Towards a Behavioural Ecological Perspective of Population Genetic Processes: Insights from Rangatira Black Robin and Florida Scrub Jay

Raazesh Sainudiin[†]

[†]School of Mathematics and Statistics, University of Canterbury,
Christchurch, New Zealand

Part I is joint work with: James Briskie, Marie Hale, Melanie Massaro, Don Merton, Anthony Poole

Part II is joint work with: Bhalchandra Thatte and Amandine Veber

Part III is joint work with Nancy Chen, Andrew Clark and John Fitzpatrick

July 23 2015,

UC SBS Seminar

Ecological History of Black Robins

Likelihood of Population Pedigrees

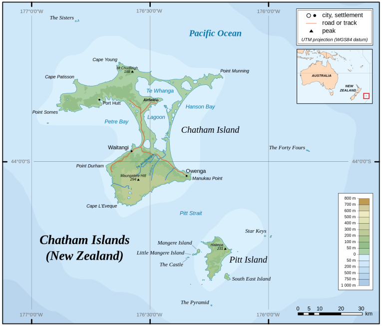
Insights from the field

Ancestries of Recombining Diploid Populations

Ecological Pedigrees

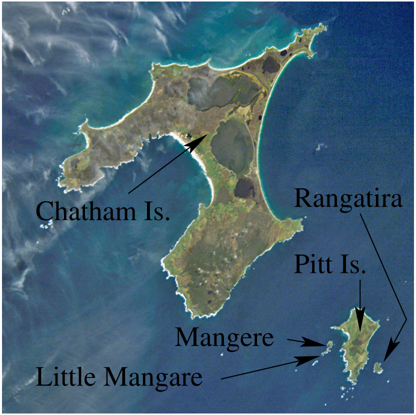
Part I

Map of Antarctica / Chatham Islands (New Zealand)



- The Chatham Islands lie 800 kms east of Christchurch.
- Two main islands – Chatham Island and Pitt Island

Chatham Islands (New Zealand)

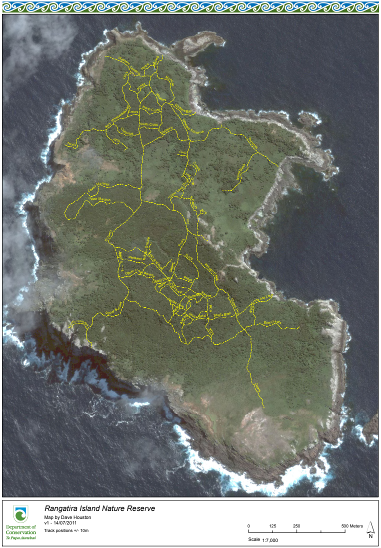


- Many smaller islands – including Little Mangere, Mangere and Rangatira
- The islands have only emerged above sea level in the last 4 million years and are part of the Chathams rise connected below ocean to NZ.
- The climate is cool and wet with salt-laden winds all year round.

Chatham Islands (New Zealand)

Rangatira Island

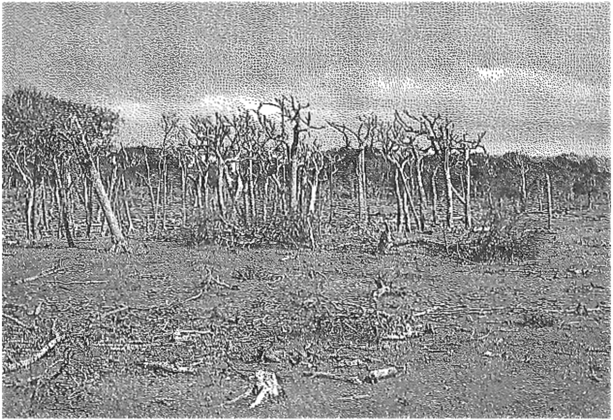
- is the third largest island in the Chatham Islands
- covers an area of 218 hectares (539 acres)
- is a gazetted nature reserve since 1953
- is now home to many endemic species
- is home to black robins – saved from near extinction by team led by Don Merton in the 1980s



A Chatham island Black Robin in Rangatira today

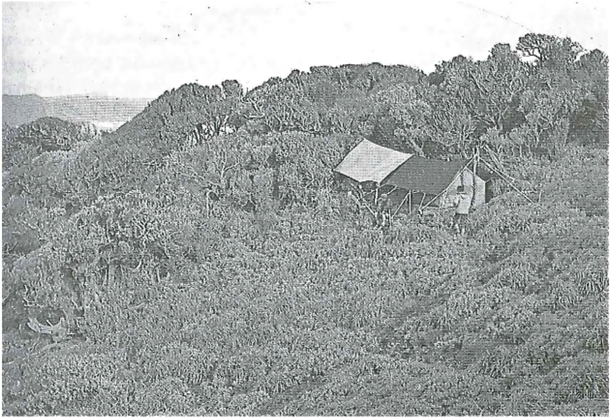
Decimation and Conservation

Whaling (< 1964), Sealing (< 1926) & Farming destroyed Ecosystems



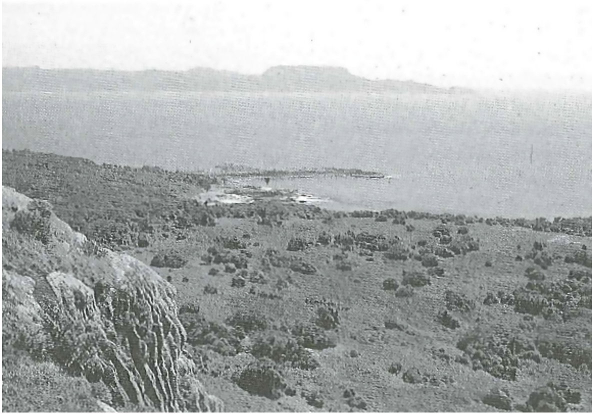
Decimation and Conservation

By 1938 black robins were wiped off from all the Chatham islands except for a few pairs on Little Mangare (a rocky outcrop of 15ha of bush)



Decimation and Conservation

Crown buys Rangatira and Magere Islands



Decimation and Conservation

A Massive Restoration Ecological Effort

Decimation and Conservation

Ecological Engineering of Native Flora and Predator-Proofing

Decimation and Conservation

Native fauna (sea-birds) start returning

Early Studies and Relocation in 1980s

Little Mangare Island

- Studies in 1970s by Douglas Flack on Little Mangare (15ha) tracked a handful of breeding pairs and
- recommended transloaction to nearby Mangere (113ha) and Rangatira islands (218ha)

Species Recovery Program in 1980s

Last 7 robins were moved.

Species Recovery Program in 1980s

Eggs were removed after being laid to encourage egg-laying

They were hatched and raised by foster species (tomtits)

Species Recovery Program in 1980s

the foster-cared chicks were transferred to black robin nests for sexual imprinting by its own species

Species Recovery Program in 1980s

Managements ceased when the population reached about 100

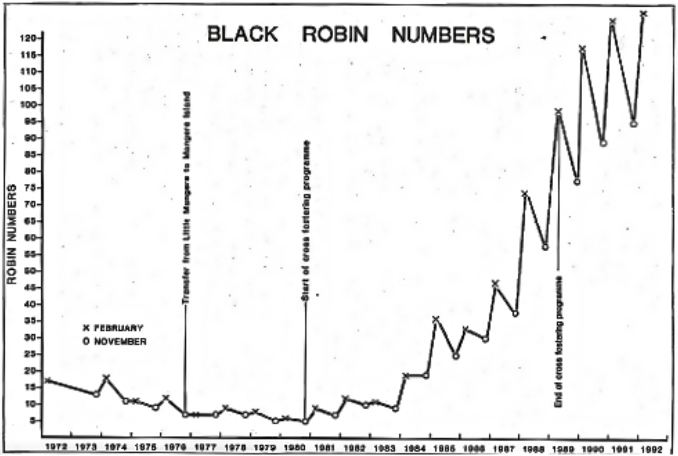


Fig. 22. Black robin numbers

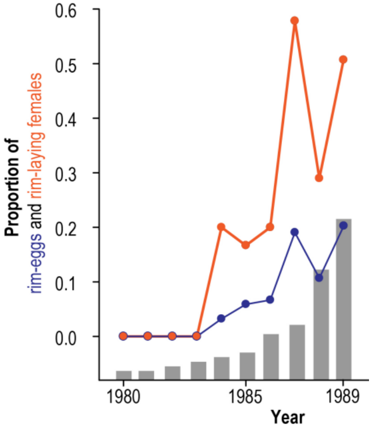
Species Recovery Program in 1980s

Well-laid Eggs

Rim-laid Egg

Unfortunately, a maladaptive bavioural trait of “rim-laying” also increased with the population

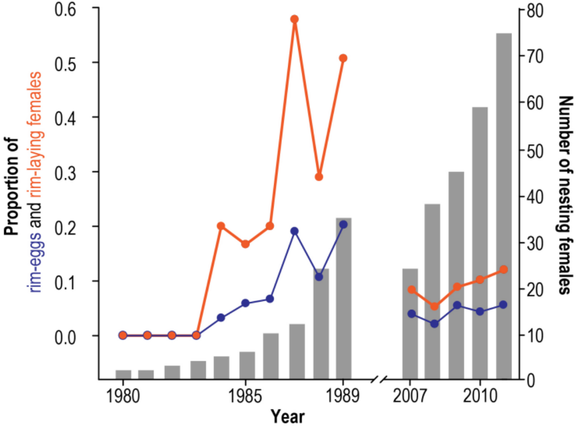
The Conservation Dilemma



Rim-laying increases in 1980s

- Conservation Dilemma – when to cease management?
- Need managed pop. size to ↑
- But rim-laying trait also ↑
- Don Merton's team ceased management when 50% of females were rim-layers by 1989
- Our Question: Is rim-laying a heritable trait?

