GENOMICS AND ADAPTATIONS PACK

Discover how genetics underpins adaptations that let creatures survive and thrive!

Age: 11+

**Extracting DNA from fruit**
Extract DNA from a sample just like it’s done in the lab.
[yourgenome.org/activities/extracting-dna-from-fruit](http://yourgenome.org/activities/extracting-dna-from-fruit)

**Genome challenge**
Are humans really more complex than other animals and plants?
[yourgenome.org/activities/genome-challenge](http://yourgenome.org/activities/genome-challenge)

**Function finders BLAST**
Run some code cracking software and unlock nature’s potential from your computer!
[yourgenome.org/activities/function-finders-blast](http://yourgenome.org/activities/function-finders-blast)

**Creature report**
Research an organism and uncover how it is genetically adapted.
[yourgenome.org/activities/creature-report](http://yourgenome.org/activities/creature-report)
A genome is an organism’s complete set of DNA instructions. All living things have genomes - each genome contains all of the information needed to create a functional organism. The Wellcome Genome Campus works in the field of ‘genomics’ (the study of genomes). The huge impact genomics has had on health science is well known but its importance in conservation and how we understand the natural world is often overlooked.

The Wellcome Genome Campus is now embarking on the Darwin Tree of Life project. This will sequence the genomes (decode the DNA) of every complex (has more than one cell) organism found in the UK and Ireland - that’s 66,000 species! This is part of a bigger global endeavour, the Earth Biogenome project, which aims to sequence everything on earth!

Through this pack you will find out more about this project– how the work will be done and how the results might be used. Work through the activities and by the end you’ll know what genomics is and how it can be used to understand the natural world.

**CHALLENGE 1: EXTRACTING DNA FROM FRUIT**

Let’s first look at what DNA – the molecule in which all genetic instructions is written – really looks like. You can find DNA in just about every cell in any living thing, but the first step in any research project is getting that DNA out!

Follow the instructions to see how scientists get DNA from a sample. For this activity we will be using strawberries as they are normally easy to find in shops but any soft fruit will be great for this experiment!
**EXTRACTING DNA FROM FRUIT**

**Instructions**

1. Mix the washing-up liquid, the water and the salt in one cup.
   - 5 ml washing-up liquid
   - 50 ml water
   - 5 g salt

2. Stir gently to mix it all together.

3. Put the fruit in a bag and add the mixture.

4. Squeeze the air out and seal the bag.

5. Squash to break up cells and release the DNA.

6. Pour the fruit mixture through the sieve and into the jug.

7. Half fill a small cup with the liquid from the jug.

8. Slowly pour very cold alcohol down the side of the small cup.
   - 10 ml alcohol (very cold)

9. You should see white clumps forming in the clear layer... that’s DNA!
Although every living thing has DNA, not all of it is the same. It is these differences in the genetic code that makes every species unique. Some of the most fundamental differences between species is the size of their genomes and how many genes (individual instructions) are contained within that genetic code.

In this next activity you are going to have a go at sorting four different organisms in order by the size of their genomes and then by the number of genes they have in their genome. Once you’ve had a guess check the answers and read about the scientific theories behind why they have the genomic features they do.

How much DNA did you get? It can be tough to hook the DNA out on a toothpick – some scientists spend years getting good at it! Because DNA is such a compact molecule, there is actually enough DNA in your cup that, if stretched out into a single strand, could go all the way from the UK to Hawaii!

