Knock on wood

The ProCoGen collaboration is exploring the genetics of adaptive response in conifers. As the research team writes on the project's blog, they are seeking answers to some particularly challenging questions about conifer genetics

Can you explain what it is that ProCoGen will accomplish through your efforts?

ProCoGen will address genome sequencing of two keystone European conifer species. Genome re-sequencing approaches will be used to obtain two reference pine genomes. Comparative genomics and genetic diversity will be closely integrated and linked to targeted functional genomics investigations to identify genes and gene networks that efficiently help to develop or enhance applications related to forest productivity, forest stewardship in response to environmental change, or conservation efforts. The development of high-throughput genotyping tools will produce an array of pre-breeding tools to be implemented in forest tree breeding programmes. ProCoGen will also develop comparative studies based on orthologous sequences, genes and markers, which will allow guiding re-sequencing initiatives and exploiting the research accumulated on each of the species under consideration to accelerate the use of genomic tools in diverse species. ProCoGen will integrate fragmented activities developed by European research groups involved in several ongoing international conifer genome initiatives and contribute to strengthening international collaboration with North American initiatives.

How important are the genome analyses that are currently available for improving conifer forest management?

Conifers are ecologically and economically the most important gymnosperms dominating many terrestrial ecosystems in the Northern Hemisphere. Fossil records show the presence of gymnosperms in the Late Carboniferous period 300 million years ago. They thrived through the age of the dinosaurs before and after mass extinction events 250–65 million years ago. Conifers are represented by 630 species belonging to eight families and 70 genera. Some of the woody plant species are the longest-living and largest organisms on this planet. The enormous size of their genome – 20–40 gigabase pairs which is seven times the human genome – was a major challenge in conifer genomics. The advancements in sequencing and bioinformatics technology have made the genome analyses possible today. The draft assemblies are providing valuable insight into the genome organisation, gene families involved in metabolic pathways, genetic basis of environment-plant interaction, and the comparative and evolutionary analysis of angio- and gymnosperms. They are also going to provide tools for better forest management strategies and markers for genome selection for tree breeding and improvement for well-adapted forests.

What kind of obstacles are there to breeding a new generation of conifers?

Genetic improvement of trees has always been a challenge, especially because of the long generation time and complex genomes. Developments in DNA marker-based technology have revolutionised plant breeding practices over the years. Molecular breeding based on DNA markers was termed marker assisted selection (MAS). By utilising markers closely linked to desirable traits one could exploit the allelic variations underlying the traits of interest. Prior identification of markers in association with the trait was a prerequisite for MAS. One of the major limitations of MAS came from the polygenic inheritance of traits. It did not consider small-effect genes underlying complex polygenic traits, thus limiting its usage to large effect alleles with association to markers.

Why are conifers so valuable when attempting to understand adaptation to changing climates?

Conifers, when they first evolved in the Permian, were warm-adapted and likely did not exhibit any of the cold-tolerance mechanisms observed in conifers today. Information on the physiological development of frost tolerance and cold hardiness of coniferous trees over geologic time is surprisingly lacking within the scientific literature. Conifers are the one plant functional group that is present in nearly all geologic periods and therefore span palaeoclimates ranging from global glaciations to hothouse temperatures.

Securing the future of Europe's forests

The ProCoGen collaboration is attempting to unravel some of the genomic secrets locked inside conifer species in order to deliver transatlantic sustainable management of the genetic resources of pine trees

Forests play a critical role in global ecosystems by providing renewable natural resources as well as helping protect soil, water and air, and conserve biodiversity. Encompassing over 40 per cent of Europe's total area, forests and other wooded land cover approximately the same amount of land as is used for agriculture. The forestry industry is important for Europe, representing around 7 per cent of the EU's manufacturing GDP. With growing awareness of global warming, Europe's forest industry is looking to understand how the economically valuable species of trees can be protected into the future. To achieve this, a better understanding is needed of how forestry species currently respond to changing climates and what adaptive measures are available to forestry managers.

A European Commission-funded research project known as ProCoGen (Promoting a functional and comparative understanding of the conifer genome-implementing applied aspects for more productive and adapted forests) aims to study the genome sequence of some of the most commercially important conifer species; look at the key economic and ecological traits; and then translate these into management and breeding programmes. This research provides timely knowledge about the molecular bases of these species, and how these genes control the response to changing climate. As the ProCoGen team write on their blog, this research will have a significant impact on forest management by building awareness of how conifer genes evolve and function: 'It would also help in discovering the potential of conifer adaptation in the face of the global environmental changes and explore various possibilities of tree improvement through marker and genomic selection'. It is hoped that this information can be relatively easily transferred to other conifer species, particularly those that are of a high value to Europe's forest industry.

