Researchers in Schools

Where do our genes come from? Finding out with maths

Key Stage 4 Programme

Pupil Name

Coursebook Designed by Dr Smith

Table of Contents

Uni Pathways launch event **3**

Introduction to Uni Pathways *3*

Pupils' experience of Uni Pathways *4*

Virtual campus tours *5*

Meet university students! *6*

Reflection *7*

Timetable and Assignment Submission **9**

Timetable – Tutorials *9*

Timetable – Homework Assignments *9*

Assignment Submission – Lateness and Plagiarism *9*

Course rationale **10**

Uni Pathways Mark Scheme **11**

Baseline Assignment: Pupil Feedback Report **13**

Final Assignment: Pupil Feedback Report **14**

Subject vocabulary **15**

Tutorial 1 Genetic inheritance and probability **17**

Who do we inherit our features from? *17*

Activity 1 – Starter 17

Activity 2 17

Our genetic code *18*

Activity 3 19

Activity 4 19

Probability *20*

Activity 5 20

Activity 6 21

Homework *22*

Tutorial 2 The Wright–Fisher model of genetic inheritance **24**

Some more probability *24*

Activity 1 – Starter 24

Activity 2 25

The Wright–Fisher model *26*

Activity 3 27

Activity 4 28

Mathematical models *29*

Homework *30*

Tutorial 3 Finding common ancestors **33**

Genealogies ... *33*
Activity 1 – Starter ... 33
Activity 2 ... 33

Finding the most recent common ancestor ... *33*
Activity 3 ... 33
Activity 4 ... 35

Calculating the mean ... *38*
Activity 5 ... 38
Activity 6 ... 38

Homework ... *39*

Tutorial 4 Finding common ancestors using probability ... 41

Activity 1 – Starter ... 41

Tree diagrams ... *41*
Activity 2 ... 43

Mathematical analysis of time of most recent common ancestor ... *44*
Activity 3 ... 44
Activity 4 ... 45
Activity 5 ... 46
Activity 6 ... 47
Activity 7 ... 47

Homework ... *49*

Tutorial 5 The Wright–Fisher model with mutation ... 53

A problem with the Wright–Fisher model ... *53*
Activity 1 ... 53

Genetic mutation ... *54*
Activity 2 ... 55
Activity 3 ... 56
Activity 4 ... 57

Final assignment ... 58

Part A ... *58*

Part B ... *58*

Tutorial 6 Feedback tutorial ... 60

Final assignment feedback from my Uni Pathways Teacher ... *60*

Tutorial 7 Final tutorial ... 62

Final assignment feedback from my Uni Pathways Teacher ... *62*

University ... *62*

Reflecting on Uni Pathways ... *63*

Appendix 1 Referencing correctly ... 64

Notes ... 66

Uni Pathways launch event

Welcome to Uni Pathways! We are delighted to be able to launch the Uni Pathways programme with you. We hope that you are ready to embark on your Uni Pathways journey and that you enjoy the video.
You will be asked to pause the video at times to complete some work in this workbook, so make sure you have a pen / pencil to hand when you start the video.

By the end of the video, you will have

- Learnt about what studying at university means
- Learnt about some of the skills that you will develop during Uni Pathways
- Heard from pupils who have participated in Uni Pathways or The Scholars Programme (which is the same programme!)
- Heard from current university students talking about what life is like at university

If you are in school your teacher will play the video. If you are at home and logged in to a session with your teacher, your teacher will play the video and show it to you. If you are participating independently your teacher will email you the link to the video or the video file. There are opportunities for you to answer some questions, and you will be told when to pause the video to answer them.

Introduction to Uni Pathways

1. Write down what you think a supra-activity is

2. Independent learning is

3. Write down in your own words what resilience means, and come up with a different example to the one mentioned in the PowerPoint.

4. Why do you think analysis and evaluation skills are useful in your academic career?

5. Why do you think analysis and evaluation skills are useful in life in general?

6. What are the different types of learning that university students do?

Pupils' experience of Uni Pathways

You will now watch some pupils talk about their experience of learning a PhD topic and producing a final assignment. Some may refer to The Scholars Programme rather than Uni Pathways. As you heard in the introduction PowerPoint, The Scholars Programme is a very similar programme to Uni Pathways. Listen carefully and then respond to the questions below. Be prepared to share some of your responses with your peers.

1. How did the pupils describe talking about writing a long essay for their final assignment?

2. Write down something that one of the pupils mentioned was particularly interesting in their course.

3. What were some challenges that the pupils met?

Virtual campus tours

You will now go on a virtual tour of some universities! Once you have seen some of the university campuses, respond to the questions below.
Note down something that you liked out of any of the campus tours you saw. It may be a particular building, space, city etc.

1. Were there any similarities between the different university campuses? If so, what are the similarities?

2. If you had to pick one of those universities to go and visit in person, which one would it be and why?

3. Note down some of the societies that you could join at different universities

Meet university students!

You are about to virtually meet or hear from some current university students.
If you are virtually meeting them, think about what questions you would like to ask the students, share your questions with a partner and note those questions down in the space below.

If you are not meeting them but watching some videos that they have made, use the space below to note down what else you would like to find out about university. You can then ask your teacher during Uni Pathways tutorials!

Reflection

Congratulations on completing the launch of the Uni Pathways programme. Before you go, take some time to reflect on what you have learnt by answering the following questions:

1. Are there any aspects of university style learning that you would look forward to? If so, what are they?

2. What challenges do you think students face when learning at university? How would you try to overcome these challenges?

3. What skills do you hope to develop during your Uni Pathways course?

4. What part of Uni Pathways seems the most challenging for you?

5. What are you most looking forward to about Uni Pathways?

Timetable and Assignment Submission

Timetable – Tutorials

Tutorial	Date	Time	Location
1	Thursday 29th April	15:20	TBC
2	Thursday 6th May	15:20	
3	Thursday 13th May	15:20	
4	Thursday 20th May	15:20	
5	Thursday 27th May	15:20	
6 (Feedback)	Thursday 17th June	15:20	
7 (Feedback)	Thursday 24th June	15:20	

Timetable – Homework Assignments

Homework Assignment	Description	Due Date
Tutorial 1	Baseline assignment	Thursday 6th May
Tutorial 2	The Wright–Fisher model	Thursday 13th May
Tutorial 3	Finding common ancestors	Thursday 20th May
Tutorial 4	Finding common ancestors using probability	Thursday 27th May
Tutorial 5	Final assignment	Thursday 10th June

Assignment Submission – Lateness and Plagiarism

Lateness	
Submission after midnight on 9th August	10 marks deducted
Plagiarism	
Some plagiarism	10 marks deducted
Moderate plagiarism	20 marks deducted
Extreme plagiarism	Automatic fail

Course rationale

We all know that our genes come from our parents. This is why we look a bit like both of them. How our bodies work and how we look is determined by our *genetic code*, called DNA, and this contains information from both of our parents. Our DNA codes for things like our hair colour, eye colour, our height, shoe size, and much much more.

DNA is one big long string of chemical compounds. We can zoom in on this and look at one very small portion called a *gene*. The combination of genes in our DNA is what determines how we look and function as humans. Different people have different combinations —- that's why we all look different —- but how many genes do we have in common exactly?

Let's think about one particular gene in your DNA. Suppose that you inherited it from your father, and that he inherited it from his mother, your grandmother, and that she from her mother too, your great-grandmother. We could keep doing this, going far back into history, and we would get a big long list of people from whom you have inherited this particular gene from.

Now let's make the same list, but this time for someone else — we could choose anyone else on Earth if we liked. Again, we get a big long list from whom they inherited this gene.

If we go back far enough in time, do these lists eventually contain the same people? In these tutorials we'll see that a famous mathematical model of inheritance, the Wright—Fisher model, predicts that *yes,* they do: eventually, if you go back enough generations (maybe thousands), you and this person will both have inherited this gene from someone alive in the past.

Using maths, we'll investigate this further to see just how much in common our genetic code is with others', and how related we really are.

Uni Pathways Mark Scheme

	Subject Knowledge	Critical Thinking	Written Communication
1st	The work shows a depth of knowledge and understanding of key concepts and mathematical methods, drawing on sources of evidence where appropriate. Knowledge is used to build and support highly effective mathematical argume nts and explanations.	Analyses the available data/evidence and identifies all the relevant mathematical information. Identifies and critically evaluates available mathematical approaches and evidence, deciding on their credibility, strength and relative significance. Makes connections between different pieces of information and puts them together to form accurate solutions and verify answers and assumptions.	Mathematical arguments are complete, and the reasoning behind each step has been clearly communicated to the reader. The work has a coherent flow and is well structured. The writing style is appropriate; mathematical language and key mathematical terms are used accurately and effectively to support the arguments and explanations made. There are no, or very few, errors in spelling or grammar. Consistent referencing, appropriate paragraphing and use of correctly labelled tables and figures; matching the style taught in the course.
2:1	The work shows an understanding of key concepts and mathematical me thods, drawing on sources of evidence where appropriate. Knowledge is used to build and support effective mathematical argume nts and explanations.	Analyses the available data/evidence and identifies all the relevant mathematical information. Engages with multiple approaches and critically evaluates their utility, forming judgements and using reasoning and evidence to inform problem solving strategy. Shows some understanding of the relative value of evidence and data.	Mathematical arguments are clearly made, with few gaps in logical reasoning. The work is well-structured. The writing style is appropriate; mathematical language and key terms are used correctly. There are few errors in spelling or grammar. Mostly consistent referencing and use of tables and figures; matching the style taught in the course.

2:2	The work shows an understanding of key concepts and mathematical methods, with no major misconceptions. Beginning to apply this knowledge to build and support effective mathematical argume nts and explanations.	Identifies and uses basic mathematical evidence and reasoning. Evaluates the quality of mathematical reasoning and evidence, forming judgements. Not yet showing understanding of the relative value of evidence and reasoning.	Mathematical arguments have been constructed, but there is little explanation from step to step. The work has some structure. The writing style can sometimes be informal; occasionally key terms are not used when it would be appropriate to do so. There are some errors in grammar and spelling, but these do not get in the way of communicating the content. Referencing has some consistency; matching the style taught in the course. Limited use of tables and figures.
3rd	Shows a developing understanding of key concepts and mathematical methods, with some misconceptions. Not yet applying this knowledge to build and support mathematical argument s and explanations.	Beginning to analyse mathematical evidence and reasoning. Describes or attempts to solve the problem without any/or little evaluation of whether the approach was reasonable or accurate.	Mathematical arguments to solve problems are poorly sequenced, with little or no explanation. The grammar, spelling, style, and structure of the work need improving in order to communicate ideas to the reader. Key terms are not always used correctly. Limited, or no use of referencing, tables and figures.

