

## **5** Non-Random Mating: Phenotypic Assortative Mating

---

### **5.1: Learning Objectives**

By the end of this chapter, the reader should be able to define positive and negative assortative mating, distinguish assortative mating from inbreeding, quantify its effect on genotype frequencies at the locus under selection, and discuss Nigerian examples such as height, skin tone, and educational assortative mating, as well as consanguinity.

### **5.2: What is Non-Random Mating?**

The Hardy–Weinberg law (Chapter 3) assumes random mating: that the choice of mate is independent of genotype at the locus under study. In reality, mating is frequently non-random. Two broad classes are recognised:

- **Assortative mating**, in which individuals choose mates on the basis of **phenotype similarity** (positive assortative mating) or dissimilarity (negative assortative mating / disassortative mating), for a trait that may or may not be genetically correlated between mates.
- **Inbreeding (consanguineous mating)**, in which mates are chosen (or, more often, simply occur) on the basis of shared ancestry/relatedness, regardless of phenotype.

This chapter focuses on **phenotypic assortative mating**; inbreeding is treated as a related but distinct phenomenon,

important in parts of Northern Nigeria where cousin marriage is culturally favoured for reasons of family and property, and increasingly documented in genetic-counselling clinics because it raises the frequency of recessive disorders such as sickle-cell disease and some inherited forms of deafness.

### 5.3: Positive Assortative Mating for a Single Locus

**Positive assortative mating** occurs when individuals with similar phenotypes mate with one another more often than expected under random mating — for example, tall individuals tending to marry tall individuals. When the trait is fully genetically determined by a single locus with complete dominance (so that phenotype = genotype in effect, i.e., “like mates with like” genotype), positive assortative mating has the following effects, in contrast to random mating:

1. **Allele frequencies are unchanged** across generations (assortative mating, unlike selection, does not by itself remove or add alleles from the gene pool).
2. **Genotype frequencies deviate from Hardy–Weinberg proportions:** homozygote frequencies increase and heterozygote frequency decreases, generation upon generation, because like-with-like mating preferentially produces  $AA \times AA$  and  $aa \times aa$  pairings (each yielding only homozygous offspring) at the expense of random  $Aa$  pairings.
3. The population becomes progressively **more genetically differentiated** into homozygous “types,” even though total allele frequency in the population is unchanged — this is a form of genetic (though not literal family) structuring analogous in its statistical footprint to inbreeding.

### 5.4: A Simple Quantitative Model

Consider a trait fully determined by an autosomal locus with alleles  $A$ ,  $a$  and suppose mating is **completely** assortative by phenotype: individuals mate only with others of the identical phenotype (a simplifying, extreme case used to illustrate the direction of the effect; in practice, assortment is partial, and is measured by a correlation coefficient between mates' phenotypic values, discussed in Section 5.6).

Under complete phenotypic assortment with dominance ( $AA$  and  $Aa$  share one phenotype,  $aa$  the other), let the population, before assortment, have HWE genotype frequencies  $p^2$ ,  $2pq$ ,  $q^2$ . Mating occurs only within phenotype class:  $\{AA, Aa\}$  individuals mate only with other  $\{AA, Aa\}$  individuals, and  $aa$  individuals mate only with other  $aa$  individuals. Within the dominant phenotype class, matings are further sub-structured ( $AA \times AA$ ,  $AA \times Aa$ ,  $Aa \times Aa$  in proportions determined by the relative frequencies of  $AA$  and  $Aa$  within that class), and each mating type produces offspring by ordinary Mendelian segregation (Chapter 2). Working through the offspring generation shows that heterozygosity ( $2pq$  at HWE) declines each generation, while the homozygote classes ( $p^2$  and  $q^2$ ) each gain the “missing” heterozygotes, roughly in proportion to their own frequency — precisely analogous, in its algebraic form, to the effect of inbreeding described by Wright's  $F$ -statistic (Chapter 6).

