How to Use Genomic Selection for Tree Improvement

June 3, 2021

SELECT superior parent trees from natural stands.
2-4 YEARS

GATHER materials to reproduce trees.

Cones

Seed

Seedlings

Cuttings

GRAFT & GROW cuttings from superior parent trees.

1ST GENERATION ORCHARD

REFOREST ~15 YEARS SINCE START

GROW & TEST progeny for desired traits/phenotypes.

Example:
- 100–400 families (family = same parent tree)
- 3–5 progeny test sites/breeding program
- 5–10 progeny/family/site

1ST GENERATION PROGENY TRIAL **~20 YEARS**

MEASURE phenotype of progeny.

For example:
- Height
- DBH (diameter at 1.3m)
- Wood density
- Disease resistance
- Drought tolerance

COLLECT Needle samples from cloned parent trees and progeny for DNA extraction.
- Example: sample ~3 progeny/family/site
- *Need personnel competent in collecting samples*

SEQUENCE DNA to find genetic variation between the families.
- Use genotyping-by-sequencing (GBS), SNP chip/target capture, or other method (see glossary)
- *Need qualified bioinformatician*

PERFORM GENOMIC ANALYSIS Digitize the DNA sequence into base pairs of G, A, C, and T.
- Identify genomic relationships
- Build genomic model to link patterns of genetic variation to patterns of phenotypic variation (the model is trained using the real-world phenotype data measured in the progeny trials)
- *Need qualified quantitative geneticist*

AGTAC =

AGCAC =

## OUTPUTS FROM THE GENOMIC ANALYSIS:

1. A GENOMIC SELECTION MODEL

A model is produced which can now be used to predict the genomic breeding value of a tree based on its DNA sequence.

$$genomic\ breeding\ value_i = \sum_{j=1}^{n} genotype_{ij} \cdot effect_j$$

VERIFICATION

The model is validated using an independent sample of mature, phenotyped trees from the same population. The phenotype predicted from the model is then compared to the actual phenotype of the tree.

MODEL

PREDICTED   **ACTUAL**
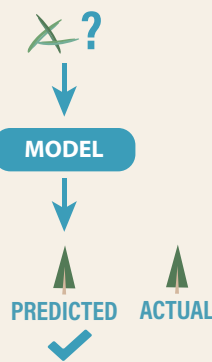
2. POPULATION STRUCTURE

The genomic analysis can reveal hidden relationships between families and seedlings.

A B C D E → A C E B D

This information can identify precise family structure and help detect unintentional errors (e.g., record keeping, nursery or planting mix-ups).

3. GENOMIC BREEDING VALUES FOR INDIVIDUAL TREES

Both pedigree analysis and genomic analysis can estimate breeding values for each family, but genomic analysis reveals a greater range of the variation within a given family.

The major advantage of genomic analysis is that it more accurately estimates the differences between individuals within a family. This helps ensure that only the best progeny are selected for further breeding and orchard development.
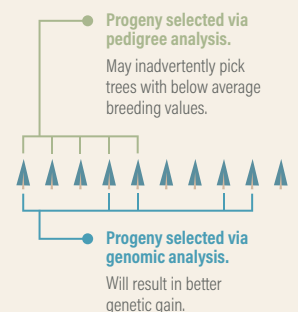
BREEDING VALUE