Baseline Assignment: Pupil Feedback Report

Name of pupil	
Name of School	
Name of RIS Teacher	
Title of Assignment	

How your assignment is graded:

Grade	Marks	What this means
1st	70+	Performing to an excellent standard at A-level
2:1	60-69	Performing to a good standard at A-level
2:2	50-59	Performing to an excellent standard at GCSE
3rd	40-49	Performing to a good standard at GCSE
Working towards a pass	0-39	Performing below a good standard at GCSE
Did not submit	DNS	No assignment received by The Brilliant Club

Lateness	
Any lateness	10 marks deducted
Plagiarism	
Some plagiarism	10 marks deducted
Moderate plagiarism	20 marks deducted
Extreme plagiarism	Automatic fail

Marks

OVERALL MARK /100		FINAL MARK /100 (including any deductions)	
DEDUCTED MARKS		FINAL GRADE	
If marks have been deducted (e.g. late submission, plagiarism) the teacher should give an explanation in this section:			

Mark breakdown and feedback

Subject knowledge	
	mark
Critical thinking	
	mark
Written communication	
	mark
Problem set	
	mark

Final Assignment: Pupil Feedback Report

Name of pupil	
Name of School	
Name of RIS Teacher	
Title of Assignment	

How your assignment is graded:

Grade	Marks	What this means
1st	70+	Performing to an excellent standard at A-level
2:1	60-69	Performing to a good standard at A-level
2:2	50-59	Performing to an excellent standard at GCSE
3rd	40-49	Performing to a good standard at GCSE
Working towards a pass	0-39	Performing below a good standard at GCSE
Did not submit	DNS	No assignment received by The Brilliant Club

Lateness	
Any lateness	10 marks deducted
Plagiarism	
Some plagiarism	10 marks deducted
Moderate plagiarism	20 marks deducted
Extreme plagiarism	Automatic fail

Marks

OVERALL MARK /100		FINAL MARK /100 (including any deductions)	
DEDUCTED MARKS		FINAL GRADE	
If marks have been deducted (e.g. late submission, plagiarism) the teacher should give an explanation in this section:			

Mark breakdown and feedback

Subject knowledge	
	mark
Critical thinking	
	mark
Written communication	
	mark
Problem set	
	mark

Subject vocabulary

Word	Definition	In a sentence
Model	A representation of an idea of system, often described in a mathematical language.	Sewall Wright and Ronald Fisher separately proposed a model for genetic inheritance.
Wright—Fisher model	A model for how genes are inherited from previous generations.	The Wright—Fisher model has been hugely influential in the study of genetics.
Probability	A probability is a number between 0 and 1 that tells you how likely something is to happen. Probabilities can be written as fractions, decimals or percentages.	The probability of getting a 6 when rolling a fair die is 1/6.
Random variable	A number *X* which takes different values with different probabilities.	The outcome when I roll a fair die is a random variable. It takes values 1, 2, 3, 4, 5, 6 each with probability 1/6.
Tree diagram	Tree diagrams are a way of showing combinations of two or more events. Each branch is labelled at the end with its outcome and the probability is written alongside the line.	Tree diagrams are used to calculate the probability of complicated events.
Independent	Two events are independent if the probability of the first event happening has no impact on the probability of the second event happening.	The outcomes of each flip of a fair coin are independent.
DNA	DNA stands for deoxyribonucelic acid. It is a chemical made up of two long molecules, arranged in a spiral. It has all the instructions that a living organism needs to grow, reproduce and function.	All of our genetic information is encoded by our DNA.
Gene	Genes are short sections of DNA.	Genes carry information for particular characteristics, such as ear shape or eye colour.
Ancestor	An ancestor is any person from whom someone is descended, eg one of their parents, grandparents and so on.	My great-grandmother is one of my ancestors.
Generation	All of the people born and living at about the same time.	You and your friend's grandparents belong to the same generation.

Genealogy	A genealogy traces all the ancestors of a single person. It often takes the form of a family tree.	My genealogy tells me all the people that I am related to in the past.
Gene frequency	The frequency of a particular gene is the proportion of individuals in a population who carry it in their DNA, expressed as a fraction or a percentage.	The frequency of the gene for white flowers on pea plants is 22%.
Mutation	A mutation occurs when a gene is copied incorrectly from parent to child, so that the child carries a different gene to their parent.	Some diseases such as cancer are caused by a gene mutating.

Tutorial 1
Genetic inheritance and probability

By the end of this tutorial, I will be able to

1. describe and explain how our genetic code is stored and how it is inherited from our parents;
2. identify the need for a mathematical model for inheritance; and
3. define and calculate the probability of simple events.

Who do we inherit our features from?

Activity 1 – Starter

Think of one of your features (for example, your eye colour, hair colour, shape of your lips) and, as far back as you can go, list whom your inherited this from, and whom they inherited that from, and so on.

For example, Dr Smith inherited his blue eyes from his mother, who inherited them from his (maternal) grandmother, who inherited them from her father (his great-grandfather). How many generations back can you go?

Activity 2

Write down two tools that we could use to look further back into history and extend the length of your list. For example, we could look at photographs if we have them.

For each, write down a limitation (or drawback) of this approach.

Our genetic code

DNA stands for **deoxyribonucleic acid**. It is a chemical made up of two long molecules. The molecules are arranged in a spiral, like a twisted ladder. We call this the **double helix** structure.

There is DNA in the **nucleus** of every cell. DNA carries genetic information. It has all the instructions that a living organism needs to grow, reproduce and function.

In a cell nucleus, DNA is organised into coiled strands called **chromosomes**. Humans have 46 chromosomes in each cell.

Half the chromosomes are inherited from one parent and half from the other. As humans, therefore, we have **23 chromosomes** from each parent.

This explains why organisms can share characteristics from both parents. A child, for example, might have red hair like their dad and long fingers like their mum.

Genes are short sections of DNA. Genes carry information for particular characteristics, such as ear shape or eye colour. Different sets of genes carry information for different characteristics. There are many genes in a chromosome.

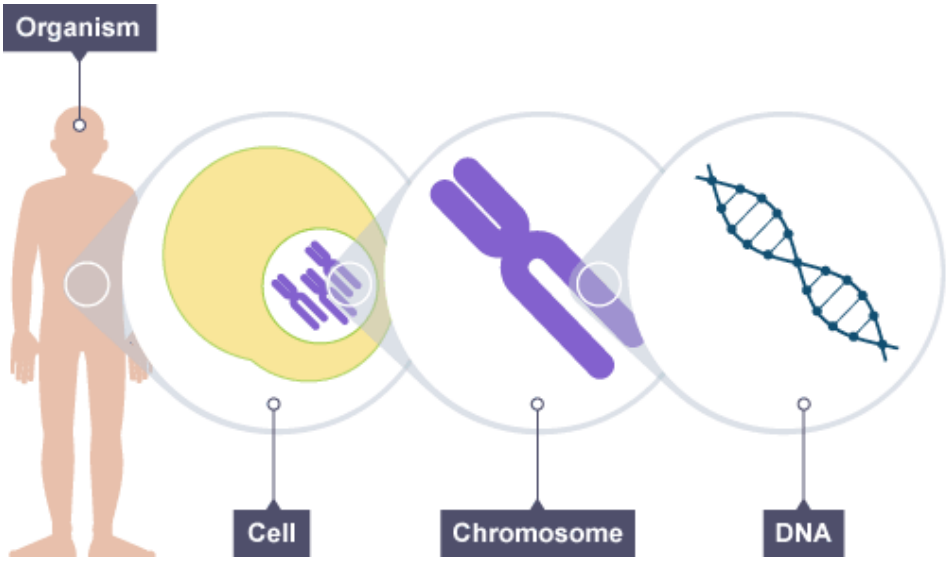


Figure 1

The diagram shows the relationship between the cell, its nucleus, chromosomes in the nucleus, and genes. Source: BBC Bitesize

Activity 3

Figure 2 shows a particular chromosome belonging to an individual.

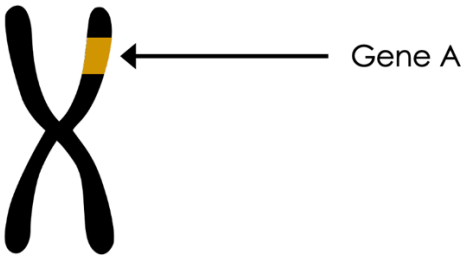


Figure 2

In your own words, describe the possible ways in which Gene A is inherited.

Activity 4

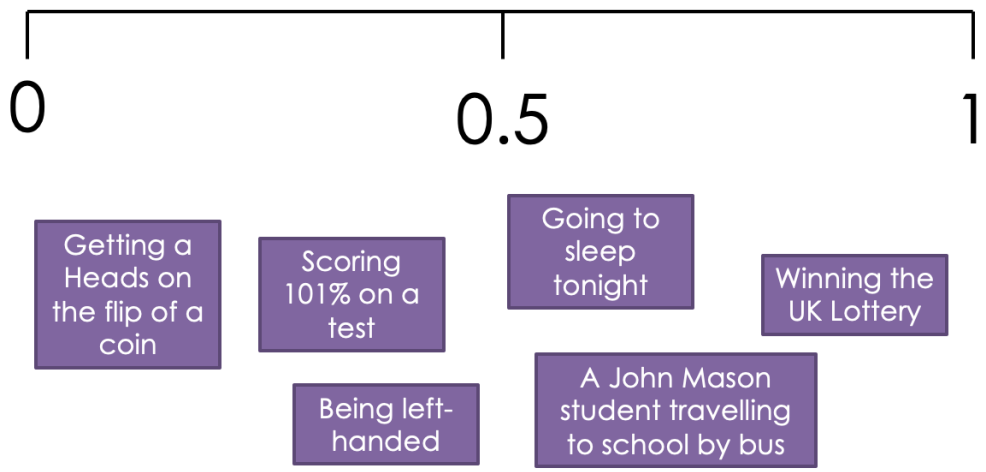
In your own words, explain how your genetic code is stored and inherited. In your answer, you should make sure that you use the key words: *cell, DNA, gene, chromosome*.

Probability

In order to use maths to calculate things about genetics we will need to remind ourselves how to calculate with probability. We will do this for the rest of this tutorial and in other tutorials, to come.

Activity 5

Match up the events below with where they should be on the probability scale. Be ready to share your answers with the group.



The **probability** of an event is

$$P(\text{Event}) = \underline{\hspace{6cm}}$$

An event with probability 1 happens with certainty. An event with probability 0 never happens. An event with probability 0.5 is equally likely to happen as not happen.

Probabilities may equally be written using **decimals**, **fractions** or **percentages**.

Activity 6

Example

What is the probability of picking a jack from a deck of cards?

$$P(\text{Jack}) = \frac{4}{52}.$$

What is the probability of seeing an even number when I roll a fair die?

Look again at Figure 2. A particular gene, A, is labelled on a chromosome. What is the probability that gene A was inherited from the individual's mother?

What is the probability that when I roll two dice, the sum of the scores is 7?

