

Types, Importance and Limitations of DNA Microarray

¹*Gemechu Berhanu and* ²*Tesfaye Itefa Dula*

¹College of Agriculture and Veterinary Medicine, Dambi Dollo University, Ethiopia

²Kellem Wollega Zone Livestock and Fisheries Development Office, Ethiopia

Abstract: DNA microarrays are a type of biotechnological instrument used to detect changes in genomic DNA and mRNA and to monitor genes and their expressions related to various functions and, they have been widely applied in different industries and researches. The invention of this device in biotechnology helped a much more significant quantitative and simultaneous monitoring of the expression of many of genes. There two types of DNA microarrays that involve oligonucleotide microarrays and cDNA microarrays. Both of them depend on correct base pairing and show many of transcripts. Different applications of DNA microarray technology include disease diagnosis, drug discovery and development, toxicological research, immunology, microbial detection and identification, comparative genomics and microbial typing and determination of virulence factors. Though DNA microarray is very important, it has several drawbacks including huge quantity of data and cost of the machine.

Key words: cDNA • DNA Microarray • Importance • Oligonucleotide

INTRODUCTION

Microarray is the technology that permits screening thousands of gene expression. The basic principle on the microarray is a base supplement, the adenine supplementary thymine and guanine is cytosine supplements [1]. In a microarray a thousand points in a square shape, anywhere contains a great number of DNA fragments on a particular gene is placed. The collection of molecular probes that are analogous to particular display of cDNA and on a solid phase, for example, glass slides have been fixed, usually considered by fluorescent materials glowing [2]. DNA microarray dependent strand is hybridized and paired DNA strands, which becomes stronger due to hydrogen bond. Final signals depend on the sequence of levels that have established a strong bond interaction [3].

Microarrays, initially used for the analysis of whole genome gene expression, are also being used for the detection of different pathogens and to the investigation of the evolutionary association in different species of bacteria [1, 3]. The advantage of microarray-based study is that it can combine powerful nucleic acid amplification methods with a substantial selection ability, which results in a high sensitivity, specificity and throughput capacity.

Microarrays allow the characterization of microorganisms by giving information for specific identification of isolates; besides this, they permit detecting the pathogenesis based on the accessibility of different virulence genes and showing how new pathogenic strains developed epidemiologically and phylogenetically. Even if, microarrays have been used principally for gene expression study, they have gradually been applied in the revealing and description of microbial pathogens [2, 4].

A DNA microarray was developed in the late 1990s and early 2000s where methods for fluorescent detection were adapted for the development of array technology [3]. It is a collection of probes fitted on solid substance that can be used for genetic studying. The probes are fragments of DNA obtained from nucleobases such as cytosine, guanine, adenine and thymine. Sometimes it is also called biochips, having different applications in medicine, genetics, molecular biology and pharmacology for gene detection, disease diagnosis, drugs inventing and toxicological analyses [1, 3].

DNA microarrays are a type of biotechnological instruments used to detect changes in genomic DNA and mRNA and to monitor genes and their expressions associated with various functions. DNA microarray assay

has been used to monitor and characterize different materials among mixtures of chemicals and extracts of natural resources such as plants [2, 3]. It has its own strength and also limitations relative to other techniques. It has been used as a diagnostic device, such as for the genotyping of drug-metabolizing genes and predicting the metastatic risk of breast cancer, which is attributable to its unique characteristic of giving sufficient complexity to differentiate the variations desired for diagnosis and the reliability needed to forecast gene expression profiles correctly [5].

The development of this tool in the biotechnology field allowed a much more efficient quantitative and simultaneous studying of the expression of thousands of genes. This tool is best developed and the most significant technique of all types of microarrays. It could be expressed in two ways, as an arrangement of known sequenced genes printed on a solid support e.g. plastics, silicon chips, nylon membrane, glass microscope slides and also machinery used to analyze and detect gene expression and mutations, respectively [1, 2, 4]. It simplifies identification of genes expressed because of alteration in environmental parameters, cellular differentiation and mutations in metabolic pathways. Due to this, the technique is taken as an important means for genome-wide expression profiling. DNA arrays are printed with DNA, cDNAs, oligonucleotides (synthetic) or PCR products that generally represent a gene it is used for studying many different genes of concern at the same time, even a whole; genome of an organism [6].

