



MONASH University

Science

# Population structure and inference of demography

Anders Gonçalves da Silva

Population and landscape genomics workshop – Day 2

25 March 2014

CBA – ANU

# Outline

- Ensure we all know what an allele, a genotype, a locus, and some other basic terminology.
- Define population genetic structure
- Explore some of the ways population genetic structure arises
- Use R to estimate population structure
- Use R to evaluate sampling strategies for population genetic structure studies

## Question?

- What is the most important factor to consider when designing a sampling strategy for a population genetic structure study?
- Is it:
  - number of individual samples/site
  - number of sites, or
  - number of loci?

## Some definitions

### Locus



 **A**

 **a**

### Allelic frequency

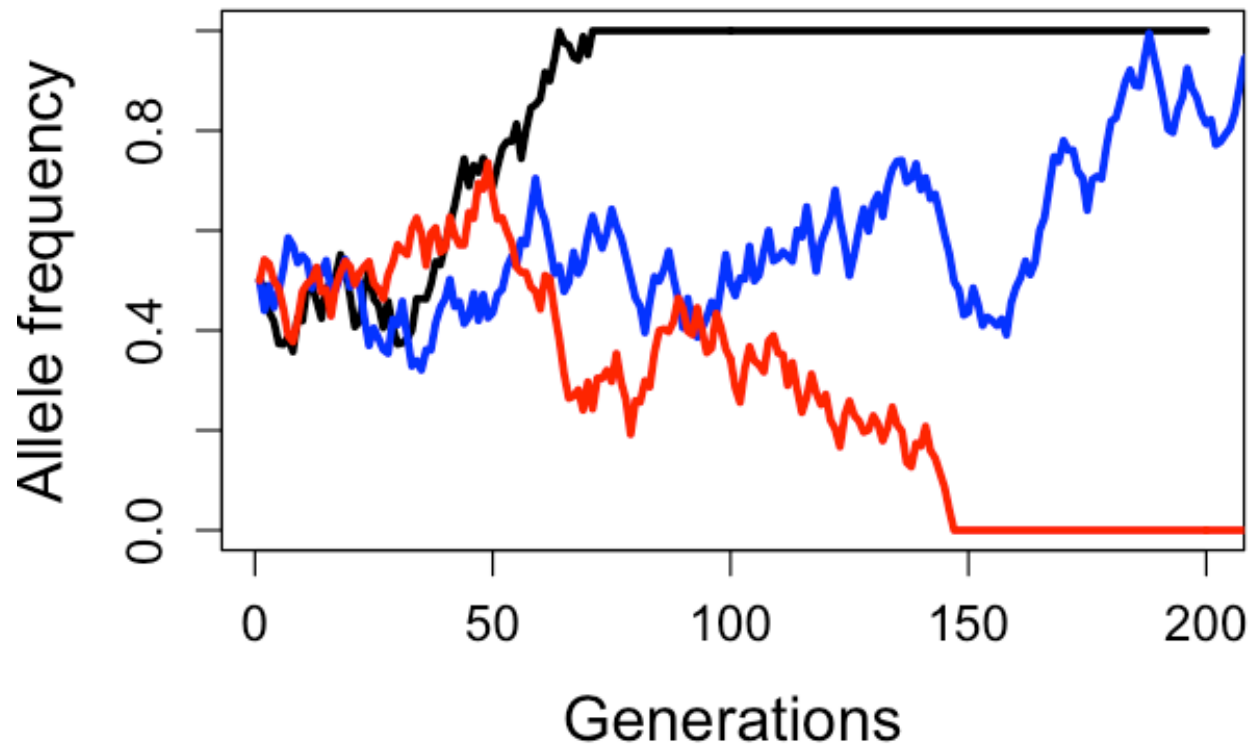
$$p = 0.6$$

$$q = 0.4$$

### Genotypic frequency

	A	a
A	0.36	
a	0.48	0.16

# What is evolution?



# Hardy-Weinberg Equilibrium

Allelic frequency:

$p$

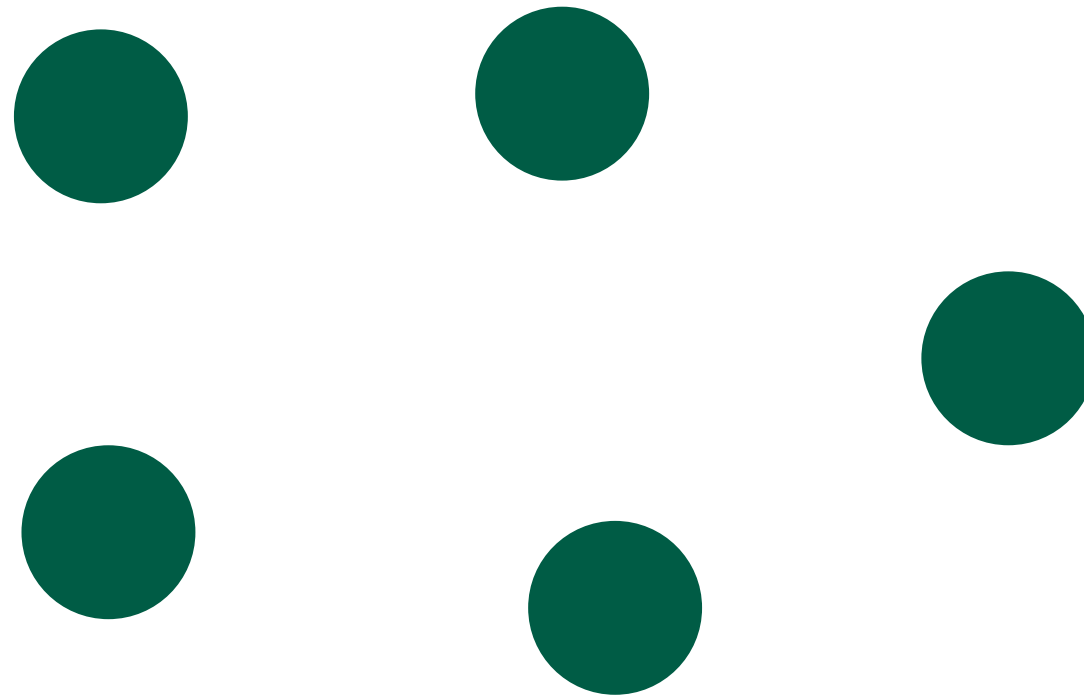
$q$

Genotypic frequency  
under random mating

	A	a
A	$p^2$	
a	$2pq$	$q^2$

Allele frequency does not change over time

# Population genetic structure



# Population genetic structure

0.2

0.2

0.2

0.2

0.2