Homework

This homework will give you the opportunity to practice some probability from Tutorial 1. Work of a high standard should include both correct calculations, but you should also carefully explain the mathematics you are doing using full written sentences.

Complete the table by listing all outcomes for each event and then computing the required probability. The first one is done for you.

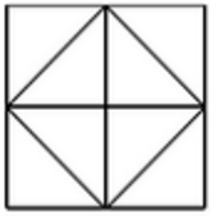
	Event	Outcomes	Probability
1	Getting one heads and one tails on the throw of two coins.	HH, **HT**, **TH**, TT	1/2
2	Getting two tails after two throws.		
3	Getting at least 2 heads after 3 throws.		
4	Getting exactly 2 heads after 3 throws.		
5	Rolling a prime number and throwing a head.		
6	In three throws of a coin, a heads never follows a tails.		
7	For a randomly chosen meal with possible starters **A**vacado, **B**eans and **C**auliflower, and possible main courses **D**og, **E**scalopes or **F**ish, ending up with neither Avacado nor Dog.		

What proportion of your father's genes do you inherit?

On one of his chromosomes, your father has a gene which means he is more likely to get to certain types of cancer. What is the probability that you have inherited this gene?

If both your mother and your father have this gene, what is the probability you inherited this gene from at least one of your parents?

Hint: there are four possible outcomes.



Challenge question

A square is divided into 8 congruent triangles, as shown. Two of these triangles are selected at random and shaded black. What is the probability that the resulting figure has at least one axis of symmetry?

Hint: it might be easier to count those which are not symmetrical!

Tutorial 2
The Wright–Fisher model of genetic inheritance

By the end of this tutorial, I will be able to

1. compute the probability of an event not happening;
2. define the Wright—Fisher model of genetic inheritance; and
3. understand the process of developing mathematical models.

Some more probability

Activity 1 – Starter

What is the probability that in 3 throws of a coin, a Heads never follows a Tails?

What is the probability of seeing exactly two heads in four throws of a coin?

If A is an event we are interested in the probability of, then we write A' for the event that A **does not happen**. We can compute the probability of the event A' using the simple formula

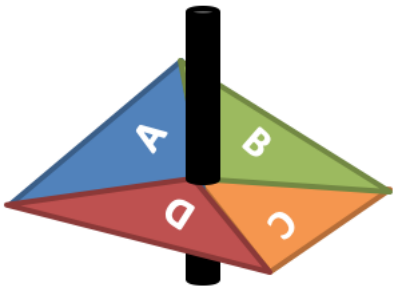
$$P(A') =$$

Activity 2

A bag consists of red, blue and green balls.

The probability of picking a red ball is 1/3 and a blue ball 1/4.

What is the probability of picking a green ball?



An unfair spinner is spun. The probability of getting A, B, C and D is indicated in the table below.

A	B	C	D
0.1	x	0.4	x

Determine x.

The Wright–Fisher model

This Photo by Unknown Author is licensed under CC BY-SA

This Photo by Unknown Author is licensed under CC BY-NC-ND

Figure 3
Sewall Wright (left) and Ronald Fisher (right)

In the last tutorial we learned that our genetic code is stored within 46 chromosomes. We inherit half of these from our father and half from our mother.

In your homework, you calculated that this means that a particular gene is equally likely to have been inherited from your mother or father. We can phrase this mathematically by saying that *the probability a particular gene is inherited from your father is 0.5*.

However, we also saw that if we go back far enough into the past, it is impossible to know who the parents of a particular individual are. This is because there simply aren't enough records to find out the ancestry of all individuals. This is even more so the case if we were interested in the ancestry of animals, plants or bacteria. Imagine trying to find out the ancestry of dandelions in the school playing field!

To get around this **Sewall Wright** and **Ronald Fisher** came up with a model, like a *best guess*, for establishing who genes were inherited from in the previous generation of individuals. They claim that if we don't know, then we should just randomly select from all available individuals.

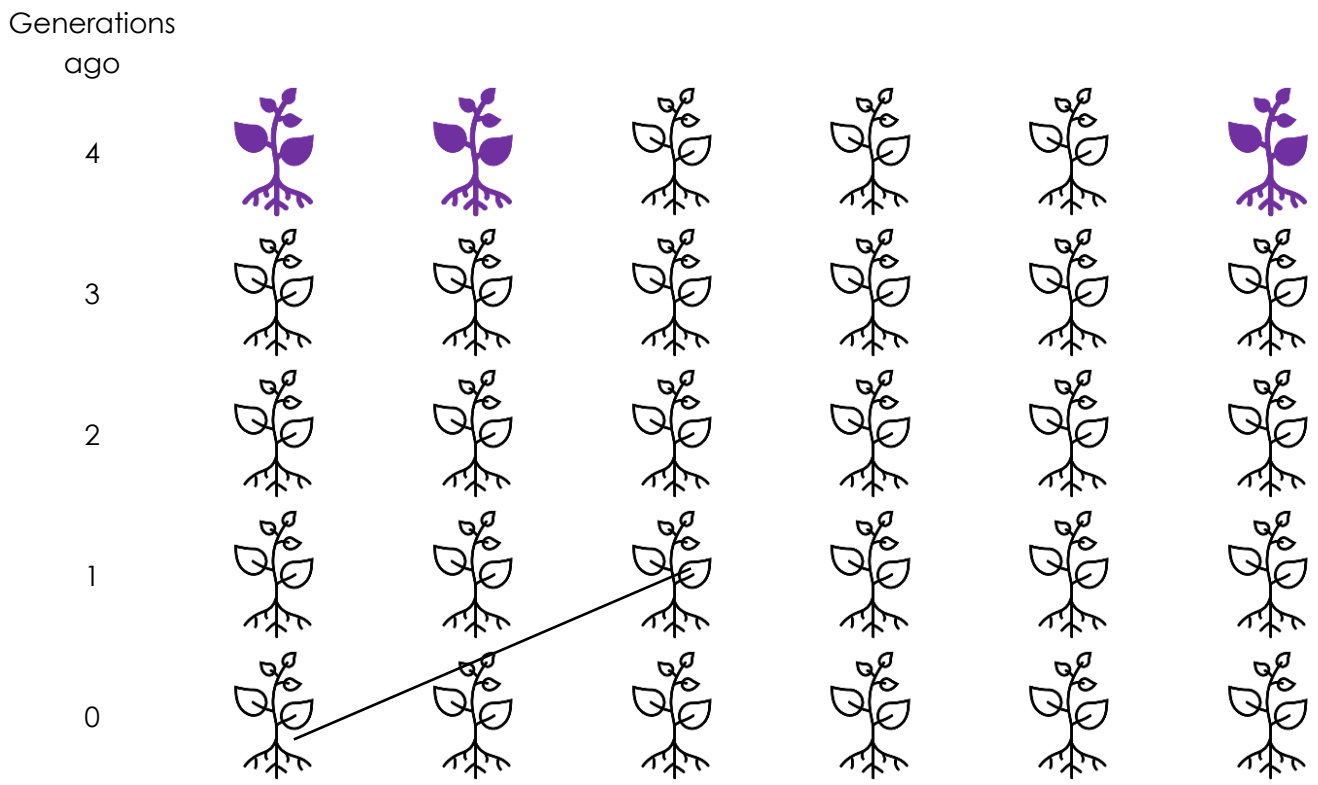
Their model is as follows.

- We model the relationships between **non-overlapping generations**. As an example, the present-day generation consists of all individuals alive today; the previous generation consists of all their parents; the generation before that consists of all their parents, and so on.
- The **population size** in each generation is assumed to be constant. We assume that the population size is N, for some whole number N.
- Each individual inherited a particular gene from someone chosen randomly from the previous generation.

This is best understood with an example, which we will look at now.

Activity 3

So that we can understand the model a bit better, let's see it in action. Below we have **5 generations** of pea plant. The **population size** is $N = 6$.



Pea plants have either white or violet flowers. In a famous experiment, Gregor Mendel discovered that the colour of the pea plant's flowers is controlled by one particular gene. We will use the Wright–Fisher model to model which plants this gene was inherited from in previous generations.

This Photo by Unknown Author is licensed under CC BY-NC-ND

Figure 4

Gregor Mendel is considered the father of modern genetics.

To do this, the Wright–Fisher model says we should select the ancestor of the gene randomly from the previous generation.

As the population size is $N = 6$, we can do this by rolling a fair die.

Let's begin by modelling where the first pea plant in the current generation inherited its gene from. If we roll a 3 on a die, then it inherited it from the third individual in the previous generation. This relationship is marked on the diagram.

1. Using a die, complete the genealogical relationships in this diagram.

2. Four generations ago, 3 of the pea plants in the population have violet flowers.

   If a plant in the next generation inherits its gene from one of these plants, then it will also have violet flowers.

   Using the relationships you just modelled, colour in the diagram to show the colour of the flowers in the 5 generations of pea plants.

3. Compare the colour of flowers in the current generation with your neighbour. Do you have the same outcome? If no, why not?

Activity 4

In your own words, explain what the Wright–Fisher model is used for and how it is formulated.

Mathematical models

A **mathematical model** is a representation of an idea of system, often described in a mathematical language.

Constructing a mathematical model is a process.

1. **Identify the problem**
   We identify something in the real world we want to know, do, or understand. The result is a question in the real world.

2. **Make assumptions and identify variables**
   We select "objects" that seem important in the real-world question and identify relations between them. We decide what we will keep and what we will ignore about the objects. The result is an idealised version of the original question.

3. **Do the maths**
   We translate the idealised version into mathematical terms and obtain a mathematical formulation of the idealised questions. This formulation is the model. We do the maths to see what insights and results we get.

4. **Analyse and assess the solution**
   We consider: does it address the problem? Does it make sense when translated back to the real world? Are the results practical, the answers reasonable, and the consequences acceptable?

5. **Iterate**
   We iterate the process as needed to refine and extend our model.

Homework

The following tables indicate the probabilities for spinning different sides, A, B, C and D, of an unfair spinner. Work out x in each case.

Show your working to the right of each table.

A	B	C	D
0.1	0.3	x	x

A	B	C	D
0.5	2x	0.2	x

A	B	C	D
x	2x	3x	4x

A	B
x	4x + 0.25

I am going on holiday to one destination this year, either France, Spain or America. I'm 3 times as likely to go to France as I am to Spain but half as likely to go to America than Spain. What is the probability that I don't go to Spain?

In the tutorial we saw what the five steps of making a mathematical model are. The first step is to identify a problem. For the Wright—Fisher model, what was the problem we identified?

The second step is to make assumptions. List four assumptions we made for the Wright—Fisher model and explain why we aren't too worried about them.

The third step of modelling is to "do the maths". This will be your task for the next few questions.

If the population size is 16 and half are men, what's the probability of inheriting a gene from a woman in the previous generation?

If 3 of these 16 have a gene for blue eyes, what's the probability of inheriting this gene?

If 7 in a population of size N have a gene which means they can roll their tongue what's the probability of inheriting this gene?