Basic Principle and Steps of DNA Microarray Technique: The principle of DNA microarrays lies on the hybridization of nucleic acid strands. The character of complementary nucleic acid array is to exclusively combine with each other by forming hydrogen bonds between complementary nucleotide base pairs. For this purpose, samples are colored by using fluorescent dyes, in which at least two samples are attached to chip. Complementary nucleic acid sequences between the sample and the probe hybridized on the chip get matched via hydrogen bonds [2, 3]. The non-specific bonding sequences while remain unattached and removed during the washing step of the process. Target sequences that are fluorescently labeled are attached to a probe sequence and create a signal. The signal depends on the hybridization circumstances (example, temperature), washing after hybridization, while the total strength of the signal, depends upon the amount of target sample present. General procedures include collection of samples,

isolation of mRNA, creation of labeled cDNA, hybridization and collection and analysis. Materials required are DNA chip, target sample (fluorescently labeled), fluorescent dyes, probes and scanner [4, 5].

Types of DNA Microarray: The two types of DNA microarrays include cDNA microarrays and oligonucleotide microarrays. Both microarrays are dependent on base complementarity and express the abundance of transcripts; always two mRNA samples are prepared, the control and experimental sample, these mRNA samples are changed into cDNAs with reverse transcriptase and are labeled with fluorophore, silver and chemiluminescence [6].

cDNA Microarrays: cDNA microarrays have long fragments of DNA (from 100 to thousands of base pairs). cDNA arrays are formed by robotically spotting each samples of selected cDNA clones onto a solid support. Some of the basic principles of the preparation of cDNA arrays include: i) selection of the targets to be attached on the array from the database storage such as Gene Bank, dbEST and UniGene or by chance from any library of interest; ii) arraying the chosen cDNA targets onto the known site of coated glass microscope slide by using a computer-based high speed robot; iii) fluorescently labeling the total RNA from both tests and suggested samples using dyes with a single round of reverse transcription; iv) pooling the fluorescent target for hybridization under strict conditions; v) measuring the laser agitated incorporated targets using a scanning confocal laser microscope; and vi) lastly, analyzing the images from scanner by importing into a software in which they are pseudo-colored and merged [2, 4, 7]. In the case of cDNA micro arrays, the samples are prepared with a different label and then are mixed to finally be marked in the microarrays where probes are ready to attach complementary sequences [6].

cDNA microarray assays are made-up on solid materials, usually microscopic glass slides. In order to boost the attachment of probes as well as to lesser background and to limit spreading of the droplets, the slides are pre-coated with materials such as poly-lysine or amino silanes. Nowadays two processes i.e. either contact printing or non-contact printing are present [3, 5]. In contact printing technique, either solid or split pins are dipped into the DNA solution and then a micro droplet consequently placed upon direct attachment with the solid surface of the array. The method uses a motion control structure that spots or prints an accurate amount

of sample probe on many surfaces in a particular operation. As the objective, one can use contact printing typically to make sub-nano liter droplets at a pitch of 100-250 nm [8].

Preparation of Probes: Starting of cDNA microarray production necessitates the assortment of appropriate probes to be applied. Generally, fluorescent probes are made to attach to the microarray. They can be obtained from mRNA from different cells or tissues to be studied. Depending on key character, needed mRNA can be separated and obtained from the cells or tissues of interest. Obtained mRNAs are helpful to make complementary DNA using enzyme known as reverse transcriptase. Then, synthesized cDNA is labeled by the inclusion of fluorescent or radioactive nucleotides into the DNA. The two samples with distinct behavior are colored or labeled by using different fluorescent dyes-say, Cy3 (red) or Cy5 (green). They are then attached to the microarray glass slide or DNA chip. Two differentially labeled cDNA mixtures allow the measurement of ratio of fluorescence; this circumvents most of the problems of hybridization kinetics [8].

Oligonucleotide Microarrays: Oligonucleotide arrays or DNA chips are small parallel analytical devices containing libraries of oligonucleotides robotically spotted or synthesized *in situ* on solid supports (glass, coated glass, silicon or plastic) in a such way that the characteristics of each oligonucleotide is defined by its position. Oligonucleotide arrays possess short DNA fragments, commonly 25 base pairs. One of the commercially existing oligonucleotide microarrays are the Gene Chips made by firms such as Affymetrix. Traditionally, Affymetrix Gene Chip Arrays are made as a single array caged in a potted cartridge with glass as a substrate [7].

