

Saimaa Ringed Seal Aids the Study of Population Genomes

The research groups of Jukka Jernvall and Petri Auvinen at the Institute of Biotechnology are investigating the genomes of different species and the structures of populations. The objective is to understand when the species arose and diverged from one another. The groups are particularly interested in the Saimaa ringed seal, whose full genome will be determined.



Foto: Suomen Luonnonsuojeluyhdistys/Juha Taskinen

The Saimaa ringed seal is an excellent research subject for the study of genetic diversity, isolation and inbreeding. The Saimaa ringed seal has not been in contact with other seal species in more than ten thousand years. Its eyes, brain and skull are different from those of other types of ringed seal. The Saimaa ringed seal developed from a seal population that probably came from the Baltic Sea to Lake Ladoga before moving to the Saimaa archipelago.

“If a Ladoga ringed seal was transferred to Lake Saimaa, it might not survive. The Saimaa ringed seal has adapted to the murky waters containing humus and the maze-like

archipelago”, says **Petri Auvinen**, Laboratory Director of the Institute of Biotechnology.

The DNA Sequencing and Genomics Laboratory of the Institute of Biotechnology specialises in gene sequencing, or determining the order of base pairs in DNA. The laboratory has sequenced the entire genome of several organisms, starting from the *Lactococcus piscium* bacteria that spoil cold food. Gene expression is also studied at the laboratory through sequencing. Key events in the evolution of organisms include cellular division and differentiation, which are highly temporally and spatially regulated.

Cellular differentiation takes place in stages. Sometimes a gene is switched on and sometimes it stops functioning. This active functioning is called gene expression. When gene expression can be measured, it is possible, for example, to monitor which genes start to function when a tree prepares for winter. The EST (Expressed Sequence Tag) technology provides information on the location and function of a gene. By identifying the base pair sequence of genes, a tag can be provided for each expressed gene. Currently, the RNA-Seq method is mainly used to study gene function.

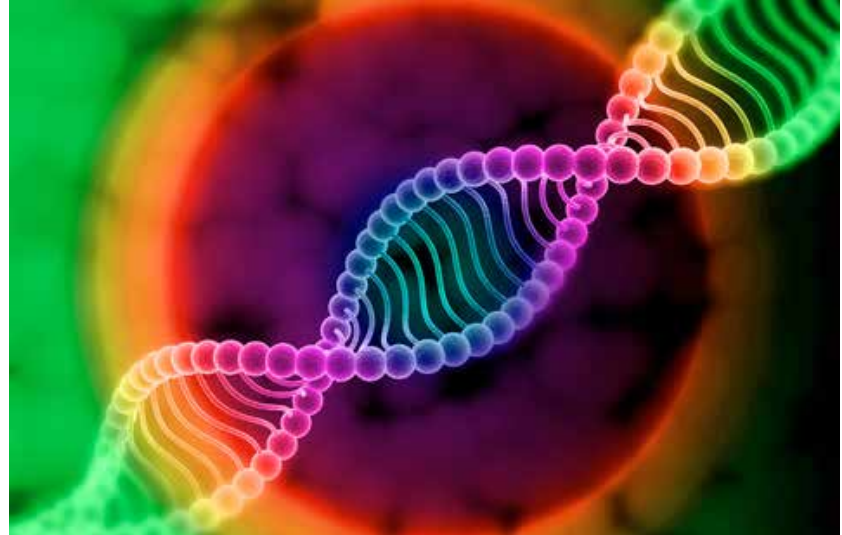
Reference genome for population research

The aim of the researchers of the Institute of Biotechnology at the University of Helsinki is to obtain a reference genome of the highest quality possible from the Saimaa ringed seal. A reference genome is a digital sequence database on the full base pair sequence of a single species, compiled from one individual in the case of the Saimaa ringed seal and from numerous genomes in the case of humans. Collecting a good reference genome requires the use of various advanced technologies.

The reference genome and deviations in individual genomes enable the efficient study of the population. In the STR (Short Tandem Repeat) method, a specific locus on DNA where a few base pairs in a row are always repeated is compared with two or more DNA samples. The DNAs of individuals are clearly distinguished with STR. Mitochondrial DNA, in turn, can be used to trace the maternal lineage of individuals back thousands of years. The rapid development of DNA sequencing technologies has enabled the identification of single-nucleotide polymorphisms (SNP), providing a very accurate estimate on the differences between individuals. This method is also used in the Saimaa ringed seal genome project. The data collection requires a lot of storage space and computing power, provided by CSC – IT Center for Science via the ELIXIR infrastructure.

The genome of the Saimaa ringed seal is 2.5 billion base pairs in length, the same as the canine genome. In determining the genome of the Saimaa ringed seal, the group of Academy Professor Jukka Jernvall focuses on studying the teeth of seals while the group of Petri Auvinen focuses on population history and genome structure. Once the genome has been determined, the genome of the Saimaa ringed seal will be compared with the genomes of the ringed seals in Lake Ladoga, the Baltic Sea and the Arctic Ocean.

The researchers are collecting data on the connections between the genotype (genetic factors) and phenotype (environmental factors) together with researchers from the universities of Oulu and Eastern Finland. A lot of data on developmental biology is obtained by analysing teeth.



Once a tooth erupts, it will no longer develop or change due to the environment. However, there is huge variation in teeth. That is why it is studied as to which genes have affected unusual teeth. The teeth of the crabeater seal, for example, have become very polymorphic due to evolution and function like the baleen of whales because the seals eat krill.