### 5.5 Genotype Frequencies Under Partial Assortative Mating

In practice, assortative mating is partial, not complete. Its strength is measured by the **marital (spousal) correlation**,  $r$ , between the phenotypic values of husband and wife for the trait — for example, for height, published studies (including in West African settings) commonly find  $r$  in the range 0.1–0.3, indicating modest positive assortative mating for stature. For a quantitative trait with total phenotypic variance  $V_P$  and additive genetic variance  $V_A$

(Chapter 9), the equilibrium effect of assortative mating with marital correlation  $r$  is to inflate the additive genetic variance transmitted across generations by an approximate factor:

$$V_A^* = V_A / (1 - rh^2)$$

where  $h^2$  is the (narrow-sense) heritability of the trait (Chapter 9). Positive assortative mating ( $r > 0$ ) thus increases the additive genetic variance at equilibrium relative to random mating ( $r = 0$ ), because “like marries like” concentrates trait-increasing and trait-decreasing alleles, respectively, into different lineages, rather than allowing them to average out as under random mating.

### 5.6: Worked Example 5.1

Suppose the narrow-sense heritability of adult height in a Nigerian population is estimated at  $h^2 = 0.80$ , and the observed spousal correlation for height is  $r = 0.20$ . Find the proportional inflation of additive genetic variance due to assortative mating, relative to random mating.

**Solution.**

$$\frac{V_A^*}{V_A} = \frac{1}{1 - rh^2} = \frac{1}{1 - (0.20)(0.80)} = \frac{1}{1 - 0.16} = \frac{1}{0.84} \approx 1.19$$

So assortative mating at this level increases the equilibrium additive genetic variance for height by about 19% relative to random mating — a modest but non-trivial effect that increases the observable variability of stature in the population, and slightly increases the resemblance between relatives (e.g., between parent and offspring, or between siblings) beyond what random mating alone would predict.

**Theorem 5.1 (Equilibrium Additive Variance Under Assortative Mating)**

*For a polygenic trait with narrow-sense heritability  $h^2$  under random mating, subject to constant positive phenotypic assortative mating with spousal phenotypic correlation  $r$  ( $0 \leq r < 1/h^2$ ), the additive genetic variance reaches a new equilibrium value*

$$V_A^* = \frac{V_A}{1 - rh^2}$$

*strictly greater than the random-mating value  $V_A$  whenever  $r > 0$ , and the inflation factor  $1/(1 - rh^2)$  is strictly increasing in both  $r$  and  $h^2$ .*

*Proof sketch.* Positive assortative mating increases the phenotypic (and hence, through the additive genetic component, the genotypic) similarity between mates beyond the random-mating baseline. Each generation, mates' shared additive genetic value induces a positive correlation between the additive genetic deviations they each transmit to their offspring; this correlation,  $r_A = rh^2$  (the spousal phenotypic correlation  $r$ , scaled down by  $h^2$  since only the heritable fraction of phenotypic similarity translates into shared *genetic* value), adds a covariance term to the offspring's additive genetic variance beyond the usual within-family segregation variance. Solving the resulting recursive variance equation,  $V_A^{(t+1)} = V_A^{(0)}(1 + r_A + r_A^2 + \dots)$  as the process iterates toward a stationary distribution, gives the geometric series  $\sum_{k=0}^{\infty} r_A^k = 1/(1 - r_A) = 1/(1 - rh^2)$  (valid since  $0 \leq rh^2 < 1$ ), yielding the stated equilibrium multiplier. Monotonicity in  $r$  and  $h^2$  follows immediately since  $1/(1 - rh^2)$  is an increasing function of the product  $rh^2$ , which is itself increasing in each factor separately.

### **Worked Example 5.2 (Assortative mating for educational attainment, Nigerian context)**

Educational attainment is a polygenic, moderately heritable trait (studies in various populations estimate  $h^2$  in the range 0.3–0.5) that also shows strong positive assortative mating in most societies, including Nigeria, where spousal correlations for years of schooling are often estimated around  $r = 0.4$ –0.5 (partners frequently meet through shared educational or professional environments — universities, workplaces requiring similar qualifications, and professional associations). Taking  $h^2 = 0.4$  and  $r = 0.45$  as illustrative values, Theorem 5.1 gives  $V_A^*/V_A = 1/(1 - 0.45 \times 0.4) = 1/(1 - 0.18) = 1/0.82 \approx 1.22$  — a 22% inflation in additive genetic variance for educational attainment, which (as with height in Worked Example 5.1) is expected to modestly increase the observed phenotypic variance in educational outcomes across the Nigerian population, and to slightly increase sibling and parent-offspring resemblance for this trait beyond the random-mating baseline, an effect of practical relevance to behaviour-genetic and educational-policy research alike.