The oligonucleotide technologies pioneered by Affymetrix Gene Chips are different from cDNA microarray, in two important respects. First, the probes are a set of 20-25 short oligonucleotides specific for each gene or exon, along with the associated set with single base mismatches included at the middle position of each oligonucleotide. These are made *in situ* on each silicon chip using genome sequence information to direct photolithographic deposition. Second, the arrays are attached to a single biotinylated amplified RNA sample and the strength of assessment for each gene is the difference between the match and mismatch measurements and averages over each oligonucleotide which are often computed [4, 5]. The variation between

cDNA microarray and gene chips is the method genes are placed on the arrays. In *in situ* photolithographic production method used by Affymetrix gene chip probe arrays, the quality of chips formed depends critically on the efficiency of photo-deprotection [7]. Both of the samples use the same color; however the labeled cDNAs of each sample are attached in separate arrays in oligonucleotide arrays [6].

DNA Microarray Design: Designing of DNA microarrays arise a hard combinatorial difficulty of arranging probes through minimizing the possibility of defects and reduction of the complication of masks used in production. DNA microarray can be produced by printing probes on a material made of glass or plastic. In the beginning, the microarray has unfilled spots, on which probes are printed by applying nucleobase after nucleobase until all probes are completed. Even though one probe is not actually one DNA molecule, but rather many copies of the similar single-stranded DNA molecule on a particular spot, it is simpler to consider about the probe as one single-stranded DNA molecule. Photolithographic masks are used to choose spots on which the needed nucleobases will be inserted. Using one mask only one type of nucleobases: cytosine (C), guanine (G), adenine (A) or thymine (T) can be put on unmasked spots. In other spots on the microarray, nucleobases are not put on the spot. These steps are continued by using correct masks in the correct array until all probes are finished [2, 4, 8]. DNA microarray assay requires a commercial array or custom designed array and cDNA labeling synthesis from each RNA sample. Following attachment of the cDNA samples on array, a scanning device takes image and measures the fluorescent levels of the microarray [3, 9].

Importance of DNA Microarray: DNA microarrays give a simple and natural vehicle for delivering the genome in both systematic and comprehensive ways. The universality of DNA microarrays as experimental tools related to the exquisite specificity and affinity of complementary base-pairing. Some DNA array techniques are appropriate techniques to detect many pathogens using a single PCR assay on the 16S rRNA region. In relation with DNA microarrays, new emerging technologies seem particularly well suited for application in food safety context since they can join together well known and standardized techniques to novel platforms and/or detection methods, capable of overcoming the limits imposed by fluorescence-based techniques [2].

DNA microarrays can investigate the transcripts of tumor cells by making comparison of gene expression with normal cells. The data obtained from the DNA microarray images can identify patterns with similar transcriptional actions and as a result to keep together genes that are related with cell proliferative state [5]. Similarly, transcript profiling and clustering are power tools for sub-classification of tumor types that may direct to a better diagnosis and therapy of cancer. Moreover, different researches have been conducted in mitosis and meiosis for analysis of changes in gene expression at different times of budding yeast. Another remarkable application of DNA microarrays is in the field of stress response and aging, gene expression of cells varies depending on the environmental factors by which they are surrounded. Myeloid cells have been considered after being treated with genotoxic agents and ionizing radiation and unsuspected genes were expressed [6]. Apart from applications of cDNA array in relative genome analysis and functional genome analysis, the recent importance include gene expression of drug-resistant small cell lung cancer cells, characterization and profiling of osmotic stress induced genes in poplar cells and veterinary diagnostics [7].

DNA microarray technique gives the possibility of high-throughput systematic studying of the transcriptome in experiments. The most revealing and probably the most advantageous importance of DNA chips is the parallel study of gene expression from different biological materials focusing on the functionally active parts of the genome [6]. DNA microarrays with groups of cDNA fragments or gene specific oligonucleotide on their surfaces are used to get a molecular fingerprint of gene expression of cells at a given time point in a comparative way gene expression monitoring means quantification of RNA in many situation of cells, mRNA number of a tissue or cell in a given condition is called transcriptome. The differences between transcriptomes because of different factors including treated cells versus non-treated or diseased state versus control, transgenic plant versus wild type is the most interesting ideas in biological systems. The action of several genes can be determined in one hybridization step using microarray techniques [10].