By extracting DNA from an organism, and sequencing and analysing the DNA code, you can understand how it functions and identify the genetic instructions (genes) behind some of the adaptations it has to help it survive.
Mouse

Human

Sugar beet

Fruit fly
Put the organisms in order of genome size:

1  2  3  4

**bigger genome**
(more DNA letters in the genetic code)

**smaller genome**
(fewer DNA letters in the genetic code)
Put the organisms in order of number of genes:

1 2 3 4

more genes
(more instructions in the genetic code)

less genes
(fewer instructions in the genetic code)
Put the organisms in order of genome size:

1. Human
   - 3,100 million base pairs
   - (more DNA letters in the genetic code)

2. Mouse
   - 2,700 million base pairs

3. Sugar beet
   - 570 million base pairs

4. Fruit fly
   - 140 million base pairs

(smaller genome)

3,100 million base pairs > 2,700 million base pairs > 570 million base pairs > 140 million base pairs
Put the organisms in order of number of genes:

1. Sugar beet (27,000 genes)
2. Mouse (23,000 genes)
3. Human (20,000 genes)
4. Fruit fly (14,000 genes)

more genes (more instructions in the genetic code)
less genes (fewer instructions in the genetic code)
Scientists were surprised to discover how relatively few genes humans and other mammals (such as mice) have compared to the size of their genomes. Further study has begun to shed some light on this observation. It now appears that mammalian genomics is much more complex than originally thought with each gene being able to make more than one protein (each gene being more than a single instruction) and added layers of genetic regulation allowing for greater complexity in the organisms.

Plant genomes can vary massively in terms of size and the number of copies of their genome contained in each of their cells. They do, however, tend to have more genes within their genomes than other organisms, relative to genome size. It is not known for sure why plants tend to have more genes than animals, but one theory is that their inability to move is a big contributor. Most animals can move towards food or away from predators so effectively only need the genes to allow them to move to deal with quite a range of situations (or external stimuli). As plants can’t move in the same way, they instead need to have genes for each situation. For example: a set of genes to enable them to use photosynthesis to make their food, a set of genes to deal with insect predators, a set of genes to deter herbivorous animals, genes to alter many plant systems in response to the weather and climate changes (seasonal and otherwise).

**Sugar Beet (**Beta vulgaris**)**

This plant now accounts for around 20% of the world’s sugar production with much of the UK production found in East Anglia. It is a great example of human-driven crop breeding in agriculture resulting from selectively breeding plants with more desirable characteristics. In this case, selecting for the most sugary plants has led to modern day crops having ~20% sucrose (a type of sugar) dry weight compared to ~5% when the plant started being used as an alternative to sugarcane for sugar production in the 18th century.

Stats rounded to 2 significant figures and correct as of 28/05/2020: plants.ensembl.org/Beta_vulgaris/Info/Annotation/
Mouse (*Mus musculus*)

Mice have been used as a model organism in scientific and medical research for around 100 years. They are used because of the relative ease of keeping them in labs, their short generation times (can have offspring very young – 6 weeks after birth), and large litter sizes. This research has led to many medical breakthroughs especially in the field of genetics. There are now numerous laws protecting the welfare of the animals with the ‘3Rs’ (ethical principles) underpinning all animal research in the UK. These principles are: to replace animal models wherever possible, to reduce the number of animals used in projects to the minimum, and to refine the processes used to ensure the animals suffer as little as possible.

Stats rounded to 2 significant figures and correct as of 28/05/2020: [asia.ensembl.org/Mus_musculus/Info/Annotation](http://asia.ensembl.org/Mus_musculus/Info/Annotation)

Human (*Homo sapiens*)

The first human reference genome was generated by the Human Genome Project. This project was started in 1990 and took 13 years to complete. Although the human genome was published in 2003, the results posed more questions than answers (e.g. what do all the genes do, can we alter the genetic code to treat certain health conditions, etc.) and led to expansion of the genomics field to better understand the genetic underpinnings of humans. Although the human genome project took a long time and cost a lot of money (~$3 billion), the technology used to sequence DNA has advanced in the last couple of decades. It now takes a matter of days and costs less than $1000 to sequence a person’s genome.

Stats rounded to 2 significant figures and correct as of 28/05/2020: [asia.ensembl.org/Homo_sapiens/Info/Annotation](http://asia.ensembl.org/Homo_sapiens/Info/Annotation)

Fruit fly (*Drosophila melanogaster*)

These insects have been used as a model organism in scientific research for around 100 years. They have relatively simple genomes with well documented links between their genotype (the instructions in their genetic code) and their phenotype (the result of those instructions that you can see in how the fly looks and functions). It has also been estimated that they share ~60% of the genes involved in human genetic conditions such as cancer. These two facts make them a great species for studying how certain genetic changes can lead to health conditions developing.