# Population genetic structure



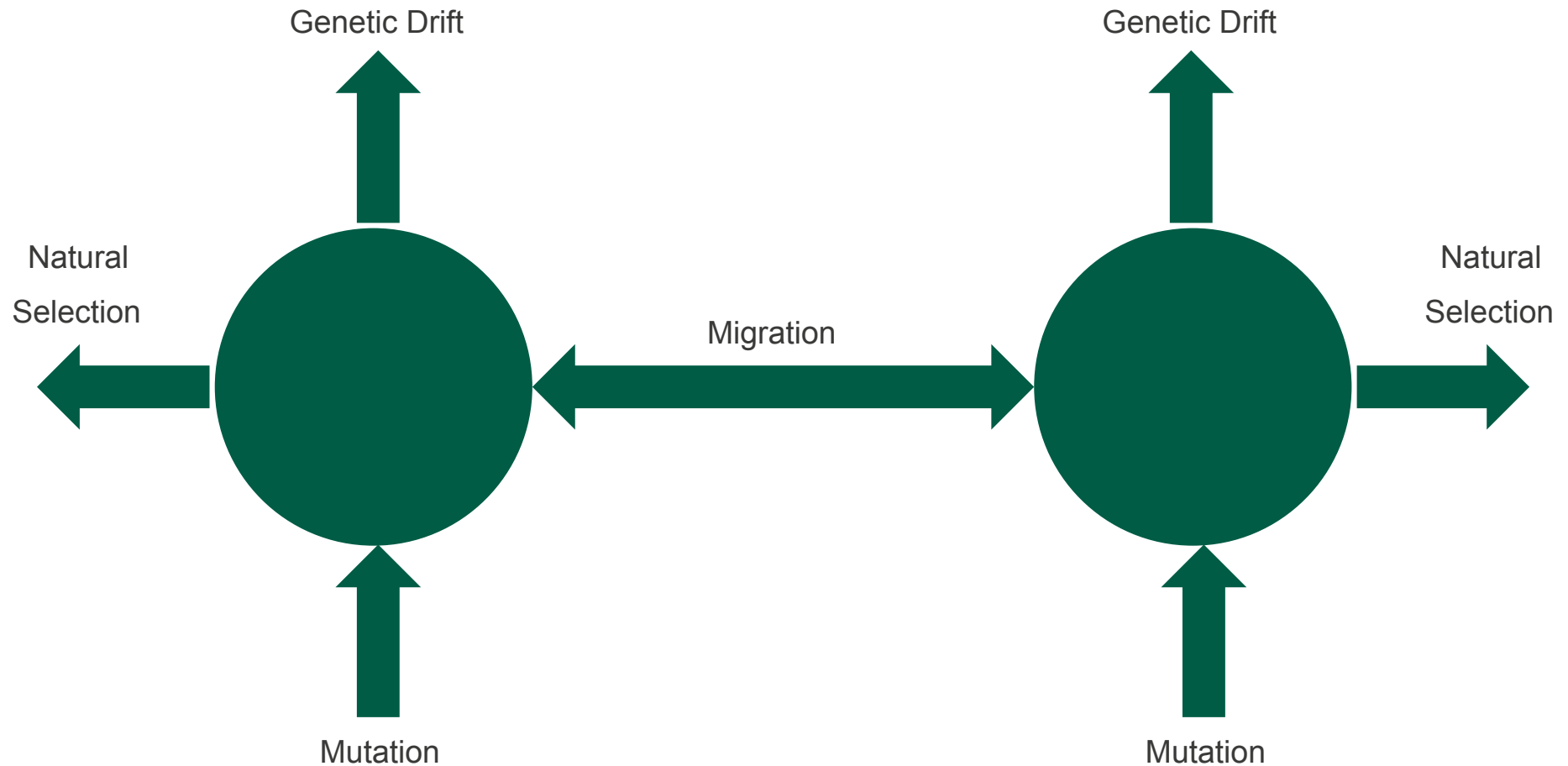
# Population genetic structure



## Population genetic structure

- Is the partitioning of individuals into units that are largely homogenous in their allele frequencies.
- It is a departure from the our Hardy-Weinberg model assumptions.