Disease Diagnosis: Microarray technology helps researchers to know more about many diseases, including mental illness, heart diseases and infectious diseases. Scientists have divided different types of cancers depending on the organs in which the tumors develop.

Using microarray technology, however, they were able to classify those types of cancers depending on the patterns of gene activity in the tumor cells. Researchers then designed different treatment strategies targeted directly to each specific type of cancer [11, 12]. Lin *et al.* [13] expressed molecular mechanism of cervical carcinogenesis depending on systems biological techniques as method of different disease diagnosis.

Drug Discovery and Development: Medicinal chemistry has employed microarrays to study both key target genes and gene networks that can adjust the effectiveness of drugs. Important application of DNA microarray technology, in relation with drug effectiveness and safety evaluation, is its use as a screening tool for the studying of biochemical pathways, possible targets for novel molecular therapeutics, for the consideration of molecular mechanisms of toxicity and know individual drug susceptibility and resistance¹¹. Some include DNA microarray analysis as a tool to consider the therapeutic mechanisms and drug development of Chinese medicinal herbs [14], importance of using DNA microarray in studying different medicinal plants [8], DNA microarray technology and drug development [15].

Toxicological Research: The important application of microarray, within the context of neurotoxicological studies, is its use as a selection tool for the recognition of molecular ways of toxicity. This method helps different researchers to recognize those genes involved in conferring resistance or sensitivity to different toxic substances [11].

Immunology: DNA microarray techniques have been used in immunological studies including development, maturation, activation and division of immunity system cells, the molecular actions of allergic reaction, the relation between genotypic and phenotypic expression, the regulation of immune responses and immunological pharmacology. It is as well be helpful in the research of the regulative mechanism of long-established medication related immune responses, the mechanism of action different drug toward allergy, equivalence of differentiation of syndrome and herbal pharmacology [11, 14].

Microbial Detection and Identification: The most important area in applying DNA microarray techniques in microbiology field is for evaluation of many of microbial genetic targets at the same time [16]. Some microbial study

include a novel tool for recognition and utilization of fish conservation in aquaculture [9], advances in DNA microarray technology for the detection of foodborne pathogens [2], accurate and rapid identification of *Pseudo-nitzschia* species and other harmful algal species [17], rapid pathogen identification by using a novel microarray-based technique with purulent meningitis in the cerebrospinal fluid [18], DNA microarray assay for the rapid detection and identification of food and water borne bacteria [19], new oligonucleotide microarray for rapid diagnosis of avian viral diseases [20].

Comparative Genomics and Microbial Typing: Genomic hybridization of a complete genome array recognizes the existence or absence of analogous DNA regions in other microorganisms, allowing genome-wide association of their genetic contents. Conducting a comparative genomic detection or study in the absence of whole genome sequences is an effective method. DNA microarrays can be used to study genome differences between different microorganisms [16].

DNA microarrays are used as sensitive techniques for exploring the formation of complex microbial ecosystems [21, 22], as well as investigation of functional gene sets [23]. Because of the growth of new generation sequencing techniques applied to metatranscriptomics and metagenomics their application has been radically minimized. Definitely, these non-targeted ways enable the recognition of different types of genes or transcripts present in the ecosystem, that have been already studied or not and that appear of significant or not.

Determination of Virulence Factors: Different genes associated with virulence factors are affected by different conditions. One way to consider the virulence factors is to study the genome-wide gene expression profiles under appropriate conditions, such as physiological changes during interaction with the host. The next one depends on comparative genomics. A group of different virulence genes was considered by their origin with a pathogenicity island in a genome association study among the *H. pylori* strains. The whole-genome microarray tools *H. pylori* was also expressed to be an effective technique to consider differences in gene content between two *H. pylori* strains that provoke distinct pathological outcomes. The power of *H. pylori* to direct epithelial cell responses linked to irritation based on undamaged *cag* Pathogenicity Island [16]. Also researches done by Strauß *et al.* [24] using whole-genome sequencing and DNA microarray technology recognized *S. aureus* virulence and resistance

genes, by Shridhar *et al.* [25] DNA microarray-based consideration of virulence potential of Shiga toxin gene carrying *Escherichia coli* O104:H7 isolated from feedlot cattle feces.

