

Team presents draft assembly of the Norway spruce genome sequence



This image shows, from left: Ove Nilsson, Per Ingvarsson and Stefan Jansson, Umea University. Credit: Johan Gunesius/Synk

Swedish scientists have mapped the gene sequence of Norway spruce (the Christmas tree) – a species with huge economic and ecological importance - and that is the largest genome to have ever been mapped. The genome is complex and seven times larger than that of humans. The results have been published in the journal *Nature*.

This major research project has been led by Umeå Plant Science Centre (UPSC) in Umea and the Science for Life Laboratory (SciLifeLab) in Stockholm.

In addition to its scientific interest this new knowledge has immense importance to the forestry industry in many countries.

"Forest tree breeding is now entering a new era, and Sweden has the potential to be in the forefront of development," says Professor Ove Nilsson from UPSC. "Newer and more effective methods can begin to be used to ensure that the over 200 million tree seedlings planted each year in Sweden are as strong, healthy and well-adapted as possible for both poor and rich soil areas in different parts of the country."

The scientists have identified about 29,000 functional genes, marginally more than humans have, but the question arises: why is the spruce genome still seven times larger than ours? According to the study an explanation is "genome obesity" caused by extensive repetitive DNA sequences, which have accumulated for several hundred million years of evolutionary history. Other plant and animal species have efficient mechanisms to eliminate such repetitive DNA, but these do not seem to operate so well in conifers.

"Team presents draft assembly of the Norway spruce genome sequence." Phys.org, 22 May 2013.

<http://phys.org/news/2013-05-team-norway-spruce-genome-sequence.html>

"It is remarkable that the spruce is doing so well despite this unnecessary genetic load," says Professor Pär Ingvarsson at UPSC. "Of course, some of this DNA has a function but it seems strange that it would be beneficial to have so very much. This appears to be something special for conifers."

The greatest challenge in the project has been to get the approximately 20 billion "letters" found in spruce's genetic code into the correct order, rather than obtaining the actual DNA sequences.

"Imagine a library with ten thousand books as thick as the bible, written in a language with only four letters," explains Professor Stefan Jansson at UPSC. "If someone took one hundred identical copies of each of the ten thousand titles, passed them all through a document shredder and mixed all the shreds, and you then were asked to piece together an accurate copy of each title, you can realize that it can be a bit problematic."

"We had to customise computers and rewrite many of the computer programmes used in similar studies in order to handle the large amount of DNA sequences," says Professor Joakim Lundeberg from SciLifeLab. The national data storage system was stretched to the limit, and there were many other practical problems that had to be solved along the way to pull through the project.

"But the timing was optimal; when the new DNA sequencing machines were unpacked at our newly established laboratory, DNA arrived from our model spruce tree. By sequencing and analysing the largest genome in the world so far, we have shown that SciLifeLab has both technical and scientific capacity for research at the highest international level," concludes Joakim Lundeberg.

More information: Nystedt, B. et al. The Norway spruce genome sequence and conifer genome evolution. *Nature*. [dx.doi.org/10.1038/nature12211](https://doi.org/10.1038/nature12211)

Provided by Umea University

This document is subject to copyright. Apart from any fair dealing for the purpose of private study, research, no part may be reproduced without the written permission. The content is provided for information purposes only.