Science & Society

OPEN ACCESS

EMBO reports

Threatened forests

As the Northern forests suffer from the effects of climate change, genomics has great potential to help them adapt

Harry Wu* & Ove Nilsson*

"The truth is: the natural world is changing. And we are totally dependent on that world. It provides our food, water and air. It is the most precious thing we have and we need to defend it," said Sir David Attenborough 10 years ago referring to human-induced climate change.

The impact of climate change on boreal forests

In the summer of 2022, most of Europe experienced unprecedented high temperatures and droughts. Elsewhere, the USA and Asia had some of the highest temperatures and severe droughts since complete records are available. Wildfires have been frequent in Europe last summer, and the areas of vegetation burned in Spain, Romania and France were about four, seven and five times more than the average respectively (https://atmosphere.copernicus.eu). Even northern countries such as Sweden have experienced warmer weather in the summers and forest fires above the Arctic Circle—in 2018, the weather was about 5 degrees warmer than normal. Warmer temperatures also helped the spread of diseases and increased insect attacks, which caused the catastrophic destruction of 28 million hectares of forests in Canada by the mountain pine beetle and in Europe by the bark beetle. The great boreal forests in the North may no longer be able to cope with the changing climate that occurs much faster than the vegetation has time to adapt.

Climate change, the associated shifts of species distribution and adaptative evolution of forest plant species have been an ongoing process ever since the first plants colonized the land. However, the rapid changes caused by anthropogenic actions are new and unprecedented and could result in large-scale maladaptation and destruction of forests. Trees are important plant species that absorb huge amounts of carbon dioxide from the atmosphere. Boreal forests alone store approximately 22% of the world's carbon, making them a critical part of any efforts to mitigate the effects of climate change. Boreal forests are also an important ecosystem and an essential resource for the global timber and pulp industry. The development of mitigation strategies is therefore essential to maintain healthy forests that act as carbon sinks and continue to supply the raw materials for a sustainable green economy that replaces fossil fuels and other materials with carbon-neutral renewable products.

> *"The great boreal forests in the North may no longer be able to cope with the changing climate that occurs much faster than the vegetation has time to adapt."*

Normally, when organisms have to cope with changing environmental conditions they react by producing more offspring so as to create larger populations on which the forces of evolution can act. Eventually, over generations, most species will adapt to this new environment. However, the major issue with the current human-induced rapid climate change is that many tree species suffer an adaptational lag for several generations—meaning hundreds of years—owing to their long reproduction time (Browne *et al*, 2019). Concerns over maladaptation to climate change have therefore motivated population genomics studies in model species and genecology studies of how an organism's genes and its environment interact to shape its traits and characteristics. By understanding its genecology, scientists can better predict how a particular tree species will respond to changes in the environment, such as climate change or habitat loss (Exposito-Alonso *et al*, 2019). This understanding and the ability to make reliable predictions in turn will enable optimized strategies to mitigate the effects of climate change on forests. In particular, it would help forestry to select populations specifically adapted to a particular, local climate for replanting and breeders to generate better adapted trees more quickly and efficiently.

> *"... the major issue with the current human-induced rapid climate change is that many tree species suffer an adaptational lag for several generations – meaning hundreds of years – owing to their long reproduction time."*

Conifers are locally adapted to existing climates

Common garden experiments with reciprocally planted populations between different environments have been used for a long time

Department of Forest Genetics and Plant Physiology, Umeå Plant Science Centre, Swedish University of Agricultural Sciences, Umeå, Sweden
*Corresponding author. E-mail: Harry.Wu@slu.se, ove.nilsson@slu.se

DOI 10.15252/embr.202357106 | EMBO Reports (2023) 24: e57106 | Published online 11 April 2023

to analyse whether local plant populations are better adapted to specific environmental conditions than non-local populations. These demonstrated that the local, native population usually has a higher fitness compared to introduced populations of the same species (Alberto *et al*, 2013). Such experiments have also detected strong signals of local adaptation to existing or recent climates, especially in widely distributed forest trees with large population sizes (Savolainen, 2011), such as northern conifer species. Major conifer species such as Norway spruce, Scots pine and lodgepole pine have a continuous distribution in the northern hemisphere but only re-occupied the American and European continents about 20,000 years ago after the last glacial period. They show local adaptation in fitness-related traits, such as earlier budding at higher latitudes to avoid early frost (Fig 1).