“We have computer models of all ringed seal skulls. We can create accurate phenotypes and look for the probable genes that caused a particular tooth. Gene function can be modelled on a computer and analyse which areas of the genome could affect the tooth.”

A different skull or teeth indicate adaptation or specification, adjustment to different conditions. Because the orbits of the Saimaa ringed seal are different from those of other, even closely related ringed seals, it can be concluded that, for example, it has adapted to murky and maze-like waters.

The groups of Auvinen and Jernvall have access to the DNA of the only known hybrid between a ringed seal and a grey seal in the world. In 1929, Skansen Zoo was the birthplace of a cub from whose tooth Auvinen managed to isolate DNA. The offspring of a huge grey seal and a small ringed seal only lived for a short time. The teeth and skull of the hybrid indicate an intermediate form. According to Auvinen, it would probably be the equivalent of a hybrid between a chimpanzee and a human. It is now

possible to compare why a specific kind of tooth or skull develops.

Auvinen considers this significant research also for human evolution because it is not known when modern humans differentiated into their own species. Hybrids have also occurred during human evolution. There have been findings of human skull fragments that are a cross between Cro-Magnon and Neanderthal. 2-5% of Europeans carry genes passed down from Neanderthals. Furthermore, the skeleton of a human subspecies named the Denisovan was found in the Denisova Cave in Siberia. It became extinct 40,000 years ago, earlier than its cousin, the Neanderthal. When DNA was isolated from the finger of the Denisovan's skeleton, it was found that Tibetans have Denisovan genes. One hereditary gene helps Tibetans survive in a high altitude climate.

Bottlenecks relate an interesting genetic history

The researchers of the Institute of Biotechnology want to find out whether the Saimaa ringed seal is its own species or a subspecies. The researchers know exactly for how many generations the ringed seal has been isolated in Lake Saimaa. The population of the Saimaa ringed seal is small. There were only 140 individuals left in the 1980s, now the number is 320. By comparing the samples from Lake Saimaa, the Baltic Sea and Lake Ladoga to the reference genome of the Saimaa ringed seal, it is pos-



sible to study what kind of a population has passed through a so-called bottleneck.

Nowadays, there are also computational methods that make it possible to determine reasonably accurately, even from a single genome, the kind of a population its ancestors have lived in. The bottleneck phenomenon faced by a population refers to an event where a large part of the population is destroyed or only a small number of individuals establish a new group, such as the people who once arrived in Finland. The reason behind the destruction may be changes in the environment or a transition to a new environment, which can prevent reproduction.

Studying the genetic history of the Saimaa ringed seal is also helpful for human genome research. Bottlenecks can increase inbreeding and thus also affect the disease heritage. In Finland, bottlenecks have given rise to about forty hereditary diseases in the population that are much common here than anywhere else. Finnish genetic bottlenecks have included the adoption of agriculture 4,000 years ago and the spread of settlements to the northern and eastern Finland in the 16th century.

“It is now possible to examine the impact of disease genes on population structure and the bottlenecks caused by nature

and humans. The Finnish disease heritage is interesting in this respect. It can be determined what the disease heritage carried by Finns was like when they went through a bottleneck”, says Auvinen.

Data can be reused

Creating a reference genome comes with many benefits. The reference genome data can always be reused. The better the reference genome is, the easier it is to analyse new data that can be compared to the data of the reference genome.

For example, analysing the reference genome of birch accelerates and enhances wood research for the needs of industry and medicine. New properties that affect the quality and quantity of wood can be looked for in the birch genome. This data can also be utilised in research on other wood species.

“Unlike birch, it will take 10 years to determine the properties of, for example, poplar and eucalyptus. Birch can be genetically modified. Since birch can be made to bloom up to three times a year, new properties can be introduced to birch in one to two years. These techniques can also be applied to other wood species. The genetic model of birch can be used, for example, in the study of eucalyptus”, says Petri Auvinen.

The birch reference genome project was also followed by industry representatives. Thanks to genetic data, birch properties can be refined and the forest industry can use the wood for purposes other than timber.

New applications include nanomaterials, wood processing industry side streams and, for example, hemicellulose. Auvinen also mentions the betulin in birch bark that has been reported to have anti-cancer and even antiviral effects. Betulin has already been used to create medicinal creams. Striving to produce birches with more betulinic acid using conventional breeding methods is also a possibility.

Ari Turunen

FURTHER INFORMATION:

Institute of Biotechnology

is an independent research unit belonging to the Helsinki Institute of Life Science (HiLIFE) at the University of Helsinki

<http://www.biocenter.helsinki.fi/bi/dnagen/index.htm>

CSC – IT Center for Science

is a non-profit, state-owned company administered by the Ministry of Education and Culture. CSC maintains and develops the state-owned, centralised IT infrastructure.

<http://www.csc.fi>

<https://research.csc.fi/cloud-computing>

ELIXIR

builds infrastructure in support of the biological sector. It brings together the leading organisations of 21 European countries and the EMBL European Molecular Biology Laboratory to form a common infrastructure for biological information. CSC – IT Center for Science is the Finnish centre within this infrastructure.

<http://www.elixir-finland.org>

<http://www.elixir-europe.org>

ELIXIR FINLAND

Tel. +358 9 457 2821 • e-mail: servicedesk@csc.fi
www.elixir-europe.org/about-us/who-we-are/nodes/finland

www.elixir-finland.org

ELIXIR HEAD OFFICE

EMBL-European Bioinformatics Institute
www.elixir-europe.org