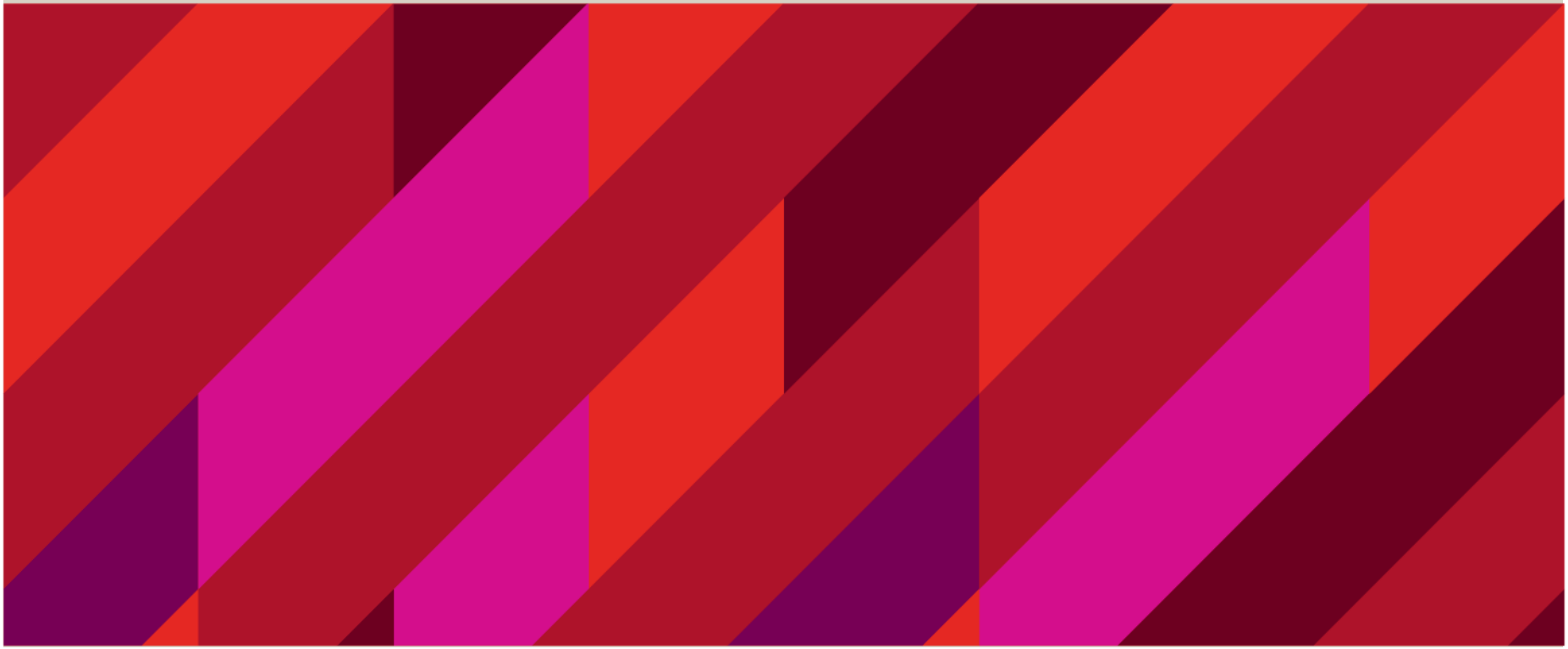
MACQUARIE University

BIOL3110 Conservation & Ecological Genetics

LECTURE 4: CHARACTERIZING GENETIC VARIATION



Genetic Diversity (V_G)

AS A "BAROMETER" OF POPULATION GENETIC HEALTH

MACQUARIE University

Useful comparisons:

(1) Among (sub)populations

(2) Over time (the same pop/species) as an index of change, i.e.:

$$\frac{H_0}{H_T}$$

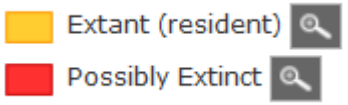
Where :

H_o = H at the time of observation

H_T = H as known for the past or predicted for the future...

Genetic Diversity (V_G)

MAURITIUS KESTREL (*Falco punctatus*)



Home > Falco punctatus (Mauritius Kestrel)

Falco punctatus

NOT EVALUATED	DATA DEFICIENT	LEAST CONCERN	NEAR THREATENED	VULNERABLE	ENDANGERED	CRITICALLY ENDANGERED	EXTINCT IN THE WILD	EXTINCT
NE	DD	LC	NT	VU	EN	CR	EW	EX

Driven down to **4** wild birds – **1 breeding pair** in 1974!!

Genetic Diversity (V_G)

MAURITIUS KESTREL (*Falco punctatus*)

MACQUARIE University

Population size:

1974: single mating pair bottleneck **N=2**

1997: N=500

Calculations using microsatellite markers extracted from museum material:

$$\frac{H_O}{H_T} = \frac{H_{1997}}{H_{1974}} = \frac{0.10}{0.23} = \mathbf{0.43}$$

Loss of more than 50% of *H* since 1974

...even despite N increasing from 2 to ~500?

