

Genomics-based selective breeding of forest trees

The selective breeding of forest trees is a long-term endeavor requiring sustained organizational and financial commitments, and follows the recurrent selection scheme with protracted repeated cycles of selection, mating and testing. This process encounters numerous challenges including long times to reach sexual maturity, fertility variation and reproductive asynchrony, mating hundreds of parents, testing hundreds of thousands of individuals in multiple locations over vast geographic territories, genotype by environment interaction, and late expression of economically important traits. Marker-Assisted Selection (MAS) has been considered a viable approach to simplify and speed this lengthy process; however, the complex nature of economic traits such as wood quality and tree size that are controlled by large numbers of genes, each with a small effect, has precluded the operational adoption of MAS. Recently, the relative affordability of high-throughput DNA sequencing, development of specialized and efficient quantitative genetics algorithms, and accessibility to high performance computational facilities, has resulted in their collective conversion and the emergence of a novel quantitative genomics approach known as Genomic Selection (GS).

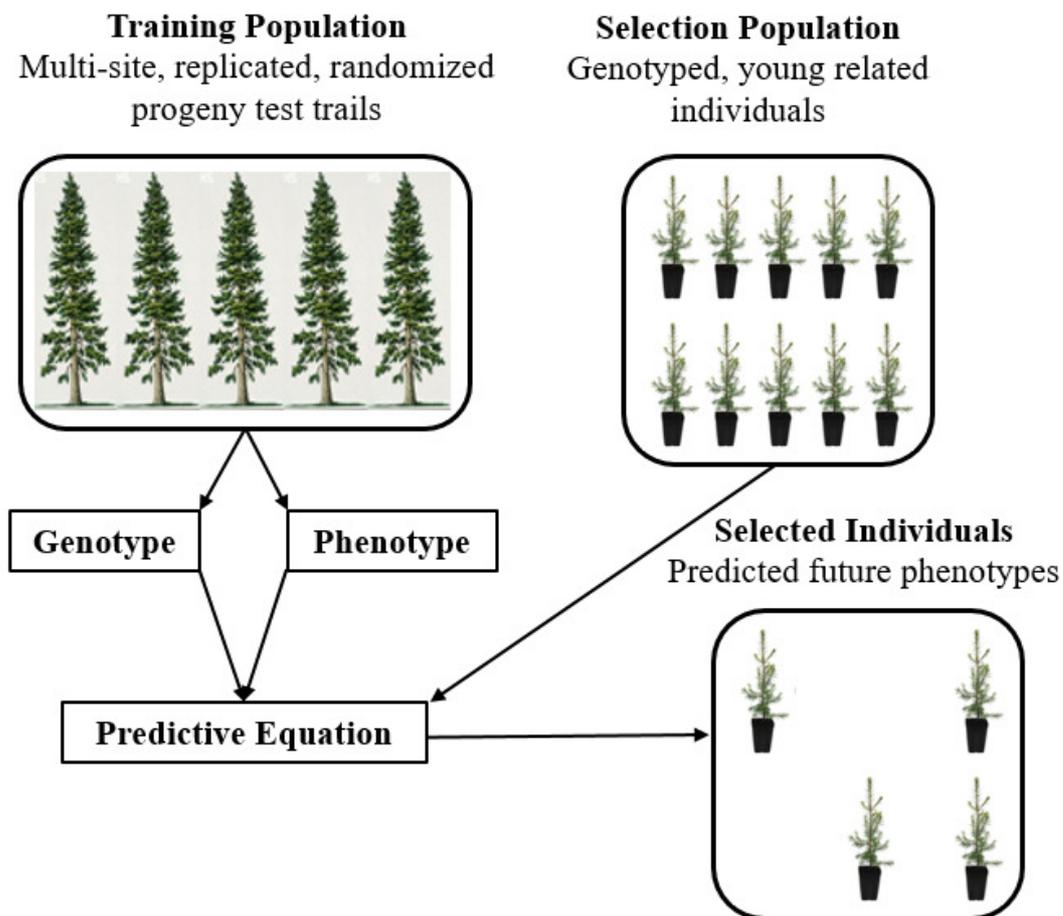


Fig. 1. Illustration of genomic selection steps.

This approach jointly considers complex-trait phenotypic data and vast numbers of DNA markers (Single Nucleotide Polymorphisms, SNPs) to predict the genetic worth of individuals by summing the collective effect of all SNPs irrespective of their significance level. For GS to work effectively, the genome needs to be saturated with DNA markers, so that the proximate distances between SNPs and the causal genes underlying a specific attribute are captured (a.k.a., linkage disequilibrium (LD), defined as the non-random segregation between SNPs and their adjacent causal genes). GS was first implemented in dairy cattle, over-turning traditional pedigree-based (line-of-descent) breeding in favor of genome-based (presence/absence of allele) breeding.

GS starts with a “training population” of individuals with known genotypes and phenotypes, which in turn, are used to develop predictive equations for their subsequent use in a “selection population” consisting of many genotyped individuals to predict their phenotypes. Desired individuals are subsequently selected, without further testing, for another breeding cycle (Fig. 1). In other words, the breeding paradigm has changed from classical phenotypic-dependent to genomic phenotypic-predicted. This drastic change can substantially shorten forest tree selective breeding cycles because the development of predictive models from older well-designed experiments can be used to predict the genetic worth for many younger genotyped, unphenotyped, and untested individuals.

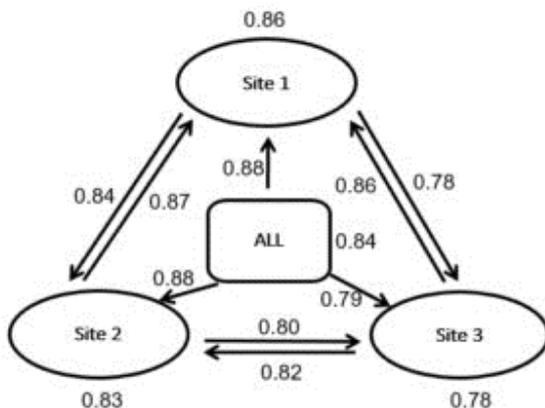


Fig. 2. GS predictive accuracies within- and across- site, and combined-sites. Arrows show predictive accuracy from one site to another.

For example, a GS predictive model developed for a difficult-to-assess attribute such as wood density of a 40-year-old population can be used to predict the anticipated future wood density, at age 40, of individual seedlings that are only a few weeks old. Additionally, the reduced time and cost of genotyping compared to traditional long-term testing permits the prediction of phenotypes for thousands of individuals, thus providing an added advantage by increasing the selection differential (the difference between the top selected individual means and the base population’s mean), resulting in additional gain.

Thistlethwaite et al. (2017) investigated the GS of wood density in a multi-site, replicated 38-year-old Douglas-fir (*Pseudotsuga menziesii*) training population in coastal British Columbia. Within- and across-site predictive accuracies were high and encouraging (Fig. 2). Further predictive accuracy improvement was also

demonstrated with combined-sites analysis to account for genotype by environment interaction (Fig. 2). These results, while preliminary in nature, are encouraging and warrant further study, specifically the validation and extension of the developed predictive models to independent and across-generational populations.

*Yousry A El-Kassaby, Frances R. Thistlethwaite
Faculty of Forestry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada*

Publication

[Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform.](#)

Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, El-Kassaby YA
BMC Genomics. 2017 Dec 2