

AARHUS UNIVERSITY

BIOINFORMATICS RESEARCH CENTER (BiRC)

A model based approach to inferring rate of genome evolution in experimental evolution lineages of *E.coli*

Oliver Kjærlund Hansen
201708342

Internal supervisor:

Thomas Bataillon (TB)

External supervisor, IGC Oeiras:

Isabel Gordo (IG)



Master's thesis in Bioinformatics

June, 2022

Abstract

Inferring the rate of adaptation is important in understanding genome evolution in natural populations. Investigating dynamic biological systems, with complex dynamics is difficult, so the need for generating sufficient data is important. Experimental designs are becoming more sophisticated when modeling natural interaction, which gives us the opportunity to measure important parameters, for adaption, more intimately.

We use an extended resource competition model to investigated several parameter's significance when fitting experimental data of *E.coli* in two distinct environments. The gut of obese and wildtype mice. Bayesian inference was used to estimate the best set of parameters, though the technique of Approximation Bayesian Computation. A method using simulations to calculate the posterior of paramters, when computing the likelihood is to computational expensive. We observed a negative linear relationship between the densities of mutation rate and mutational effect, a phenomenon observed preciously[1]. A selection of different mutation rates fit the rate of adaptation in populations of *E.coli* differently. This suited our interpretation of the observed data, which indicated *E.coli* had undergone more adaptation in obese mice, compared to wildtype mice.

Contents

1	Introduction	1
1.1	Population genetics	1
1.2	Molecular evolution	2
1.3	Experimental evolution	3
1.4	Adaptation in asexual populations	4
1.5	Aim of this study	5
2	Methods and models	6
2.1	Data and experimental design	6
2.2	Resource competition model	6
2.3	Parameters of interest	8
2.4	Approximation Bayesian Computation (ABC)	9
2.5	Summary Statistics	11
2.6	Simulations	12
3	Results	13
3.1	Mutation frequencies	13
3.2	Parameter and model investigation	14
3.3	Cross validation	22
4	Discussion	23
4.1	<i>E.coli</i> in obese and wildtype mice	23
4.2	Model fit	24
4.3	Parameter inference and interpretation	24
4.4	ABC	25
5	Conclusion	26
6	Bibliography	27
7	Supplementary	31