Definition of Evolution

The Operational Definition of Evolution at the Level of a Deme is a Change in Allele or Gamete Frequency In the Gene Pool.

Evolutionary Force

A Factor or Process That Can Change The Frequency of an Allele In the Gene Pool.
\[ A \]
\[ p = 1 \]
\[ A \]
\[ p = 1 - 1/(2N) \]
\[ q = 1/(2N) \]

**Deme of N Individuals**

**Gene Pool Before Mutation**

**Gene Pool After Mutation**

**Mutation Is an Evolutionary Force**

---

**Genetic Drift**

Genetic Drift Occurs When Sampling Error Alters Allele Frequencies.

Sampling Error Occurs When Populations Are Finite in Size.

Therefore, Finite Population Size is An Evolutionary Force.
Mendel’s Ratios Were Not “Perfect” Because They Are Based On A Finite Number of Observations. A Frequency In A Sample Only Converges To the Probability As The Sample Size Gets Larger and Larger.

A Deme Is A Collection of Such Crosses, Each Subject to Random Sampling Error in Its Mendelian Ratios
Computer Simulation of Genetic Drift

Gene Pool

Sample 10 Gametes to Create 5 Individuals

Do This 20 Times To Show Sampling Variation

Property 1 of Genetic Drift: No Direction

20 Simulations of a population of size = 5
Property 2 of Genetic Drift: It Is Cumulative

Let This Be The Sample That Actually Occurs
Property 2 of Genetic Drift: It Is Cumulative

Simulations of $N = 50$, $p = 0.5$

$\text{BottleneckSim}[50, 50, 20, 40, .5]$
$\text{MultiSim}[50, 50, 20, 40, .5]$
Property 3 of Genetic Drift: Strength $\alpha \frac{1}{2N}$

Simulations of $p=0.5$ with $N=25$, 100 and 1000

MultSim[N, N, 20, 40, .5]
Property 4 of Genetic Drift: Loss of Alleles

![Graph showing the number of occurrences in simulation for alleles A and a over 10 gametes.]

Properties 3 & 4 of Genetic Drift: Rate of Loss of Alleles

- **Property 3:**
  - Large Population: 10,000
  - Allele frequency: \( \frac{1000}{10000} = 10\% \)
  - Population survival: 50% survival, including 450 allele carriers
  - Allele frequency: \( \frac{450}{5000} = 9\% \)
  - Little change in allele frequency (no alleles lost)

- **Property 4:**
  - Small Population: 10
  - Allele frequency: \( \frac{1}{10} = 10\% \)
  - Population survival: 50% survival, with no allele carrier among them
  - Allele frequency: \( \frac{0}{2} = 0\% \)
  - Dramatic change in allele frequency (potential to lose one allele)

Rate of Loss = \( \frac{1}{2N} \)
Properties 3 & 4 of Genetic Drift: Loss of Alleles = Ultimate Fixation If Have No Mutation

DriftSim[N, .5]

Property 5 of Genetic Drift: Isolated Demes Become Genetically Differentiated (From Property 1)

2N = 20

4 Isolated Demes Started From One Ancestral Deme With p = 0.5

p

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

0 2 4 6 8 10

Generation
Properties of Genetic Drift

1. Has No Direction
2. Is Cumulative
3. Strength is Proportional to $\frac{1}{2N}$
4. Leads to Loss (and Fixation and Coalescence) of Alleles Within Demes
5. Leads to Genetic Differentiation Between Isolated Demes
6. Creates $|D| > 0$
Although Strength of Genetic Drift is Proportional to $1/2^N$, Drift Can be Important in Large Populations

1. Founder Effects -- A Large Population Today Was Founded By A Small Number of Founders in the Past.
3. Coalescence – all copies of a gene ultimately coalesce to a common ancestral allele due to drift in all populations.
4. Neutral Alleles -- Alleles With No Impact on Any Phenotype Related to Reproductive Success. Their Fate is Determined by Drift and Mutation.

MultiSim[500, 2, 20, 40, .5]
A Human Founder Event

- The Village Was Founded By A Handful of People 7 Generations Before
- One Founder, Altagracia Carrasco, Had Many Children by Four Women
- The Alleles Carried by Him Were Therefore in High Frequency in the Founder Population Gene Pool
- Subsequent Population Growth Reduced the Force of Drift But “Freezes In” The Allele Frequencies Created by the Initial Founder Event So His Alleles Remain In High Frequency Even Today

Altagracia Carrasco, Like Most People, Was A Heterozygous Carrier For an Autosomal Recessive Genetic Disease: 5-α Steroid Reductase Deficiency

5-α Steroid Reductase

\[\text{testosterone} \xrightarrow{\text{5-α Steroid Reductase}} \text{dihydrotestosterone}\]
Linkage Disequilibrium Is Created By Population Subdivision In A Manner Not Related To Recombination (Creates Serious Problems For Disequilibrium Mapping)

