bacterial water born diseases which includes tuberculosis, typhoid, cholera etc. Water born diseases are growing rapidly due to change in the genomics for the adaptability of an organism. Bacteria's show some mutations in their genomics as the environmental conditions change. We have chosen four water-borne bacteria's which are known globally for their diseases, and some are under research because of the unavailability of suitable treatment. World health organization has shown the reports of death and infected individuals. Water born disease-account for 3.6% of DALY (disability-adjusted life year) and cause 1.5 million deaths every year. The reason for this is the lack of availability of safe drinking water, sanitation, hygiene and loss of awareness. This study highlights the characterization of the gene of all four species and their life cycle along with ecological studies.

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

Introduction

A report by WHO has given an indication that there is very less use of freshwater in developing and the current developed nations. It also indicated that the usage of water across the world is getting increased three times than the current growth of population, with the rapid increase in population the pollution and the waste is increasing which is depleting the freshwater resources and directly affecting the whole population, agriculture and also affecting the economy of the nations. Current research and analysis are suggesting that huge amount of waste is being generated especially from the cities which needs proper disposal and proper methods of disposal should be implemented to avoid disposal of generated waste into water resources. The public is getting affected because of some measure disposal of toxins, chemicals, metals which leads to the rise of new strains from the water and give rise to some fatal pathogenic agents. The impact of chemicals on the non-pathogenic strains or pathogenic strains imposes an environmental change for bacteria which forces them to adapt themselves and led to some unexpected mutations in genomics and increases the level of virulence. It also gives rise to some other issues like designing and production of new drug molecules because of change in genomics or because of drug resistance [1]. Table-1 and table-2 demonstrate some bacterial and viral disease.

Table1 List of the pathogenic bacterium

Disease	**Microbial agent**	**Source of agents**	**Symptoms**
Botulism	*Clostridium botulinum*	Contaminated drinking water	Blurred or double vision, dry mouth, vomiting, diarrhoea, muscle weakness, respiratory failure causes death
Cholera	*Vibrio cholera*	Contaminated water containing the bacterium	Watery diarrhoea, cramps, nausea, Hypovolemic shock
E. coli infection	*Certain strains of E. coli*	Contaminated water	Frequent diarrhoea leads to dehydration
Dysentery	*Shigella, salmonella*	Contaminated water	Excretion of faeces or mucus with blood
Salmonellosis	*Salmonella*	Contaminated drinking water, commonly foodborne	Abdominal cramps, vomiting and fever
Typhoid	*Salmonella typhi*	Contaminated water containing faecal matter	Diarrhoea, enlargement in liver and spleen causes death
Leptospirosis	*Leptospira*	Water contaminated by animal urine	Renal failure, liver damage, initial

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

			symptoms are flu-like
Campylobacteriosis	*Campylobacter jejuni*	Contaminated drinking water with faeces	High fever lasts for 2 days
Dermatitis	*Pseudomonas*	Dermal contact with water	Infection of skin, inflammation

Table 2 List of waterborne viral diseases

Disease	**Viral agents**	**Source of agent**	**Symptoms**
Severe acute respiratory syndrome	Coronavirus	Improperly treated water	Cough, sore throat, gastrointestinal symptoms
Poliomyelitis	Poliovirus	Enter water through infected individuals faecal matter	Headache, fever, spastic paralysis
Polyomavirus infection	JC virus and BK virus	Widespread, found in water	Respiratory infections affect kidney
Hepatitis A	Hepatitis A virus	Contaminated water and food	Depression, weight loss, abdominal fever pain, jaundice
Diarrheal illness	Norovirus	Contaminated water associated with ships	Diarrhoea

Use of treated water has shown good benefits in irrigation because of the presence of nitrogen and salinity, which reduces the production costs. So it has been proved that the use of treated water for irrigation is better than groundwater. Pathogenic microbes found in wastewater can be divided into three groups that are viruses, bacteria, protozoans or helminths. Majority of the pathogenic bacterium is enteric and gets excreted from one host to water through faecal matter and because of usage of wastewater gets ingested into another host. A bacterium is the most commonly found in water and almost are enteric with some exceptions like *Mycobacterium tuberculosis*. Some of the diseases like gastrointestinal infection caused by bacteria, other diseases like diarrhoea by *Vibrio cholera*, salmonellosis by *Salmonella*, dysentery by *Shigella*, stomach ulcers by *Helicobacter pylori* and typhoid by *Salmonella typhii*. Some of the salmonella strains were found in some food items manufactured with wastewater. Usage of Contaminated water in food items lead to some serious problems because of the presence of pathogenic bacterium like clostridium, staphylococcus, and salmonella [1]. *M. tuberculosis* found in wastewater cause problems due to contact with wastewater leads to ulcerous lesions on the body.