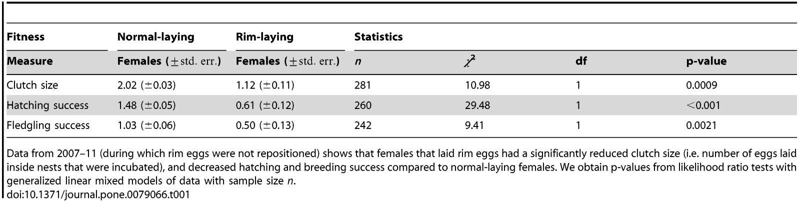
The Conservation Dilemma



Rim-laying increases in 1980s

but between 2007 and 2011 (unmanaged phase) rim-laying is lower

The Conservation Dilemma



Fitness	**Normal-laying**	**Rim-laying**	**Statistics**			
Measure	**Females** (±std. err.)	**Females** (±std. err.)	*n*	χ^2	**df**	**p-value**
Clutch size	2.02 (±0.03)	1.12 (±0.11)	281	10.98	1	0.0009
Hatching success	1.48 (±0.05)	0.61 (±0.12)	260	29.48	1	<0.001
Fledgling success	1.03 (±0.06)	0.50 (±0.13)	242	9.41	1	0.0021

Data from 2007–11 (during which rim eggs were not repositioned) shows that females that laid rim eggs had a significantly reduced clutch size (i.e. number of eggs laid inside nests that were incubated), and decreased hatching and breeding success compared to normal-laying females. We obtain p-values from likelihood ratio tests with generalized linear mixed models of data with sample size *n*.

doi:10.1371/journal.pone.0079066.t001

Rim-laying females have reduced

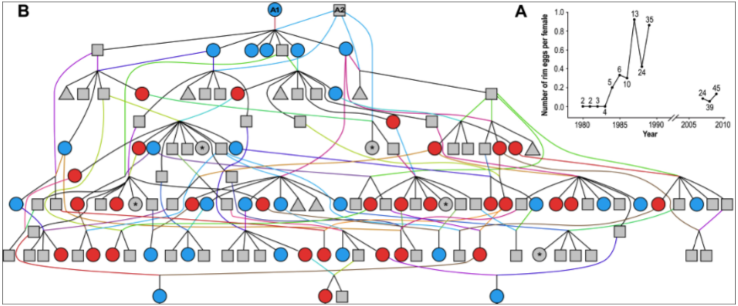
- clutch size
- hatching success
- fledgling success

during 2007-2011 (unmanaged phase)

Our Hypothesis: If rim-laying is heritable then natural selection would purge it from the population during unmanaged phase.

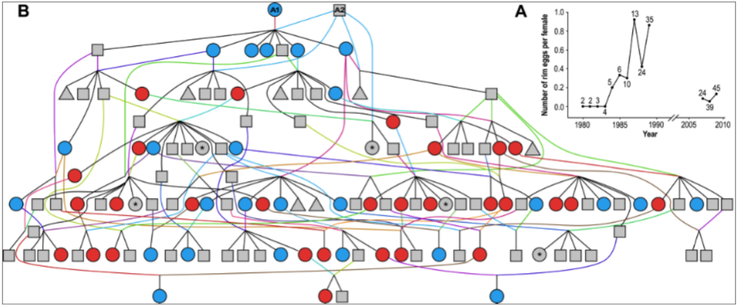
Need to show: rim-laying is heritable

Highly Looped Black Robin Pedigree in 1980s



cicle = female, square = male, traingle = unknown sex,
red circle = rim-layer, blue circle = non-rim-layer

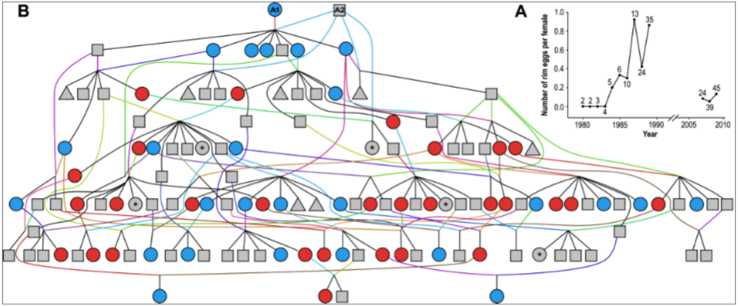
Highly Looped Black Robin Pedigree in 1980s



cicle = female, square = male, traingle = unknown sex,
red circle = rim-layer, blue circle = non-rim-layer

Want: The likelihood of the observed rim-laying phenotypes on this population pedigree.

Highly Looped Black Robin Pedigree in 1980s



cicle = female, square = male, traingle = unknown sex,

red circle = rim-layer, blue circle = non-rim-layer

Want: The likelihood of the observed rim-laying phenotypes on this population pedigree. Trouble: The likelihood via Peeling Algorithm (Cannings,Thompson & Skolnick 1978) in standard pedigree software runs out of RAM!

Likelihood of the Phenotypes on the Pedigree

For the likelihood we need:

1. simplest Model for Pedigree

Likelihood of the Phenotypes on the Pedigree

For the likelihood we need:

1. simplest Model for Pedigree
2. Phenotypes and Genotypes

Likelihood of the Phenotypes on the Pedigree

For the likelihood we need:

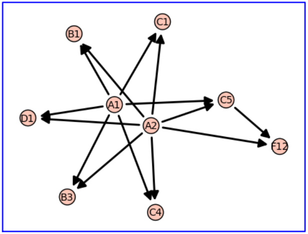
1. simplest Model for Pedigree
2. Phenotypes and Genotypes
3. Model for Mendelian Inheritance

Likelihood of the Phenotypes on the Pedigree

For the likelihood we need:

1. simplest Model for Pedigree
2. Phenotypes and Genotypes
3. Model for Mendelian Inheritance
4. Likelihood = the proportion of joint genotypic configurations that are compatible with the observed phenotypes under the model of inheritance on the observed pedigree

Simple Pedigree

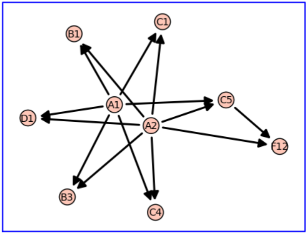


A pedigree is a directed acyclic graph $\mathcal{G} := (\mathbb{V}, \mathbb{E})$ in which $\mathbb{V}$ partitions a set of individuals (males,females,unknown-sex):

$$\mathbb{V} = \mathbb{M} \cup \mathbb{F} \cup \mathbb{U}$$

and where each vertex either has no-incoming arc or two incoming arcs, with one from a vertex in $\mathbb{M}$ and the other from a vertex in $\mathbb{F}$.

Simple Pedigree



Let the female and male founders be

$$\mathbb{A} := \{\mathrm{A1}, \mathrm{A2}\}$$

For an individual

$$u \in \mathbb{V} \setminus \mathbb{A}, \text{ let } M_u \text{ and } F_u$$

denote its mother and father, respectively.