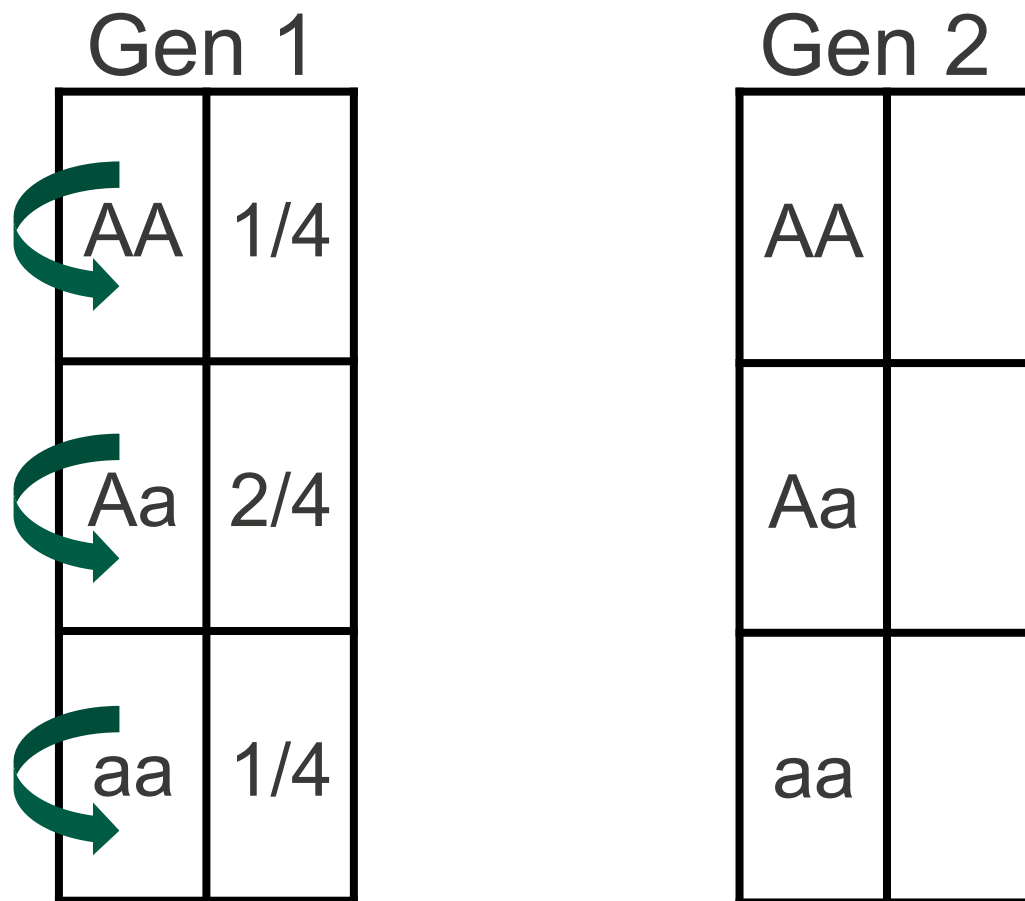
# How does PGS arise?



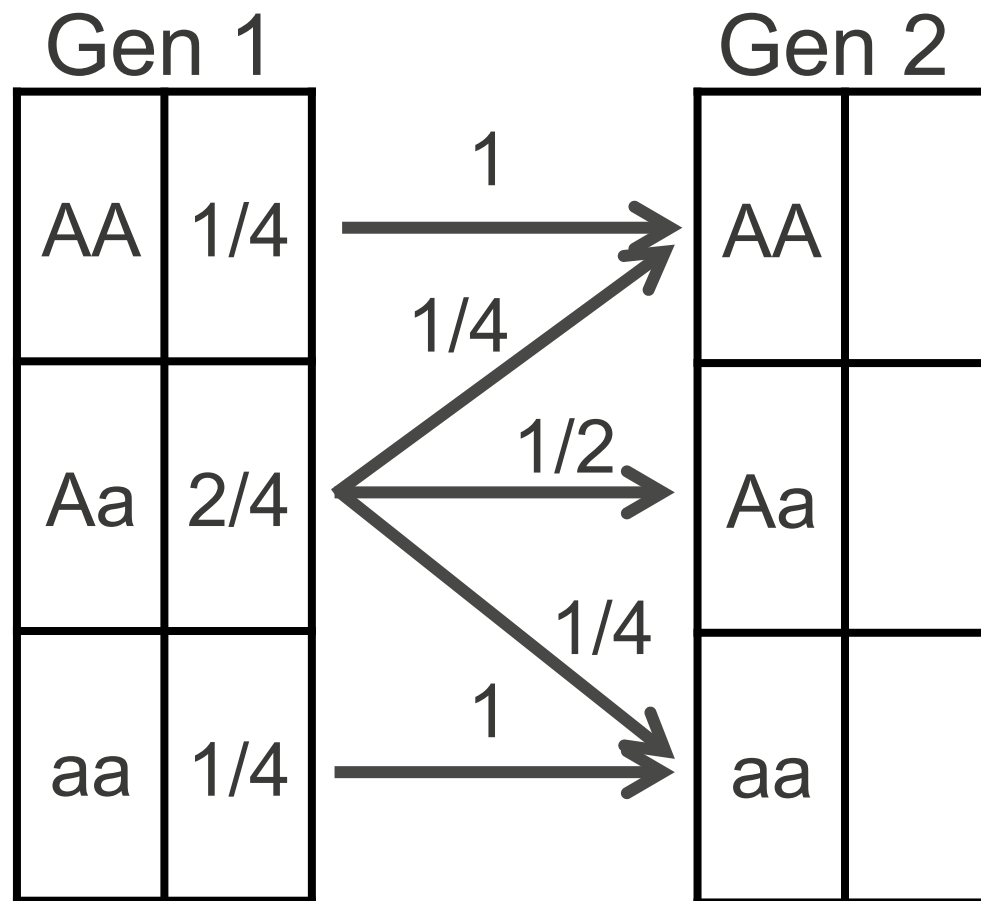
## How do we measure PSG?