Trees not only have very long generation times but also have high levels of genetic variation. This variation is great potential for evolving responses to climate change. However, the genomic basis of such genetic variation and trees' potential capacity to adapt to new and changing climates is almost unknown. A better understanding of the genomic basis of this adaptation would help answer questions such as whether local adaptation can be maintained under rapid climate changes and whether mitigation strategies, such as delineating adaptive seed and breeding zones, assisted migration and selection of breeding populations can be implemented in sufficient time.

Genetic and genomic base of local adaptation

The genetic architecture for adaptive traits was initially mapped using gene association studies. The results indicate that adaptive traits are controlled by many genes, hence the identification of individual genes that govern variation in adaptive traits is a challenging task, especially with the limited genomic resources available for trees, such as high-quality reference genomes, SNP-arrays for genotyping and so on. Finding the genomic sites related to climatic adaptation by combining genome-wide analysis of single nucleotide polymorphisms (SNPs) and estimates of adaptive traits from multiple-site experiments has yielded more robust results.

The first genome-wide study using multiple reciprocal experiments was reported in the model plant *Arabidopsis thaliana*. It characterized fitness trait variation of hundreds of accessions at multiple sites and genotyped ~250 K SNPs (Fournier-Level *et al*, 2011). Alleles that confer higher fitness in northern or southern Europe were more common in the northern or southern population, respectively. SNPs associated with fitness occurred at different loci in different environments, and amino-acid-changing SNPs were found to have a higher relation with climate variables and to account for adaptive variation. A recent genome re-sequencing of 517 *A. thaliana* lines from Spain and Germany discovered that current strong local adaptation was caused by past climate-driven natural selection. It follows that many native *A. thaliana* populations will become maladapted to rising temperatures in Europe (Exposito-Alonso *et al*, 2019).

Reference genomes and genomic selection for mitigating the effects of climate change

Improved genomic tools for tree species are critically needed to allow genome-wide studies of non-model species that show different patterns of adaptation compared with traditional model species, which are often short lived. The first reference genome assemblies of Norway spruce and white spruce (Birol *et al*, 2013; Nystedt *et al*, 2013) yielded only highly fragmented assemblies, but chromosome-level high-quality assemblies are now available thanks to the development of long-read sequencing technologies (Niu *et al*, 2022). We are currently working on high-quality sequencing assemblies of Norway spruce and Scots pine. These genomes are huge (20–25 Gb) and extremely complex, but enable us to detect genomic variation, including genome structure variation, that is shaped by historical climate changes; to construct genome-wide dense SNP arrays for fast and