Phenotypes

Phenotypes

Let the set of females who laid eggs be:

$$\mathbb{L} \subset \mathbb{F}$$

Phenotypes

Let the set of females who laid eggs be:

$$\mathbb{L} \subset \mathbb{F}$$

Let $\phi : \mathbb{L} \to \{0, 1\}$ denote a binary phenotype map:

$$\text{for every } u \in \mathbb{L}, \quad \phi(u) = \begin{cases} 1 & \text{if } u \text{ lays rim eggs} \\ 0 & \text{if } u \text{ lays normal eggs} \end{cases} \tag{1}$$

Phenotypes

Let the set of females who laid eggs be:

$$\mathbb{L} \subset \mathbb{F}$$

Let $\phi : \mathbb{L} \to \{0, 1\}$ denote a binary phenotype map:

$$\text{for every } u \in \mathbb{L}, \quad \phi(u) = \begin{cases} 1 & \text{if } u \text{ lays rim eggs} \\ 0 & \text{if } u \text{ lays normal eggs} \end{cases} \tag{1}$$

Let the phenotypes of a set of individuals

$$\mathbb{U} := \{u_1, \ldots, u_m\}$$

be

$$\phi(\mathbb{U}) := \{\phi(u) : u \in \mathbb{U}\} = \{\phi(u_1), \ldots, \phi(u_m)\}$$

Genotypes

Genotypes

The set of ordered biallelic genotypes of an individual is:

$$\Gamma = \{(a, a), (a, A), (A, a), (A, A)\}$$

Genotypes

The set of ordered biallelic genotypes of an individual is:

$$\Gamma = \{(a, a), (a, A), (A, a), (A, A)\}$$

We denote the genotype of an individual u by

$$\gamma(u) \in \Gamma$$

Genotypes

The set of ordered biallelic genotypes of an individual is:

$$\Gamma = \{(a, a), (a, A), (A, a), (A, A)\}$$

We denote the genotype of an individual u by

$$\gamma(u) \in \Gamma$$

and its maternal and paternal copies by

$$\gamma_m(u) \text{ and } \gamma_f(u), \text{ i.e., } \gamma(u) = (\gamma_m(u), \gamma_f(u))$$

Genotypes

The set of ordered biallelic genotypes of an individual is:

$$\Gamma = \{(a, a), (a, A), (A, a), (A, A)\}$$

We denote the genotype of an individual u by

$$\gamma(u) \in \Gamma$$

and its maternal and paternal copies by

$$\gamma_m(u) \text{ and } \gamma_f(u), \text{ i.e., } \gamma(u) = (\gamma_m(u), \gamma_f(u))$$

Let the genotypes of a set of individuals $\mathbb{U} := \{u_1, \ldots, u_m\}$ be

$$\gamma(\mathbb{U}) := \{\gamma(u) : u \in \mathbb{U}\} = \{\gamma(u_1), \ldots, \gamma(u_m)\} \in \Gamma^{\mathbb{U}}$$

Genotypes

The set of ordered biallelic genotypes of an individual is:

$$\Gamma = \{(a,a),(a,A),(A,a),(A,A)\}$$

We denote the genotype of an individual u by

$$\gamma(u) \in \Gamma$$

and its maternal and paternal copies by

$$\gamma_m(u) \text{ and } \gamma_f(u), \text{ i.e., } \gamma(u) = (\gamma_m(u), \gamma_f(u))$$

Let the genotypes of a set of individuals $\mathbb{U} := \{u_1, \ldots, u_m\}$ be

$$\gamma(\mathbb{U}) := \{\gamma(u) : u \in \mathbb{U}\} = \{\gamma(u_1), \ldots, \gamma(u_m)\} \in \Gamma^{\mathbb{U}}$$

and that of the founders be

$$\gamma(\mathbb{A}) \in \Gamma^{\mathbb{A}}$$

Genotypes

The set of ordered biallelic genotypes of an individual is:

$$\Gamma = \{(a, a), (a, A), (A, a), (A, A)\}$$

We denote the genotype of an individual u by

$$\gamma(u) \in \Gamma$$

and its maternal and paternal copies by

$$\gamma_m(u) \text{ and } \gamma_f(u), \text{ i.e., } \gamma(u) = (\gamma_m(u), \gamma_f(u))$$

Let the genotypes of a set of individuals $\mathbb{U} := \{u_1, \ldots, u_m\}$ be

$$\gamma(\mathbb{U}) := \{\gamma(u) : u \in \mathbb{U}\} = \{\gamma(u_1), \ldots, \gamma(u_m)\} \in \Gamma^{\mathbb{U}}$$

and that of the founders be

$$\gamma(\mathbb{A}) \in \Gamma^{\mathbb{A}}$$

So, with $\sim$ 100 individuals in 1980-89: $|\Gamma^{\mathbb{V}}| = 4^{|\mathbb{V}|} > 10^{60}$

Mendelian Inheritance – Prob. of Genotype

Let $\mathcal{M}_h^c$ denote a family of four models of Mendelian inheritance

Mendelian Inheritance – Prob. of Genotype

Let $\mathcal{M}_h^c$ denote a family of four models of Mendelian inheritance where $h \in \{D, R\}$ allows for a dominant (D) allele A or for a recessive (R) allele a

Mendelian Inheritance – Prob. of Genotype

Let $\mathcal{M}_h^c$ denote a family of four models of Mendelian inheritance where $h \in \{D, R\}$ allows for a dominant (D) allele A or for a recessive (R) allele a and $c \in \{\alpha, z\}$ allows for the locus to be on an autosome (α) or the sex chromosome (z).

Mendelian Inheritance – Prob. of Genotype

Let $\mathcal{M}_h^c$ denote a family of four models of Mendelian inheritance where $h \in \{D, R\}$ allows for a dominant (D) allele A or for a recessive (R) allele a and $c \in \{\alpha, z\}$ allows for the locus to be on an autosome (α) or the sex chromosome (z).

$$\Pr\left(\gamma(u) \mid \gamma(M_u), \gamma(F_u), \mathcal{M}_h^c\right) = \begin{cases} \frac{1}{4} & \text{if } c = \alpha, \gamma_m(u) \in \{\gamma_m(M_u), \gamma_f(M_u)\}, \\ & \quad \gamma_f(u) \in \{\gamma_m(F_u), \gamma_f(F_u)\} \\ \frac{1}{2} & \text{if } c = z, \gamma_m(u) = \gamma_f(M_u), \\ & \quad \gamma_f(u) \in \{\gamma_m(F_u), \gamma_f(F_u)\}, u \in \mathbb{M} \\ \frac{1}{2} & \text{if } c = z, \gamma_m(u) = \gamma_m(M_u), \\ & \quad \gamma_f(u) \in \{\gamma_m(F_u), \gamma_f(F_u)\}, u \in \mathbb{F} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Mendelian Inheritance – Prob. of Phenotype

Mendelian Inheritance – Prob. of Phenotype

Finally the probability of $\phi(u)$, the phenotype for an individual u,

Mendelian Inheritance – Prob. of Phenotype

Finally the probability of $\phi(u)$, the phenotype for an individual u, given its genotype $\gamma(u)$

Mendelian Inheritance – Prob. of Phenotype

Finally the probability of $\phi(u)$, the phenotype for an individual u, given its genotype $\gamma(u)$ and the model of inheritance $\mathcal{M}_h^c$ is the following:

Mendelian Inheritance – Prob. of Phenotype

Finally the probability of $\phi(u)$, the phenotype for an individual u, given its genotype $\gamma(u)$ and the model of inheritance $\mathcal{M}_h^c$ is the following:

$$\Pr\left(\phi(u) \mid \gamma(u), \mathcal{M}_h^c\right) = \begin{cases} 1 & \text{if } \phi(u) = 1, \gamma(u) \in \Gamma \setminus \{(a,a)\}, h = D, u \in \mathbb{L} \\ 1 & \text{if } \phi(u) = 0, \gamma(u) \in \{(a,a)\}, h = D, u \in \mathbb{L} \\ 1 & \text{if } \phi(u) = 0, \gamma(u) \in \Gamma \setminus \{(a,a)\}, h = R, u \in \mathbb{L} \\ 1 & \text{if } \phi(u) = 1, \gamma(u) \in \{(a,a)\}, h = R, u \in \mathbb{L} \\ 1 & \text{if } u \notin \mathbb{L} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Likelihood, Prior and Posterior

We want probability of the inheritance model $\mathcal{M}_h^c$

Likelihood, Prior and Posterior

We want probability of the inheritance model $\mathcal{M}_h^c$ and the founder genotypes $\gamma(\mathbb{A})$

Likelihood, Prior and Posterior

We want probability of the inheritance model $\mathcal{M}_h^c$ and the founder genotypes $\gamma(\mathbb{A})$ given $\mathcal{G}$, the observed pedigree,

Likelihood, Prior and Posterior

We want probability of the inheritance model $\mathcal{M}_h^c$ and the founder genotypes $\gamma(\mathbb{A})$ given $\mathcal{G}$, the observed pedigree, and $\phi(\mathbb{V})$, the observed phenotypes over the nodes in $\mathcal{G}$:

Likelihood, Prior and Posterior

We want probability of the inheritance model $\mathcal{M}_h^c$ and the founder genotypes $\gamma(\mathbb{A})$ given $\mathcal{G}$, the observed pedigree, and $\phi(\mathbb{V})$, the observed phenotypes over the nodes in $\mathcal{G}$:

$$\begin{aligned}\Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A}) \mid \mathcal{G}, \phi(\mathbb{V})\right) \\ = \frac{\Pr\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right) \Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A})\right)}{\sum_{\substack{c \in \{a,z\} \\ h \in \{D,R\} \\ \gamma(\mathbb{A}) \in \Gamma^{\mathbb{A}}}} \Pr\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right) \Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A})\right)}\end{aligned} \tag{4}$$