PEDIGREE ANALYSIS

GENOMIC ANALYSIS
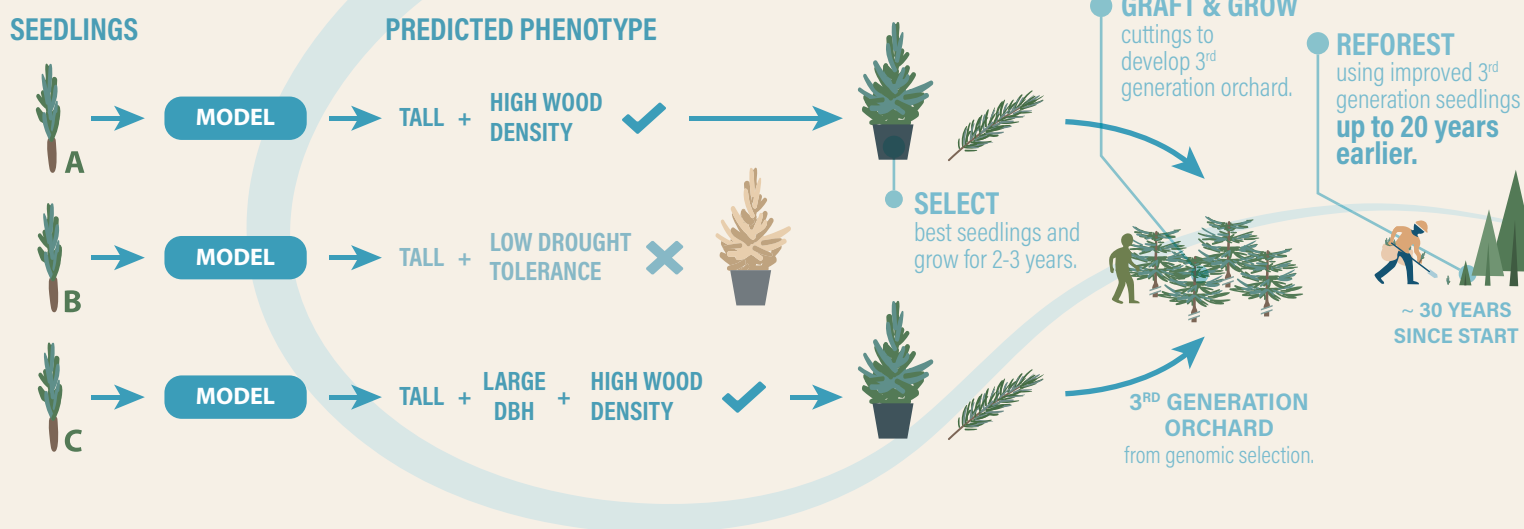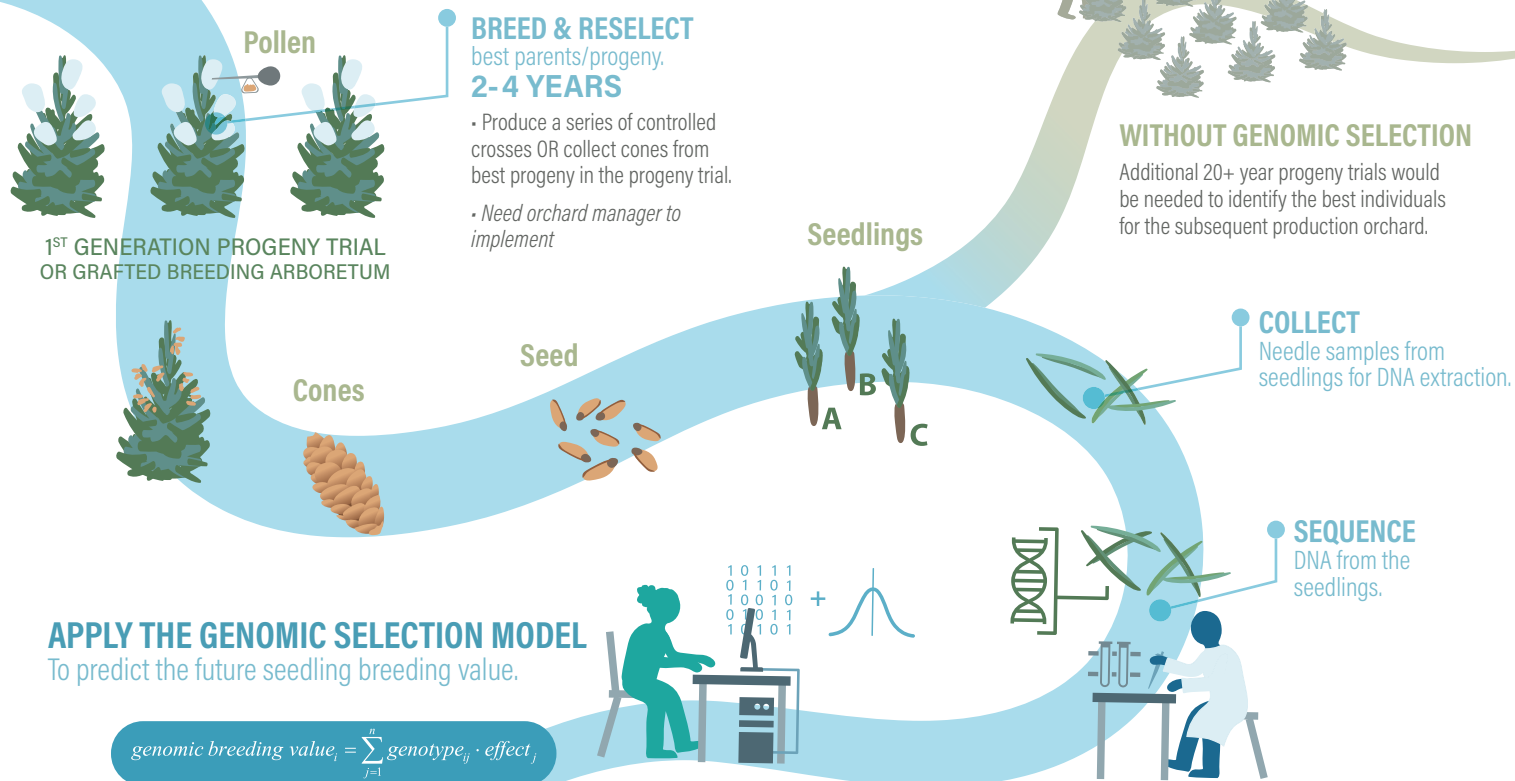
A   B   C
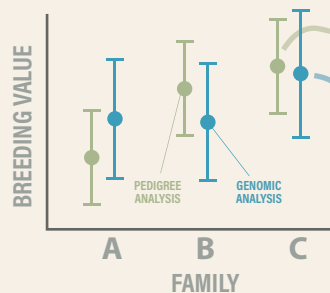
FAMILY