### **5.7: Assortative Mating versus Inbreeding: A Key Distinction**

Although both increase homozygosity relative to random mating, assortative mating and inbreeding differ in an essential way: **inbreeding increases homozygosity at every locus in the genome simultaneously** (because it is based on shared ancestry, which affects the whole genome), whereas **phenotypic assortative mating increases homozygosity mainly at loci that influence the trait being assorted on** (and at loci in linkage disequilibrium with them — Chapter 15), leaving the rest of the genome close to Hardy–Weinberg proportions. This distinction matters in Nigerian genetic epidemiology: elevated homozygosity

confined to disease-relevant loci (e.g., skin-tone loci if mate choice correlates with complexion) points toward assortative mating, whereas genome-wide elevated homozygosity, together with an elevated rate of rare recessive disease in the offspring of related couples, points toward consanguinity/inbreeding.

### **Theorem 5.2 (Assortative Mating Preserves Allele Frequency)**

*Under any mating system — including positive or negative assortative mating — in which every individual in the population contributes gametes and mating does not otherwise alter viability or fertility, the population allele frequency in the offspring generation equals the population allele frequency among the parental gametes, regardless of the pattern of assortment.*

*Proof.* The offspring generation's allele frequency is, by definition, the total count of A alleles among all offspring gametes divided by twice the number of offspring. Every offspring receives exactly one allele from its mother and one from its father, and every parental individual contributes gametes to the offspring pool regardless of whom it mates with — assortative mating changes which genotypes pair up, not whether or how many gametes each genotype contributes to the next generation's gamete pool. Summing over all mating pairs, the total number of A alleles transmitted equals the total number of A alleles present in the parental gamete pool (since each allele is transmitted, just to a partner chosen non-randomly). Hence the offspring allele frequency equals the parental gamete-pool allele frequency exactly, proving that assortative mating alone — in the absence of differential fertility or viability — changes only genotype frequencies (as shown in Section 5.5), never allele frequencies.

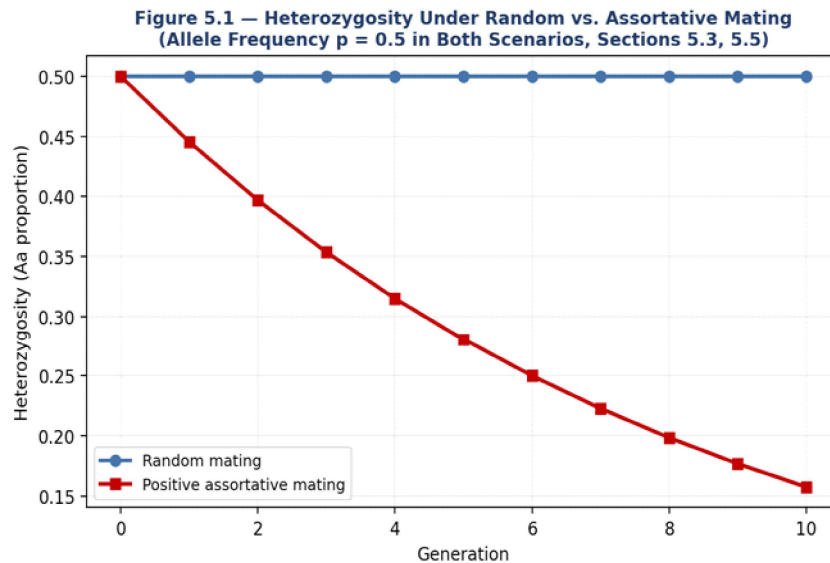


*Figure 5.1 — Heterozygosity under random vs. positive assortative mating (Sections 5.3, 5.5).*

## 5.8: Chapter Summary

- Assortative mating is non-random mating based on phenotype similarity (positive) or dissimilarity (negative); it does not, by itself, change allele frequencies but does distort genotype frequencies away from Hardy–Weinberg proportions.
- Positive assortative mating increases homozygosity and decreases heterozygosity at loci influencing the assorted trait, generation upon generation.
- For quantitative traits, assortative mating (measured by spousal phenotypic correlation  $r$ ) inflates the equilibrium additive genetic variance by a factor  $1/(1 - r^2)$ .
- Assortative mating is locus/trait-specific in its genomic footprint, unlike inbreeding, which affects the whole genome.