Stats rounded to 2 significant figures and correct as of 28/05/2020: [asia.ensembl.org/Drosophila_melanogaster/Info/Annotation](http://asia.ensembl.org/Drosophila_melanogaster/Info/Annotation)
No matter the organism, knowing how big the genome is and how many genes it has is only the first step! By reading the DNA code, we can work out what proteins the genes code for and what those proteins do.

Proteins are like tiny molecular machines that do lots of jobs inside our bodies whether it is forming muscle, breaking down food or facilitating the complex biochemical reactions that keep us alive. Every three letters of DNA makes what we call an amino acid. These amino acids bond together, like beads on string, then fold into specific shapes to make the whole protein.

**CHALLENGE 3: FUNCTION FINDERS BLAST**

Your task is to decode some DNA sequences to find out what protein they produce. Use the codon wheel on page 13 to translate the DNA codes on your worksheet into a protein sequence. Once you have your protein codes, follow the instructions to carry out a BLAST search using the UniProt database where you will find out what your protein is and what it does.

When you have completed the challenge, think about these questions, and if you can discuss them with your teacher, friends or family:

- What was the most interesting protein you found?
- Are there other features or adaptations in the natural world that you would like to find out the DNA code for?
- How do you think you could use this type of research for conservation projects?
Use the codon wheel to translate DNA codons into amino acids:

To decode a codon find the first letter of your sequence in the inner circle and work outwards to see the corresponding amino acid. For example: CAT codes for H (Histidine).

Please note that this wheel uses the sense DNA codons (5’ to 3’).
## Translate the DNA sequences to find the matching protein using Uniprotein BLAST search:

<table>
<thead>
<tr>
<th>DNA sequence 1</th>
<th>atg</th>
<th>aag</th>
<th>tca</th>
<th>gct</th>
<th>att</th>
<th>tta</th>
<th>acc</th>
<th>ggt</th>
<th>ttg</th>
<th>ctt</th>
<th>ttc</th>
<th>gtc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DNA sequence 2</th>
<th>atg</th>
<th>agt</th>
<th>aaa</th>
<th>gga</th>
<th>gaa</th>
<th>gaa</th>
<th>ctt</th>
<th>ttc</th>
<th>act</th>
<th>gga</th>
<th>gtc</th>
<th>gtt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DNA sequence 3</th>
<th>gaa</th>
<th>aac</th>
<th>atg</th>
<th>gag</th>
<th>aac</th>
<th>gat</th>
<th>gaa</th>
<th>aat</th>
<th>att</th>
<th>gtg</th>
<th>tat</th>
<th>ggt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DNA sequence 4</th>
<th>ggt</th>
<th>tgg</th>
<th>gct</th>
<th>ttg</th>
<th>cgg</th>
<th>atc</th>
<th>atg</th>
<th>ttt</th>
<th>cta</th>
<th>cat</th>
<th>ctg</th>
<th>tac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

14/27
### Translate the DNA sequences to find the matching protein using Uniprotein BLAST search:

<table>
<thead>
<tr>
<th>DNA sequence 5</th>
<th>cct</th>
<th>ggg</th>
<th>gag</th>
<th>aac</th>
<th>cta</th>
<th>tgc</th>
<th>tat</th>
<th>aga</th>
<th>aag</th>
<th>atg</th>
<th>tgg</th>
<th>tgc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA sequence 6</td>
<td>ccc</td>
<td>aga</td>
<td>gag</td>
<td>atc</td>
<td>cag</td>
<td>acc</td>
<td>gcc</td>
<td>gtg</td>
<td>aga</td>
<td>ctg</td>
<td>tta</td>
<td>ctc</td>
</tr>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA sequence 7</td>
<td>gag</td>
<td>aag</td>
<td>aga</td>
<td>aag</td>
<td>ctg</td>
<td>ttt</td>
<td>atc</td>
<td>cgt</td>
<td>tcc</td>
<td>atg</td>
<td>ggt</td>
<td>gaa</td>
</tr>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA sequence 8</td>
<td>atg</td>
<td>gag</td>
<td>ttt</td>
<td>act</td>
<td>ttg</td>
<td>agg</td>
<td>caa</td>
<td>gag</td>
<td>gct</td>
<td>tta</td>
<td>gtt</td>
<td>aat</td>
</tr>
<tr>
<td>Translated sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HOW TO COMPLETE THE WORKSHEETS

1. Use the codon wheel to **translate the DNA sequences** on the worksheet to amino acids.

2. **Type the amino acid sequence** in to the Uniprot Blast search [www.uniprot.org/blast/](http://www.uniprot.org/blast/).

   Press **Run BLAST** to get results (it may take a few seconds for results to appear).

3. When the search results appear, **filter the results** to only show reviewed entries (gold file icon with a star). Each result is known as a “hit”.

![Uniprot Blast screenshot]

---

**ConnectingScience**

16/27

[YourGenome.org](http://yourgenome.org)
4. After filtering the hits should look like this. **Scroll down** to the Overview section.

Look at the info column. This will give you an idea of how reliable your hits are. The Expect value (E-value) indicates the number of random hits you would expect by chance for the given query sequence and the size of the sequence database against which the BLAST is performed.

For example, an E-value of 1.0 means that you would expect on average to get one match in the database for the submitted query simply by chance. The lower the E-value, or the closer it is to zero, the more “significant” the match is. In general, the E-values should be in the range of 0.01 to 0.1 to be statistically significant.

The identity % describes how similar your sequence is to the hit, i.e. whether the amino acids are in the same position when aligned. 100% means the sequences match exactly.
5. **Click on the hit** that you think best matches your sequence. Find out the name of the protein which species the sequence is from (common and species name) and what the protein does. Is the protein found in other species?

You can reduce the amount of information on the screen by unticking the blue display categories on the left hand side of the screen.

If you cannot find all the information you need, try using a Google search or Wikipedia to find out more.

6. To start a new BLAST search click **Edit and resubmit**, and enter your next set of amino acids. Repeat steps 3 to 5.
Hopefully those examples showed the potential DNA sequencing has on uncovering the natural world. By understanding adaptations on a genetic level, we can begin to better understand the impact environmental changes are having on endangered species.

There can be other applications too. Once we know the DNA code, it is possible to reproduce some proteins that we find in nature- these could be useful for biotechnological and biomedical purposes. Can you think of any proteins from the natural world that could be useful?

**CHALLENGE 4: CREATURE REPORT**

The Darwin Tree of Life project aims to sequence and better understand the genomics of all living creatures in the UK, a big task! Scientists have to start somewhere and on the next few pages is some information about the first 25 species the Wellcome Genome Campus sequenced with some information as to why they are of interest. Your task is to write an information poster on one of these species. There is a worksheet with some ideas of what to include after the information pages but remember the more focus on how the organism's genomics makes it adapted for its environment the better!
Scientists from the Wellcome Genome Campus are involved in the Darwin Tree of Life project. This aims to decode the DNA of all species living in the British Isles – that’s a lot of life forms! From the robin and golden eagle to squirrels, voles, otters and trout, there are a wide range of species being studied.

Every life form has a genome, a set of DNA instructions that make up all living things. DNA sequencing is a method used to decode DNA, taking it from a gloopy substance in a tube and translating it into a sequence of DNA letters. Studying the DNA of different plants, animals and fungi can help us understand how they are adapted to the environments in which they live and how they have evolved over time.