If 30% of the population have a gene for blonde hair, what's the probability of inheriting this gene? Why does it not matter what the population size is?

If 121 pea plants in the previous generation carry the gene for violet flowers and the probability of inheriting the gene for violet flowers is 0.4, what is the total population size?

Additional writing space:

Tutorial 3
Finding common ancestors

By the end of this tutorial, I will be able to

1. construct graphically the genealogical relationships between individuals under the Wright—Fisher model and investigate how these depend on the population size;
2. find the time until the most recent common ancestor of two individuals under the Wright—Fisher model; and
3. calculate the mean time until the most recent common ancestor for different population sizes.

Genealogies

Activity 1 – Starter

Using a die, construct the genealogical relationships for a fixed gene in a population of $N = 6$ finches on the next page.

The gene determines whether the finch has a long or a short beak.

Activity 2

Complete the following definitions.

An ancestor is

The **time of the most recent common ancestor** for a set of individuals alive today is

Finding the most recent common ancestor

Activity 3

On your diagram of genealogies of finches above, find the time of most recent common ancestor for five distinct pairs of finches in generation 0.

Record your's and everyone else's findings in the tables on the following pages.

Generations ago

11

10

9

8

7

6

5

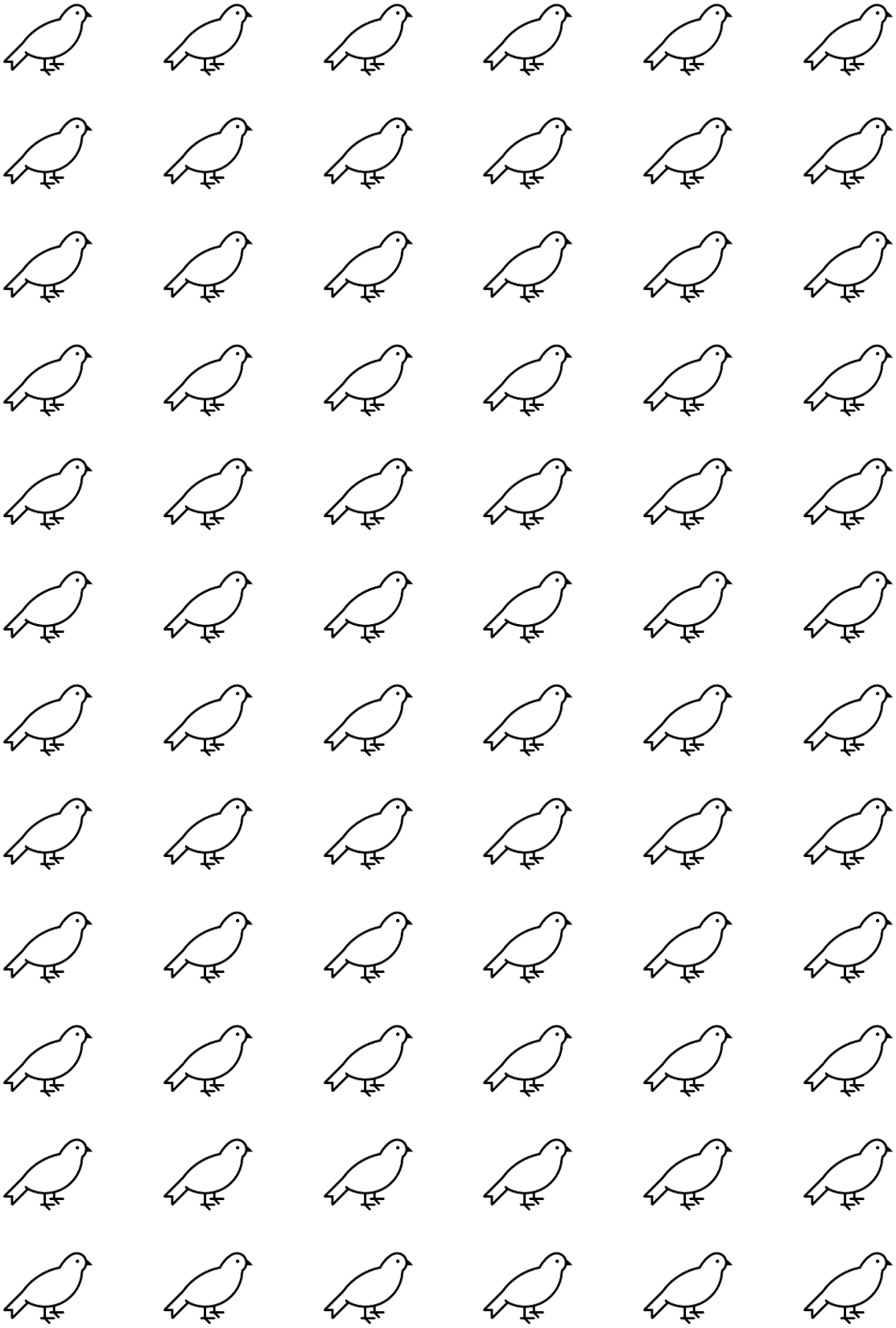
4

3

2

1

0



Pair	Time of most recent common ancestor for a population of size $N = 6$	Pair	Time of most recent common ancestor for a population of size $N = 6$
1		16	
2		17	
3		18	
4		19	
5		20	
6		21	
7		22	
8		23	
9		24	
10		25	
11		26	
12		27	
13		28	
14		29	
15		30	

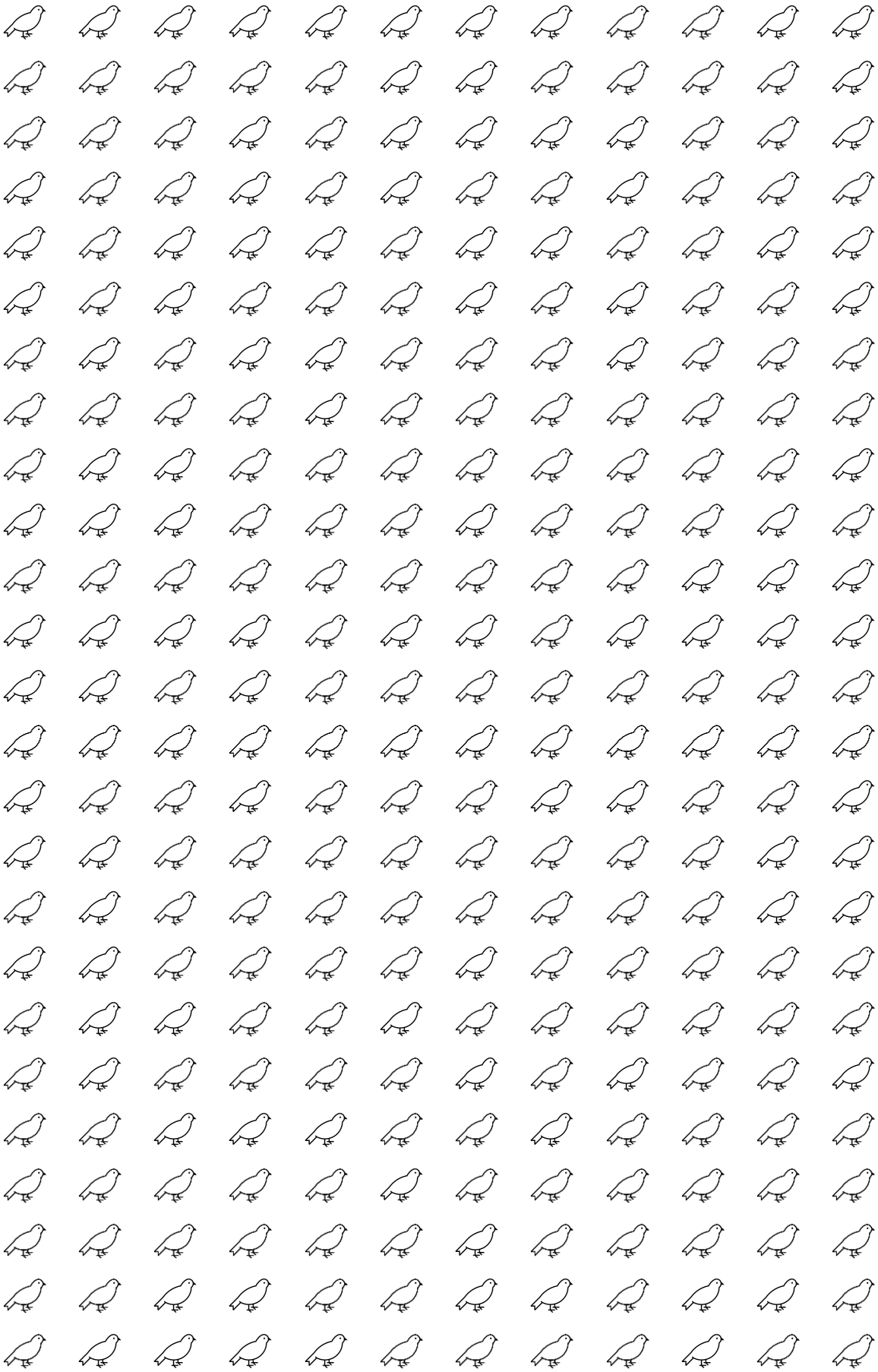
To find the most recent common ancestor, we don't actually need to find the genealogical relationships for all individuals in the population. In fact, we can just do it for the pair we are trying to find the common ancestor of, and only need to go far enough back in time to see this ancestor.

Activity 4

For a population of size $N = 12$ finches, find the time of most recent common ancestor for 5 distinct pairs. **You do not need to construct the genealogical relationships for all individuals**. Record your findings in the table below and copy in the data discovered by others too.

Generations ago

24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0



Pair	Time of most recent common ancestor for a population of size *N* = *12*	Pair	Time of most recent common ancestor for a population of size *N* = *12*
1		16	
2		17	
3		18	
4		19	
5		20	
6		21	
7		22	
8		23	
9		24	
10		25	
11		26	
12		27	
13		28	
14		29	
15		30	

Calculating the mean

Activity 5

With your partner, discuss how you compute the mean of a set of data. Write down a formula for the mean of a set of data.

Activity 6

Compute the mean time of most recent common ancestor for the population 6 finches and the population of 12 finches. **You may use a calculator**.

Compare your answers in the two cases.

Rachel says that

"In a small population, two individuals are more likely to be closely related."

Does the evidence you have collected today support Rachel's statement? Explain your answer.

Homework

On the same set of axes, draw a bar chart for the time of most recent common ancestor for the two population of finches above. Graph paper is provided below. You may wish to draw your graph horizontally on the page.

List two similarities and two differences between these graphs. Be prepared to share these in the next tutorial.

In each of the cases $N = 6$ and $N = 12$, compare the mean time to most recent common ancestor that you computed in the tutorial and the population sizes N. What do you notice?