## 5.9: Computational Corner: Simulating Assortative Mating

### R.

```
# Simulate genotype-frequency change over generations under
# partial positive
# assortative mating for a single bi-allelic locus (Section 5.5)
simulate_assortative <- function(p0, alpha, generations = 10) {
  # alpha: degree of assortment, 0 = random mating, 1 = complete
  # assortment
  p <- p0
  freqs <- data.frame(gen = 0, AA = p^2, Aa = 2*p*(1-p), aa =
    (1-p)^2)
  for (g in 1:generations) {q <- 1 - p
    AA <- p^2 + alpha * p * q      # heuristic enrichment of
    homozygotes
    aa <- q^2 + alpha * p * q
    Aa <- 1 - AA - aa
    freqs <- rbind(freqs, data.frame(gen = g, AA = AA, Aa = Aa,
    aa = aa))
    p <- (AA + Aa/2) # allele frequency unchanged by assortment
    alone}
  freqs}
result <- simulate_assortative(p0 = 0.5, alpha = 0.3)
print (result)
```

### Python.

```
import pandas as pd
```

```
def simulate_assortative(p0, alpha, generations=10):
  p = p0
  rows = [{ "gen": 0, "AA": p**2, "Aa": 2*p*(1-p), "aa": (1-
    p)**2}]
```

```

for g in range(1, generations + 1):
    q = 1 - p
    AA = p**2 + alpha * p * q
    aa = q**2 + alpha * p * q
    Aa = 1 - AA - aa
    rows.append({"gen": g, "AA": AA, "Aa": Aa, "aa": aa})
    p = AA + Aa / 2
return pd.DataFrame(rows)

result = simulate_assortative(0.5, 0.3)
print(result)

```

The simulation confirms the key qualitative result of Section 5.5: positive assortative mating leaves the allele frequency,  $p$ , essentially unchanged generation to generation, while steadily enriching the population for the two homozygote classes (AA and aa) at the expense of heterozygotes (Aa) — the opposite genotypic signature from inbreeding-driven homozygosity, though both processes reduce heterozygosity (contrast with Section 5.7).

### Mini-Project

#### Mini-Project 5: Comparing Assortative Mating to Random Mating by Simulation

*Build a discrete-generation simulation comparing genotype frequency evolution under random mating versus positive assortative mating.*

- Step 1:** Implement a simulated mating population under random mating for a single biallelic locus, tracking genotype frequencies over 10 generations.
- Step 2:** Implement a second simulation under positive assortative mating (Section 5.3), using the same starting allele frequency.

- Step 3:** Plot heterozygosity over time for both scenarios on the same axes.
- Step 4:** Confirm numerically (Theorem 5.2, if added) that allele frequency remains constant across generations in both scenarios.
- Step 5:** Write a short paragraph distinguishing what assortative mating changes (genotype frequencies) from what it leaves unchanged (allele frequencies).

**Deliverable:** *Simulation code, a comparative heterozygosity plot, and a short written explanation distinguishing the two mating systems' effects.*

### 5.10: Practice Questions

- Q1.** Explain, in your own words, why complete positive assortative mating at a single bi-allelic locus leaves allele frequencies unchanged but reduces heterozygosity over successive generations.
- Q2.** The heritability of systolic blood pressure in a Nigerian adult cohort is estimated at  $h^2 = 0.40$ , and the spousal correlation for blood pressure is  $r = 0.15$ . Compute the proportional inflation of additive genetic variance due to assortative mating.
- Q3.** A genetic counsellor in Kano observes elevated homozygosity across many unrelated loci genome-wide in a set of married couples' children, together with a higher-than-expected rate of rare autosomal recessive disorders. Is this pattern more consistent with phenotypic assortative mating or with consanguinity? Justify your answer.
- Q4.** If spousal correlation for a trait were  $r = 0$  (random mating with respect to that trait), what does the formula  $V_A^*/V_A = 1/(1 - rh^2)$  reduce to, and why does this make sense?
- Q5.** A Nigerian population geneticist finds elevated homozygosity at a skin-pigmentation-associated locus, but

genome-wide homozygosity (measured at hundreds of unrelated background markers) is close to Hardy–Weinberg expectation. Using the distinction in Section 5.7, what does this pattern suggest about the mating pattern in this population, and why?

- Q6.** Using the R or Python code of Section 5.9, increase the assortment parameter *alpha* from 0.3 to 0.8 (stronger positive assortative mating) and report how the equilibrium heterozygote proportion changes after 10 generations compared with the original simulation.