DNA Microarray Data Analysis: During the process of DNA microarray assay, the collection of data is taken place by using a microarray scanner. The scanner consists of computer, laser and camera, for which the laser excites fluorescence of the cDNA generating signals during the assay. The laser scans the array tools, the camera captures the images produced and the computer keeps the data results immediately for analysis [4]. After imaging of the microarray, the raw data observed is first analyzed by subtracting the background fluorescence from the fluorescence of each particular spot. Common normalization method, such as the non-linear locally weighted scatter plot smoothing technique, is applied to the background-adjusted intensities the methods for fluorescent detection were adapted for the progression of array technique [3, 17].

Limitations of DNA Microarray: Even though DNA microarrays enable researchers to simultaneously consider the expression of many different types of genes, the technology has several limitations. Different studies have shown that, for relatively available transcripts, the presence and direction, but not the quantity, of expression changes can be reliably observed. Conversely, correct measurements of complete expression levels and the reliable recognition of low abundance genes are not easy to achieve. The most important troubles seem to be the suboptimal design or selection of probes and some incorrect probe annotations and preparations [14].

Translation levels of many different genes are affected by posttranscriptional monitoring and different proteins are grossly affected by post-translational alteration such as glycosylation, acetylation, proteolysis and phosphorylation. Apparently, a nucleic acid-based array is blind to such property and for certain uses, the tedious sample preparation requirements of DNA microarray assay make them unfeasible. The solution for such problem is to study proteins rather than cause inferences based on RNA amounts directly which can be done through protein microarrays techniques [7].

The technique only identifies sequences that the array was supposed to detect. Additionally, non-coding RNA's that are not yet predicted as expressed are typically not represented on an array. Furthermore, for highly variable different genomes such as those from

bacteria, arrays are specifically determined using information from the genome of a reference strain. Such arrays can miss a large part of the genes in a given isolate of the similar species. For example, in bacterial species *Aggregatibacter actinomycetemcomitans*, the gene content differs by as much as 20% between any two different isolates [26]. Consequently, an array designed using gene annotation from a reference isolate may not contain different types of the genes found in other isolates [11, 18].

CONCLUSION

DNA Microarray is a technology that enables the researchers to consider and concentrate on issues that were once thought to be non-traceable. The usage of cDNA or oligonucleotide microarrays depends on the equipment, resources and the time available for the investigator. The technology of DNA microarrays is widely being used to measure and detect levels of gene expression. It has many applications in different fields of study including veterinary and medical microbiology which includes drug discovery and development, disease diagnosis, toxicological research, microbial detection and identification, immunology, comparative genomics and microbial typing, determination of virulence factors. Even though it is a very useful tool in various applications, this technology has a number of limitations.