Likelihood, Prior and Posterior

We want probability of the inheritance model $\mathcal{M}_h^c$ and the founder genotypes $\gamma(\mathbb{A})$ given $\mathcal{G}$, the observed pedigree, and $\phi(\mathbb{V})$, the observed phenotypes over the nodes in $\mathcal{G}$:

$$\begin{aligned}
&\Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A}) \mid \mathcal{G}, \phi(\mathbb{V})\right) \\
&\qquad = \frac{\Pr\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right) \Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A})\right)}{\sum_{\substack{c \in \{a,z\} \\ h \in \{D,R\} \\ \gamma(\mathbb{A}) \in \Gamma^{\mathbb{A}}}} \Pr\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right) \Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A})\right)}
\end{aligned} \tag{4}$$

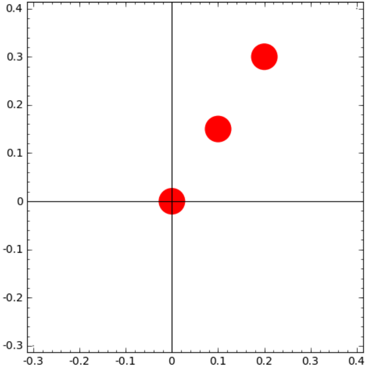
We assume a uniform prior probability $\Pr\left(\mathcal{M}_h^c, \gamma(\mathbb{A})\right) = \frac{1}{64}$ over the $2 \times 2 \times 4^2 = 64$ models in the family:

$$\left\{ \left(\mathcal{M}_h^c, \gamma(\mathbb{A})\right) : c \in \{a, z\}, h \in \{D, R\}, \gamma(\mathbb{A}) \in \Gamma^{\mathbb{A}} \right\} \ .$$

Likelihood Computation - A Pedagogical Example

Sequential Monte Carlo (SMC) Algorithm for Rare Event Simulation

(Johansen, Del Moral & Doucet, 2006)

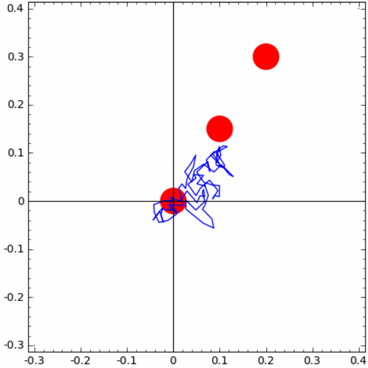


What is Prob. of Hitting First Red dot and then Second Red Dot?

Likelihood Computation - A Pedagogical Example

Sequential Monte Carlo (SMC) Algorithm for Rare Event Simulation

(Johansen, Del Moral & Doucet, 2006)

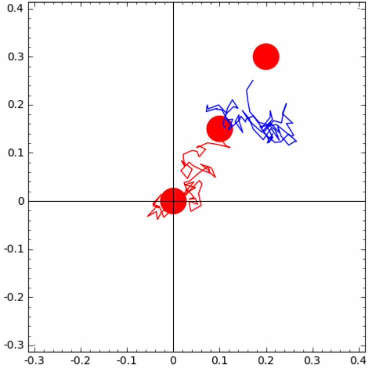


Estim. Prob. Hitting First Red dot $= 1/9$

Likelihood Computation - A Pedagogical Example

Sequential Monte Carlo (SMC) Algorithm for Rare Event Simulation

(Johansen, Del Moral & Doucet, 2006)



Estim. Prob. Hitting First and Second Red Dots $= 1/9 \times 1/9 = 1/81$

Likelihood Computation - Adapting SMC to Pedigrees

Likelihood Computation - Adapting SMC to Pedigrees

$\Pr\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right)$, the likelihood in (4), is computed using a sequential Monte Carlo Algorithm.

Likelihood Computation - Adapting SMC to Pedigrees

Pr $\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right)$, the likelihood in (4), is computed using a sequential Monte Carlo Algorithm.
The strategy involves sequentially growing the pedigree as a nested and increasing family of sub-graphs via a breadth-first expansion from the founders A1 and A2.

Likelihood Computation - Adapting SMC to Pedigrees

Pr $\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right)$, the likelihood in (4), is computed using a sequential Monte Carlo Algorithm.
The strategy involves sequentially growing the pedigree as a nested and increasing family of sub-graphs via a breadth-first expansion from the founders A1 and A2.
Let $\mathcal{G}[\mathbb{V}_t]$ denote the sub-pedigree of $\mathcal{G}$ induced by vertices $\mathbb{V}_t$.

Likelihood Computation - Adapting SMC to Pedigrees

$\Pr\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right)$, the likelihood in (4), is computed using a sequential Monte Carlo Algorithm.

The strategy involves sequentially growing the pedigree as a nested and increasing family of sub-graphs via a breadth-first expansion from the founders A1 and A2.

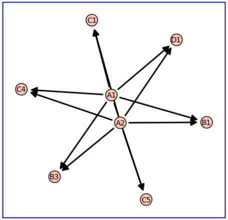
Let $\mathcal{G}[\mathbb{V}_t]$ denote the sub-pedigree of $\mathcal{G}$ induced by vertices $\mathbb{V}_t$.

Let $\mathbb{V}_0 = \{\text{A1}, \text{A2}\}$ be the vertex set of the founding male and female.

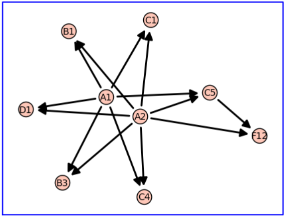
Likelihood Computation - Adapting SMC to Pedigrees

Pr $\left(\phi(\mathbb{V}) \mid \mathcal{G}, \mathcal{M}_h^c, \gamma(\mathbb{A})\right)$, the likelihood in (4), is computed using a sequential Monte Carlo Algorithm.
The strategy involves sequentially growing the pedigree as a nested and increasing family of sub-graphs via a breadth-first expansion from the founders A1 and A2.
Let $\mathcal{G}[\mathbb{V}_t]$ denote the sub-pedigree of $\mathcal{G}$ induced by vertices $\mathbb{V}_t$.
Let $\mathbb{V}_0 = \{\text{A1}, \text{A2}\}$ be the vertex set of the founding male and female.
Let $\mathbb{A} = \mathbb{V}_0 \subset \mathbb{V}_1 \subset \cdots \subset \mathbb{V}_T = \mathbb{V}$ denote the increasing sequence of vertices in $\mathcal{G}$ obtained by a breadth-first expansion from $\mathbb{V}_0$

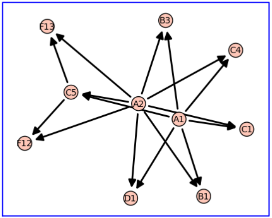
Sequential Monte Carlo on Nested Combinatorial Structures



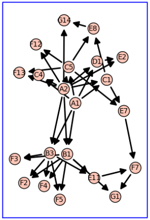
→



↓



← · · ·

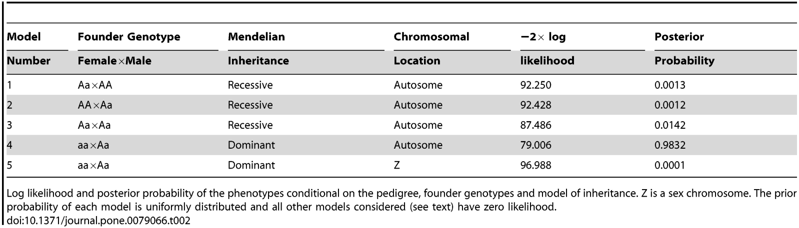


← · · ·

Sequential Monte Carlo on Nested Combinatorial Structures

- At $t = 0$, initialize all particles at the founder nodes
  $\texttt{FOR}\ i = 1, \ldots, N;\ \texttt{DO}$:
  - $\gamma_0^{(i)}(u) \leftarrow \gamma(u)$, for each founder $u \in \mathbb{V}_0 = \mathbb{A}$
  - $W_0^{(i)} \leftarrow \prod_{u \in \mathbb{V}_0} \Pr(\phi(u) \mid \gamma_0^{(i)}(u), \mathcal{M}_h^c)$
- $\texttt{FOR}\ t = 1, \ldots, T;\ \texttt{DO}$:
  - Resample to obtain $\left\{ \frac{1}{N}, \left(\hat{\gamma}_{t-1}^{(i)}(u) : u \in \mathbb{V}_{t-1} \right) \right\}_{i=1}^{N}$
  - $\texttt{FOR}\ i = 1, \ldots, N;\ \texttt{DO}$:
    - Extend $\left\{ \hat{\gamma}_{t-1}^{(i)}(u) : u \in \mathbb{V}_{t-1} \right\}$ to $\left\{ \gamma_t^{(i)}(u) : u \in \mathbb{V}_t \right\}$ by proposing genotypes for each new offspring node in $\mathbb{V}_t \setminus \mathbb{V}_{t-1}$ according to $\mathcal{M}_h^c$ and the genotypes of its parent nodes in $\mathbb{V}_{t-1}$ specified by the sub-pedigree $\mathcal{G}[\mathbb{V}_t]$.
    - $W_t^{(i)} \leftarrow \prod_{u \in \mathbb{V}_t \setminus \mathbb{V}_{t-1}} \Pr(\phi(u) \mid \gamma_t^{(i)}(u), \mathcal{M}_h^c)$
- Estimate $\Pr\left(\mathcal{G}, \phi(\mathbb{V}) \mid \mathcal{M}_j^c, \gamma(\mathbb{A})\right)$ by $\prod_{t=0}^{T} \hat{Z}_t$, $\hat{Z}_t = \frac{1}{N} \sum_{i=1}^{N} W_t^{(i)}$.