$$F = \frac{H_{exp} - H_{obs}}{H_{exp}}$$

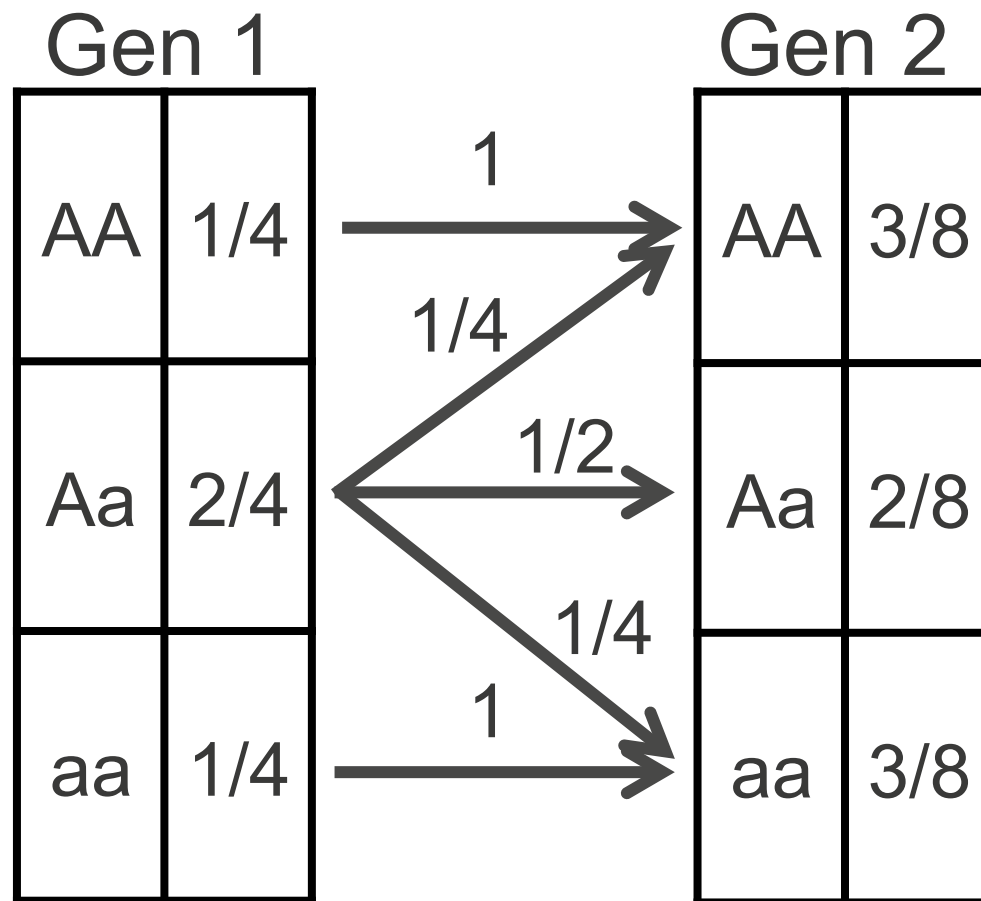
# Inbreeding or how to loose heterozygotes