The analysis of waterborne pathogenic bacteria's includes four species, i.e. *E. coli*, *M. tuberculosis*, cholera and *M. marinum*. All these bacteria's are very pathogenic especially

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

Mycobacterium which causes tuberculosis. The treatment with higher potency is still not available for tuberculosis. *E. coli* is having hundreds of strains; some are pathogenic, and some are non-pathogenic. Some strains live in human and animal intestines which are non-pathogenic, the strain 0157:H7 of *E. coli* can produce a harmful toxin that can cause severe illness which leads to death and is found in small cattle farms. It is found in the grounded meat and contaminated water. Bacteria vibrio cholera causes cholera leads to diarrhoea, vomiting, leg cramps, the rapid loss of fluids from the body leads to dehydration and shock. It is typically ingested with contaminated water. The contamination occurs when untreated sewage is supplied to the users. Cholera rarely spreads through person to person; the main source is contaminated water. *M. marinum* infection (aquarium granuloma) caused by *M. marinum*. It mostly infects immune-compromised individual's cause of exposure to swimming pools, or aquariums leads to lesions on elbows and knees that may be painless or painful on the other hand the lesions on the hands appear because of exposure to aquariums. The objectives of our study also related the die removal organism and to find the alternative microorganism for dye removal.

Tuberculosis caused by *M. tuberculosis* having different strains. It is mentioned among the top cause of death worldwide and leads to the death of many people across the globe. Scientists have failed to tackle this disease which gives rise to another issue of multiple drug resistance. Due to the lack of availability of efficient treatment and use of multiple drugs, allowed different strains to mutate. The WHO report has reported that 92 countries have patients with MDR and is predominantly higher in European and central Asian countries. The treatment rate varies between 36%-79%.The guidelines for treatment from WHO is of 20 months with a combination of four drugs against *M. tuberculosis* to which the microorganism shows sensitivity. Due to the rise of multiple drug resistance European countries and the whole WHO region have a plan to prevent drug resistance. The resolution passed by the WHO European union has 61 sessions with proper strategy and directions [2-4].

M. tuberculosis has different strains namely *M. Africana*, *M. bovis*, *M. pinnipedii*, *M. microti* are the strains causing tuberculosis in domestic and wild mammals. M.Tb has evolved as pathogenic to humans 70,000 years ago then it spreads with the migration of humans. It is now believed that it has evolved from *Mycobacterium* and still can isolate it from immunocompromised individuals of east Africa. One of the main structural strength of M.Tb is its cell wall which provides impermeability of chemicals and drug compounds and supports the virulence. Protein secretion plays a major role in the pathogenesis they need to remain in

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

control. In M.Tb dormancy regulon control more than 50 genes which are responsible for hypoxic response [5, 6]. M Tb shows the sensitivity against change in the environment in macrophages which gets separated and identified by a decrease in levels of oxygen and depletion available nutrients and M Tb shows the response by activating a dormancy state where the bacteria stops multiply and reduction in regulation of metabolism and shows the activation of anaerobic metabolic activities. These proteins can change from inactive to active conditions at any time to regulate metabolic activities. M Tb remains in dormant conditions inside the tissues of a host, and when the cascade gets the stimulation, it gets active. The infection occurs when the *Mycobacterium* gets transferred from a pulmonary patient and enters into the host alveolar sac. It gets attacked by alveolar macrophages but if it survives the first line of defence. It starts to grow actively inside macrophages and also enter into surrounding cells like endo and epithelial cells. In one week time, it divides and reaches to exponential growth rate, and during the infection, it can diffuse to other organs as well. The treatment to M Tb requires the use of biological samples for the detection of bacteria using advanced biological assays. Non-pulmonary patients it is very difficult to treat and identify the specific sample because of the low presence of bacteria and can be detected from urine sample and stools [7, 8].