ADVANCED BIOINFORMATICS

By breaking down the project into seven distinctly separate work packages the group are able to address many of the discrete technical investigations needed. This includes a group looking at de novo genome sequences and extensive catalogues of genetic variation for two species; a package studying the functional regulation of growth and adaptive responses to abiotic stresses; and a group translating the basic scientific discoveries from the other work packages into practical applications to support more accurate and efficient breeding.

There is a large body of information available on the sequencing of angiosperm genomes, such as those of oaks and eucalypts; however this cannot be transferred to conifer species.

Therefore investigations into the molecular basis for the productive and adaptive traits of conifers are needed. This is at the heart of the research questions posed by ProCoGen, which is built around an integrative modelling programme on species chosen for their economic and ecological importance for Europe, including the maritime pine found in Atlantic and Mediterranean Europe (*Pinus pinaster*), the Scots pine found in Central and Northern Europe (*Pinus sylvestris*), the Norway spruce found in Central and Northern Europe (*Picea abies*) and the Sitka spruce (*Picea sitchensis*), which is of high economic interest in the United Kingdom. This work utilises some of the most cutting-edge genome sequencing technologies available to the investigators.

Existing knowledge from other projects will be used to help build a clearer picture about conifer genetics, which is particularly challenging due to the large genome sizes and high frequency of repetitive sequences. The project team note on their blog that they are able to draw on advances in molecular technologies and use high-throughput platforms to sequence, genotype and undertake functional analysis of conifer species: 'Owing to the large genome size (18–35 gigabase pairs) of conifers their genome sequencing had remained a

> ' Traditional methods of breeding based on progeny testing and seed orchards may take from 10 to 20 years before producing commercial varieties improved for such novel traits. ProCoGen will develop tools for advanced faster breeding '

from other conifer species: 'Preliminary transcriptomic analysis has shown differential gene expression in 14 different cells and tissues under study. Epigenetic studies on somatic embryos of *Pinus sylvestris* showed temperature dependent alteration of gene expression.' In addition, a pilot exome capture was designed and tested on different conifer species. They have also prepared a preliminary catalogue of SNPs from *Pinus pinaster*. Through the work package on translational genomics the team has developed a proof-of-concept experiment for the implementation of genomic selection in maritime pine, which has delivered some new knowledge about genome wide diversity patterns and identified where there are gaps in the breeding populations.

A key aim of ProCoGen is to support improved decision making and policy development around how regional climate changes will potentially impact the growth of conifer forests and forest stewardship in general. In addition, the team consider that conservation efforts for forests will benefit from the new knowledge they are generating. In terms of the bigger picture, this work will support European compliance with international climate change policy, including obligations on climate change mitigation and adaptation of the Kyoto Protocol.

Throughout the project the team are working closely with other European programmes that are involved with conifer sequencing initiatives. It is hoped that all of these efforts initiated by the ProCoGen programme will draw together research efforts to build a much more robust knowledge base on conifer genomics and bioinformatics.

Sharing of the work and technology transfer of the results has been an important focus for the project. This has been achieved through activities such as training workshops and staff exchange programmes where training in emerging technological and translational approaches is supported. In addition, dissemination material has been prepared and a website set up which includes a dedicated section where stakeholders can access legal documents, minutes of meetings, and presentations. Through all of these initiatives it is hoped that a pool of technology providers can be created and supported with the most up-to-date information on conifer forest management to sustain Europe's commercial forestry interests into the future.

challenge until recent advancements in technology made it feasible to a great extent.'

Ultimately, this work helps to identify which genes control important traits, like how conifers adapt to drought and cold. The ProCoGen summary report describes how the scientists are developing novel breeding selection tools, whose design is based on associations between genomic and phenotypic variants, and improving the efficiency of breeding techniques: 'Traditional methods of breeding based on progeny testing and seed orchards may take from 10 to 20 years before producing commercial varieties improved for such novel traits. ProCoGen will develop tools for advanced faster breeding.'

TECHNOLOGY TRANSFER
The group have seen some encouraging results. Within the first 18 months of the programme they prepared a number of libraries using haploid tissue from two of the target conifer species. They have been using exome capture probes and total RNA extracts to help identify genetic variation based on single nucleotide polymorphisms (SNPs). Laser-capture microdissection and tissue-specific transcriptome libraries have been constructed. The ProCoGen project team reported in their annual newsletter that the following year preliminary results from genome sequencing were able to be used for comparative studies using data available