Expected time to coalescence and F_{ST} under a skewed offspring distribution among individuals in a population
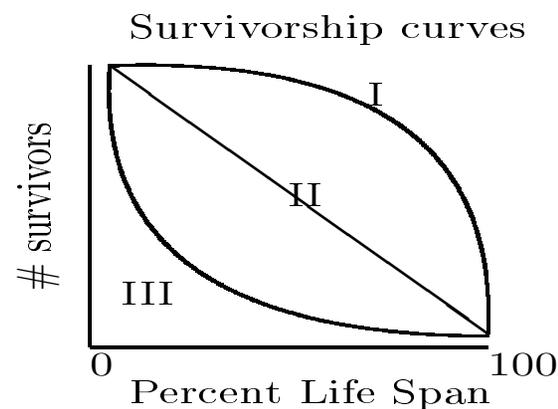
Bjarki Eldon (with John Wakeley)

High variance in offspring distribution

- broadcast spawning and external fertilization

- type III survivorship curves

- very large population sizes

- low genetic variation

- large number of singleton genetic variants

Survivorship curves

I

II

survivors

III

0   100

Percent Life Span

Standard reproduction models

have finite variance in offspring number

Moran model of overlapping generations:

a single randomly chosen individual

produces one offspring

Coalescent timescale: $\dfrac{N^2}{2}$ generations

A modified Moran model

A special case of the models of Pitman (1999) and Sagitov (1999)

The number U of offspring is a random variable

The probability $G_{n,x}$ of an x-merger ($2 \leq x \leq n$):

$$G_{n,x} = \sum_{u=2}^{N} P_U(u) \frac{\binom{u}{x}\binom{N-u}{n-x}}{\binom{N}{n}}$$

$$P_U(u) = \begin{cases} 1 - \phi N^{-\gamma}/2 & \text{if } u = 2 \\ \\ \phi N^{-\gamma}/2 & \text{if } u = \psi N, \, 0 < \psi < 1 \end{cases}$$

Population subdivision with migration

Conservative migration between finite number D of subpopulations
gives convergence to the structured coalescent
if Nm is finite as $N \to \infty$

$$N_\gamma \equiv \min\left(N^\gamma, N^2\right), \gamma > 0$$

N_γ is the coalescence timescale;

$$m_\gamma N_\gamma < \infty \text{ as } N \to \infty$$

m_γ is rescaled migration;

$$\lambda_\gamma = I_{\gamma \geq 2} + \phi\psi^2 I_{\gamma \leq 2}, \phi > 0, 0 < \psi < 1$$

λ_γ is the rate of coalescence of two lines

Time to coalescence for two lines sampled from *same* (T_0) or *different* (T_1) subpopulations

$$E(T_0) = \frac{D}{\lambda_\gamma} < \frac{D}{\lambda_\gamma} + \frac{D-1}{N_\gamma m_\gamma} = E(T_1)$$

$$\lambda_\gamma = I_{\gamma \geq 2} + \phi \psi^2 I_{\gamma \leq 2}, \; N_\gamma \equiv \min\left(N^\gamma, N^2\right)$$

Indicators of population subdivision - F_{ST} and N_{ST}

F_{ST} defined in terms of probabilities of identity

N_{ST} defined in terms of average numbers of pairwise differences

$$F_{ST} = \frac{1}{1 + \dfrac{N_\gamma m_\gamma}{\lambda_\gamma} \dfrac{D^2}{(D-1)^2} + \dfrac{\theta/2}{\lambda_\gamma} \dfrac{D}{D-1}}, \quad N_{ST} = \frac{1}{1 + \dfrac{N_\gamma m_\gamma}{\lambda_\gamma} \dfrac{D}{D-1}}$$

In a Wright-Fisher population:

$$F_{ST} = \frac{1}{1 + 4Nm\dfrac{D^2}{(D-1)^2} + \theta\dfrac{D}{D-1}}, \quad N_{ST} = \frac{1}{1 + 4Nm\dfrac{D}{D-1}}$$

$$\lambda_\gamma = I_{\gamma \geq 2} + \phi\psi^2 I_{\gamma \leq 2}, \; N_\gamma \equiv \min\left(N^\gamma, N^2\right)$$

SUMMARY

(i) multiple mergers coalescent processes may better apply to some marine organisms

(ii) coalescent times are shorter than in the standard coalescent

(iii) patterns indicating population subdivision can be observed in DNA sequence data even if the usual migration rate Nm is very, very large