Model Selection – Results



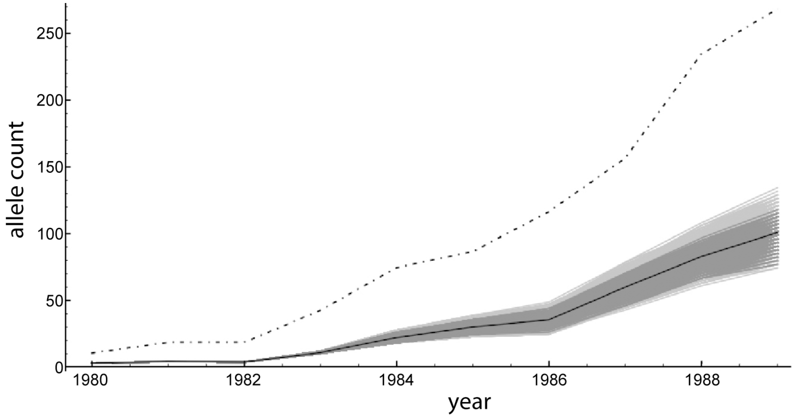
Model Number	Founder Genotype Female×Male	Mendelian Inheritance	Chromosomal Location	−2× log likelihood	Posterior Probability
1	Aa×AA	Recessive	Autosome	92.250	0.0013
2	AA×Aa	Recessive	Autosome	92.428	0.0012
3	Aa×Aa	Recessive	Autosome	87.486	0.0142
4	aa×Aa	Dominant	Autosome	79.006	0.9832
5	aa×Aa	Dominant	Z	96.988	0.0001

Log likelihood and posterior probability of the phenotypes conditional on the pedigree, founder genotypes and model of inheritance. Z is a sex chromosome. The prior probability of each model is uniformly distributed and all other models considered (see text) have zero likelihood.

doi:10.1371/journal.pone.0079066.t002

The best model (Model 4) is the simple autosomal dominant model with homozygous founding female (aa) and heterozygous founding male (Aa).

Allele Trajectory – Results



Inferred trajectories of allele A under Model 4 between 1980 and 1989. The mean trajectory (black); 50% and 95% confidence sets (dark & light grey); $2N$ (dash-dotted).

Phenotype Permutation Test of Heritability

- Under Null Hyp.: Environmental (non-genetic) Basis to Rim-Laying
- Permuting female phenotypes would leave the null likelihood unchanged
- Use likelihood under Model 4 as the test statistic
- p-value = fraction of 10000 phenotype permutations have a likelihood greater than the observed likelihood
- Result: p-value = 0.03

Conclusions - Part I

- Black robins were managed from just one successful breeding pair in the wild

Conclusions - Part I

- Black robins were managed from just one successful breeding pair in the wild
- The maladaptive rim-laying trait also increased with the populaiton in the absence of natural selection

Conclusions - Part I

- Black robins were managed from just one successful breeding pair in the wild
- The maladaptive rim-laying trait also increased with the populaiton in the absence of natural selection
- Such human-assisted spread of maladaptive traits poses the conservation dilemma

Conclusions - Part I

- Black robins were managed from just one successful breeding pair in the wild
- The maladaptive rim-laying trait also increased with the populaiton in the absence of natural selection
- Such human-assisted spread of maladaptive traits poses the conservation dilemma
- In the case of black robins the managers seemed to have succeeded by stoppping management when half the population were rim-layers

Conclusions - Part I

- Black robins were managed from just one successful breeding pair in the wild
- The maladaptive rim-laying trait also increased with the populaiton in the absence of natural selection
- Such human-assisted spread of maladaptive traits poses the conservation dilemma
- In the case of black robins the managers seemed to have succeeded by stoppping management when half the population were rim-layers
- Merton suspected rim-laying was heritable
  (persn. commn. to Massaro)

Conclusions - Part I contd...

- Our research group at Canterbury (with Briskie, Hale, Massaro and Poole) have developed tests to establish that rim-laying during 1980s is likely to be a heritable trait

Conclusions - Part I contd...

- Our research group at Canterbury (with Briskie, Hale, Massaro and Poole) have developed tests to establish that rim-laying during 1980s is likely to be a heritable trait
- We have also shown that the trait is at a low frequency in the unmanaged population between 2007-2011 perhaps due to strong natural selection since management ceased in 1990

Conclusions - Part I contd...

- Our research group at Canterbury (with Briskie, Hale, Massaro and Poole) have developed tests to establish that rim-laying during 1980s is likely to be a heritable trait
- We have also shown that the trait is at a low frequency in the unmanaged population between 2007-2011 perhaps due to strong natural selection since management ceased in 1990
- One can also develop a theory of periodic environmental extremes under which rim-layers will have an advantage by successfully raising only one chick as opposed to none (but difficult to test without long-term data)

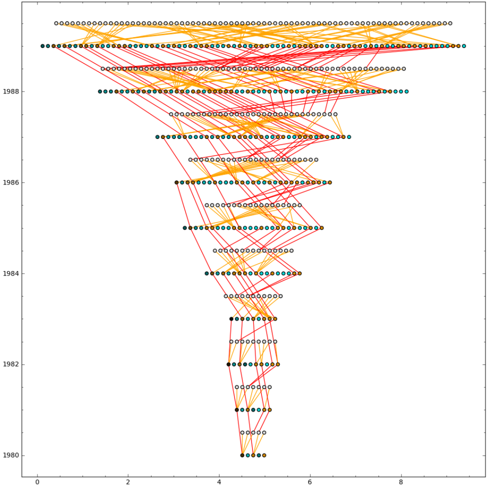
Conclusions - Part I contd...

- Our research group at Canterbury (with Briskie, Hale, Massaro and Poole) have developed tests to establish that rim-laying during 1980s is likely to be a heritable trait
- We have also shown that the trait is at a low frequency in the unmanaged population between 2007-2011 perhaps due to strong natural selection since management ceased in 1990
- One can also develop a theory of periodic environmental extremes under which rim-layers will have an advantage by successfully raising only one chick as opposed to none (but difficult to test without long-term data)
- To appreciate the intensity of field operations and limitations of biomathematical models read: *The black robin: saving the world's most endangered bird*, David Butler and Don Merton, Oxford University Press, 1994.

Part IIa

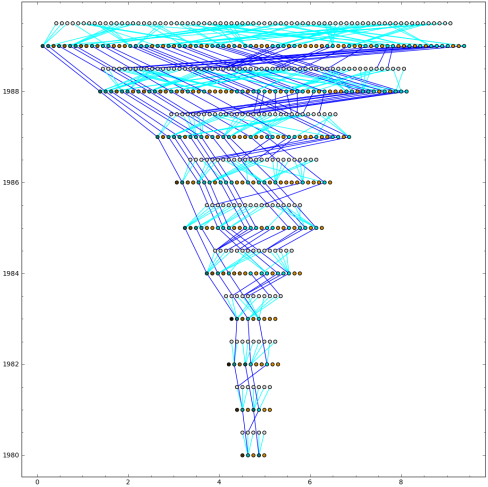
Insights from field work in November 2011

Temporal Pedigrees with Overlapping Generations



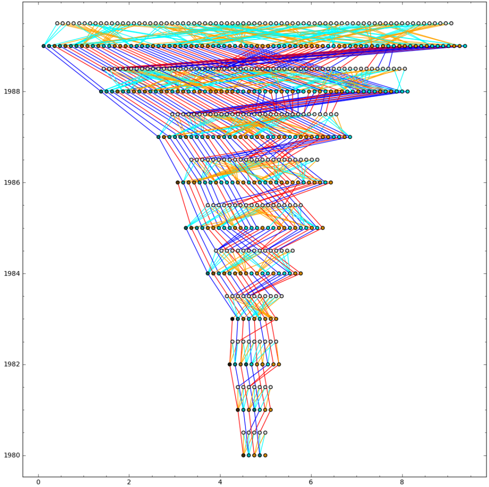
Female Population Tree during Field Conservation in 1980s