REFERENCES

1. Dadkhah, K., S. Hossein, M. Anijdan, M. Karimi, M.K Abolfazli, F. Rezaei, F. Adel and G. Ataei, 2015. DNA microarray, types and its application in medicine. *Scholars Academic Journal of Bioscience*, 3(7): 598-602.
2. Severgnini, M., P. Cremonesi, C. Consolandi, G. De Bellis and B. Castiglioni, 2011. Advances in DNA Microarray Technology for the Detection of Foodborne Pathogens. *Food Bioprocess Technology*, 4: 936-953.
3. Narrandes, S. and W. Xu, 2018. Review: Gene Expression Detection Assay for Cancer Clinical Use. *Journal of Cancer*, 9(13): 2249-2265.
4. Ivkovic, N., M. Golub and D. Jakobovic, 2016. Designing DNA Microarrays with Ant Colony Optimization. *Journal of Computers*, 11(6): 528-536.
5. Kiyama, R., 2017. DNA Microarray-Based Screening and Characterization of Traditional Chinese Medicine. *Microarrays*, 6(1): 4.
6. Herrera, H.J. and M. Gancino, 2017. DNA microarrays: Recent Advances. *Bionatura*, 2: 2-7.
7. Naidu, C. and Y. Suneetha, 2012. Review Article: Current Knowledge on Microarray Technology - An Overview. *Tropical Journal of Pharmaceutical Research*, 11(1): 1-4.
8. Kumar, A., M. Asthana, S. Sharma, P. Roy, S. Amdekar, V. Singh, H. Deval, O. Parkash and R. Sharma, 2012. Importance of Using DNA Microarray in Studying Medicinal Plant. *Webmed Central Molecular Biology*, 3(12): WMC003876.
9. Malaka, A.K., K.V. Singh and S. Srivastava, 2013. A Review on DNA Microarrays: A Novel Tool for Identification and Exploitation of Fish Conservation in Aquaculture. *World Journal Fish and Marine Science*, 5(1): 26-34.
10. Zvara, A., K. Kitajka, F. Nóra and L.G. Puskás, 2015. Review: Microarray technology. *ActaBiologica Szegediensis*, 59: 51-67.
11. Bumgarner, R., 2013. DNA microarrays: Types, Applications and their future. *Current Protocols in Molecular Biology Journal*, 22: 1.
12. Kim, C., K.J. Kim, J. Bok and E.J. Lee, 2012. Microarray-based mutation detection and phenotypic characterization in Korean patients with retinitis pigmentosa. *Molecular Vision Journal*, 18: 2398-2410.
13. Lin, M., Y. Miaomiao and Z. Junhan, 2019. Recent Advances on the Molecular Mechanism of Cervical Carcinogenesis Based on Systems Biology Technologies. *Computational and Structural Biotechnology Journal*, 17: 241-250.
14. Li, C., H.Y. Lo, C.Y. Hsiang and H. Tin-Yun, 2012. DNA microarray analysis as a tool to investigate the therapeutic mechanisms and drug development of Chinese medicinal herbs. *BioMedicine*, 2(1): 10-16.
15. Sana, K., S. Chaturvedi, N. Goel, S. Bawa and S. Drabu, 2010. DNA Microarray Technology and Drug Development. *Chronicles of Young Scientists*, 1(1): 1-5.
16. Nsofor, C.A., 2014. Review: DNA microarrays and their applications in medical microbiology. *Biotechnology and Molecular Biology Reviews*, 9(1): 1-11.
17. Charlotte, N., A. Anne T. Lidwine, A.L. Véronique and D. Catherine, 2015. Phytochip: Development of a DNA-microarray for rapid and accurate identification of Pseudo-nitzschia species and other harmful algal species. *Journal of Microbiological Methods*, 112: 55-66.

18. Hou, Y., X. Zhang and X. Hou, 2018. Rapid pathogen identification using a novel microarray-based assay with purulent meningitis in cerebrospinal fluid. *Scientific Reports*, 8: 15965.
19. Ranjbar, R., P. Behzadi, A. Najafi and R. Roudi, 2017. DNA Microarray for Rapid Detection and Identification of Food and Water Borne Bacteria: From Dry to Wet Lab. *The Open Microbiology Journal*, 11: 330-338.
20. Sultankulova, K.T., S. Nurlan and M.K. Vitaliy, 2017. New oligonucleotide microarray for rapid diagnosis of avian viral diseases. *Virology Journal*, 14: 69.
21. Kang, S.H., P. Evans, M. Morrison and C. McSweeney, 2013. Identification of metabolically active proteobacterial and archaeal communities in the rumen by DNA- and RNA-derived 16S rRNA gene. *Journal of Applied Microbiology*, 115: 644-653.
22. Tottey, W., J. Denonfoux and F. Jaziri, 2013. The human gut chip “HuGChip”, an explorative phylogenetic microarray for determining gut microbiome diversity at family level. *PLoSOne*, 8: e62544.
23. Tu, Q., H. Yu, Z. He and Y. Deng, 2014. A functional gene-array-based high-throughput environmental technology for microbial community analysis. *Molecular Ecology Resources*, 14: 914-928.
24. Strauß, L., U. Ruffing, S. Abdulla and A. Alabi, 2016. Detecting *Staphylococcus aureus* virulence and resistance genes: a comparison of whole-genome sequencing and DNA microarray technology. *Journal of Clinical Microbiology*, 54: 1008-1016.
25. Shridhar, P.B., I.R. Patel, J. Gangiredla and L.W. Noll 2018. DNA microarray-based assessment of virulence potential of Shiga toxin gene-carrying *Escherichia coli* O104: H7 isolated from feedlot cattle feces. *PLoSOne*, 13(4): e0196490.
26. Kittichotirat, W., R.E. Bumgarner, S. Asikainen and C. Chen, 2011. Identification of the pangenome and its components in 14 distinct *Aggregatibacter actinomycetemcomitans* strains by comparative genomic analysis. *PLoSOne*, 6(7): e22420.