WHO treatment and diagnosis approach is to implement quick treatments which can be directly observed within a short course of time. It involves some of the main parameters. (1) The government of a particular country should be highly serious and committed to treating the infection. (2) Microscopic observations for every case to detect drug susceptibility. (3) The standard treatment for the patients by a well trained professional. (4) the drug supply to the patient should not get disturbed and continues direct observation. (5). Standard records of a patient should be maintained properly. These regimes have reduced the number of patients and have shown huge improvement, but this approach has not shown any decline in drug-resistant patients. There are a number of drugs available, and some are under pipeline which may reduce the drug resistance and can have some decline in the increase in drug resistance patients. Agonists and antagonist used in drug designing for M Tb have screened huge class of different compounds

Methodology and Results

Gene expression analysis of waterborne pathogens

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

The GSE numbers for the selected bacteria which cause disease and affect the human health and the metabolic cycles very severely. Table-3 explains the gene regulation in M.Tb.

Table 3. Represent three up-regulated and three down-regulated genes

ID	Adj.P.val	P.Value	T	B	logFC
1759996_s_at	0.183	0.0000179	7.67	-0.9	4.405
1766235_s_at	0.511	0.0001294	6.04	-1.39	2.095
1766570_s_at	0.511	0.0001513	5.92	-1.43	3.695
1763962_s_at	1	0.0273392	-2.59	-3.5	-1.35
1764719_s_at	1	0.0273496	2.58	-3.5	1.808
1760341_s_at	1	0.0275967	-2.58	-3.51	-0.808

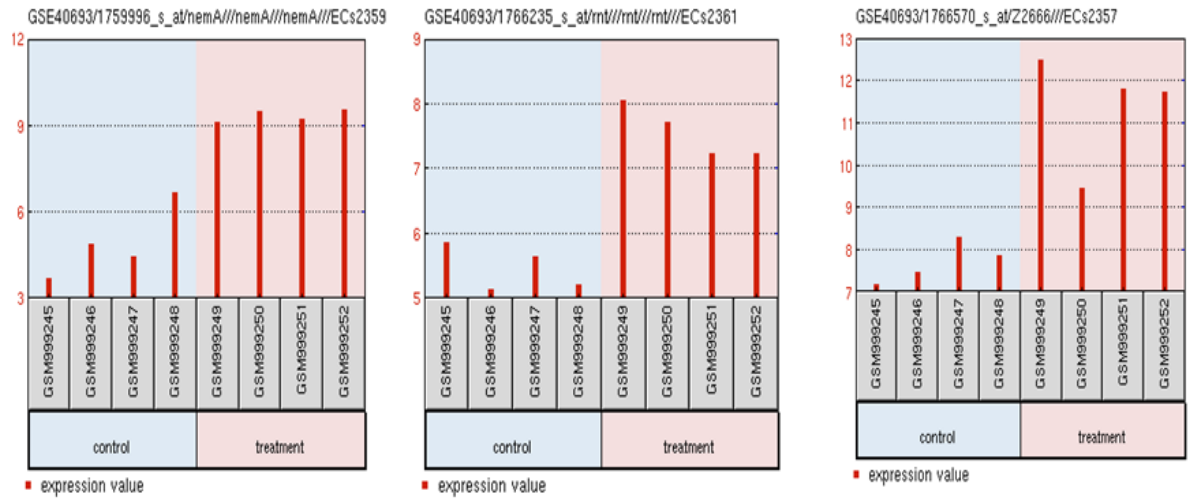


Figure 1 Upregulated gene expression between control and treatment under different conditions

Figure-1 is showing the gene expression in *E. coli* strain which is pathogenic. Some genes are expressing at higher levels when compared with the control samples. The treated genes in above figure GSM999252, GSM999250 are expressing little higher than the other two genes while as the gene expression of control samples is very low in comparison to treated genes. Above three graphs are showing the expression of upregulated genes in *E. coli* and treatment control gene expression comparison analysis. In GSE40693, the highest expression shown by GSM999249 and other samples are also showing relatively suitable gene expression while as the expression level in control sample genes is low.