Below is a quick introduction to 25 of the species having their DNA decoded for the first time. The species have been grouped into five categories; iconic, dangerous, cryptic, floundering and flourishing. Find out more below!
CONNECTIONS SCIENCE

CREATURE REPORT
Creature information

**ICONIC**
These plants and animals are Iconic to the British Isles.

- **Blackberry** (*Rubus ulmifolius*)
  Blackberries are found in hedgerows, parks and gardens all over the UK. There are hundreds of different types, which can be really difficult to tell apart. Millions and millions of Blackberries are grown every year and can be used to make jams and desserts. Understanding the plant’s DNA could help fruit farmers breed bigger and tastier blackberries and improve the way they grow this soft fruit.

- **European Robin** (*Erithacus rubecula*)
  The robin is a common sight in our gardens and parks and is an iconic British bird, featuring on many Christmas cards and decorations. Like many birds, some robins migrate to warmer countries during the winter months. Sequencing and studying its DNA will help us to understand how these birds navigate using the magneto receptors in their eyes to “see” the Earth’s magnetic fields.

- **Lesser Spotted Catshark** (*Scyliorhinus canicula*)
  You might not have seen a lesser-spotted catshark but it’s likely that you’ve found mermaid’s purses on the beach – these are the shark’s egg cases that often wash up from the seabed. Sharks are important for studying how vertebrates (animals with back bones) have evolved over time. Sequencing the sharks DNA will help with this research. The catshark won the public vote to be sequenced in this category.

- **Red Mason Bee** (*Osmia rufa*)
  Unlike honey bees and bumble bees, the red mason bee is a solitary bee, living and making nests on its own. It is found in lowland England and Wales. The Mason Bee is an excellent pollinator (actually better than the honey bee), so understanding more about its DNA will be useful especially when comparing it to honey bee DNA which has already been studied.

- **Golden Eagle** (*Aquila chrysaetos*)
  The majestic golden eagle has suffered from hunting and pesticide poisoning in the past; it is now a protected species and there are around 500 breeding pairs in the UK. However, to increase their numbers in the wild, breeding programmes are being set up. Sequencing and studying the DNA of eagles will help scientists understand more about the eagles biology but also enable conservationists to pick the right birds for breeding and release into new habitats.
These species are potentially harmful to other species living in the UK.

**Asian Hornet (Vespa velutina)**
The Asian hornet or yellow legged hornet is found in South East Asia, however it has spread to Europe and scientists are worried it could spread to the UK. The Asian hornet is dangerous because it kills honey bees to feed to its larvae. Honey bees are important pollinators of our crops. If Asian hornets settle in the UK they could destroy our bee population and have a serious impact on beekeeping and agriculture. Genome sequencing will help researchers stay one step ahead of the hornet.

**King Scallop (Pecten maximus)**
The King Scallop is a type of shellfish that lives in UK waters and is a popular seafood dish. Scallops feed on microscopic algae that floats in the water, however some of these algae produce an acid that can build up inside the scallop. This acid is toxic and can poison humans if eaten. Studying the DNA of the scallop could help us understand more about how the scallop survives the toxic algae.

**Giant Hogweed (Heracleum mantegazzianum)**
A close relative of cow parsley, giant hogweed was introduced to the UK in the 19th century because of its impressive size. It now out-competes native plants and its sap can cause serious burns if it comes in contact with our skin. Studying its DNA will help to answer questions about how it can grow so big and why it is so toxic.

**Himalayan Balsam (Impatiens glandulifera)**
Himalayan balsam is a fast-growing and very successful plant species that was introduced to UK gardens in the 19th century. In the summer it produces purplish pink helmet shaped flowers. It has now spread across large areas of the UK and is found around rivers and waterways. It grows to over 2.5 m so shades out other plants meaning they cannot grow. It is now seen as a “problem plant”.

**New Zealand Flatworm (Arthurdendyus triangulatus)**
The New Zealand flatworm arrived in the UK in the 1960s, probably on imported plants. It is a ribbon-like creature, purple-brown in colour with a white edge around it. The flatworm can grow up to 15 cm long and feeds exclusively on earthworms. Very few animals eat the flatworm so its population has increased in the last few decades. There are concerns that it could have an impact on soil health but also earthworm eating mammals such as badgers and moles.
CRYPTIC
Species that are out of sight or are unidentifiable from others based on looks alone.

