Notes

David Helekal

August 27, 2020

Contents

Chapter 1

Introduction

In epidemiology, it is often desired to be able to reconstruct the history of a pathogen population and it's structure. The problem of reconstructing the hisotry of a pathogen population can be tackles using phylodynamics. Phylodynamics utilises genomic data to assemble phylogenies, which are then used to infer the population size history. This is possible by viewing a phylogeny as a realisation of a coalescent process, in which the rates are dependent on the total population. This claim can be justified by viewing the coalescent as a Moran model, run backwards in time with the time rate equal to the population size.
Within this report we will first introduce the coalescent process for phylodynamic inference, review it's inhomogenous generalisation, and finally introduce the main result of this work, a new model capable of doing local phylodynamic inference, i.e. on a subset of the whole population capable of detecting and modelling clonal expansions.
Clonal expansions are a process in which a particular subsest of a given bacterial strain undergoes explosive population growth that can be traced back to a particular individual [1]. The presence of clonal expansions in bacterial populations have been of long-standing interest and is implicated in epidemic processes, were an outbreak can be traced to a single ancestor [1, 2, 3, 4]. This often happens when a particular strain or individual obtains a variant of a particular gene that confers evolutionary advantange, for example, antibiotic resistance [5, 6, 4].
The presence of clonal expansions leaves an imprint in the overall population structure of a given bacterial strain, the particular topology associated with this often being referred to as star-like [1, 2]. The problem of detecting hidden population structure corresponding to clonal expansions has become a problem of interest in epidemiology and outbreak surveillance [7].
While methods to detect inhomogeneities in the population structure and size have been of interest since the early days of genetic sequencing [1, 2], the interest in the problem increased with whole genome sequencing becoming more accessible and affordable [5, 8, 9].

Despite the problems of inferring population size from a genealogy and detecting heterogeneities in the population size of the entire population being intrinsically tied, all but one method [7], to our knowledge, rely either on manual detection or indirect detection. We aim to propose a simulation for the formation of clonal expansions in genealogy using the structured coalescent process, and devise a fully bayesian method for joint estimation and detection of relative population size and clonal expansions.

Chapter 2

Bibliography

Bibliography

[1] J. M. Smith et al. "How clonal are bacteria?" In: *Proceedings of the National Academy of Sciences* 90.10 (May 15, 1993), pp. 4384–4388. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.90.10.4384. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.90.10.4384 (visited on 07/29/2020).

[2] Brian G. Spratt et al. "Displaying the relatedness among isolates of bacterial species – the eBURST approach". In: *FEMS Microbiology Letters* 241.2 (2004). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1016/j.femsle.2004.11.015, pp. 129–134. ISSN: 1574-6968. DOI: 10.1016/j.femsle.2004.11.015. URL: https://onlinelibrary.wiley.com/doi/abs/10.1016/j.femsle.2004.11.015 (visited on 07/29/2020).

[3] Christophe Fraser, William P. Hanage, and Brian G. Spratt. "Neutral microepidemic evolution of bacterial pathogens". In: *Proceedings of the National Academy of Sciences of the United States of America* 102.6 (Feb. 8, 2005), pp. 1968–1973. ISSN: 0027-8424. DOI: 10.1073/pnas.0406993102. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC548543/ (visited on 07/29/2020).

[4] Alice Ledda et al. "Re-emergence of methicillin susceptibility in a resistant lineage of Staphylococcus aureus". In: *Journal of Antimicrobial Chemotherapy* 72.5 (May 1, 2017). Publisher: Oxford Academic, pp. 1285–1288. ISSN: 0305-7453. DOI: 10.1093/jac/dkw570. URL: https://academic.oup.com/jac/article/72/5/1285/2930201 (visited on 07/29/2020).

[5] Matthew T. G. Holden et al. "A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant Staphylococcus aureus pandemic". In: *Genome Research* 23.4 (Apr. 2013), pp. 653–664. ISSN: 1549-5469. DOI: 10.1101/gr.147710.112.

[6] Li-Yang Hsu et al. "Evolutionary dynamics of methicillin-resistant Staphylococcus aureus within a healthcare system". In: *Genome Biology* 16.1 (Apr. 23, 2015), p. 81. ISSN: 1465-6906. DOI: 10.1186/s13059-015-0643-z. URL: https://doi.org/10.1186/s13059-015-0643-z (visited on 07/29/2020).

[7] Erik M. Volz et al. "Identification of Hidden Population Structure in Time-Scaled Phylogenies". In: *Systematic Biology* (). DOI: 10.1093/sysbio/syaa009. URL: https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syaa009/5734655 (visited on 07/01/2020).

[8] Bethany L. Dearlove and Simon D. W. Frost. "Measuring Asymmetry in Time-Stamped Phylogenies". In: *PLOS Computational Biology* 11.7 (July 6, 2015). Publisher: Public Library of Science, e1004312. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004312. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004312 (visited on 07/29/2020).

[9] Vegard Eldholm et al. "Four decades of transmission of a multidrug-resistant Mycobacterium tuberculosis outbreak strain". In: *Nature Communications* 6.1 (May 11, 2015). Number: 1 Publisher: Nature Publishing Group, p. 7119. ISSN: 2041-1723. DOI: 10.1038/ncomms8119. URL: https://www.nature.com/articles/ncomms8119 (visited on 07/29/2020).