Downregulated genes

The figure-2 below is showing the expression level is higher in control samples relative to treated samples. The genes of the control sample in the below graphs are clearly showing the expression of genes which is higher in control genes.

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

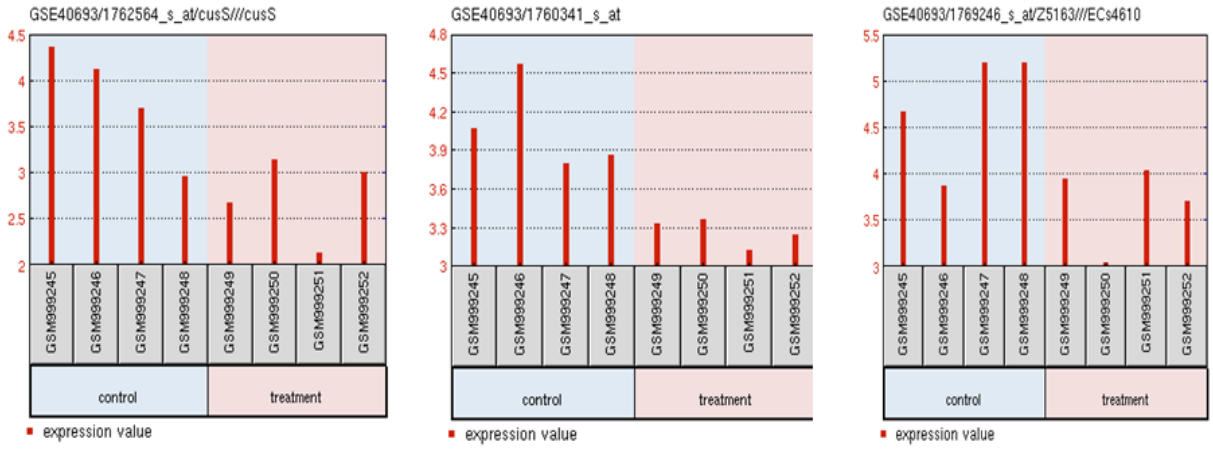


Figure 2 Down-regulated gene expression

Value distribution presentation of control and treated

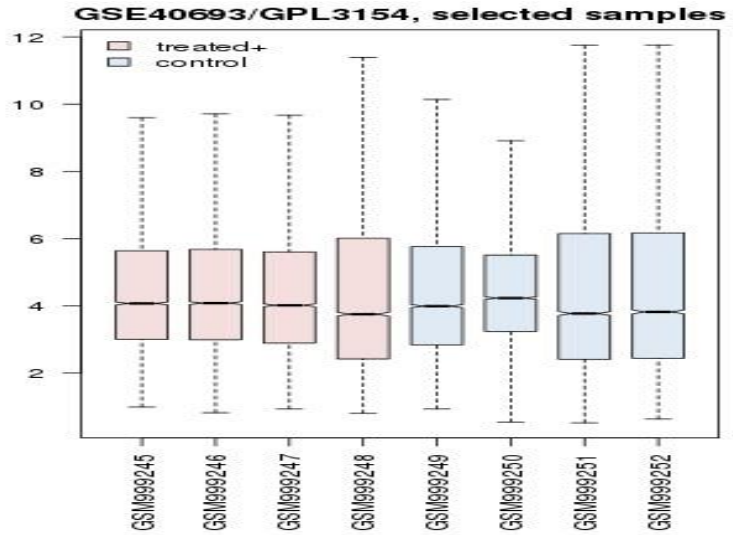


Figure 3 Showing distribution of different genes

In *E. coli* strain gene expression, figure-3 is indicating the value distribution of genes, some are distributed at higher levels and some at lower levels. We have mentioned six genes, three upregulated and three down-regulated showing different gene expression. The genes in the up-regulated system, when analysed with the control system usually the gene expression in control genes, is relatively lower than the treated samples. While as the gene expression in downregulated genes is relatively low and the control samples are showing higher gene expression. From different calculation and logfc values, the upregulation and downregulated

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

genes can be identified, and the expression level can be observed separately by observing the bar graphs.