Gene Pool for Population 1

\[ g_{AB} = 1 \]

\[ D = 0 \]

Gene Pool for Population 2

\[ g_{ab} = 1 \]

\[ D = 0 \]

Gene Pool for Pooled Populations

\[ g_{AB} = \frac{1}{2} \]

\[ g_{ab} = \frac{1}{2} \]

\[ D = g_{AB}g_{ab} = \frac{1}{4}, \quad D' = 1 \]
Problem!
Population Structure or Historical Isolates Can Create Spurious Phenotypic Associations. E.g., in Quebec there are French and English Speaking Canadians. French Canadians Have Been Strongly Influenced by a Past Founder Event and Show Allele Frequency Differences At Many Loci From the English Population. Therefore, A Mapping Study of the “Quebec” Population Would Reveal A Strong Association Between Many Loci and the Language One Spoke. Similarly, A Candidate Locus Study Would Find An Association With Language If The Candidate Locus Showed Haplotype Frequency Differences Between English and French Canadians.

Avoiding Problem of Hidden Population Structure
1. Use founder or bottleneck populations (but must make sure they truly are and have been highly isolated since the drift event)
2. Use several loci to reconstruct recent evolutionary history and population structure prior to initiating association study, and then choose populations accordingly or use as a control set of loci in the association study.
**Founder & Bottleneck Events**

- Can Drastically Alter Allele Frequencies, Including Making Certain Genetic Disease Allele or Disease Risk Alleles Common (makes obtaining pedigrees for linkage mapping much easier)
- Leads to pedigree inbreeding (Speke’s gazelles; humans on Tristan da Cunha)
- Creates Linkage Disequilibrium, Which Rarely Extends Over 1 cM in Large Demes (makes disequilibrium mapping much easier)
- Reduce Overall Genetic Variation, Creating A Simpler Genetic Background
- For The Above Reasons, Such Populations Are Important In Biomedical Research & Conservation

---

**E.g., Positional Cloning & QTL’s**

- The First Case of Positional Cloning Was the Gene for Huntington’s Chorea
- Nancy Wexler Realized That The Key Was to Find a Founder Population With A High Frequency of HD.
- She Found Such A Population On Lake Maracaibo
- Now, Founder Populations Such As This Are Regarded As Commercially Valuable Assets.
E.g., Positional Cloning & QTL’s

About 200 years ago, a single woman who happened to carry the Huntington's allele bore 10 children — and today, many residents of Lake Maracaibo trace their ancestry (and their disease-causing gene) back to this lineage.

Effective Population Size

• Founder And Bottleneck Events Show That The Current Size Of A Population May Not Be A Good Indicator Of The Impact Of Genetic Drift Upon That Population

• The Concept of EFFECTIVE POPULATION SIZE Solves This Problem.
Effective Population Size

measures the strength of genetic drift in influencing some population genetic feature of interest relative to how that same feature evolves through genetic drift in an idealized population over the same number of generations.

The Idealized Reference Population

- a diploid population of hermaphroditic, self-compatible organisms
- constant size of $N$ breeding Adults
- random mating
- complete genetic isolation (no contact with any other population)
- discrete generations with no age structure
- all individuals contribute the same number of gametes on the average to the next generation (no natural selection)
- the sampling variation in the number of gametes contributed to the next generation by an individual is given by a Poisson probability distribution.
The Most Common Parameters Used To Monitor Genetic Drift are:

- **The Average Level of Identity by Descent (inbreeding effective size)**
- **The Variance In Allele Frequency Induced By Genetic Drift (variance effective size)**

### Impact of Drift On Average F In An Idealized Population

\[
F(t) = \left( \frac{1}{2N} \right) + \left( 1 - \frac{1}{2N} \right) F(t-1)
\]

- **Average Probability Of Identity By Descent At generation t**
- **Probability The 2 Gametes From The Same Individual Are Identical**
- **Probability Randomly Draw 2 Gametes From The Same Individual**
Impact of Drift On Average $F$ In An Idealized Population

\[
F(t) = \frac{1}{2N} + \left( 1 - \frac{1}{2N} \right) F(t-1)
\]

Can Use The Above Equation Recursively To Obtain:

\[
F(t) = 1 - \left( 1 - \frac{1}{2N} \right)^t \quad [F(0) = 0]
\]
Impact of Drift On Average F In An Idealized Population

\[ F(t) = 1 - \left( 1 - \frac{1}{2N} \right)^t \]

If A Real Population Has An Observed Average F of \( F(t) \) After \( t \) Generations From the Reference Generation With \( F = 0 \); Then The Inbreeding Effective Size Is Given By:

\[ F(t) = 1 - \left( 1 - \frac{1}{2N_{ef}} \right)^t \quad \text{or} \quad N_{ef} = \frac{1}{2\left\{1-\left[1-F(t)\right]^{1/2}\right\}} \]