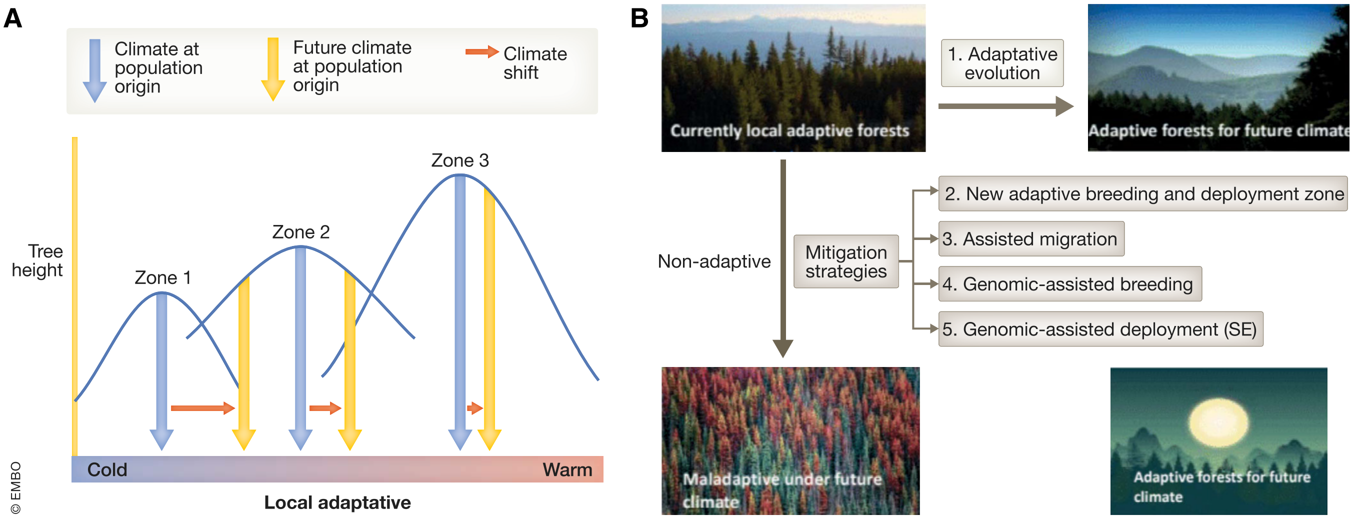

Figure 1. Detecting local adaptation by the response of adaptive traits to assisted migration or climate change.
(A) The response curves of different populations to the annual temperature indicate that each population has an adaptive peak that usually coincides with its site of origin before climate change. After climate shift, population from Zone 1 is not adaptive to Zone 1 any more while population from Zone 2 is more adaptive in the new climate of Zone 1 (A). Several adaptive strategies are possible to mitigate the impact of climate change on boreal forests (B).

inexpensive genotyping to study the genomic base of climate adaptation; and to implement genomic selection for accelerating tree breeding programmes for future climate adaptation. The immediate use of genome data in relation to climate change would encompass five approaches (Fig 1).

> *"Improved genomic tools for tree species are critically needed to allow genome-wide studies of non-model species that show different patterns of adaptation…"*

Understanding adaptation of local populations under rapid climate change

A dense genome-wide marker map allows a highly accurate estimation of the genetic variance of any natural population, which makes it possible to study the intensity of selection for different traits required to cope with climate change (Pritchard & Di Rienzo, 2010). As our understanding of the genomic basis of adaptive trait variation increases, it becomes possible to predict longer-term responses to climate change at the genomic level. Polygenic selection models that predict gradual changes at the causative loci of adaptive traits—which have been detected in many empirical studies—could be used to predict rates of selection. The expected effect on allele frequency is a function of the number of generations under selection, the selective pressure per allele per generation and the ancestral allele frequency. This enables us to estimate the strength of past selection and predict future adaptations under climate change. However, trees with a long generation time will experience a rapidly changing climate within their lifetimes. Alternative mitigation strategies are therefore needed to speed up the selection of adaptive traits, such as accelerated breeding programmes and deployment of better adapted trees.

Adaptive seed and breeding zones

A proper delineation of seed and breeding zones for reforestation and restoration is critical to ensure that seedlings are optimally adapted to the local environmental conditions. Seed and breeding zones have traditionally been assigned based on phenotypic traits associated with climate, an approach that requires long-term field experiments. In Sweden, for instance, breeding zones were traditionally delineated based on temperature and photoperiod data only, dividing the country into 22 Norway spruce-breeding zones.

With genomic data, it is now possible to apply a landscape genomics approach to delineate seed and breeding zones without the need for long-term field testing (Yu *et al*, 2022). By combining genotype and climate data of local populations, spatial genomic variation can be modelled using a gradient forest (GF) model. This analyses data from genotypes and environmental variables to identify patterns and relationships so as to predict how a particular plant species might be affected by a changing environment. Based on the GF models, adaptive seed or breeding zones can be delineated and bench-marked against breeding zones defined by climate or phenotype only. The verified models can be used to improve the assignment of existing seed or breeding zones under future climates.

Assisted migration to mitigate climate change

Assisted migration involves moving seeds or seedlings from one location to another area. At the level of single species, phenotype data have been used to construct species distribution modelling (SDM), which models the suitable climatic envelope of a species from historic records and climatic data but without considering genetic potential variation of local populations to adapt to new climates.

