The Hardy-Weinberg equilibrium

There are some concepts in biology that are best learned and understood by problem-solving. One such concept is the Hardy-Weinberg equilibrium, which states that allele and genotype frequencies within a population will remain constant from generation to generation unless some external influences act directly on those alleles or genotypes. Here, geneticist Kevin O'Dell takes you on an imaginary journey to the fictitious Scottish Island of Eilean nan Caorach Geala (White Sheep Island)

One of the most successful Scottish businesses of recent years is Eilean nan Caorach Geala Farm. The farm occupies the whole of Eilean nan Caorach Geala (see Figure 1), a small, fictitious 25 km^2 island in the Outer Hebrides, off the coast of northwest Scotland. The owners and brains behind the enterprise are Farmer Collogie, a former genetics graduate, and his partner Simon, an animal behaviour expert who prefers to be called Si. It is this unique combination of genetic and animal behaviour skills of Farmer Collogie and Si that appears to have given Eilean nan Caorach Geala Farm the edge over all of their competitors. Understandably they are very proud of their unique stock of pure-bred Caorach Geala sheep, which are the only sheep on the island.

Figure 1 Map of Eilean nan Caorach Geala (White Sheep Island)

The success of Eilean nan Caorach Geala Farm was severely compromised when, early in 2010, the whole of western Scotland was struck by the 'storm of the century'. Caught up in the carnage of storm Elizabeth was the French ship *The Auld Alliance* and its cargo of around a hundred pure-bred Mouton Blanc (French White sheep). Unfortunately, in the chaos of storm Elizabeth, nobody noticed that the wreck of *The Auld Alliance* reached Eilean nan Caorach Geala, potentially depositing panic-stricken French sheep onto the island.

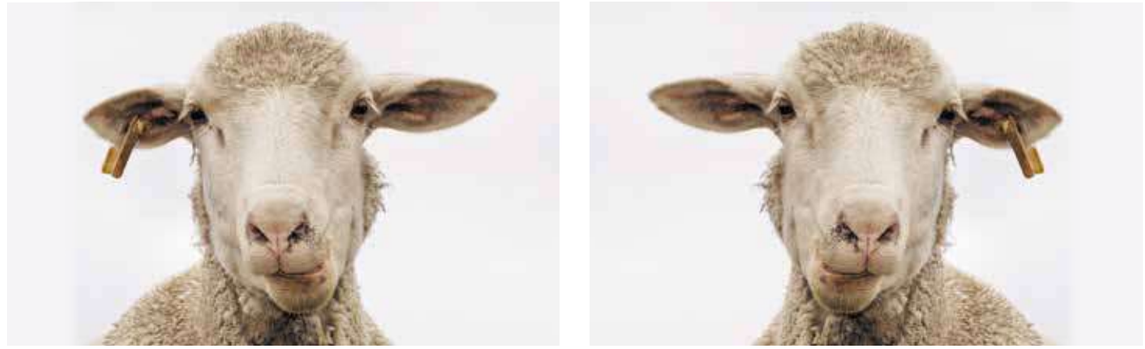

Figure 2 Mature Caorach Geala sheep (left) and Mouton Blanc sheep (right) illustrating how phenotypically similar they are. To help researchers distinguish them, Caorach Geala sheep are given identification tags in their right ear, whereas Mouton Blanc sheep are given identification tags in their left ear

The Caorach Geala and Mouton Blanc breeds of sheep are very difficult to tell apart (see Figure 2). Both breeds are almost entirely white, though some of the Caorach Geala sheep have a dark blaze (essentially a broad line of dark pigmentation) on their faces. The dark blaze has never been found in pure-bred Mouton Blanc sheep. In 2009, the last year prior to storm Elizabeth for which data are available, a census of the sheep on Eilean nan Caorach Geala revealed that 607 (74.4%) had entirely white faces, whereas 209 (25.6%) had white faces with dark blazes.