Tutorial 4
Finding common ancestors using probability

By the end of this tutorial, I will be able to

1. calculate the probability of complicated events using tree diagrams;
2. compute the probability that two individuals have the same parent in the previous generation under the Wright—Fisher model; and
3. derive the probability that two individuals have a common ancestor exactly *n* generations ago.

Activity 1 – Starter

In pairs, discuss your graphs and the observations you made in your homework. Be ready to share your observations with the rest of the group. You may use the following space for notes.

Tree diagrams

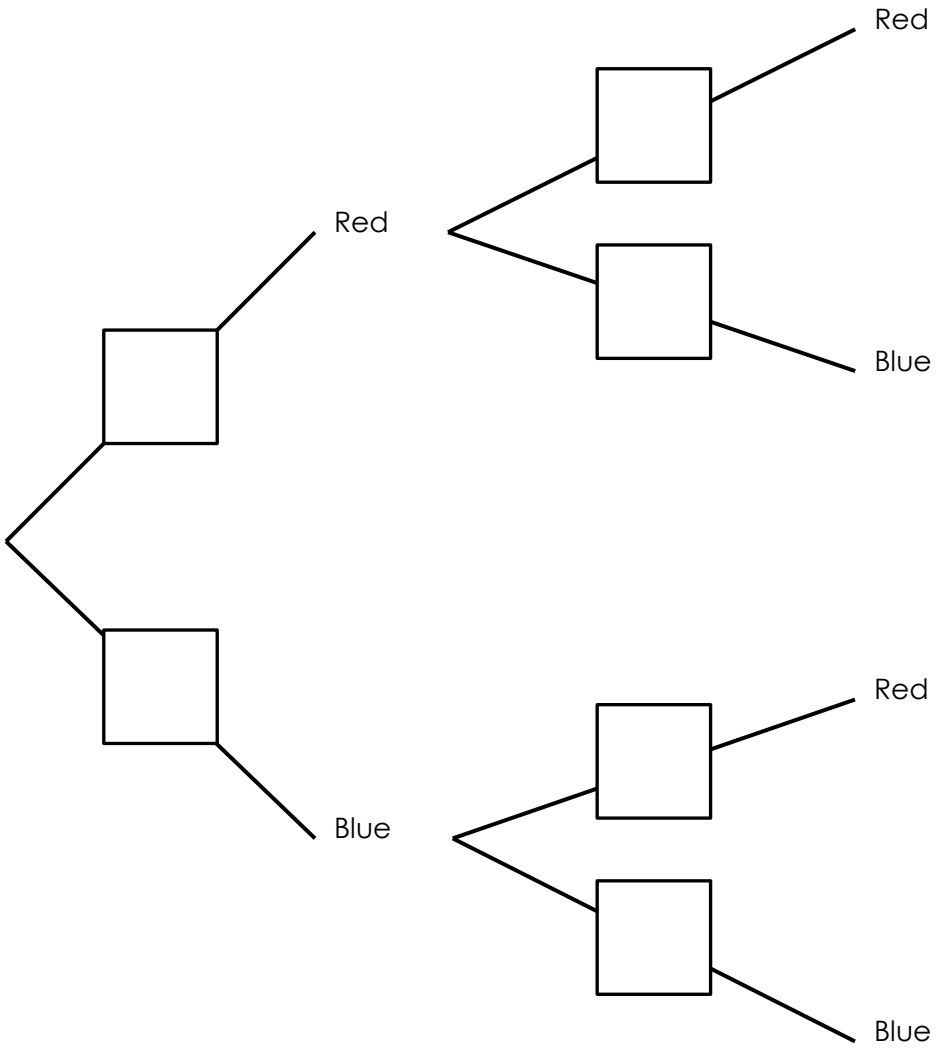
In the last tutorial we constructed the genealogical relationships between generations under the Wright–Fisher model. By looking at these relationships, we could work out when two individuals had a common ancestor in the past.

For different pairs this was a different number of generations. Because the genealogical relationships are themselves random, the number of generations until the most recent common ancestor is too.

Today, we will calculate the probability that it was *n* generations ago, where *n* is a positive integer (1, 2, 3, …).

To do this, first we need to learn how to compute the probability of more complex events using **tree diagrams**.

Question: Given there are 5 red balls and 2 blue balls in a bag, what's the probability that after two picks we have **a red ball and a blue ball**?



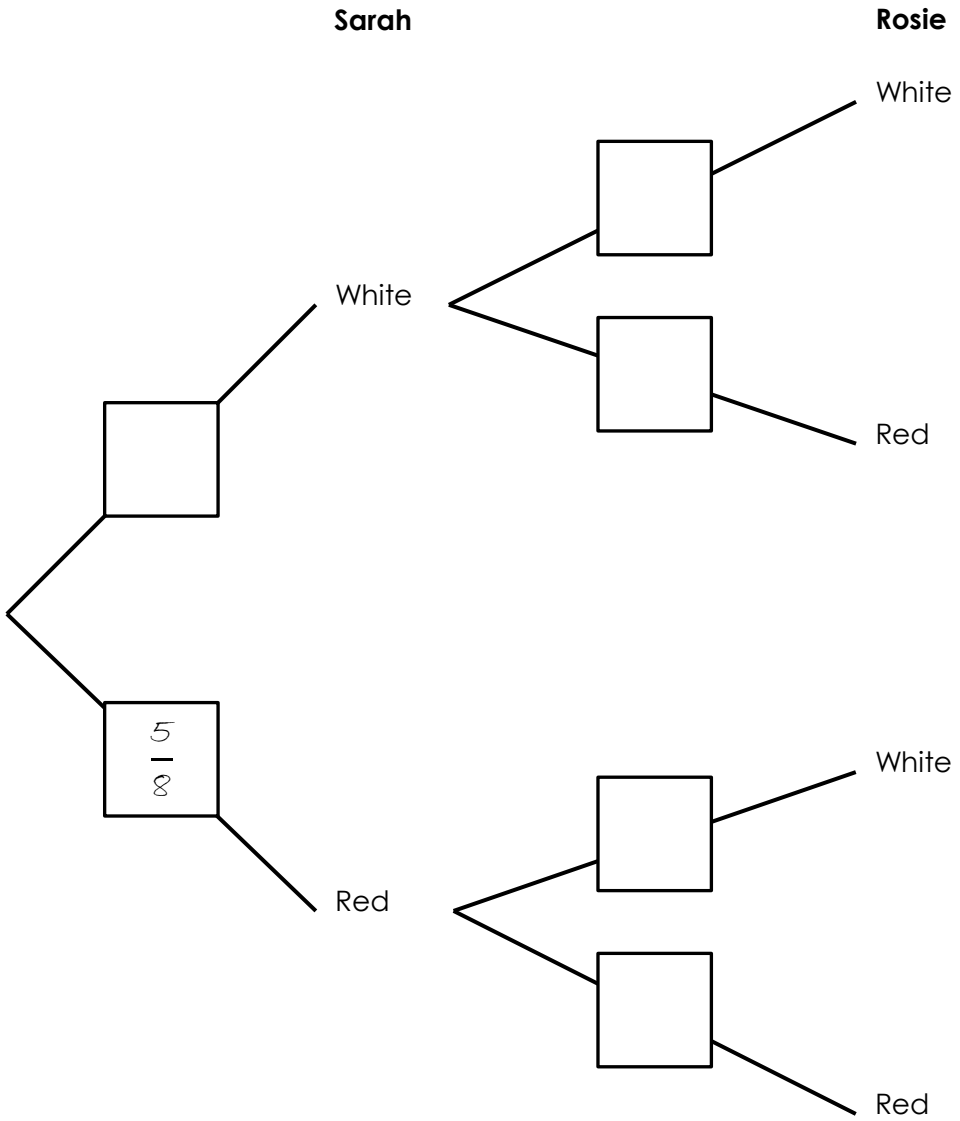
We multiply across the matching branches, then add these values.

Activity 2

Now it's your turn.

A bag contains 5 red counters and 3 white counters. Sarah takes a counter at random from the bag. She notes its colour and replaces it. Rosie then takes a counter from the bag.

Complete the probability tree diagram shown below.



Work out the probability that both Sarah and Rosie take a red counter.

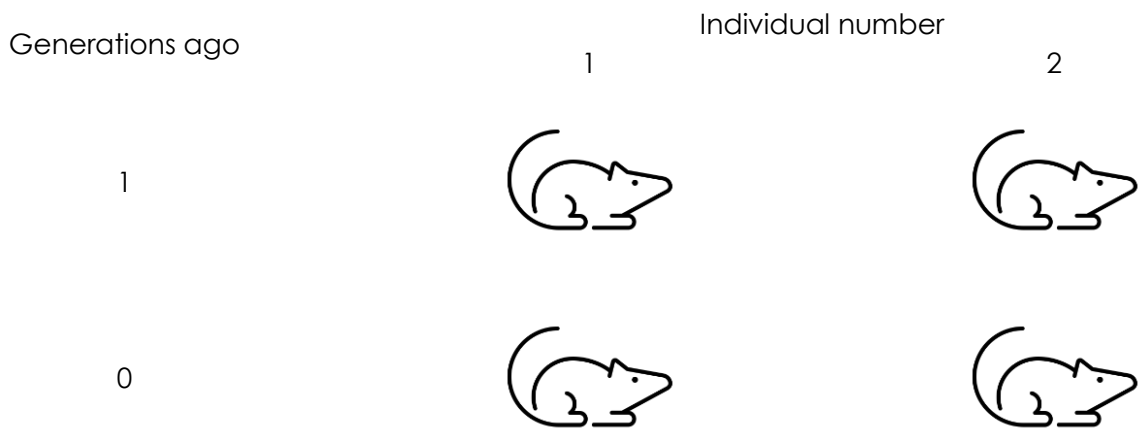
Work out the probability that only one of the counters taken is red.