# Inbreeding or how to loose heterozygotes

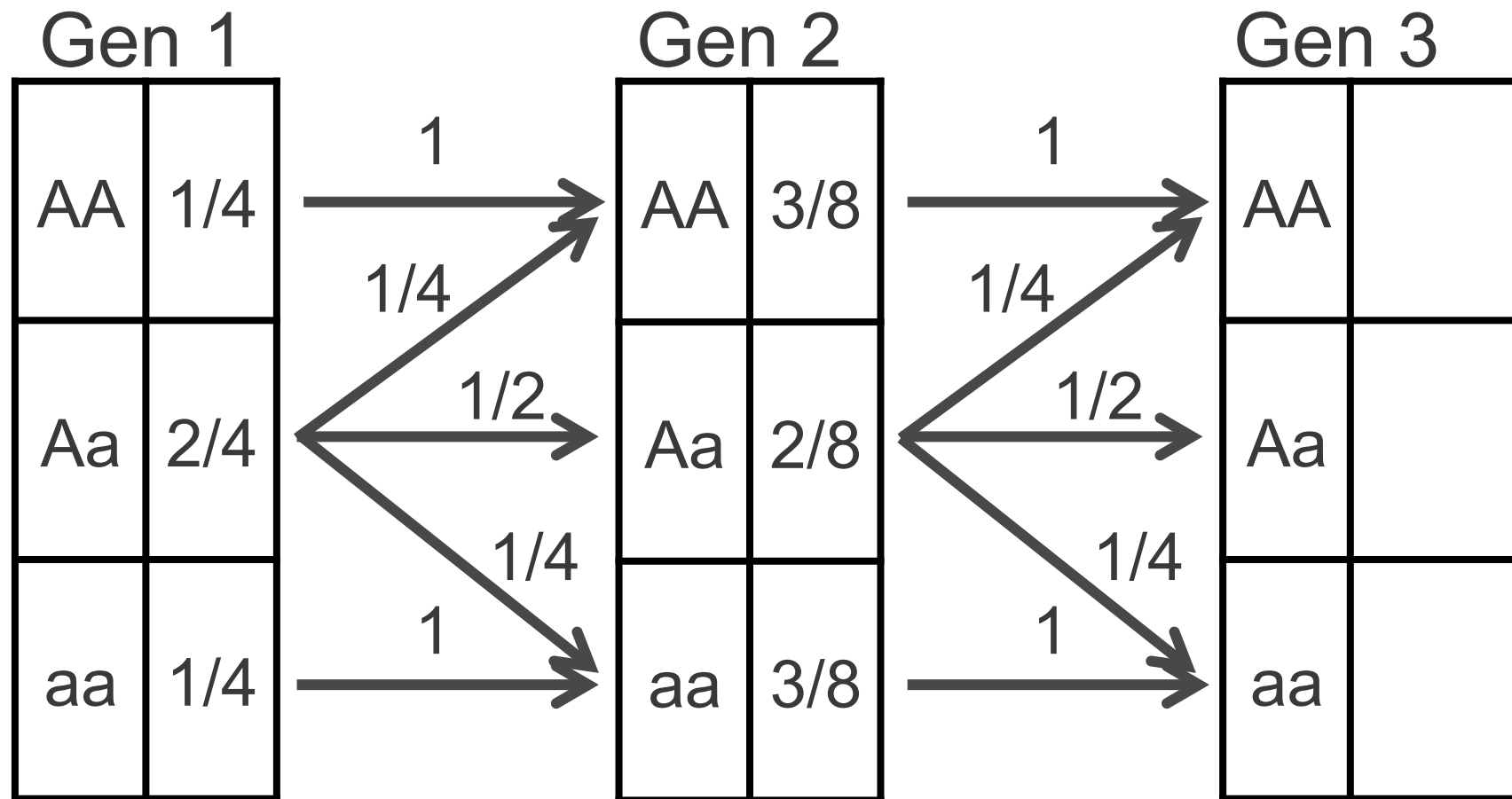


# Inbreeding or how to loose heterozygotes

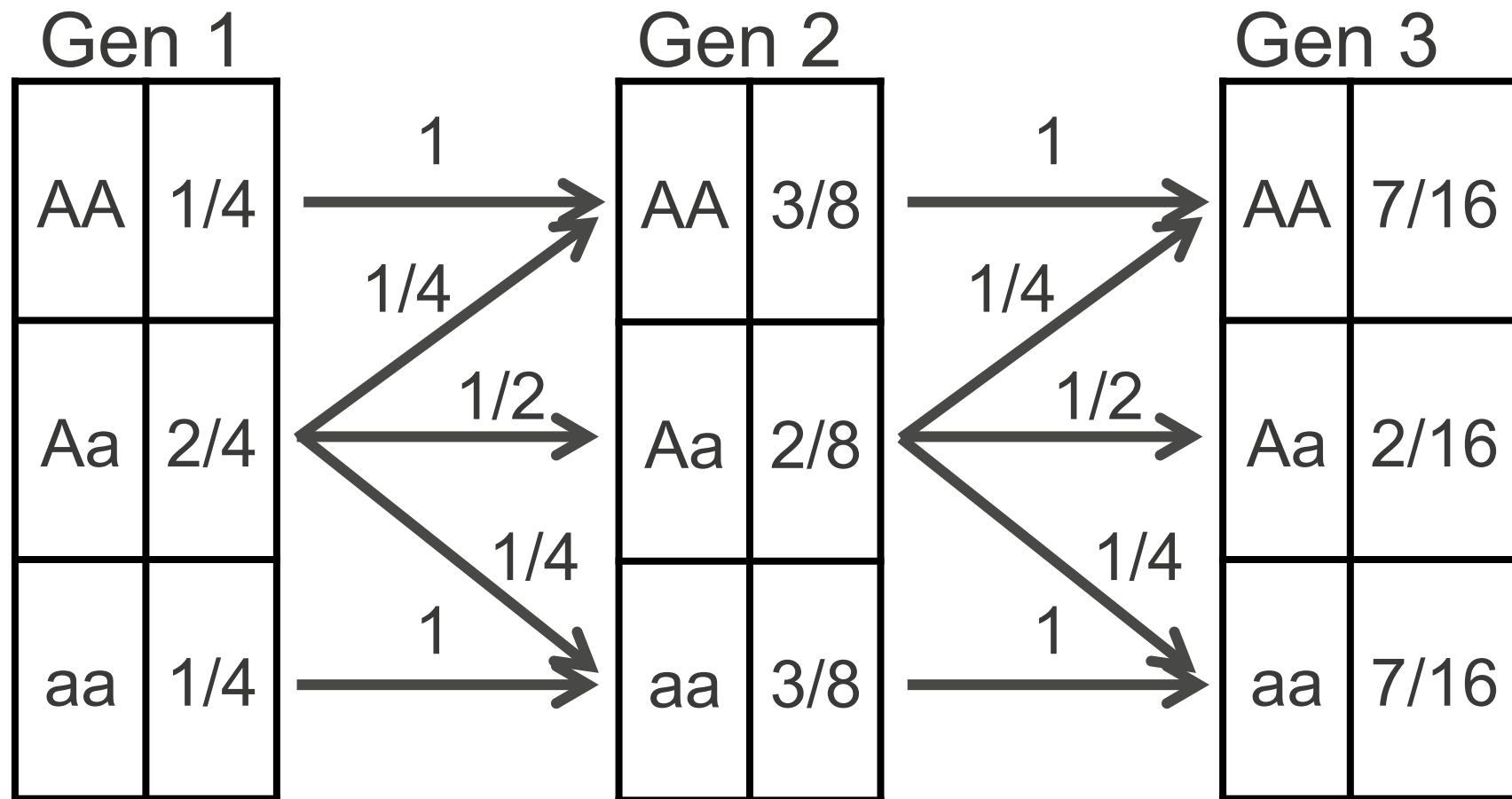




# Inbreeding or how to loose heterozygotes



# Inbreeding or how to loose heterozygotes



## Inbreeding...

- Excellent way of **LOOSING** heterozygotes...
- BUT...
- Has **NO EFFECT** on the allele frequency...

$$F = \frac{H_{exp} - H_{obs}}{H_{exp}}$$

## Inbreeding...

- Excellent way of **LOOSING** heterozygotes...
- BUT...
- Has **NO EFFECT** on the allele frequency...

$$F = \frac{2pq - \frac{n_{het}}{n_{total}}}{2pq}$$

# Wright's insight...

- Inbreeding is a form of **NON-RANDOM MATING**...
- A population divided in to  $n$  sub-populations, where individuals randomly mate within sub-populations, is also a form of **NON-RANDOM MATING**...