The time lag to adapt to climate changes can be shortened by genome-informed assisted migration. This approach uses genomic information to identify genotypes with faster growth rates or other fitness traits under warmer temperatures and optimally matches these to future climates based on genotype–phenotype–environment associations. Population-level genomic data can be used to estimate the magnitude of disruption in gene–environment relationships caused by a change in climate. We can relate current patterns of adaptive genomic composition to climate; the predicted optimal genomic compositions are then projected across a species' range (space), onto future climatic conditions (time), to estimate the magnitude of genetic shift in allele frequencies required by populations to maintain the current fitness under different climates (Capblancq *et al*, 2020). If natural selection and gene flow are too slow to match the climatic changes, assisted migration may be needed to move individuals from one population/region to another population/region (Aitken & Bemmels, 2015; Fig 1).

Genomic selection to accelerate breeding cycles

Genomic selection has revolutionized cattle breeding by shortening breeding cycles by more than 50%. It is, therefore, particularly attractive to conifer breeding which typically requires long-term field testing for 6–15 years to obtain reliable values to select optimally adapted seedlings. With genomic selection, tree breeders can already make selections from 1-year-old seedlings. This will substantially shorten breeding cycles and allow breeding of two to three generations within the time formerly required for one cycle (Fig 2).

Genomic selection employs prediction models using a "training" population and estimating the effects of genomic markers on phenotype in a dense, genome-wide panel. Typically, SNP markers are used. A few hundred or thousand progeny trees that have mature phenotype data are usually sampled from a target breeding population. The prediction models are validated against the samples and the model that delivers a high correlation between observed and predicted values is subsequently used for genomic selection in related breeding populations. The genomic breeding values for the candidates are predicted based on the genomic data only. Both training and validation populations derive from a single population, mainly from a progeny trial (Fig 2). The genomic models built for the selection of Norway spruce and Scots pine elite trees are now almost as effective as progeny testing (Chen *et al*, 2018; Calleja-Rodrguez *et al*, 2020).

Somatic embryogenesis for quicker deployment

Even if a genomic selection-based breeding programme is in place, a limiting factor for providing forestry with adapted seeds and seedlings is the time needed to raise a new generation of nursery stock. Today, planted