Genetic Diversity (V_G)

THE HARDY-WEINBERG (H-W) EQUILIBRIUM

MACQUARIE University

H-W equilibrium provides the "null expectation" for allele/genotype frequencies in a pop. Genotypes will accord to HWE unless:

- **Inbreeding** (>>homozygotes)
- **Selection** (not exp at neutral loci)
- **Non-random mating** (e.g. MHC)
- **Small populations** (Drift)
- **Mutation** – takes a long time!

Genetic Diversity (V_G)

THE HARDY-WEINBERG (H-W) EQUILIBRIUM

MACQUARIE University

For 2 alleles (A_1 and A_2) segregating at a locus, if we have:

- A large, closed population (no migration)
- Random mating (i.e. no selective mate choice)
- Equal fitness of genotypes
- Negligible mutation

Then we expect alleles to assort among genotypes as:

Genotypes:	A_1A_1	A_1A_2	A_2A_2
Frequencies:	p^2	$2pq$	q^2

Where p = frequency of A_1 and q = frequency of A_2

Genetic Diversity (V_G)

THE HARDY-WEINBERG (H-W) EQUILIBRIUM

MACQUARIE University

Where p = frequency of A_1 and q = frequency of A_2

Genotypes:	A_1A_1	A_1A_2	A_2A_2
Frequencies:	p^2	$2pq$	q^2

For a locus with **2** alleles:

$$H_{exp} = 2pq$$

Genetic Diversity (V_G)

ALLELE COUNTING

MACQUARIE University

Absolute frequencies	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
Individuals:	**38**	**60**	**2**	**100**
Alleles:	2 copies of A_1	1 copy of A_1 1 copy of A_2	2 copies of A_2	**200**

Genetic Diversity (V_G)

ALLELE COUNTING

MACQUARIE University

Absolute frequencies	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
Individuals:	**38**	**60**	**2**	**100**
Alleles:	2 copies of A_1	1 copy of A_1 1 copy of A_2	2 copies of A_2	**200**
A_1:	**76**	**60**	**0**	**136**
A_2:	**0**	**60**	**4**	**64**

Genetic Diversity (V_G)

ALLELE COUNTING

MACQUARIE University

Absolute frequencies	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
Individuals:	**38**	**60**	**2**	**100**
Alleles:	2 copies of A_1	1 copy of A_1 1 copy of A_2	2 copies of A_2	**200**
A_1:	**76**	**60**	**0**	**136**
A_2:	**0**	**60**	**4**	**64**

Relative frequencies (~%) of each allele:

$A_1 = p = 136/200 = \mathbf{0.68}$

$A_2 = q = 64/200 = \mathbf{0.32}$ (...or simply 1-p)

Genetic Diversity (V_G)

MACQUARIE University

USE p & q TO CALCULATE H-W EXPECTATION

---

p (A_1) = 0.68, q (A_2) = 0.32

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
H-W exp frequencies:	p^2	$2pq$	q^2	

Genetic Diversity (V_G)

MACQUARIE University

USE ***p*** & ***q*** TO CALCULATE H-W EXPECTATION

$p\ (A_1) = 0.68,\ q\ (A_2) = 0.32$

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
H-W exp frequencies:	p^2	$2pq$	q^2	
For p=0.68 & q =0.32:	0.68^2	2 x 0.68 x 0.32	0.32^2	**1.0**
=	0.462	0.435	0.102	**1.0**

Genetic Diversity (V_G)

MACQUARIE University

USE ***p*** & ***q*** TO CALCULATE H-W EXPECTATION

$p\ (A_1) = 0.68,\ q\ (A_2) = 0.32$

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
H-W exp frequencies:	p^2	$2pq$	q^2	
For p=0.68 & q =0.32:	0.68^2	2 x 0.68 x 0.32	0.32^2	**1.0**
=	0.462	0.435	0.102	**1.0**

Expected genotypes given a sample of **N=100**	**46.2**	**43.5**	**10.2**	**100**

Genetic Diversity (V_G)

MACQUARIE University

USE ***p*** & ***q*** TO CALCULATE H-W EXPECTATION

$p\ (A_1) = 0.68,\ q\ (A_2) = 0.32$

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Total
H-W exp frequencies:	p^2	$2pq$	q^2	
For p=0.68 & q =0.32:	0.68^2	2 x 0.68 x 0.32	0.32^2	**1.0**
=	0.462	0.435	0.102	**1.0**
Expected genotypes given a sample of **N=100**	**46.2**	**43.5**	**10.2**	**100**
What we **observed**:	**38**	**60**	**2**	**100**

DO THESE DISTRIBUTIONS DIFFER?

Genetic Diversity (V_G)

THE "CHI-SQUARED" TEST FOR SIGNIFICANCE

MACQUARIE University

$$\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp}$$

Does your observed distribution of genotypes **differ** from the H-W expected distribution?