Mathematical analysis of time of most recent common ancestor

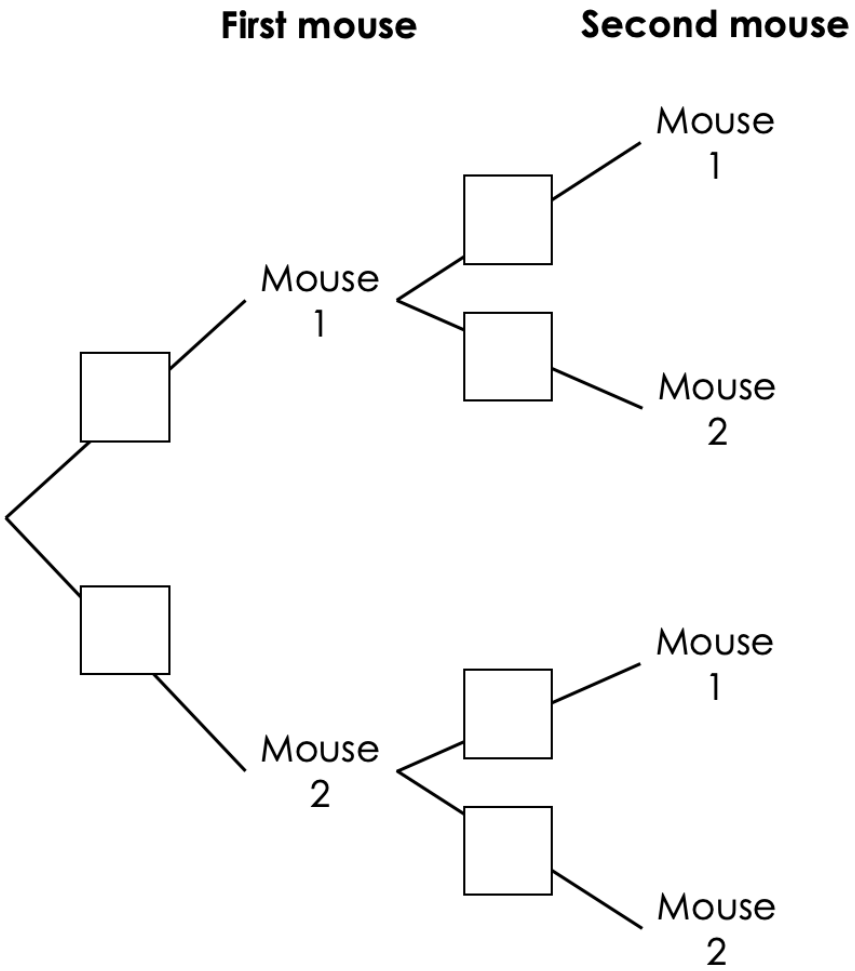
Finally, we are in a position to use all the maths you have learned to mathematically analyse the time of most recent common ancestor.

Activity 3

Consider a population of size $N = 2$. It might help to consider a picture as follows.



The tree diagram below shows the mouse, in the previous generation, which the first and second mouse in the current generation inherit their gene from.



Using the tree diagram, what is the probability that both mice in the current generation inherit their gene from the same mouse in the previous generation?

Activity 4

Now consider a population of size ***N* = 3**. By creating your own tree diagram, what is the probability that two mice both inherit their gene from the same mouse in the previous generation.

Hint: this time, your tree should have **three** branches.

Activity 5

Without drawing a tree diagram, explain why in a population of size $N = 4$, the probability that a pair of mice both inherit their gene from the same mouse in the previous generation is 0.25.

Activity 6

Now consider a population of size N, where N is **any number**. Explain why the probability that a pair of mice both inherit their gene from the same mouse in the previous generation is $1/N$.

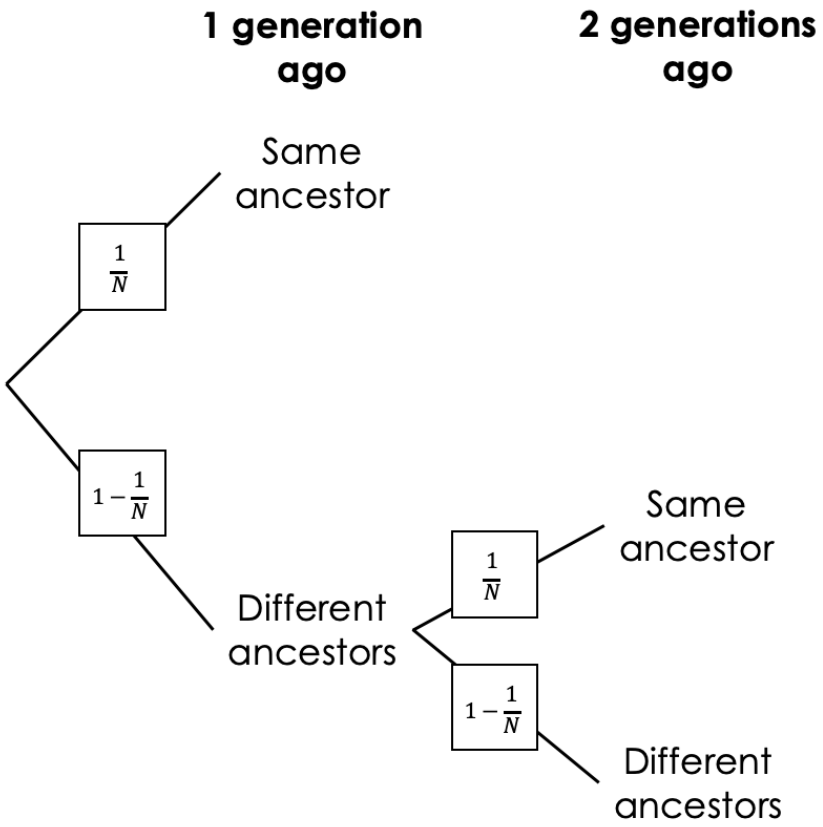
Activity 7

In Activity 6, you showed that the probability a pair of mice both inherit their gene from the same mouse in the previous generation is $1/N$.

Let T be the number of generations in the past when this pair have a common ancestor.
Explain why

$$P(T = 1) = \frac{1}{N}.$$

We can use a tree diagram to find the probability that $T = 2$.



Multiplying across the only matching branch, we can work out that

$$P(T = 2) = \left(1 - \frac{1}{N}\right) \times \frac{1}{N}.$$

Explain why for any whole number n we have the formula

$$P(T = n) = \left(1 - \frac{1}{N}\right)^{n-1} \times \frac{1}{N}.$$

Homework

For $N = 6$, compute $P(T = n)$ and $30 \times P(T = n)$ for different values of n and enter it into the following table. You may use a calculator.

n	$P(T = n)$	$30 \times P(T = n)$
1		
3		
5		
7		
9		
11		
13		
15		
17		
19		
21		
23		

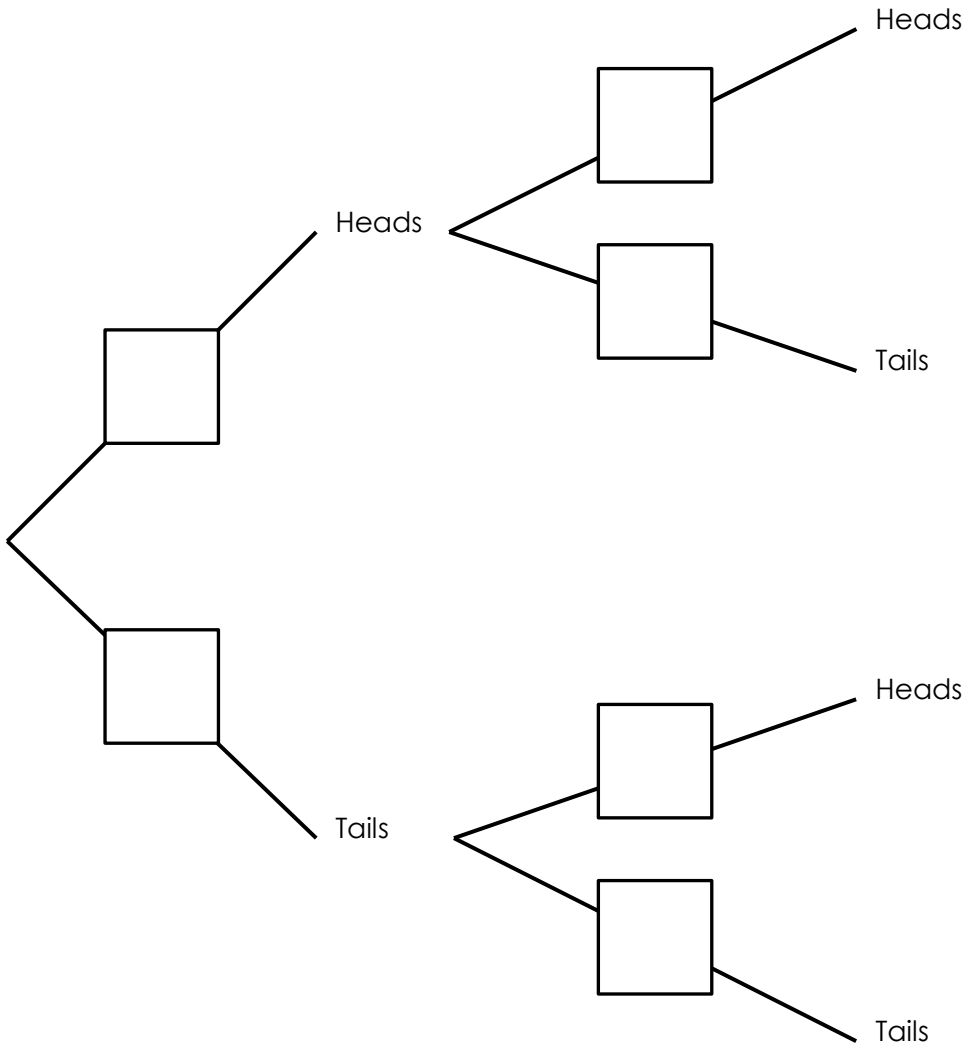
For your homework last week, you plotted a bar chart of the time until the most recent common ancestor for a population of 6 finches. Add to this graph a line graph of $P(T = n)$.

Compare the shape of the line graph and the bar chart for the population of size $N = 6$.

The next two questions will test your understanding of probability tree diagrams.