Table 4 Represents logfc and other values for one up-regulated and down-regulated genes

ID	Adj.P.val	P.Value	T	B	logFC
RY2624c-2950487-29	1.42e-12	9.80e-16	80.7	26.6	3.312
PE_PGRS-336558-33	1.22e-07	7.06e-09	17.1	10.5	0.836

The above table-4 showing the gene expression of *M. tuberculosis* genes by analysing the samples against control. Some of the genes show less value of distribution and some show higher distribution in a sample. The expression is tested, and from the graph of different genes, it can be easily evaluated which gene among the sample is expressing at higher levels. In the above figure GSM1070317, GSM1070316, GSM1070315 are distributed in higher among than the other respective genes in control and treatment

Discussion

Few strains trigger diarrhoea which may be very dangerous the particular strain of *E. coli* O157:H7, it produces some powerful toxin that damages the small intestine lining which leads to bloody diarrhoea. Some can recover from *E. coli* infection within days, and some develop a life-threatening form of kidney failure called haemolytic uremic syndrome. As far as treatment is concerned, there are some measures taken by the scientists, but no vaccine has yet developed.

Currently, there is no unified technique available for the detection of all the pathogens in the water because of their variations in morphology and the amount, a small number cannot be detected easily due to which still water contamination is checked by culture technique only, and this technique is particular for a particular microorganism. But the non-culturable state such as *E. coli*, helicobacter, V. cholera leads to false result from the culture methods. Molecular methods are highly specific and also provide the phylogenetic relations of a particular species. Instead of using *E. coli* host different alternative indicators can be used for the detection. Molecular methods are highly sensitive and better suited for detection such as nucleic acid amplification.

Further PCR can be used for *E. coli* with some limitations as PCR is unable to differentiate between viable and non-viable [1]. Methods like PCR and micro-array can be coupled

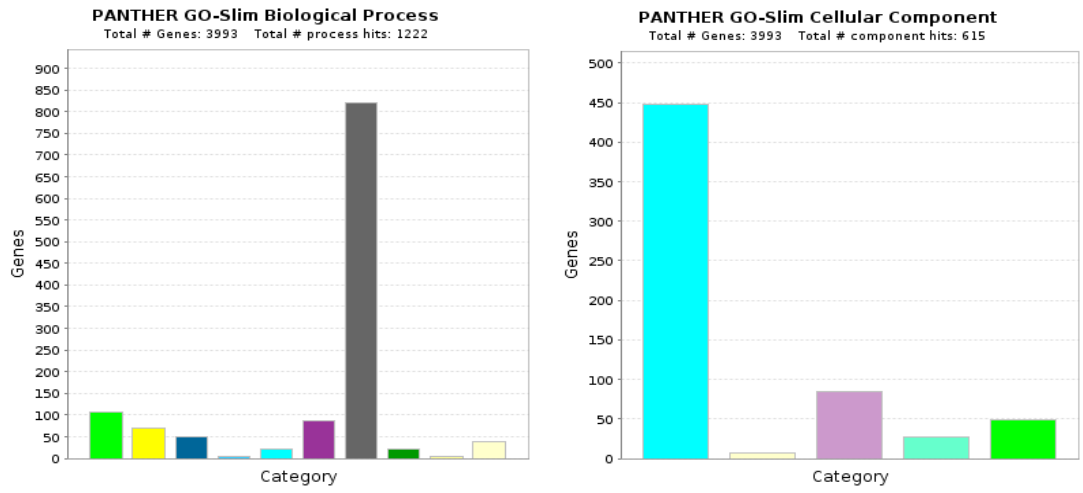
ISSN: 0971-1260
Vol-22-Issue-17-September-2019

together to increase the signal sensitivity for the detection of pathogens. Signal sensitivity increased up to 106 folds. PCR -microarray might lack sensitivity because of less availability of sample but here the amplification followed by PCR. Other methods can be used like sensor detection, pyrosequencing, immunological methods, FISH etc. Pathogen indicators need to be modified continuously because the involvement of new pathogenic bacterium is increasing in water [9].