- Proposal:
  - If we have  $n$  sub-populations (where  $n$  is somewhere close to infinity), each with a finite number of individuals, we can expect the same pattern observed within sub-populations to emerge because of inbreeding should emerge across sub-populations.



## Wright's insight...

- Let us prove it in  $R$ ...

## Wright's equations

$$F_{IS} = \frac{\bar{H}_S - \bar{H}_I}{\bar{H}_S} \quad \text{Inbreeding effect}$$
$$F_{ST} = \frac{H_T - \bar{H}_S}{H_T} \quad \text{Genetic drift effect}$$
$$F_{IT} = \frac{H_T - \bar{H}_S}{H_T} \quad \text{Combined effect}$$

# Wright's equations

$$\underbrace{(1 - F_{IT})}_{\text{Combined effect}} = \underbrace{(1 - F_{IS})}_{\text{Inbreeding effect}} \underbrace{(1 - F_{ST})}_{\text{Genetic drift effect}}$$



# Estimators of PGS

Wright 1930	Nei 1973	Weir and Cockerham 1984	Excoffier 1992	Hedrick 2005
$F_{IS}$	$G_{IS}$	$f$	$\phi_{IS}$	$G'_{IS}$
$F_{ST}$	$G_{ST}$	$\theta$	$\phi_{ST}$	$G'_{ST}$
$F_{IT}$	$G_{IT}$	$F$	$\phi_{IT}$	$G'_{IT}$

<http://www.molecular ecologist.com/2011/03/should-i-use-fst-gst-or-d-2/>



# ESTIMATING PGS IN *R*...

## Further reading

- Waples, R. S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, 89(5), 438–450. doi:10.1093/jhered/89.5.438
- Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15(6), 1419–1439.
- Whitlock, M. C., & McCauley, D. E. (1999). Indirect measures of gene flow and migration:  $F_{ST}$  not equal  $1/(4Nm+1)$ . *Heredity*, 82, 117–125.
- Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, 19(15), 3038–3051.
- Jost, L. (2008).  $G_{ST}$  and its relatives do not measure differentiation. *Molecular Ecology*, 17(18), 4015–4026.