

Gene Genealogies and Large Offspring Numbers

Bjarki Eldon

*Department of Statistics
University of Oxford, Oxford, United Kingdom*

Population models in which individuals can have very many offspring with some probability are considered. These selectively neutral “large offspring” models may better apply to species with high fecundity and high early mortality than the usual Wright-Fisher or the Moran models. Marine species with broadcast spawning, such as Atlantic cod (*Gadus morhua* L.) or Pacific oyster (*Crassostrea gigas*), are prime examples of species for which large offspring number models may apply.

Gene genealogies of samples of DNA sequences obtained from natural populations with large offspring numbers can differ in shape from gene genealogies that follow Kingman’s coalescent. The coalescent process that results from large offspring numbers can allow for multiple mergers of ancestral lines. In contrast, Kingman’s coalescent only allows two lines to coalesce each time. Large offspring numbers can therefore predict a high number of singleton genetic variants, often observed among marine populations like Atlantic cod.

A prediction about linkage disequilibrium is shown to be a function of the parameters controlling the size and frequency of large offspring numbers. High linkage disequilibrium can be predicted despite high recombination rate, depending on parameter values. Low linkage disequilibrium can also be predicted despite low recombination rate, contrary to common intuition.

Estimates of migration rates based on identity by descent are also shown to be confounded by the reproduction parameters of large offspring numbers. Genetic heterogeneity between subpopulations can therefore be observed despite high gene flow.