Question 1 The 2009 census reveals that 607 Caorach Geala sheep had white faces, while 209 had dark blazes, which looks suspiciously like a 3:1 ratio. However, this doesn't tell you anything about the genetic basis of dark blazes in Caorach Geala sheep. Why?

You had become aware of the fascinating story of Eilean nan Caorach Geala earlier this year, when your biology teacher had encouraged you to read Biological Sciences Review, a magazine that focuses on contemporary issues in biology. Your teacher also suggested you try to find a summer internship in a biology research laboratory. So you wrote a speculative email to Farmer Collogie and now, to your great surprise, you are sailing to Eilean nan Caorach Geala to see how the local sheep population is coping with the alleged French invasion.

When you arrive on Eilean nan Caorach Geala and set to work, you discover there is a whole army of researchers studying the island's sheep and their environment. While much of the island is grassland, they have discovered that pure-bred Caorach Geala sheep have a particularly efficient metabolism that allows them to convert more of their favourite food, seaweed, into muscle. Seaweed is found in vast quantities south of the farmhouse on the coast along Wool Bay. Researchers believe that it is the seaweed that gives Caorach Geala sheep meat its high-quality flavour, but this is as yet unproven. The researchers have also discovered that the metabolism of pure-bred Mouton Blanc sheep is significantly less efficient than their Caorach Geala cousins.

Figure 3 The Eilean nan Caorach Geala Research Centre. The building itself is in need of some substantial renovation

You soon discover that the Eilean nan Caorach Geala research centre (see Figure 3) in the west of the island is not among the best-funded research institutes in the world. As a consequence, some of the more expensive experiments you might like to do, such as whole genome sequencing, simply are not financially possible. However, the research centre has access to *Baanet*, the Sheep Genome Project database. Whole genome sequences of typical pure-bred Caorach Geala sheep and typical pure-bred Mouton Blanc sheep are available in the *Baanet* database.

Farmer Collogie is keen to know whether the sheep on Eilean nan Caorach Geala in the spring of 2020 are all descended from the original pre-2010 Caorach Geala sheep stock, or whether some of the Mouton Blanc sheep survived the horrors of Storm Elizabeth and subsequently bred on the island.

Question 2 Bearing in mind the funding limitations at the Eilean nan Caorach Geala research centre, design a simple genetic test that would allow you to identify whether a sheep was of pure-bred Caorach Geala, pure-bred Mouton Blanc or mixed ancestry. What, if any, are the limitations of your proposed test?

Your tests reveal that there are indeed sheep on Eilean nan Caorach Geala that have Mouton Blanc ancestry, indicating that at least some of the sheep from *The Auld Alliance* managed to survive, thrive and reproduce on the island. The full details are shown in Table 1.

Question 3 What are the frequencies of the Caorach Geala-derived and Mouton Blanc-derived alleles in the island population?

Table 1 Number of sheep on Eilean nan Caorach Geala that, according to the simple genetic test that you have designed, are of pure-bred Caorach Geala (CG/CG), pure-bred Mouton Blanc (MB/MB) or mixed ancestry (CG/MB). These data were collected in the spring of 2020. CG is the Caorach Geala-derived allele and MB is the Mouton Blanc-derived allele

Ancestry	Caorach Geala only	Mixed	Mouton Blanc only
Genotype	CG/CG	CG/MB	MB/MB
Number of sheep	666	188	146

You sensibly decide to investigate whether the genotypes identified by the single gene test you have developed reveal whether the Eilean nan Caorach Geala sheep are in Hardy-Weinberg equilibrium. Remember, it has got something to do with p^2, $2pq$ and q^2. The chi-squared formula (Biological Sciences Review, Vol. 33, No. 2, pp. 30–33) is:

$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

and for 1 degree of freedom, the critical value for the 5% significance level is 3.84.

Question 4 Undertake a chi-squared test to determine whether the genotypes identified by your single gene analysis reveal whether the sheep on Eilean nan Caorach Geala are in Hardy-Weinberg equilibrium.

