

Tree of Life

60,000

species in Britain and Ireland will be sequenced as part of the Darwin Tree of Life



In this section

- 1 Darwin Tree of Life sets down roots
- 2 Sequencing an entire ecosystem
- 3 Brown trout genome will help explain species' genetic superpowers
- 4 Genome could hold the key to red squirrel survival

1

Darwin Tree of Life sets down roots

In 2019 the Sanger Institute welcomed Professor Mark Blaxter as the Tree of Life Programme Lead and the Darwin Tree of Life project secured £9.4 million in funding from Wellcome.

The Darwin Tree of Life project is one of several initiatives across the globe working towards the ultimate goal of sequencing the genomes of all complex life (eukaryotes) on Earth. The project is the first effort to sequence the species of one geographical area – Britain and Ireland. The project joins other global initiatives in a historic venture known as the Earth BioGenome project, which aims to sequence 1.5 million species of plants, animals, fungi and protists.

Exploring the genomes of these organisms will give an unprecedented insight into how life on Earth evolved and uncover new genes, proteins and metabolic pathways as well as new drugs for infectious and inherited diseases. At a time when many species are under threat from climate change and human development, these data will also help characterise, catalogue and support conservation of global biodiversity for future generations.

The Sanger Institute will serve as one of the hubs for sequencing and assembling the genomes of 60,000 species from Britain and Ireland. Professor Mark Blaxter joined the Institute to set up and lead the new Tree of Life Programme.

The Sanger Institute will collaborate with partners at the Universities of Cambridge, Edinburgh and Oxford, the Earlham Institute, EMBL's European Bioinformatics Institute (EMBL-EBI), The Marine Biological Association, Natural History Museum, and the Royal Botanic Gardens at Kew and Edinburgh. Together, teams will identify and collect species, set up new pipelines and workflows to process large numbers of species through DNA preparation, genome sequencing and assembly, gene finding and annotation.

New methods will be developed for high-throughput and high-quality assembly of genomes and their annotation, and data will be shared openly through existing data sharing archives and project specific portals.

The £9.4 million funding from Wellcome will support researchers to launch the first phase of sequencing 60,000 species in Britain and Ireland. This will see the teams collect and barcode around 8,000 key species, and deliver high-quality genomes of 2,000 species.

These data will be of enormous value to the international scientific community, including those working in life sciences, medicine, alternative energy and climate research. The data will also act as a global resource for public engagement experts, naturalists, citizen scientists, university students and schools.



Link
Darwin Tree of Life project news story

The Darwin Tree of Life project is part of a global project to sequence

1.5m

species

2

Sequencing an entire ecosystem

Species living in Wytham Woods in Oxfordshire will be among the first to be sequenced for the Darwin Tree of Life project.

Wytham Woods is one of the most studied areas of woodland in the world. Owned and maintained by the University of Oxford, its 1,000 acres are a designated Site of Special Scientific Interest (SSSI). Home to a rich diversity of species, there is a wealth of long-term biological data available about the ecosystem.

As part of the Darwin Tree of Life project, the species living there will be collected and catalogued, and their genome sequences will add to the wealth of data available on the ecosystem.

It's thought there are approximately 4,000 eukaryotic species in the woods, though this is likely to be an underestimate. There are over 700 species of Lepidoptera (butterflies and moths), and these will be among the first species to have their genomes sequenced.

The goal over the first two years of the project will be to sequence the dominant players in the ecosystem, and a broad taxonomic spread of species.

For most species, good quality genetic or genomic information is unavailable, so the data will be hugely valuable to the research community worldwide. Having data from an ecosystem will allow scientists to discover how different species have adapted to living in the same environment. For example, genome sequence data could give insights into how different insects break down the same defensive plant toxin. It may be that there are different solutions to the same problems; alternatively species may have evolved the same solutions independently.

The data will also provide a basis for analysing genetic responses to environmental changes and human disruptions to habitats, as researchers will be able to study and compare how gene expression changes over time.



Link

Tree of Life Programme at the Sanger Institute



3

Brown trout genome will help explain species' genetic superpowers

The newly sequenced brown trout genome will help settle the debate around whether it is one species or several and aid conservation efforts as we face a warming planet.

Brown trout (*Salmo trutta*) is one of the most genetically diverse vertebrates. Taxonomists once classified brown trout as up to 50 distinct species, though now they are considered one. Different populations have adapted to exploit particular biological niches, with some living their whole lives within a 200 metre stretch of freshwater stream while others migrate from the stream where they were born to the open sea.

Researchers at the Sanger Institute and their collaborators have sequenced the brown trout genome for the first time. The data will allow scientists and conservationists to better understand the genetic roots of this highly specialised species. The genome will also help settle the long-standing debate about whether the physically-varied brown trout is a single species or several.

The brown trout genome will enable scientists to compare DNA from different populations of trout to the reference genome sequence – giving insights into how particular genetic variations allow certain trout to live in habitats that would be fatal to others. Pinpointing genetic variations that allow Scottish loch trout to adapt to living in relatively acidic waters, for example, may be useful in guiding conservation efforts to protect populations affected by increasing acidity in rivers and oceans as a result of global warming.



Link

Brown trout genome news story

Taxonomists once
classified brown
trout as up to

50

distinct species





Squirrel pox can kill up to

80%

of a red squirrel population in an affected area

4

Genome could hold the key to red squirrel survival

Sanger researchers and their collaborators have sequenced the genomes of red and grey squirrels, boosting research and aiding red squirrel conservation efforts.

Red squirrels are native to Britain and Ireland, but they are under serious threat of extinction in the UK because of habitat loss, competition with grey squirrels and fatal squirrel pox. While grey squirrels are immune to the disease, squirrel pox outbreaks can kill up to 80 per cent of a red squirrel population in an affected area.

Red squirrels are now confined to isolated pockets in northern England and Wales, with a more connected population in Scotland and Ireland. They are at great risk of being pushed out of their territories by greys, outbreaks of squirrel pox and inbreeding.

Sanger scientists extracted DNA from red and grey squirrel samples sent by collaborators. They used PacBio SMRT® and Illumina sequencing technology to generate the first, high-quality red and grey squirrel reference genomes.

It is hoped that the genomes will be used to identify the genetic basis of immunity to squirrel pox in grey squirrels, as well as in the reds that survive outbreaks. The genome can also be used to help ensure that populations are genetically mixed, helping to select animals for reintroduction based on their genetic compatibility with existing populations to give them the best chance of survival through successful breeding.

The red and grey squirrels are two of the species to be sequenced as part of the Sanger Institute's 25 Genomes project and will contribute to the ambitious Darwin Tree of Life project.



Reference

Mead D. *et al.* The genome sequence of the Eurasian red squirrel *Sciurus vulgaris* Linnaeus 1758. *Wellcome Open Research*. DOI: 10.12688/wellcomeopenres.15679.1.

Stories
3-4

25 Genomes for 25 Years – Institute's anniversary project

The brown trout, the red squirrel and the grey squirrel are some of the UK species to have been sequenced as part of the Sanger Institute's 25th anniversary 25 Genomes project. The project laid the groundwork for the ambitious Darwin Tree of Life project, which will sequence 60,000 complex species in the UK.