Question 1

The probability of getting Heads on a biased coin is 0.9

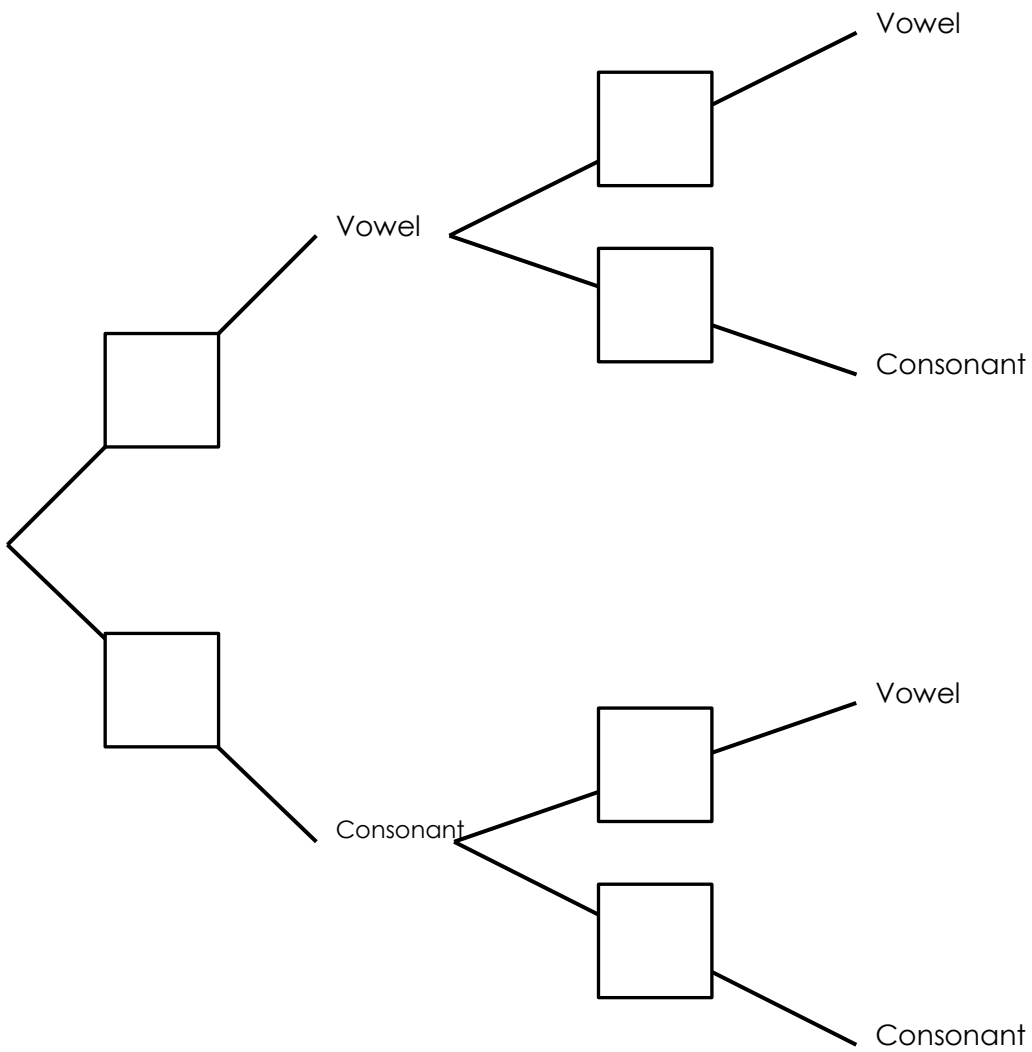


a) Complete the tree diagram.
b) What is the probability of getting Heads twice?

c) What is the probability of seeing one each of Heads or Tails?

Question 2

A bag contains the letters of 'TRANSFORMATION'. I pick two letters from the bag <u>without</u> replacing them afterward.



a) Complete the probability tree.

b) What is the probability of getting two consonants?

c) What is the probability of getting exactly one vowel?

Additional note space:

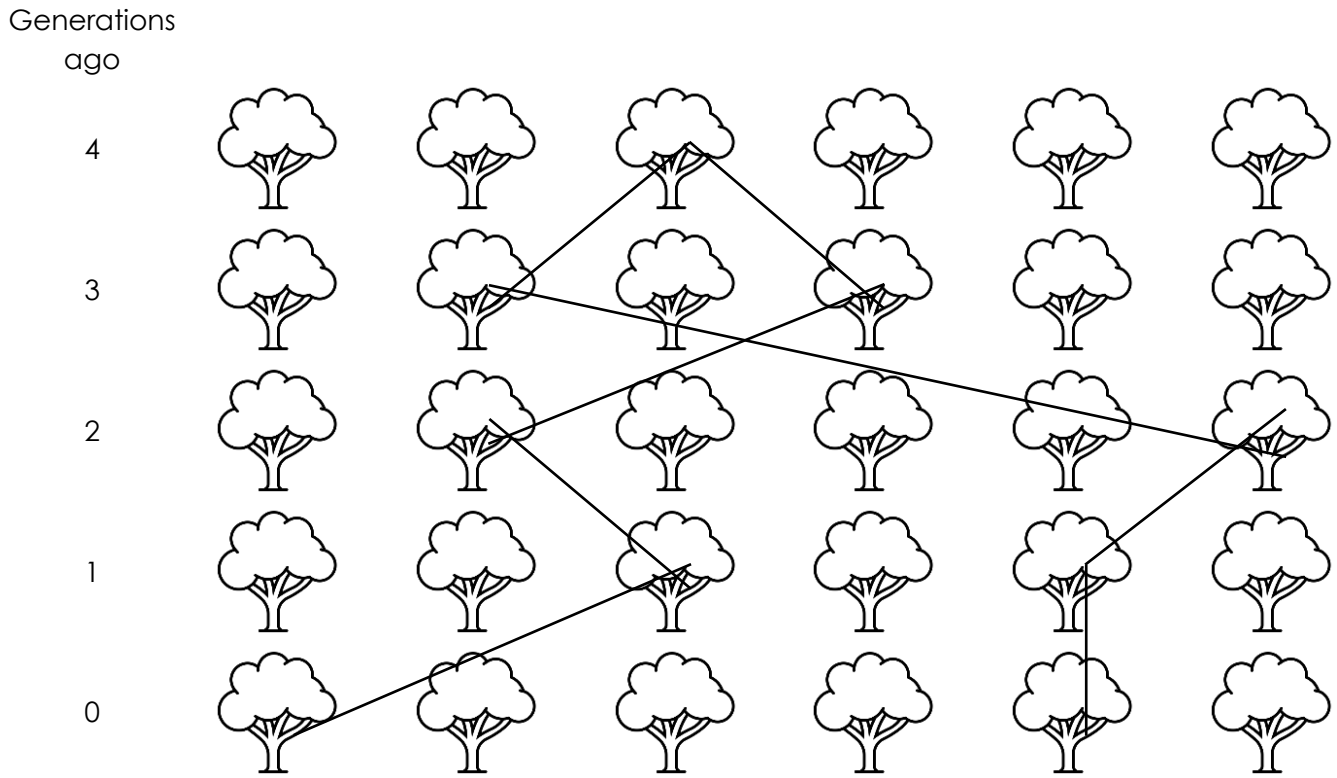
Tutorial 5
The Wright–Fisher model with mutation

By the end of this tutorial, I will be able to

1. describe the Wright—Fisher model with mutation; and
2. identify graphically whether two individuals are identical by descent.

A problem with the Wright–Fisher model

The following picture shows the genealogical relationship between two oak trees alive today. They find a common ancestor after 4 generations.



Activity 1

Under the Wright–Fisher model, if we go far enough back in time, we will find a common ancestor of **any two** individuals alive today. These two individuals therefore **carry the same gene**.

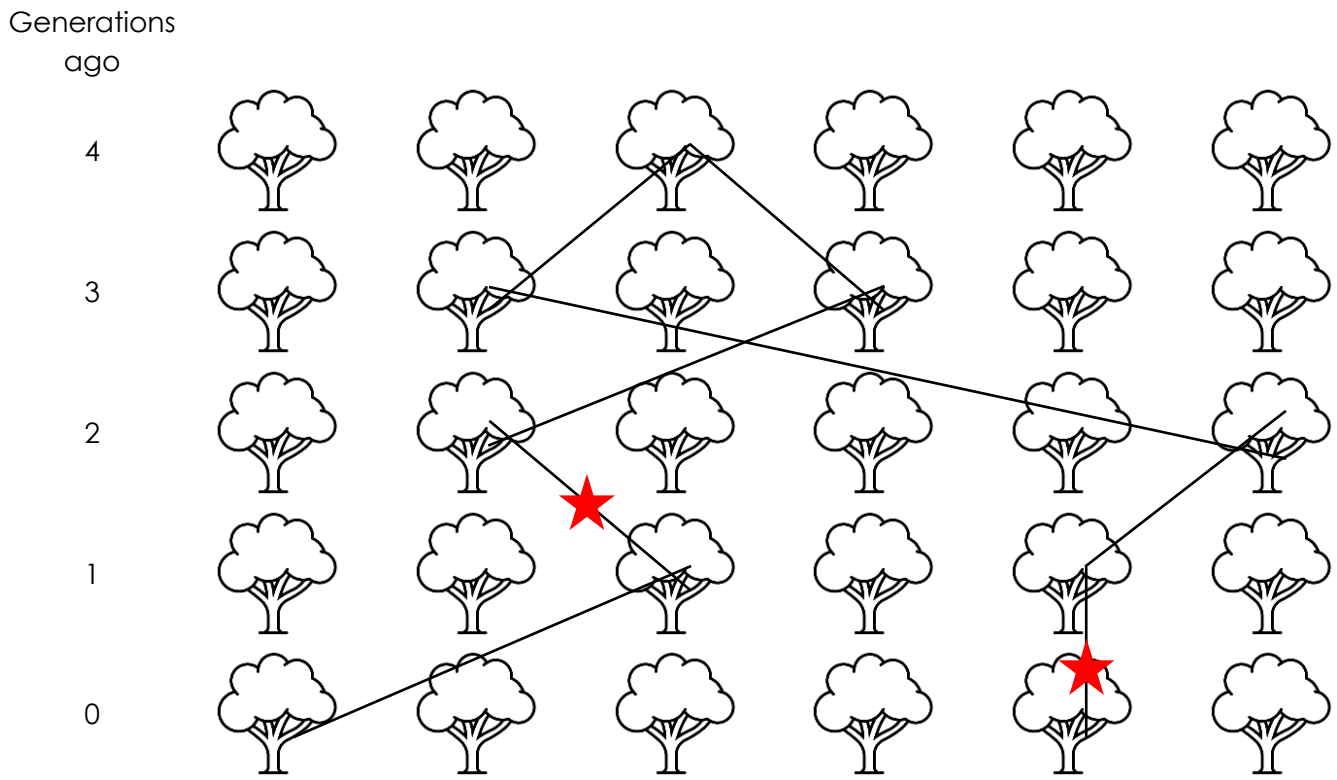
With your partner, what is the Wright–Fisher model **failing to capture** which means that we don't all carry the same genes?

Be ready to share your thoughts with the class.

Genetic mutation

When a gene is copied from parent to child, there is a small probability μ that it is copied incorrectly. If this happens, we say there has been a **mutation**.

A better version of the diagram above would look like this.



