Contents

0.1 Preface

Coalescent Theory has gone from an obscure corner of population genetics to a central concept for anybody that studies variation at the sequence level.

Besides filling the obvious need for such a book, it is also our wish to present this theory in a straightforward and elementary manner that could dispel the misconception that Coalescent Theory is inherently very difficult and needs a strong mathematical background to understand. The key issues needed for data analysis only needs basic combinatorics. Despite the present prominence of Coalescent Theory, it also belongs to the future. ¿From an application point of view, human evolution and association mapping/fine scale mapping are two areas that are bound to grow enormously in the next few years. And to make optimal use of the coming flood of data, theoretical advances will be needed. There are areas, where present theory fails (or is impractically slow) in presence of real data and if empirical researcher are to use Coalescent based method, there are plenty of challenges for the theoretician both in modelling and in improvement of simulation algorithms.

The present book is definitely not exhaustive, but is only meant to provide a good basis for further study.

0.1.1 Chapter outline

The book consists of eight Chapters:

Chapter 1 provides the basics for understanding the assumptions behind and derivation of the basic coalescent model, and some simple properties of the resulting genealogies.

Chapter 2 introduces the models of alleles and sequences and associated mutation processes. Prominent models are: Infinite alleles, infinite sites and finite sites models. When these models and mutation processes are combined with genealogies, data can either be simulated or the probability of data can be evaluated.

Chapter 3 gives some more examples of statistics that can be calculated on coalescent genealogies and mutations on such genealogies.

Chapter 4 relaxes some of the assumptions of the basic coalescent model by introducing extensions necessary in the analysis of real data. These include population size changes, population subdivision, bottlenecks, balancing selection, and directional selection.

Chapter 5 extends the coalescent model to include genetic exchange in the form of recombination and gene conversion. An introduction to the biological features of the process, the model and algorithms used and