Temporal Pedigrees with Overlapping Generations



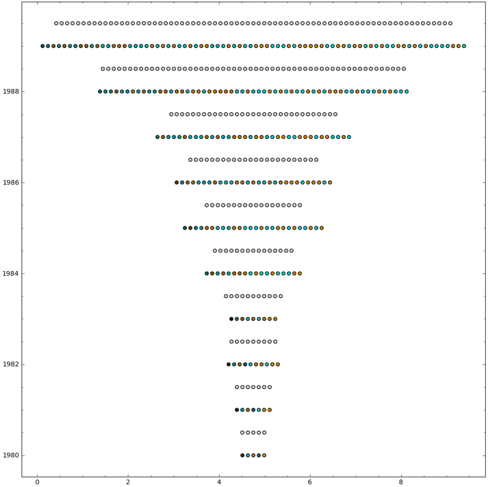
Male Population Tree during Field Conservation in 1980s

Temporal Pedigrees with Overlapping Generations



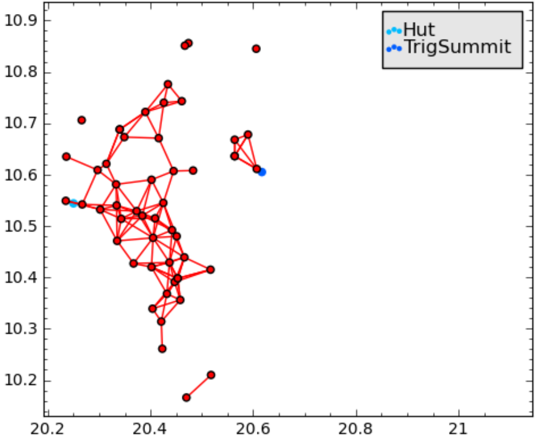
Population Pedigree during Field Conservation in 1980s

Temporal Pedigrees with Overlapping Generations



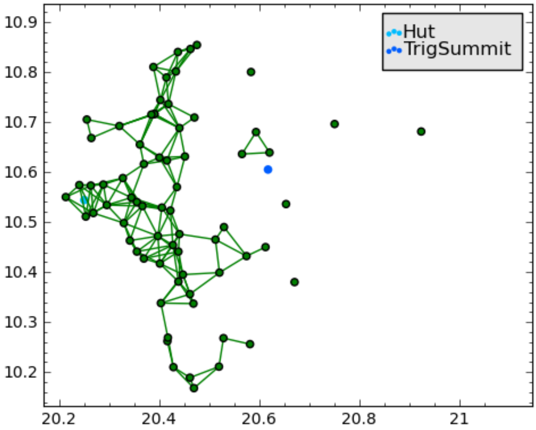
Coalescent Sub-Pedigree of the 1989 population $\Longrightarrow$ Part IIb

Male Territory Graphs by year



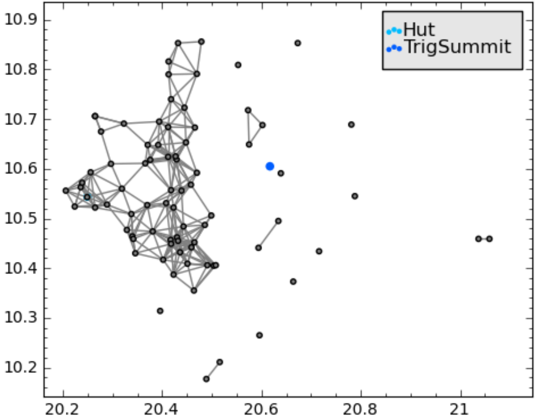
2009 (eye-balled by a ball graph)

Male Territory Graphs by year



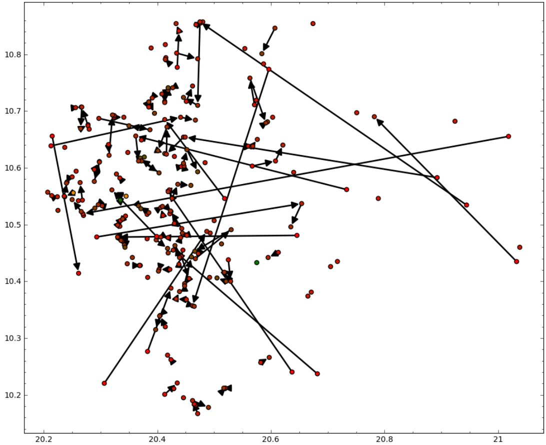
2010 (eye-balled by a ball graph)

Male Territory Graphs by year



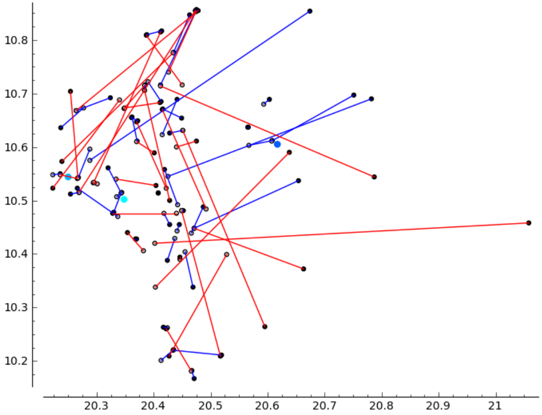
2011 (eye-balled by a ball graph)

Male Territorial Dynamics – yearly nest site relocation



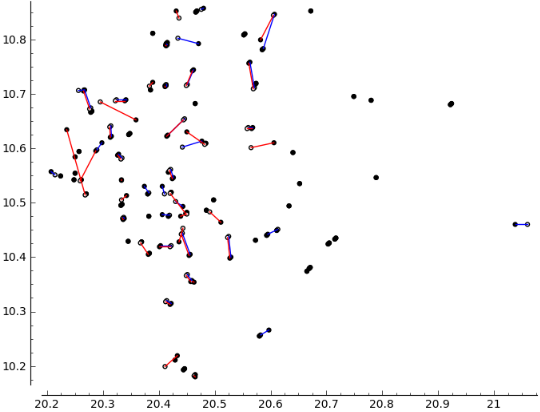
Male / Female yearly nest to nest distance

sex-specific distance between home nest and first nest

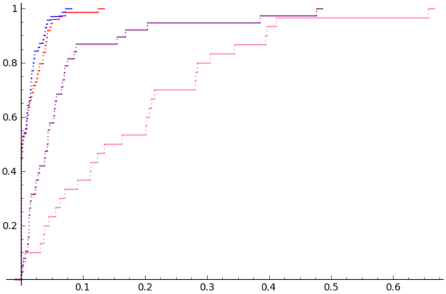


Male / Female yearly nest to nest distance

sex-specific distance between nest at age $\geq$ 1 and next nest



Male / Female yearly nest to nest distance

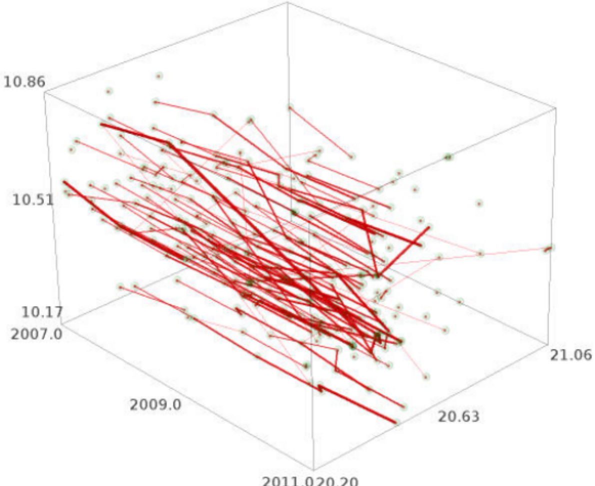


- Permutation Test to Reject H_0 : male dist = female dist (10^5 MC samples)
- Strongly reject H_0 for home to first nest (pv = 0.00025)
- Fail to reject H_0 for nest at age ≥ 1 to next nest (pv = 0.14)

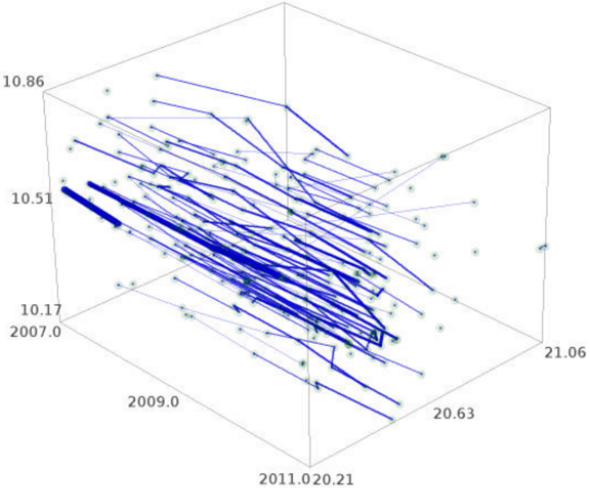
male / female distance from home nest to first nest

male / female distance from nest at age $\geq$ 1 to next nest

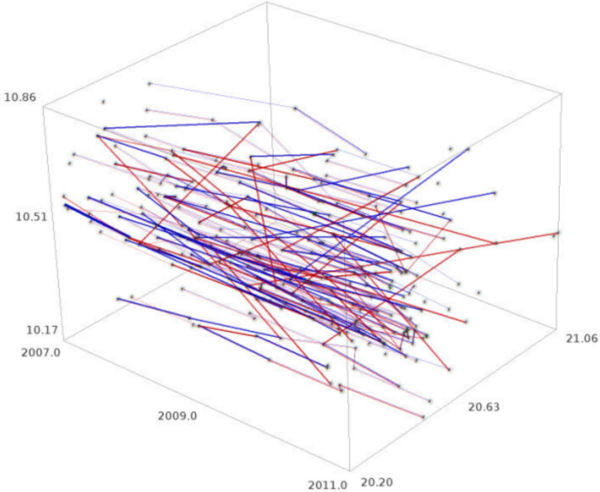
Need Models of Spatio-temporal Behavioural Population Pedigrees