Calculate using **absolute genotype frequencies** (the numbers that we just calculated)

Use a significance table to look up a **P-value**

Genetic Diversity (V_G)

CHI-SQUARED" TEST FOR SIGNIFICANCE

$$\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp}$$

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Sum
H-W Exp:	46.2	43.5	10.2	100
Obs:	38	60	2	100

Genetic Diversity (V_G)

CHI-SQUARED" TEST FOR SIGNIFICANCE

MACQUARIE University

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Sum
H-W Exp:	46.2	43.5	10.2	100
Obs:	38	60	2	100
$\frac{(Obs-Exp)^2}{Exp}$	$\frac{(38-46.2)^2}{46.2}$	$\frac{(60-43.5)^2}{43.5}$	$\frac{(2-10.2)^2}{10.2}$	

$$\chi^2 = \sum \frac{(\boldsymbol{Obs} - \boldsymbol{Exp})^2}{\boldsymbol{Exp}}$$

Genetic Diversity (V_G)

CHI-SQUARED" TEST FOR SIGNIFICANCE

$$\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp}$$

	Genotypes			
	A_1A_1	A_1A_2	A_2A_2	Sum
H-W Exp:	46.2	43.5	10.2	100
Obs:	38	60	2	100
$\frac{(Obs - Exp)^2}{Exp}$	$\frac{(38 - 46.2)^2}{46.2}$	$\frac{(60 - 43.5)^2}{43.5}$	$\frac{(2 - 10.2)^2}{10.2}$	
=	1.49	6.29	6.59	**= 14.34**

Chi Square statistic = **14.34**, **P < 0.05** (from a table)
Hence, conclude a significant deviation from HW…

CHI-SQUARE TABLE

MACQUARIE University

Critical values of the Chi-square distribution with *d* degrees of freedom

	Probability of exceeding the critical value						
d	0.05	0.01	0.001	*d*	0.05	0.01	0.001
1	3.841	6.635	10.828	11	19.675	24.725	31.264
2	5.991	9.210	13.816	12	21.026	26.217	32.910
3	7.815	11.345	16.266	13	22.362	27.688	34.528
4	9.488	13.277	18.467	14	23.685	29.141	36.123
5	11.070	15.086	20.515	15	24.996	30.578	37.697
6	12.592	16.812	22.458	16	26.296	32.000	39.252
7	14.067	18.475	24.322	17	27.587	33.409	40.790
8	15.507	20.090	26.125	18	28.869	34.805	42.312
9	16.919	21.666	27.877	19	30.144	36.191	43.820
10	18.307	23.209	29.588	20	31.410	37.566	45.315

***INTRODUCTION TO POPULATION GENETICS*, Table D.1**

Degrees of freedom = number of genotypes -1- number of alleles -1

Or more simply – number of genotypes – number of alleles

Df = 1

Genetic Diversity (V_G)

H-W EQUILIBRIUM FOR 2+ ALLELES

MACQUARIE University

2 ALLELES:

A_1 (p)
A_2 (q)

Genotypes		
A_1A_1	A_1A_2	A_2A_2
p^2	$2pq$	q^2

3 ALLELES:

A_1 (p)
A_2 (q)
A_3 (r)

Genotypes					
A_1A_1	A_1A_2	A_1A_3	A_2A_2	A_2A_3	A_3A_3
p^2	$2pq$	$2pr$	q^2	$2qr$	r^2

Genetic Diversity (V_G)

SEX-LINKED LOCI

MACQUARIE University

Loci on the **Sex Chromosomes**

Humans, primates,
mammals & most insects:
Females XX, **males XY**
(females double-dose of X-linked alleles)

Birds & butterflies:
Females ZW, **males ZZ**

This changes the allele counting process…
see **Frankham et al. Table 4.6 & Example 4.7**

Genetic Diversity (V_G)

AS A "BAROMETER" OF POPULATION GENETIC HEALTH

MACQUARIE University

Heterozygosity versus allelic diversity…

A: average allelic diversity

- Sensitive to sample size (N)
- Statistical control for unequal N

$$A = \frac{\sum(N_A)}{N_T}$$

Where:

N_A = number of different alleles across all loci

N_T = total number of loci examined

n_e: effective number of alleles

- Less sensitive to sample size
- Less influenced by rare alleles
- Highest when more alleles present but at equal frequencies

$$n_e = \frac{1}{\sum P_i^2}$$

Where:

P_i = frequency of each allele

Genetic Diversity (V_G)

LINKAGE DISEQUILIBRIUM

- Non-random associations between alleles at different loci is called **linkage disequilibrium (*D*)**

e.g.: 2 alleles at **A & B loci**:			
A_1B_1	A_1B_2	A_2B_1	A_2B_2

- Can result from strong selection (e.g. mate preference & trait alleles in sexual selection)
- Can indicate bottlenecks, recent pop mixing, etc.

see Frankham et al. Ch. 4 (Box 4.3 & 4.4, Table 4.9)

Coming up:

MACQUARIE University

Genetics from the phenotypic level