Brown Trout (Salmo trutta)
The brown trout is a fish species that takes three forms, the key difference being how and when they migrate. One form of the fish remains local to its birth where it will live out its life, spawn and die. The second migrates from lakes to fresh water rivers to spawn. The third migrates to the sea, only returning to fresh water to spawn. There appears to be no visible difference between them, so a reference genome will help to study the species in more detail.

Carrington’s Featherwort (Plagiochila carringtonii)
Carrington's featherwort is a liverwort (an ancient type of plant) only found in areas of high rainfall. In the UK it is found in just a few sites including the mountains of North West Scotland, Ireland, the Faeroes. It is also found in the Himalayas. Interestingly the Scottish plants are all male, and the females are in the Himalayas. These plants developed millions of years before flowering plants. Studying this easily overlooked plant can reveal a whole world of novel genetics.

Common Pipistrelle Bat (Pipistrellus pipistrellus)
The common pipistrelle bat is one of the most common bat species in the UK. It weighs the same as a 20 pence piece but can eat 3,000 insects in one night! It was thought to be just one species until 1999 when scientists discovered there were actually two: the common and the soprano. The two bat species can identified by differences in their echolocation frequencies. Studying the genome of the bat will help to reveal more about how and when this split occurred.

Summer Truffle (Tuber aestivum)
Fungi are a separate kingdom from animals and plants. When we think of fungi, mushrooms and toadstools come to mind. Truffles are a type of underground mushroom that forms a symbiotic relationship with the root systems of some trees - that is one that benefits both the fungus and the tree by sharing nutrients. The growing interest to harvest truffles for food means that studying the Summer Truffle’s genome will reveal more about its reproduction and life cycle.

Common Starfish (Asterias rubens)
If you go to any UK seashore you are likely to see a common starfish in a rock pool. The orange, five armed starfish is more like us than we might think. They produce a hormone linked to puberty that is helping us to understand our evolutionary development. Starfish can also regenerate limbs, so studying their genome could help with medical advances for our species.
CREATURE REPORT
Creature information

FLOUNDERING
Species that are endangered and declining in the wild.

Turtle Dove (Streptopelia turtur)
Perhaps most familiar from the Twelve Days of Christmas, turtle doves have declined by 93% since 1970, and it is on the Global Red List for Endangered Species. The turtle dove is now only found in eastern England, where farmers are working with the RSPB to create feeding habitats, the destruction of which are blamed for the bird’s decline. Understanding its genetics could help with its conservation.

Water Vole (Arvicola amphibius)
The water vole is an iconic species native to the UK but its numbers are in serious decline due to habitat loss and predation by the non-native American mink. There are active conservation and breeding programmes reintroducing voles back into the wild. Sequencing the vole’s genome will help with this conservation effort.

Eurasian Otter (Lutra lutra)
The Eurasian otter or common otter is a large semi-aquatic mammal found living by rivers and waterways in the UK. It is a great swimmer and hunts for fish underwater. However, the number of otters has drastically declined in the UK, largely because of polluted water, habitat loss and hunting. Creating a reference genome for this species will help speed up further study to better understand population genetics, the impact of environmental pollutants and the wider effects these may have on the freshwater food chain.

Northern February Red Stonefly (Brachyptera putata)
The Northern February red stonefly is only found in the UK, in North East Scotland and the Highlands. These insects have nymphs (young) that live in the water of rivers and have flying adults. Stoneflies thrive in clean, high quality water. If water quality drops they cannot survive. Due to a number of habitat changes, including acidification and chemical pollution in rivers and streams, the stonefly numbers are declining and they are becoming increasingly rare. Like other at-risk species, studying its genome could help with its conservation.

