

Microbial forensics and computational biology

Nida Tabassum Khan

Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology Engineering and Management Sciences, (BUIITEMS), Quetta, Pakistan

Email address:

nidatabassumkhan@yahoo.com (NIDA TABASSUM KHAN)

To cite this article:

NIDA TABASSUM KHAN. Microbial forensics and computational biology. *Journal of research in health science*. Vol. 2, No. 2, 2017, pp. 4-7. DOI 10.26739/2523-1243



<http://dx.doi.org/10.26739/2523-1243/-2017-2-2-1>

Abstract: Microbial forensics is focused on characterization of evidences from an act of bioterrorism, biocrime, deception, or an unintentional release of a biological or chemical agent or toxin for ascription purposes. Computational biology supports the rapid identification of microbial samples at variety of levels of resolution such as family, species, strain and isolates by using different softwares to assist in forensic analysis of microbes.

Keywords: PhyloSift; Phyla-AMPHORA; Multilocus sequence typing.

INTRODUCTION

Microbial Forensics is a branch of digital forensics. It involves the study and examination of useful structured data (information from underlying databases), obtained from crime incidents. Microbial forensics targets to discover and analysed the patterns of deceitful actions [1].

Forensic examination of microbes

Microbial forensics employed scientific methods to inspect evidences related to bioterrorism events, bio crimes, deceptions, or the unpremeditated release of biological toxins or for ascription reasons [2, 3].

Bioinformatics implications in forensic examination of microbes:

Bioinformatics applications involves the use of techniques for recognition and identification of microbes for sequence analysis of microbial genome [4]. Bioinformatics supports the rapid identification of microorganism at various levels such as family, species, strain and isolate [5]. Wet lab techniques involves sequencing, assembly and analysis of genomic data to obtain refined forensic data. Special algorithms responsible for large scale sequence alignments and genome comparison along with modelling tools to

predict protein structure[6].These techniques are based on authentic and accurate examination and comparison of microbial genome to identify gene composition, predict protein motif/domain and other regulatory networks involved in virulence and transmission [7,8].

Bioinformatics tools and methods used in forensic analysis of microbes:

- **Massively parallel sequencing (MPS):** MPS is a disruptive technology that has overcome the identification of unknown pathogens, hoax microorganism even in complex mixture samples. This next generation sequencing (NGS) involves the analysis of microbes and genetically engineered genomes with high throughput and speed employing powerful bioinformatics tools for data assessment [9]. MPS use culture-independent methods to clinically diagnose and monitor infections and identify disease epidemics in real-time [10].Therefore reducing cost and turnaround time for characterizing any microbe either abundant, trace, degraded or intact [11, 12].

- **Metagenomics:** it involves the direct assessment of genomes present in an environment sample [13].The discipline emerges with the cloning of environmental DNA, followed by gene expression later accompanied by direct random shotgun sequencing [14,15].Exploration of functional gene composition of microbial communities gives a much detailed description than phylogenetic surveys based on the diversity of a single gene such as 16S rRNA gene [16].Bioinformatics is employed in metagenomics to identify novel biocatalysts or proteins, linkages analysis between function and phylogeny for uncultured microorganisms, evolutionary profiling of community

function and structure, expression profiling etc[17, 18, 19].

- **Multilocus sequence typing (MLST):** A type of bacterial typing which uses seven housekeeping genes, whereas locus sequence is used for comparison purposes [20]. MLST generated data is stored in MLST library which consists of several databases used for identification and storage purposes [21, 22]. However MLTS has some limitations like non universal genes panels and insufficient species level resolution. In such case MLVA provide additional resolution of specific species [23].

- **Multilocus variable-number tandem-repeat analysis (MLVA):** A ribosomal multilocus sequence typing which exploits 53 housekeeping genes to type subspecies of bacteria [24] generating sufficient biological data that is utilized by bioinformaticians in specie identification. [25, 26].

- **AMPHORA (AutoMated PHylogenOmic inference):** A panel of 31 housekeeping genes developed as genetic markers to enhance taxonomic resolution. These genes are common in all bacteria and found to be resistant to horizontal gene transfer [27].

- **MetaPhyler:** Another bioinformatics program which aimed to analyse the data from taxonomics profiler which uses marker genes as a taxonomic reference [28].

- **Phyla-AMPHORA:** it is a flavour of AMPHOR2 that performs large-scale phylogenetic inference using thousands of phylum-specific bacterial phylogenetic markers to enhance phylogenomic resolution [29].

- **PhyloSift:** it enables phylogenetic analysis of metagenomes using 37 PhyEco markers to provide proficiencies for the use of protracted and convention marker

sets. Besides offers data simulation to examine newly generated markers [30].

- **mOTU**: Metagenomic profiling using 40 marker genes with resolution at specie level. Furthermore includes 11 phylogenetic markers for identification in human microbiome samples [31].

Conclusion

Thus microbial forensics is an emerging field that is facilitated with the tools and softwares of computational biology. These techniques provide an authentic and accurate examination and comparison of microbial genomic sequence to find gene composition, protein motif/domain and other regulatory networks.