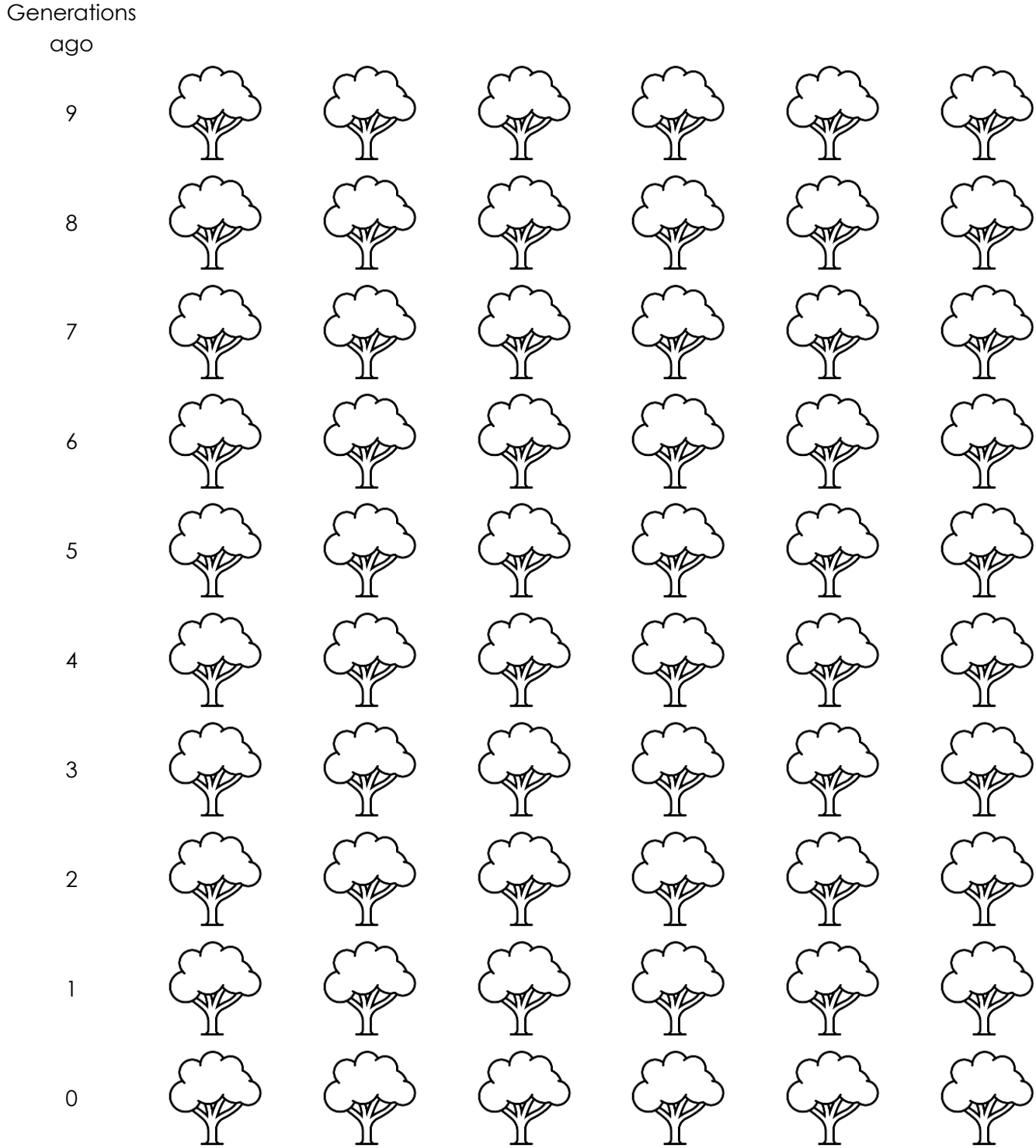
Here, red stars have been used to indicate that a mutation has occurred when the gene was copied from parent oak tree to child. Because there has been at least one mutation in their genealogy, the two individuals today **carry different genes**. In fact, the gene is also different from their common ancestor.

In Tutorial 2 we spoke of mathematical modelling as a process. The final stage was to "iterate", meaning to refine the model and add extra complexity. This is what we are doing by adding mutations to the Wright–Fisher model.

In the **Wright–Fisher model with mutation**, a mutation occurs on each separate branch of the tree with probability μ.

Activity 2

In the following population of oak trees, mutations occur with probability $\mu = 0.5$. By rolling a die, construct the genealogy of a pair of trees in the present-day generation and then, by flipping a coin, add mutations to their genealogical tree.



Do your two chosen individuals carry the same gene? Explain your answer.

Would they be more or less likely to carry the same gene if μ is bigger or small? Explain your answer.

Activity 3

In terms of the number of mutations, what needs to happen for both individuals to carry the **same** gene? When this happens, we say that the two individuals are **identical by descent**.

In terms of the number of mutations, what needs to happen for the individuals to carry **different** genes?

Activity 4

Take two individuals alive in the present-day population. What is the probability that there has been at least one mutation when copying their genes from the previous generation?

Final assignment

This assignment consists of two parts. There are 50 marks available for each part. Part A consists of the written part of the assignment and Part B will assess your understanding of the mathematical content of this course.

You should provide answers to both Part A and Part B.

There are a **total of 100 marks** available on this assignment.

Part A

With reference to the Wright–Fisher model of genetic inheritance, discuss the role and limitations of mathematical models. **[50 marks]**

You should answer this question by producing an essay of at most 1000 words. Your essay should include the following sections:

- the purpose of mathematical models;
- a description of each stage of the mathematical modelling process;
- an evaluation of the limitations of mathematical models; and
- the definition, use and limitations of the Wright–Fisher model of genetic inheritance. You should carefully state the assumptions of the Wright–Fisher model and explain whether or not you think they are realistic.

Your workbook, homework assignments and your baseline assignments should help you with answering this essay question.

Make sure your answer is coherent (each section should lead into the next). Remember you will also be marked on your spelling, punctuation, grammar and the correct use of subject vocabulary.

Part B

This section contains the mathematical questions of this assessment. In it, you will be guided to compute the probability that a randomly chosen person carries the same gene as you at a fixed point on a chromosome.

You will need to use the mathematics that you have learned in tutorials and so you are encouraged to refer back to your workbook in order to help you answer these questions.

You may use a calculator.

You **must** show all your working.

1. Explain what a mutation is and why the probability that an individual correctly copies their gene from the previous generation is $1 - \mu$. **[3 marks]**

2. Explain why the probability that two individuals both copy their genes correctly (ie, there is **no** mutation) is $(1 - \mu)^2$.

   You may wish to use a tree diagram to explain your answer. **[6 marks]**

3. Explain what it means for two individuals to be identical by descent. **[3 marks]**

4. Two individuals have the same gene if before their common ancestor all genes have been copied correctly.

   By drawing an 'infinite' probability tree, explain why the probability that two individuals have the same gene is given by

$$(1-\mu)^2 P(T=1) + (1-\mu)^4 P(T=2) + \cdots + (1-\mu)^{2n} P(T=n) + \cdots.$$

   [10 marks]

5. By expanding out the brackets, explain why

$$(1-a)(a + a^2 + a^3 + \cdots + a^n + \cdots) = a$$

   and use this to show that

$$a + a^2 + a^3 + \cdots + a^n + \cdots = \frac{a}{1-a},$$

   whenever $-1 < a < 1$.

   (Expressions of this kind are called infinite sums.) **[7 marks]**

6. Show that the probability that two individuals have the same gene in Question 4 can be written in terms of *N* and *μ* by

$$\frac{1}{N}\frac{(1-\mu)^2(N-1)}{N - (1-\mu)^2(N-1)},$$

   where *N* is the population size. You will need to use a result we proved in Tutorial 4. **[15 marks]**

7. The population on Earth is approximately *N* = 7 500 000 000 or 7.5 billion. Compute the probability that two individuals chosen on Earth carry the same gene for this value of *N* and *μ* = 0.001.

   You should give your answer correct to three significant figures. **[3 marks]**

8. Do you expect this probability to be bigger or smaller for increased *μ*?

   Explain your answer carefully. **[3 marks]**

Tutorial 6
Feedback tutorial

What is the Purpose of Tutorial 6?

- To receive feedback on your final assignment
- To respond to the feedback from your Uni Pathways teacher
- To write targets for improvement on your final assignment

Final assignment feedback from my Uni Pathways Teacher

(Remember to look at the mark scheme to help you understand what you have done well so far, and how you can do even better in your final assignment)

Here are three things that my Uni Pathways Teacher thought I did well in my draft assignment
•   •   •

Here are three things that my Uni Pathways Teacher thinks that I could do to get a higher mark in my final assignment
•   •   •

Tasks from my Uni Pathways Teacher to do during the feedback tutorial to help me improve
•   •   •

My response:

Actions I will take to improve my final assignment after this tutorial…

Hand in date for final assignment:	

Tutorial 7
Final tutorial

What is the Purpose of Tutorial 7?

- To receive feedback and a grade on your final assignment.
- To reflect on the programme including what you enjoyed and what was challenging.
- To ask any questions you may have about university.

Final assignment feedback from my Uni Pathways Teacher

Final mark:	University style grade:
Here are three things that my Uni Pathways Teacher thought I did well in my final assignment	
•	
•	
•	

Here are three things that my Uni Pathways Teacher thinks that I should remember for when I am doing this kind of study in the future
•
•
•

University

What questions do you still have about University after taking part in Uni Pathways?
•
•
•

Reflecting on Uni Pathways

What did you most enjoy about Uni Pathways?
•   •   •

What did you find most challenging about Uni Pathways?	How did you overcome these challenges?
• • •	• • •

Appendix 1
Referencing correctly

When you get to university, you will need to include references in the assignments that you write, so we would like you to start getting into the habit of referencing in your Brilliant Club assignment. This is really important, because it will help you to avoid plagiarism. Plagiarism is when you take someone else's work or ideas and pass them off as your own. Whether plagiarism is deliberate or accidental, the consequences can be severe. In order to avoid losing marks in your final assignment, or even failing, you must be careful to reference your sources correctly.

What is a reference?

A reference is just a note in your assignment which says if you have referred to or been influenced by another source such as book, website or article. For example, if you use the internet to research a particular subject, and you want to include a specific piece of information from this website, you will need to reference it.

Why should I reference?

Referencing is important in your work for the following reasons:

- It gives credit to the authors of any sources you have referred to or been influenced by.
- It supports the arguments you make in your assignments.
- It demonstrates the variety of sources you have used.
- It helps to prevent you losing marks, or failing, due to plagiarism.

When should I use a reference?

You should use a reference when you:

- Quote directly from another source.
- Summarise or rephrase another piece of work.
- Include a specific statistic or fact from a source.

How do I reference?

There are a number of different ways of referencing, and these often vary depending on what subject you are studying. The most important to thing is to be consistent. This means that you need to stick to the same system throughout your whole assignment. Here is a basic system of referencing that you can use, which consists of the following two parts:

1. **A marker in your assignment**: After you have used a reference in your assignment (you have read something and included it in your work as a quote, or re-written it your own words) you should mark this is in your text with a number, e.g. [1]. The next time you use a reference you should use the next number, e.g. [2].

2. **Bibliography**: This is just a list of the references you have used in your assignment. In the bibliography, you list your references by the numbers you have used, and include as much information as you have about the reference. The list below gives what should be included for different sources.

a. **Websites** – Author (if possible), title of the web page, website address, [date you accessed it, in square brackets].

   E.g. Dan Snow, 'How did so many soldiers survive the trenches?', http://www.bbc.co.uk/guides/z3kgjxs#zg2dtfr [11 July 2014].

b. **Books** – Author, date published, title of book (in italics), pages where the information came from.

   E.g. S. Dubner and S. Levitt, (2006) *Freakonomics*, 7-9.

c. **Articles** – Author, 'title of the article' (with quotation marks), *where the article comes from* (newspaper, journal etc.), date of the article.

   E.g. Maev Kennedy, 'The lights to go out across the UK to mark First World War's centenary', Guardian, 10 July 2014.

Notes