Red Squirrel (Sciurus vulgaris)
The red squirrel is the only native squirrel species in the UK. The introduction of the grey squirrel from North America has had a dramatic impact on red squirrel numbers. Today it is only found in Scotland, the north of England, Northern Ireland, as well as, isolated populations in Wales and on islands off England’s southern coast. Sequencing and comparing its genome with the grey squirrel may help to find new ways to protect the red squirrel against squirrel pox virus, which can be deadly for the red squirrel.
**Flourishing**

These species numbers are increasing in the wild.

**Grey Squirrel** (*Sciurus carolinensis*)

The grey squirrel was introduced to the UK from North America in the 19th century. Its larger size, more flexible diet and better breeding numbers means that it very successfully out-competes the native red squirrel. They can also carry squirrel pox virus, which can be deadly to reds. Sequencing and comparing its genome with the red squirrel may help to understand how it has adapted so well.

**Ringlet Butterfly** (*Aphantopus hyperantus*)

This velvety chocolate brown butterfly is unmistakeable when resting on a leaf or flower because of the ring patterns on its wings. Numbers of the ringlet butterfly have increased around 400% in the last 50 years, whereas many other UK butterflies have declined. A reference genome for the species will be a useful tool to compare with other butterflies to understand how it has managed to go against the trend of declining numbers.

**Oxford Ragwort** (*Senecio squalidus*)

The yellow heads of Oxford ragwort are a familiar sight from train windows. The introduction of the railways in the 19th century helped the plant to spread. It is a hardy plant that seems to be able to grow well in many locations. How has this non-native species become so established? Studying its genome will help scientists to understand how invasive species can succeed in new habitats.

**Fen Raft Spider** (*Dolomedes plantarius*)

The fen raft spider is the UKs biggest spider, measuring up to 7 cm across! This amazing arachnid lives in fens and other wetlands in Southern England and Wales. The population is under serious threat due to habitat loss. However, conservation efforts are beginning to help increase these numbers again. Studying the DNA of this spider will help breeding and conservation efforts. It could also reveal more about its semi-aquatic lifestyle, making it a useful bridge between land and water-based spiders.

**Roesel’s Bush-Cricket** (*Metrioptera roeselii*)

The Roesel’s bush-cricket is an insect found in the UK. It is brown or yellow in colour and grows up to 26 mm in length. Once restricted to saltmarsh habitats, the chirp of Roesel’s bush-cricket can now be heard in many parts of England. Researchers have questioned whether our increased use of salt on our roads is enabling the bush-cricket to spread to new habitats. By studying its genome, researchers may be able to discover genes that enable the bush-cricket to have an increased tolerance to salt.
Scientists at the Wellcome Genome Campus are decoding the DNA of all animals and plants found living in the UK. From squirrels to spiders, blackberries to bees, lots of different living things are being investigated. What will their DNA tell us about them?

**Your task is to create an information poster on one of the species being investigated.**

You can find information about 25 of these species on the creature information sheet to get you started. You can draw your poster or use a computer.

You can be as creative as you like but it must include the following information and features:

1. Name of the species
2. A picture of the species
3. Is it an animal, plant or fungi?
4. Where is it found?
5. What environment or habitat does it live in?
6. What adaptations does it have?
7. Is it rare or endangered?
8. A fascinating fact or “did you know” box
9. What do you think DNA could tell us about this species?
Would we know as much about these species if we didn’t sequence their genomes? What do you think?

Genomics gives us a way to explore the natural world around us and enable us to figure out how it all works. Even though DNA is really small, it can hold a lot of data – a lot of which we still don’t fully understand. As technology continues to improve, larger scale projects (such as the Darwin Tree of Life) can be carried out to unlock more of nature’s potential. Every year more and more creatures will have their genome sequenced. It is a growing field with huge potential to help us better understand the biodiversity of the planet and hopefully help us to protect it.

WHAT DO YOU THINK?

Now that you have finished the pack have a go at answering these questions!

www.surveymonkey.co.uk/r/Genomics_and_Adaptations