WITHOUT GENOMIC ANALYSIS

Individual breeding values of progeny within the family are imprecise.
- Progeny are ranked by pedigree

WITH GENOMIC ANALYSIS

Individual breeding values are predicted for each tree within the family with high precision.
- Progeny are ranked according to their predicted breeding value

Progeny selected via pedigree analysis.
May inadvertently pick trees with below average breeding values.

Progeny selected via genomic analysis.
Will result in better genetic gain.

USING THE GENOMIC SELECTION MODEL FOR FORWARD SELECTION

2ND GENERATION PROGENY TRIAL ~ 20 YEARS

Pollen

BREED & RESELECT best parents/progeny.
2-4 YEARS
• Produce a series of controlled crosses OR collect cones from best progeny in the progeny trial.
• *Need orchard manager to implement*

1ST GENERATION PROGENY TRIAL OR GRAFTED BREEDING ARBORETUM

Seedlings

Cones

Seed

WITHOUT GENOMIC SELECTION
Additional 20+ year progeny trials would be needed to identify the best individuals for the subsequent production orchard.

COLLECT Needle samples from seedlings for DNA extraction.

SEQUENCE DNA from the seedlings.

APPLY THE GENOMIC SELECTION MODEL
To predict the future seedling breeding value.

$$genomic\ breeding\ value_i = \sum_{j=1}^{n} genotype_{ij} \cdot effect_j$$

SEEDLINGS

PREDICTED PHENOTYPE

A → MODEL → TALL + **HIGH WOOD DENSITY** ✔

B → MODEL → TALL + **LOW DROUGHT TOLERANCE** ✘

C → MODEL → TALL + **LARGE DBH** + **HIGH WOOD DENSITY** ✔

SELECT best seedlings and grow for 2-3 years.

GRAFT & GROW cuttings to develop 3rd generation orchard.

REFOREST using improved 3rd generation seedlings **up to 20 years earlier.**

3RD GENERATION ORCHARD from genomic selection.

~ 30 YEARS SINCE START

GLOSSARY:

Base pair: complementary pairs of nucleic acids (A with T, G with C) which form the unique DNA sequence of an individual/tree.

Bioinformatician: a scientist that processes and analyzes large complex biological datasets into an interpretable format.

GBS (genotyping-by-sequencing): a simplified approach to DNA sequencing used when a reference genome sequence is unavailable.

Genomic breeding value: the expected phenotype of an individual, based on their DNA sequence.

Genetic Gain: the difference in the average target phenotype between unimproved and improved populations.

Genomic selection: a method that predicts associations between genetic markers and phenotypic traits of interest based on models.

Pedigree: a record of the genetic relationships between individuals in a population.

Quantitative geneticist: a scientist who studies the genetic basis of phenotypic variation among individuals in a population.

SNP chip: a DNA microarray that identifies the differences in the DNA sequence at many (100,000's) specific locations across a genome.

Target capture: sequencing method that focuses on specific genomic regions of interest.

To learn more about genomic selection, its potential applications in tree improvement, and the history of its use, check out these additional resources available on the RES-FOR website:

Genomic Selection in Forestry

Context of Genomic Technology

RES\FOR
https://resfor.ualberta.ca/