M. tuberculosis

Mycobacterium exists from the Jurassic period, and now they found in every habitat and ecosystem, ubiquitous except polar regions, now it has a total of 169 distinct species. *M. tuberculosis* is a causative agent for tuberculosis having no defined environment reservoirs. They are strict intracellular pathogens to humans and animals. It is estimated that 2 billion people, or one-fourth of the world's population, are infected with *M. tuberculosis*. This pathogen produces nearly 9 million new infections and 1.5 million deaths every year (one-quarter of which are deaths of TB patients co-infected with human immunodeficiency virus [HIV]), ranking second, only to HIV, as the leading cause of death from an infectious agent. TB is a dangerous health and economic problem in developed and developing countries and also high-income countries due to TB-HIV co-infection, which leads to multiple drug-resistant tuberculosis strains. Phenotype based methods for the identification of *M. tuberculosis* rely on strains phenotypic characteristics which include colony morphology, susceptibility to antimicrobial agents, biochemical and serological reactivity [10]. Figure 4,5 and 6 demonstrates gene ontology, pathway details and gene expression in M.Tb.

Gene ontology



ISSN: 0971-1260
Vol-22-Issue-17-September-2019

Figure 4 Cellular components and genes in *M. tuberculosis* using gene ontology PANTHER GO analysis

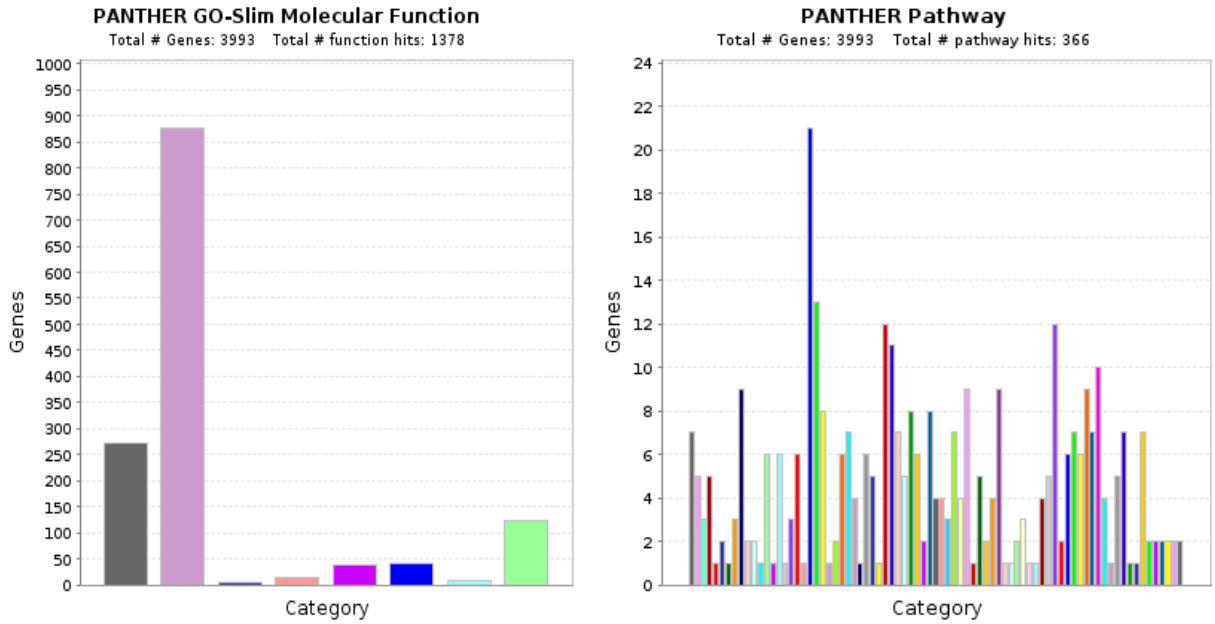


Figure 5 A. Showing molecular function frequency with genomic frequency in *M. tuberculosis*. B. Showing number of genes and the number of pathways in *M. tuberculosis*