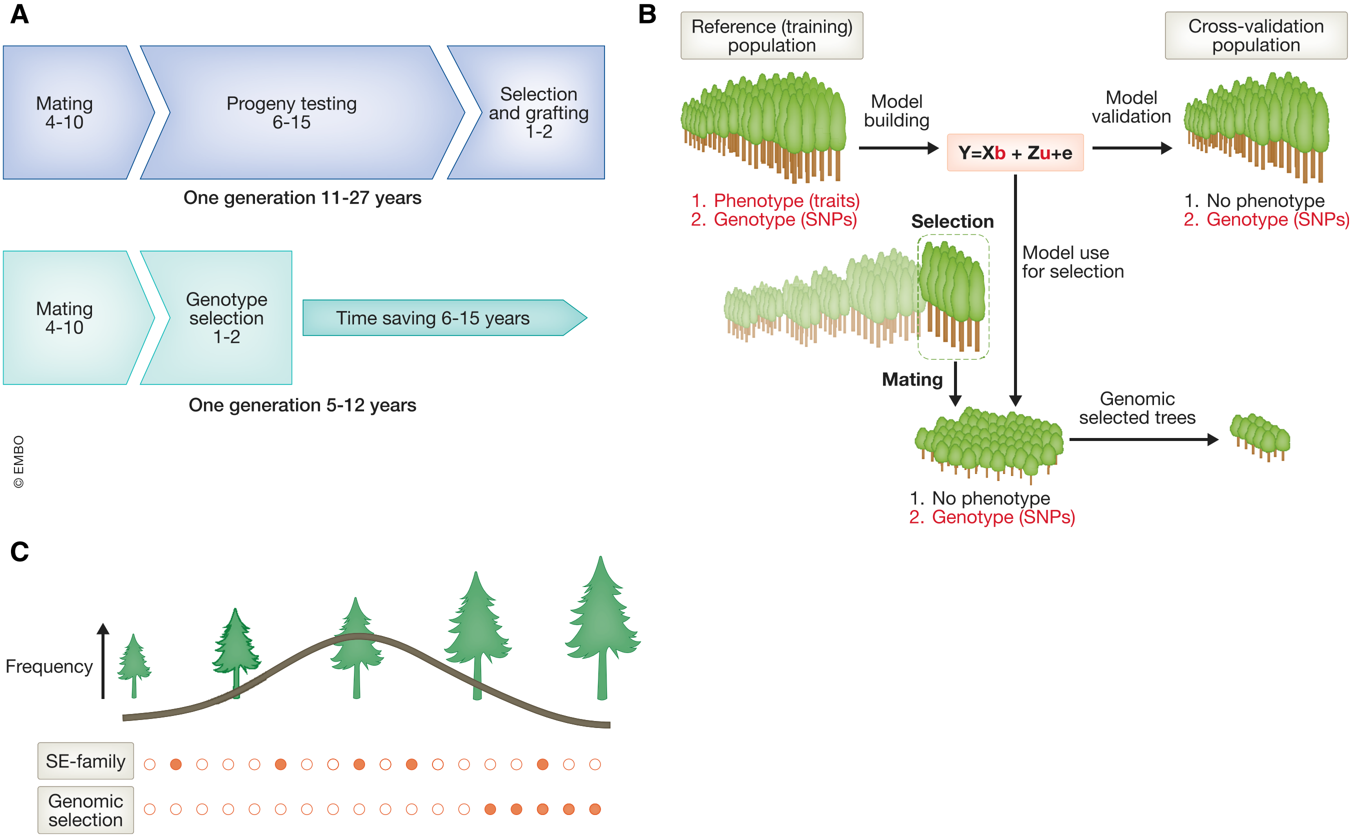

Figure 2. Genomic selection as a tool to shorten breeding cycles (A) and genomic selection as a strategy in operational tree breeding (B). Comparison of the genetic gain from propagation of SE families by either random selection of siblings (top row, red dots) or genomic selection of the best trees (bottom row, red dots). GS would allow us to identify and select the best adapted varieties (e.g. height) even if these have a low allele frequency (black line) (C).

seedlings are produced by seed orchards using grafted clones of progeny-tested parents. The problem is that seed orchards lag 20–40 years behind the breeding front because of the time needed for the trees to mature and start producing sufficient amount of seed. The genetic gain from seed orchards is also reduced by external pollen contamination from "wild" trees.

An alternative and more efficient solution is the use of vegetative propagation techniques, the most promising for large-scale deployment being somatic embryogenesis (SE). In this process, embryos isolated from seeds from a breeding programme are used to establish cell cultures from which individual cells can be stimulated to mature into an almost unlimited amount of new embryos. These clones can then be germinated into new seedlings to be deployed in a tree nursery.

These embryogenic cultures are typically cryo-preserved for decades, which allows field testing of a wide range of varieties to identify the best ones for mass amplification and planting. An alternative is SE amplification from elite crosses, and using genomic models to select a mix of the best individual embryos or seedlings within family (Fig 2). This could lead to an average genetic gain of maybe 15–20% more than family-based plantation without the time needed for field testing the best clones, and the genetic diversity would be very similar to that produced by a seed orchard. It would also give access to this genetic gain 20–40 years earlier because one would not have to wait for the establishment and maturation of the seed orchard. One of the most promising applications would therefore be to combine propagation of SE families with genomic tools to select for productivity and other characteristics, for instance, resistance to various diseases. Another big advantage of the SE technology is that it is suitable for automatization; indeed, several initiatives currently develop large-scale automated SE propagation platforms.

Future needs and prospects

Even with a breeding and propagation system using genomic selection and SE technology in place, the limiting step for producing better adapted trees is reproduction time. Trees are the latest flowering plants that we know of and conifer trees such as Norway spruce can take 20–25 years before they begin to produce reproductive structures. Tree breeders can somewhat shorten generation times by top grafting or hormone treatments, but tools for more further reductions of generation times are still lacking. However, due to a recent breakthrough in our understanding of the mechanism controlling cone development in conifers (Akhter *et al*, 2022), we could eventually create tools for inducing early flowering and shorten conifer generation times down to a few years.

Forest health and productivity are changing in response to climate change. The warming climate increases the risk and intensity of major disturbances such as drought, insect outbreaks or wildfires, which reduce productivity and drastically change the ecology of forests and distribution of tree species. However, warmer temperatures can also lead to prolonged growing seasons giving better-adapted boreal trees more time to photosynthesize. This can, with time, enhance forest growth and carbon sequestration.