Need Models of Spatio-temporal Behavioural Population Pedigrees

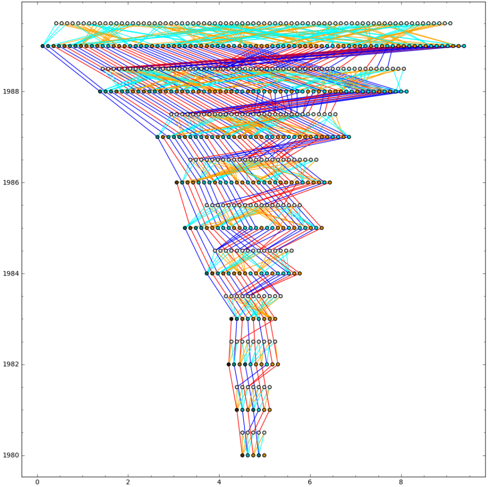


Need Models of Spatio-temporal Behavioural Population Pedigrees

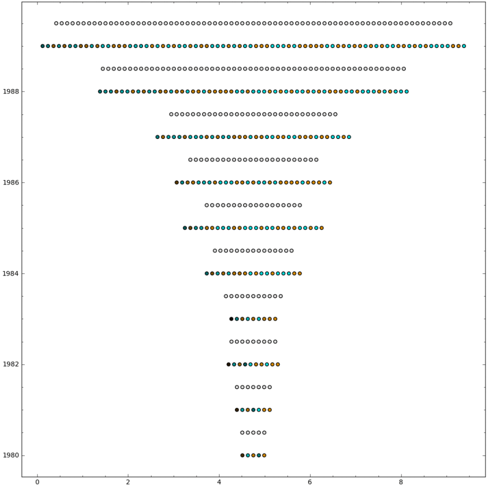


Part IIb
Ancestries of Recombining Diploid Populations

Raazesh Sainudiin, Bhalchandra Thatte, and Amandine Véber,
Journal of Mathematical Biology, 46 pages, 2015



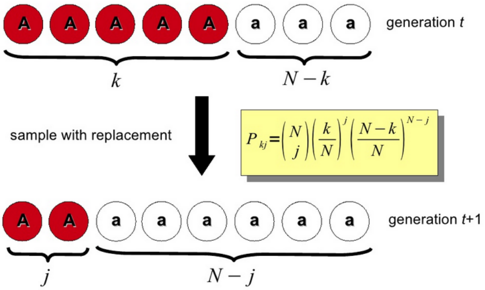
Motivation Part II: Diploid Population Pedigree (forwards in time)



Coalescent Sub-Pedigree (backwards in time) “inheritance topology of diploid zygotes?”

Three Notions of Molecular continuums within W-F Population Pedigrees

One *r* to rule them all!

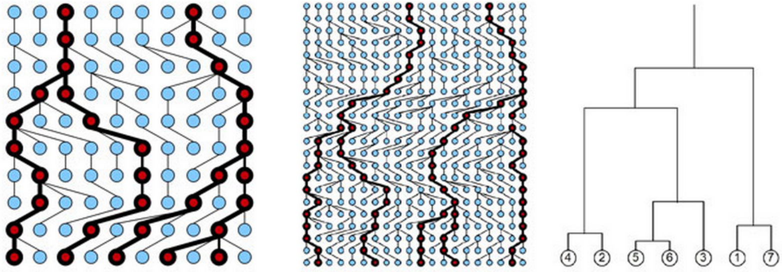


Wright-Fisher Model (forwards in time)

where $r = 0$ is the probability of recombination per locus per meiotic generation

Three Notions of Molecular continuums within W-F Population Pedigrees

One *r* to rule them all!

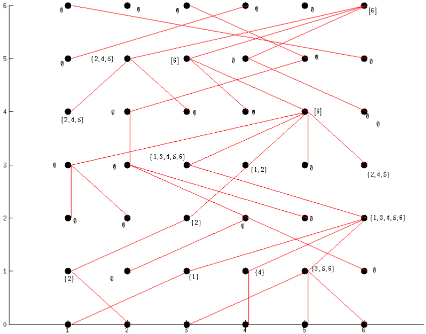


Wright-Fisher Genealogy with Large N $\Longrightarrow$ Kingman's Coalescent Tree for small n (backwards in time)

where r is the probability of recombination per locus per meiotic generation ($r = 0$ for Kingman's model)

Three Notions of Molecular continuums within W-F Population Pedigrees

One *r* to rule them all!

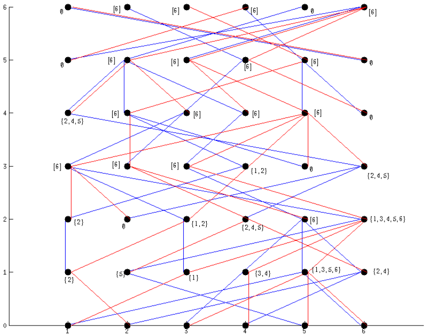


(1) Kingman's Discrete Genealogy at a non-recombining locus with $r = 0$ (forwards in time)

Observe: Time to most recent ancestor (TMRCA) = 4 (can you locate the MRCA?)

Three Notions of Molecular continuums within W-F Population Pedigrees

One *r* to rule them all!



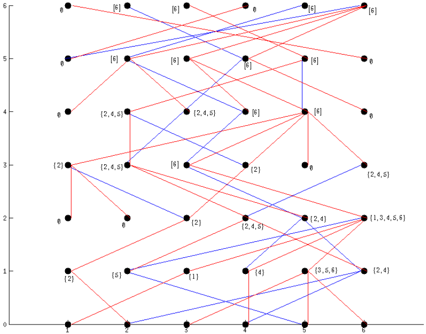
(2) Chang's Discrete Pedigree as the Genealogy of an always-recombining locus with $r = 1$ (forwards in time)

Observe: Time to most recent ancestor (TMRCA) = 2 (can you locate the MRCA? how about other CAs?)

Observe: Time to be or not to be (TTBONTB) = 5 (can you say why?)

Three Notions of Molecular continuums within W-F Population Pedigrees

One *r* to rule them all!



(3) Hudson-Griffiths Discrete Ancestral Recombination Graph at a recombining locus with $0 < r < 1$ (fwds in time)

SIMPLEST UNIFICATION PROBLEM: What is TMRCA and TTBONTB for any *r*?

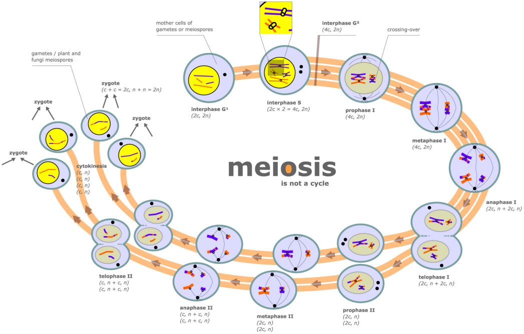
Unifying these three notions of molecular ancestry

Consider the **ancestral graphs** of these three genealogical processes for a given sample taken today (backwards in time).

- Uni-parental Ancestry of sample : Kingman's Discrete Genealogy at a non-recombining locus (1982)
- Bi-parental Ancestry of sample : Chang's Discrete Pedigree (1999)
- Uni/Bi-parental Ancestry of sample: Hudson-Griffiths Discrete Ancestral Recombination Graph (1983-1989)

Unifying these three notions of molecular ancestry

consistent with the zygotic topology in sync with meiosis:



This unification was an open problem since 1983

This topologically consistent notion of diploid ancestry (RS, Thatte & Veber, 2015) in a common probability space is a pre-requisite for more realistic models of population pedigree processes, including:

- non-overlapping generations

This unification was an open problem since 1983

This topologically consistent notion of diploid ancestry (RS, Thatte & Veber, 2015) in a common probability space is a pre-requisite for more realistic models of population pedigree processes, including:

- non-overlapping generations
- dioecious (non-hermaphroditic) diploidy

This unification was an open problem since 1983

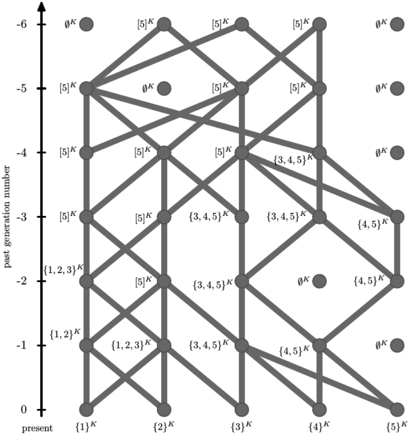
This topologically consistent notion of diploid ancestry (RS, Thatte & Veber, 2015) in a common probability space is a pre-requisite for more realistic models of population pedigree processes, including:

- non-overlapping generations
- dioecious (non-hermaphroditic) diploidy
- gender-specific behavioural effects on spatio-temporal pedigree through which genes flow under natural selection...