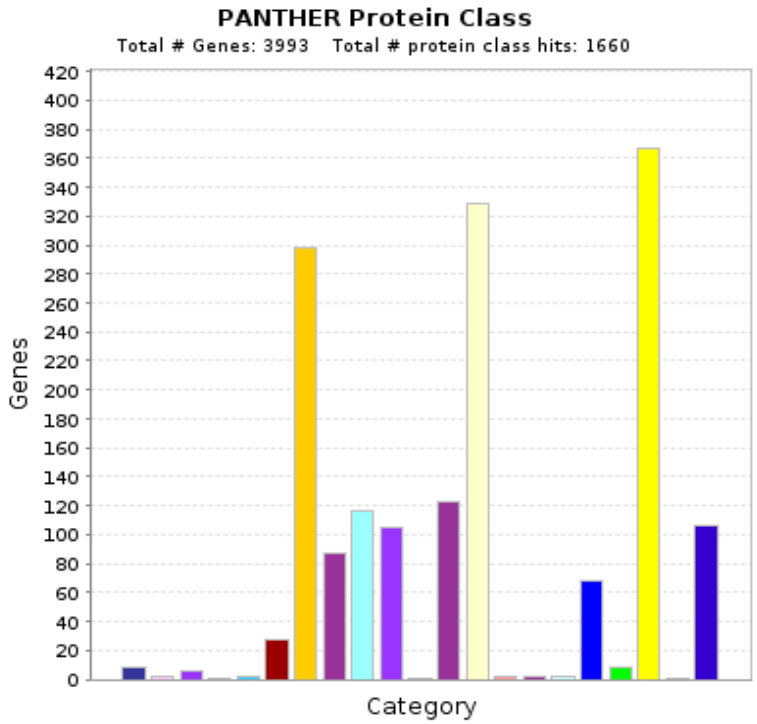


Figure 6 Showing number of genes and the classes of proteins produced by these genes in *M. tuberculosis*

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

Vibrio cholera

Vibrio cholera form biofilms like aggregates during infection which can be alternatively critical for pathogenesis and also disease transmission. Vibrio is the main causing agent for a waterborne disease called diarrhoea. It is motile and gram-negative bacteria having some serogroup and families. It is endemic to some regions and the disease there is seasonal. *Vibrio cholera* O1 in non-endemic is because of poor sanitation. It can be divided into two biotypes classical and El Tor, which differ in the severity of clinical symptoms and the expression and regulation of major virulence factors. Humans have experienced seven cholera pandemics. The seventh and current pandemic is characterized by the predominance of the O1 serogroup of the El Tor biotype, with periodic emergence of serogroup O139, which originated from the El Tor biotype and exhibited a new lipopolysaccharide (LPS) and a capsule [11, 12]. There are a number of evidence which suggests the capacity of V.cholera for the development of biofilms is particular to intestinal colonization [10]. The elements required for V.cholera adoption lifestyle are expressed only during infections. Some biofilm genes like vpsa, rbma express at a higher level when compared with LB medium expression. The in vivo expression was much higher. Deletion of rmbA significantly diminished intestinal colonization, while inactivation of rbmc and/bap1 did not affect. The research suggests the formation of biofilms during infection consist of cell aggregates that do not progress beyond the RbmA-dependent clustering stage or that other factors can functionally replace Bap1 and RbmC for biofilms formed in vivo [13,14].

M. marinum

A genetically close relative of *M. tuberculosis*. It causes chronic disease in zebrafish in a very unique dose-dependent fashion. As the dosage to the fishes increases the more quickly the fishes die. The infections are very rare and very less studied, resistance to antibiotics and chemotherapy helps in the characterization of mycobacterial infections. The treatment may include traditional which may take little longer to eradicate the infection. There is an increase in ant tubercular multiple drug resistance in case of sporotrichosis lymphangitic dissemination, due to use of multiple drugs for treatment leads to MDR [15]. The infection can be treated with multiple drug combination showing positive effects minocycline, clarithromycin, and ethambutol. In addition, there is a need for radical synovectomy of the lesion. The treatment of *M. marinum* clinically varies because it triggers the infection at some optimal growth conditions like temperature.