Question 5 Propose two different hypotheses why the genotypes you have been investigating in the sheep on Eilean nan Caorach Geala are, or are not, in Hardy-Weinberg equilibrium.

Question 6 Design an experiment that will distinguish between the two hypotheses you proposed in your answer to question 5.

You are excited by the success of your first big experiment, and immediately visit Farmer Collogie to explain what you have discovered. However, when you arrive at the farmhouse you are astonished to see police and agents of the Scottish Highlands & Islands Testing Executive surrounding the building. They believe, quite correctly as it later transpires, that Eilean nan Caorach Geala Farm is not the idyllic and successful enterprise that everyone had originally thought. Farmer Collogie has been systematically injecting his sheep with Tastine, a banned drug known to enhance the quantity, quality and flavour of muscle in sheep.

Farmer Collogie's appalling and illegal behaviour had only been discovered when the research team discovered the carcasses of six dead sheep on Flocking Hill in the northwest of the island. The research team quickly established that the six dead sheep were pure-bred Mouton Blanc animals. They speculate that the indigenous Caorach Geala sheep are tolerant to Tastine, whereas the invading Mouton Blanc sheep are not.

Suddenly a lot of money is available for a full-scale investigation into the Eilean nan Caorach Geala Farm Tastine scandal. You correctly suggest that there must be a gene that determines whether sheep are tolerant or susceptible to Tastine, and that pure-bred Caorach Geala sheep are tolerant, whereas pure-bred Mouton Blanc sheep are not. But how are you going to identify which gene confers Tastine tolerance, and how are you going to determine *how* the gene confers Tastine tolerance?

Question 7 Design an experiment that will allow you to discover whether the Tastine-tolerant allele found in pure-bred Caorach Geala sheep is dominant or recessive.

The new funds allow the research team to undertake whole genome sequencing on as many sheep as they can capture. Quite surprisingly the team is also given ethical approval that allows team members to inject the sheep with low doses of Tastine, to determine which sheep are tolerant and which are susceptible to the drug.

Question 8 Using the whole genome sequence data, and data on Tastine tolerance and susceptibility, describe a strategy that will allow you to find the sheep gene that determines tolerance or susceptibility to Tastine.

Your strategy identifies a gene that may well be the cause of Tastine tolerance in pure-bred Caorach Geala sheep. As Tastine tolerance has, as yet, never been found in any other sheep population, it is safe to conclude that Tastine sensitivity is the ancestral version of the gene. According to *Baanet*, the Sheep Genome database, no significant research has ever been done on this gene and comparative sequence analysis reveals absolutely nothing about its function. However, the comparative sequence analysis does reveal several differences between what you suspect are the Tastine-tolerant and Tastine-susceptible alleles in sheep currently living on Eilean nan Caorach Geala.

Question 9 Taking each mutation in turn, speculate on which of the mutations listed in Table 2 could cause the Tastine tolerance found in pure-bred Caorach Geala sheep.

Table 2 Mutations in or around the gene you have identified that may cause dominant Tastine tolerance in any sheep fortunate enough to be homozygous or heterozygous for it. Note that codons are numbered from the beginning of the gene, so codon 1 is always the start codon ATG

Number	DNA change in Tastine-tolerant sheep
1	T to A 12 562 base pairs upstream of transcriptional start
2	Codon 41: histidine (CAT) to arginine (CGT)
3	Codon 104: serine (TCG) to threonine (ACG)
4	12bp deletion within intron 3
5	Codon 326: serine (TCA) to STOP (TAA)
6	43bp insertion within intron 4

Question 10 How would you prove which of the mutations listed in Table 2 results in Tastine tolerance in sheep?

BiologicalSciencesReview*Extras*

Go online for the answers at:www.hoddereducation.co.uk/bioreviewextras

Kevin O'Dell is professor of behaviour genetics at the University of Glasgow. He is also an editor for Biological Sciences Review. If you like this style of problem solving, you might like his textbook *Genetics? No Problem!* which was published in 2017.