> *“Even with a breeding and propagation system using genomic selection and SE technology in place, the limiting step for producing better adapted trees is reproduction time.”*

Therefore, it is critically important to improving our understanding of the genomic basis of local climate adaptation because of its relevance to fundamental evolutionary theory, conservation of genetic resources and the survival of forests in response to climate change. Species across the globe are experiencing drastic changes in environmental conditions as a result of human activities, and understanding and predicting how they respond to climate change is therefore both a major scientific challenge and fundamental for effective management of population adaptation and conservation.

So far, only model species were used to analyse local adaptation at the genome level. With high-quality genomes and other tools becoming available, long-term field experiments and short-term experiments under simulated climate regimes will enable research in tree species to understand adaptation patterns and genetic mechanisms. To do that, we need efficient and accurate phenotyping tools; fast and inexpensive genotyping tools; accurate climate data and modelling; and new methods and bioinformatic tools to integrate these data. In order to apply this knowledge in accelerated and efficient breeding programmes and to quickly deploy a new generation of healthy and climate-adapted trees, we also need efficient amplification and deployment methods, ideally combined with methods to shorten generation times.

> *“. . . it is critically important to improve our understanding of the genomic basis of local climate adaptation because of its relevance to fundamental evolutionary theory, conservation of genetic resources and the survival of forests in response to climate change.”*

References

Aitken SN, Bemmels JB (2015) Time to get moving: assisted gene flow of forest trees. *Evol Appl* 9: 271 – 290

Akhter S, Westrin KJ, Zivi N, Nordal V, Kretzschmar WW, Delhomme N, Street NR, Nilsson O, Emanuelsson O, Sundström JF (2022) Cone-setting in spruce is regulated by conserved elements of the age-dependent flowering pathway. *New Phytol* 236: 1951 – 1963

Alberto FJ, Aitken SN, Alía R, González-Martínez SC, Hänninen H, Kremer A, Lefèvre F, Lenormand T, Yeaman S, Whetten R *et al* (2013) Potential for evolutionary responses to climate change – evidence from tree populations. *Glob Change Biol* 19: 1645 – 1661

Birol I, Raymond A, Jackman SD, Pleasance S, Coope R, Taylor GA, Yuen MMS, Keeling CI, Brand D, Vandervalk BP *et al* (2013) Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29: 1492 – 1497

Browne L, Wright JW, Fitz-Gibbon S, Gugger PF, Sork VL (2019) Adaptational lag to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted gene flow. *Proc Natl Acad Sci USA* 116: 25179 – 25185

Calleja-Rodrguez A, Pan J, Funda T, Chen ZQ, Baison J, Isik F, Abrahamsson S, Wu HX (2020) Evaluation of the efficiency of genomic versus pedigree predictions for growth and wood quality traits in scots pine. *BMC Genomics* 21: 796

Capblancq T, Fitzpatrick MC, Bay RA, Exposito-Alonso M, Keller SR (2020) Genomic prediction of (Mal)adaptation across current and future climatic landscapes. *Annu Rev Ecol Evol Syst* 51: 245 – 269

Chen Z-Q, Baison J, Pan J, Karlsson B, Andersson B, Westin J, García-Gil MR, Wu HX (2018) Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* 19: 946

Exposito-Alonso M, 500 Genomes Field Experiment Team, Burbano HA, Bossdorf O, Nielsen R, Weigel D (2019) Natural selection on the *Arabidopsis thaliana* genome in present and future climates. *Nature* 573: 126 – 129

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science* 334: 86 – 89

Niu S-H, Li J, Bo W-H, Yang W-F, Zuccolo A, Giacomello S, Chen X, Han F, Yang J, Song Y *et al* (2022) The Chinese pine genome and methylome unveil key features of conifer evolution. *Cell* 185: 204 – 217

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A *et al* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature* 497: 579 – 584

Pritchard J, Di Rienzo A (2010) Adaptation – not by sweeps alone. *Nat Rev Genet* 11: 665 – 667

Savolainen O (2011) The genomic basis of local climatic adaptation. *Science* 334: 49 – 50

Yu Y, Aitken SN, Rieseberg LH, Wang T (2022) Using landscape genomics to delineate seed and breeding zones for lodgepole pine. *New Phytol* 235: 1653 – 1664