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

Conclusion

This work demonstrates the gene expression of *M. tuberculosis* genes by analysing the samples against control. Some of the genes showed less value of distribution and some showed the higher distribution in a sample. The expression is tested, from the results, it can be easily evaluated which gene among the sample is expressing at higher levels and which one is expressing at a lower level in control and treatment

References

1. Askun, T., Tumen, G., Satil, F., & Ates, M. (2009). In vitro activity of methanol extracts of plants used as spices against *M. tuberculosis* and other bacteria. *Food Chemistry*, *116*(1), 289-294.
2. Atassi, M. Z., & Smith, J. A. (1978). A proposal for the nomenclature of antigenic sites in peptides and proteins. *Immunochemistry*, *15*(8), 609-610.
3. Balganesh, T. S., & Furr, B. J. A. (2007). Molecular Approaches to Target Discovery:-Evaluating Targets for Antituberculosis Drug Discovery Programmes. *Infectious Disorders-Drug Targets (Formerly Current Drug Targets-Infectious Disorders)*, *7*(2), 120-126.
4. Beaver, J. E., Bourne, P. E., & Ponomarenko, J. V. (2007). EpitopeViewer: a Java application for the visualization and analysis of immune epitopes in the Immune Epitope Database and Analysis Resource (IEDB). *Immunome research*, *3*(1), 3.
5. Blaser, M. J., & Kirschner, D. (2007). The equilibria that allow bacterial persistence in human hosts. *Nature*, *449*(7164), 843.
6. Blythe, M. J., & Flower, D. R. (2005). Benchmarking B cell epitope prediction: underperformance of existing methods. *Protein Science*, *14*(1), 246-248.
7. Blythe, M. J., Doytchinova, I. A., & Flower, D. R. (2002). JenPep: a database of quantitative functional peptide data for immunology. *Bioinformatics*, *18*(3), 434-439.
8. Bogatcheva, E., Hanrahan, C., Nikonenko, B., De los Santos, G., Reddy, V., Chen, P., ... & Protopopova, M. (2011). Identification of SQ609 as a lead compound from a library of dipiperidines. *Bioorganic & medicinal chemistry letters*, *21*(18), 5353-5357.
9. Borrero, N. V., Bai, F., Perez, C., Duong, B. Q., Rocca, J. R., Jin, S., & Huigens III, R. W. (2014). Phenazine antibiotic inspired discovery of potent bromophenazine antibacterial agents against Staphylococcus aureus and Staphylococcus epidermidis. *Organic & biomolecular chemistry*, *12*(6), 881-886.

ISSN: 0971-1260
Vol-22-Issue-17-September-2019

10. Brusic, V., Rudy, G., & Harrison, L. C. (1994). MHCPEP: a database of MHC-binding peptides. *Nucleic Acids Research*, *22*(17), 3663-3665.
11. Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., ... & Tekaia, F. (1998). Deciphering the biology of *M. tuberculosis* from the complete genome sequence. *Nature*, *393*(6685), 537.
12. Cynamon, M. H., Klemens, S. P., Sharpe, C. A., & Chase, S. (1999). Activities of Several Novel Oxazolidinones against*M. tuberculosis* in a Murine Model. *Antimicrobial agents and chemotherapy*, *43*(5), 1189-1191.
13. Douglas, J. D., Senior, S. J., Morehouse, C., Phetsukiri, B., Campbell, I. B., Besra, G. S., & Minnikin, D. E. (2002). Analogues of thiolactomycin: potential drugs with enhanced anti-mycobacterial activitya. *Microbiology*, *148*(10), 3101-3109.
14. Drouillon, V., Delogu, G., Dettori, G., Lagrange, P. H., Benecchi, M., Houriez, F., ... & Fadda, G. I. O. V. A. N. N. I. (2009). Multicenter evaluation of a transcription-reverse transcription concerted assay for rapid detection of *M. tuberculosis* complex in clinical specimens. *Journal of clinical microbiology*, *47*(11), 3461-3465.
15. EAST, A. (1977). Results at 5 years of a controlled comparison of a 6-month and a standard 18-month regimen of chemotherapy for pulmonary tuberculosis. *American Review of Respiratory Disease*, *116*(1), 3-8.