This unification was an open problem since 1983

This topologically consistent notion of diploid ancestry (RS, Thatte & Veber, 2015) in a common probability space is a pre-requisite for more realistic models of population pedigree processes, including:

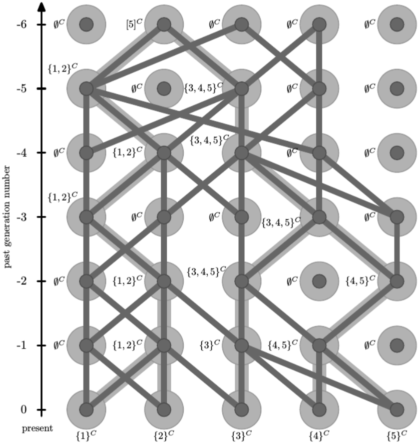
- non-overlapping generations
- dioecious (non-hermaphroditic) diploidy
- gender-specific behavioural effects on spatio-temporal pedigree through which genes flow under natural selection...
- etc. (think of Florida Scrub Jay, feral cat populations, heroms of Mujahidiin warlords post Soviet withdrawal in 1980s)



karyotic ancestral graph of a bi-parental Wright-Fisher population with five eukaryotic individuals

Can you find the MRC-karyotic-A (MRCkA)? TMRCkA? TT-karyotically-BONTB (TTkBONTB)?

Chang's Main results say that TMRCkA and TTkBONTB $\sim O(\log_2(n))$, where $n =$ population size

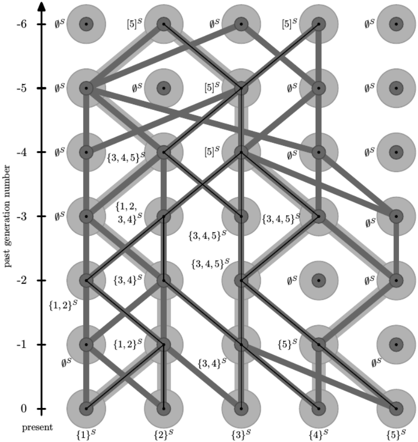


Zygotic pedigree containing the cytoplasmic tree (Kingman's discrete coalecent) and karyotic ancestral graph of a bi-parental Wright-Fisher population

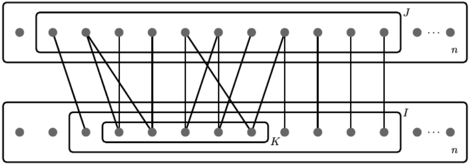
Can you find the MRC-cytoplasmic-A (MRCcA)? TMRCcA?

Kingman's Main results say that TMRCcA and TTcBONTB $\sim O(n)$, where $n =$ population size

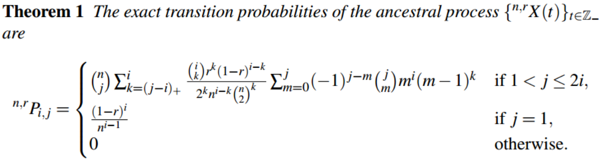
so: $O(\log_2(n))$ versus $O(n)$ is a huge difference!



Putting it all together: Zygotic, cytoplasmic, karyotic and sub-karyotic ancestral graphs (discrete ARGs of Hudson and Griffiths) of a bi-parental Wright-Fisher population with five eukaryotic diploid individuals



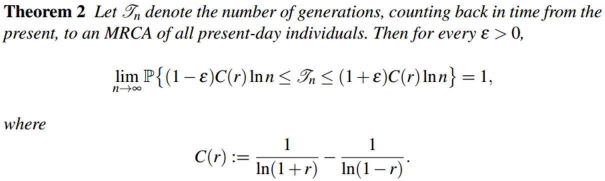
Combinatorial Structure Diagram



Theorem 1 *The exact transition probabilities of the ancestral process* $\{{}^{n,r}X(t)\}_{t\in\mathbb{Z}_-}$ *are*

$$
{}^{n,r}P_{i,j} = \begin{cases} \binom{n}{j}\sum_{k=(j-i)_+}^{i} \frac{\binom{i}{k}r^k(1-r)^{i-k}}{2^k n^{i-k}\binom{n}{2}^k} \sum_{m=0}^{j}(-1)^{j-m}\binom{j}{m}m^i(m-1)^k & \text{if } 1<j\leq 2i, \\ \frac{(1-r)^i}{n^{i-1}} & \text{if } j=1, \\ 0 & \text{otherwise.} \end{cases}
$$

Kingman-Chang-Hudson-Griffiths Unified Ancestral Chain



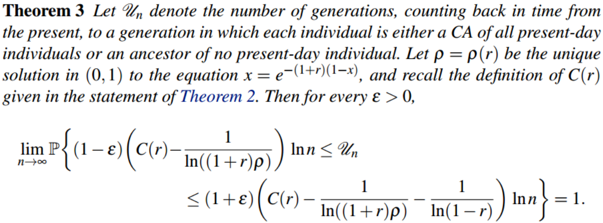
Theorem 2 *Let $\mathscr{T}_n$ denote the number of generations, counting back in time from the present, to an MRCA of all present-day individuals. Then for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left\{(1-\varepsilon)C(r)\ln n \le \mathscr{T}_n \le (1+\varepsilon)C(r)\ln n\right\} = 1,$$

where

$$C(r) := \frac{1}{\ln(1+r)} - \frac{1}{\ln(1-r)}.$$

Time to MRCA for a possibly recombining locus with intra-locus recombination probability per non-overlapping meiotic generation in a population of *n* diploid individuals.



Theorem 3 *Let* $\mathscr{U}_n$ *denote the number of generations, counting back in time from the present, to a generation in which each individual is either a CA of all present-day individuals or an ancestor of no present-day individual. Let* $\rho = \rho(r)$ *be the unique solution in* $(0,1)$ *to the equation* $x = e^{-(1+r)(1-x)}$*, and recall the definition of* $C(r)$ *given in the statement of Theorem 2. Then for every* $\varepsilon > 0$*,*

$$\lim_{n\to\infty} \mathbb{P}\left\{(1-\varepsilon)\left(C(r) - \frac{1}{\ln((1+r)\rho)}\right)\ln n \le \mathscr{U}_n \le (1+\varepsilon)\left(C(r) - \frac{1}{\ln((1+r)\rho)} - \frac{1}{\ln(1-r)}\right)\ln n\right\} = 1.$$

Time to Be or Not to Be for a possibly recombining locus with intra-locus recombination probability per non-overlapping meiotic generation in a population of *n* diploid individuals.

Conclusions - Part II

- Empirical Population Pedigree Processes should be informed by animal behavior

Conclusions - Part II

- Empirical Population Pedigree Processes should be informed by animal behavior
- Behavioral data for animal populations is far from idealized limiting probability models:
  random mating, non-overlapping generations,...

Conclusions - Part II

- Empirical Population Pedigree Processes should be informed by animal behavior
- Behavioral data for animal populations is far from idealized limiting probability models:
  random mating, non-overlapping generations,...
- Need “combinatorially rich enough” statistical models

Conclusions - Part II

- Empirical Population Pedigree Processes should be informed by animal behavior
- Behavioral data for animal populations is far from idealized limiting probability models:
  random mating, non-overlapping generations,...
- Need “combinatorially rich enough” statistical models
- Such models may pave a way for the **fusion** of behavioural ecology and evolutionary genetics
  A paradigm for "Population Structure From Behavior"

Conclusions - Part II

- Empirical Population Pedigree Processes should be informed by animal behavior
- Behavioral data for animal populations is far from idealized limiting probability models:
  random mating, non-overlapping generations,...
- Need “combinatorially rich enough” statistical models
- Such models may pave a way for the **fusion** of behavioural ecology and evolutionary genetics
  A paradigm for "Population Structure From Behavior"
- PLAN: continue to work on Part III with FSJ @Cornell (2014–2016)

Part III

Behavioural Ecological Models of Population Pedigrees

current work with Nancy Chen, Andrew Clark and John Fitzpatrick
(Cornell Univ., USA) just blabber straight-up

Thank you!

Funds for Theory:

- School of Maths & Stats (solar panels for remote computing in 2011; Sabbatical grant 2014),
- Research Chaire in 2013: Vieola Environment - French Natural History Museum - Ecole Polytechnique.