Impact of Drift On Allele Freq. Variance In An Idealized Population

\[ \sigma^2(t) = pq\left\{1 - \left( 1 - \frac{1}{2N} \right)^t\right\} \]

If A Real Population Has An Observed Variance of \( \nu(t) \) After \( t \) Generations From the Reference Generation; Then The Variance Effective Size Is Given By:

\[ \nu(t) = pq\left\{1 - \left( 1 - \frac{1}{2N_{ev}} \right)^t\right\} \quad \text{or} \quad N_{ev} = \frac{1}{2\left\{1-\left[1-\nu(t)/(pq)\right]^{1/2}\right\}} \]
There Is No Such Thing As *The* Effective Size of a Population

- The effective size depends upon which genetic parameter you are using
- The effective size depends upon which reference generation you are using
- Therefore, a single population can have many different effective sizes associated with it, all biologically meaningful but distinct

Example: Speke’s Gazelle

- Herd Started in 1969 With 4 Animals
- By 1979 There Were 19 Animals With An Average $F$ of 0.1283 After 1.7 Generations
- Therefore, $N_{ef}$ Relative to the Founders is $6.4 < 19$ (Founder Effect)
- In 1979, Management Was Changed, and 15 New Animals Bred with $F = 0.149$ and $t = 2.7$, yielding $N_{ef} = 8.6 < 15$ (Founder Effect & $f < 0$)
- Using the parents of the 19 Animals in 1979 as Reference Generation, then $F = 0.0207$ and $t = 2$, yielding $N_{ef} = 96.1 > 15$ (Effect of Avoidance of Inbreeding in System of Mating Sense)
Example: Speke’s Gazelle

- Herd Started in 1969 With 4 Animals
- In 1979, Management Was Changed, and 15 New Animals Bred with \( v/(pq) = 0.135 \) and \( t = 2.7 \) (computer simulation of exact pedigree), yielding \( N_{e0} = 9.6 < 15 \) (Founder Effect)
- The same 15 animals have
  - \( N_n = 9.6 < 15 \) (relative to founder generation)
  - \( N_q = 8.6 < 15 \) (relative to founder generation)
  - \( N_q = 96.1 > 15 \) (relative to the management change generation)
- **WHAT IS THE EFFECTIVE SIZE OF THIS POPULATION?**

In Most Cases, Do Not Have Complete Pedigree Information, Precluding the Calculation of Various Effective Sizes. Many Formulae Have Been Derived as Estimators or Approximations to Effective Size.

The Literature Is A Mess, Because Many Do Not Distinguish Among The Various Effective Sizes, and Often Mix Inappropriate Formulae
Interactions of System of Mating with Genetic Drift via Effective Size

• The ideal reference population assumes random mating.
• Suppose mating is non-random, either due to inbreeding or assortative mating such that \( f > 0 \).
• Then:

\[
\bar{F}(t) = f + (1 - f) \left[ \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\bar{F}(t - 1) \right]
\]

I by D created by system of mating beyond random mating expectations.

\[ N_{ef} = \frac{N}{1 + f(2N - 1)} \]
Interactions of System of Mating with Genetic Drift via Effective Size

- Suppose mating is non-random, either due to inbreeding or assortative mating such that $f > 0$.
- Then:

$$\text{Variance in Allele Frequency} = (1 - f) \frac{pq}{2N} + f \frac{pq}{N} = \frac{pq(1 + f)}{2N}$$

The ideal reference population assumes random mating.

$$N_{ev} = \frac{N}{1 + f}$$
Interactions of System of Mating with Genetic Drift via Effective Size

\[ N_{ev} = 0.1 \]

Interactions of Population Growth with Genetic Drift via Effective Size

\[ N_{ef} = \frac{2N - 1}{k - 1 + (1 - \frac{k}{2N})} \]

\[ N_{ev} = N \]

Where \( N \) is an idealized population in every way except that each individual has an average of \( k \) offspring (\( k=2 \) corresponds to a constant sized population)
Interactions of Population Growth with Genetic Drift via Effective Size

Effective Sizes in African Rhinos
(Braude and Templeton, African J. Ecol. 47: 546-555, 2009)

<table>
<thead>
<tr>
<th></th>
<th>Census Size, 1997</th>
<th>Inbreeding Effective Size</th>
<th>Variance Effective Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Rhinoceros</td>
<td>N=2,600</td>
<td>Ne =18,840</td>
<td>Ne_e =4,189</td>
</tr>
<tr>
<td>Diceros bicornis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern White Rhinoceros</td>
<td>N=8,440</td>
<td>Ne =106</td>
<td>Ne_e =240</td>
</tr>
<tr>
<td>Ceratotherium simum simum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern White Rhinoceros</td>
<td>N=23</td>
<td>Ne =69</td>
<td>Ne_e =41</td>
</tr>
<tr>
<td>Ceratotherium simum cottoni</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>