References

- [1] Budowle, B., Murch, R., & Chakraborty, R. (2005). Microbial forensics: the next forensic challenge. *International Journal of Legal Medicine*, 119(6), 317-330.
- [2] Keim, P. (2003). Microbial forensics: a scientific assessment.
- [3] Cummings, C. A., & Relman, D. A. (2002). Microbial Forensics--" Cross-Examining Pathogens". *Science*, 296(5575), 1976-1979.
- [4] Breeze, R. G., Budowle, B., & Schutzer, S. E. (2005). *Microbial forensics* (pp. 1-25). Burlington, MA: Elsevier Academic Press.
- [5] Murch, R. S. (2003). Microbial forensics: building a national capacity to investigate bioterrorism. *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, 1(2), 117-122.
- [6] Valdivia-Granda, W. A. (2010). Bioinformatics for biodefense: challenges and opportunities. *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, 8(1), 69-77.
- [7] Joseph, S. J., & Read, T. D. (2010). Bacterial population genomics and infectious disease diagnostics. *Trends in biotechnology*, 28(12), 611-618.
- [8] Pattnaik, P., & Sekhar, K. (2008). Forensics for tracing microbial signatures: Biodefence perspective and preparedness for the unforeseen.
- [9] Berglund, E. C., Kiialainen, A., & Syv?nen, A. C. (2011). Next-generation sequencing technologies and applications for human genetic history and forensics. *Investigative Genetics*, 2(1), 23.
- [10] Børsting, C., & Morling, N. (2015). Next generation sequencing and its applications in forensic genetics. *Forensic Science International: Genetics*, 18, 78-89.
- [11] Van Dijk, E. L., Auger, H., Jaszczyszyn, Y., & Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in genetics*, 30(9), 418-426.
- [12] Nguyen, L., & Burnett, L. (2014). Automation of molecular-based analyses: a primer on massively parallel sequencing. *The Clinical Biochemist Reviews*, 35(3), 169.
- [13] Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and molecular biology reviews*, 68(4), 669-685.
- [14] Tringe, S. G., Von Mering, C., Kobayashi, A., Salamov, A. A., Chen, K., Chang, H. W., ... & Bork, P. (2005). Comparative metagenomics of microbial communities. *Science*, 308(5721), 554-557.
- [15] Streit, W. R., & Schmitz, R. A. (2004). Metagenomics-the key to the uncultured microbes. *Current opinion in microbiology*, 7(5), 492-498.
- [16] Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., & Wilkening, J. (2008). The metagenomics RAST server-a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), 386.
- [17] Hugenholtz, P., & Tyson, G. W. (2008). Microbiology: metagenomics. *Nature*, 455(7212), 481-483.
- [18] Tringe, S. G., & Rubin, E. M. (2005). Metagenomics: DNA sequencing of environmental samples. *Nature reviews. Genetics*, 6(11), 805.

- [19] Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., & Hugenholtz, P. (2008). A bioinformatician's guide to metagenomics. *Microbiology and molecular biology reviews*, 72(4), 557-578.
- [20] Jolley, K. A., Chan, M. S., & Maiden, M. C. (2004). MlstdbNet-distributed multi-locus sequence typing (MLST) databases. *BMC bioinformatics*, 5(1), 86.
- [21] Enright, M. C., & Spratt, B. G. (1999). Multilocus sequence typing. *Trends in microbiology*, 7(12), 482-487.
- [22] Healy, M., Huong, J., Bittner, T., Lising, M., Frye, S., Raza, S., & Woods, C. (2005). Microbial DNA typing by automated repetitive-sequence-based PCR. *Journal of clinical microbiology*, 43(1), 199-207.
- [23] Vergnaud, G., & Pourcel, C.(2006).Multiple locus VNTR (variable number of tandem repeat) analysis. *Molecular identification, systematics, and population structure of prokaryotes*, 83-104.
- [24] Van Belkum, A. (2007). Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA). *Pathogens and Disease*, 49(1), 22-27.
- [25] Vergnaud, G., & Pourcel, C.(2009).Multiple locus variable number of tandem repeats analysis. *Molecular Epidemiology of Microorganisms: Methods and Protocols*, 141-158.
- [26] Fricke, W. F., Rasko, D. A., & Ravel, J. (2009). The role of genomics in the identification, prediction, and prevention of biological threats. *PLoS biology*, 7(10), e1000217.
- [27] Wu, M., & Eisen, J. A. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome biology*, 9(10), R151.
- [28] Liu, B., Gibbons, T., Ghodsi, M., & Pop, M. (2010, December). MetaPhyler: Taxonomic profiling for metagenomic sequences. In *Bioinformatics and Biomedicine (BIBM), 2010 IEEE International Conference on* (pp. 95-100). IEEE.
- [29] Wang, Z., & Wu, M. (2013). A phylum-level bacterial phylogenetic marker database. *Molecular biology and evolution*, 30(6), 1258-1262.
- [30] Darling, A. E., Jospin, G., Lowe, E., Matsen IV, F. A., Bik, H. M., & Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ*, 2, e243.
- [31] Taberlet, P., PRUD'HOMME, S. M., Campione, E., Roy, J., Miquel, C., Shehzad, W., & Melodelima, C. (2012). Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, 21(8), 1816-1820.