# IRC1/Project 2.3 - Fall 2021:Updated

# SNP-BASED ANALYSIS FOR THE REGION G1 PROGRAM

## Overview (2019-2021)

Accurate monitoring of genetic diversity levels and mating patterns in seed production populations are crucial to ensure that tree breeding programs are long-lasting and will deliver anticipated genetic gains. We used a single nucleotide polymorphism (SNP) genotyping method to characterize a founder population, seed orchard seedlots, and trees from progeny trials to assess pollen contamination. The impact of roguing 65% of the genotypes on effective population size and parental contributions in the G1 first-generation open-pollinated white spruce seed orchard was also studied. The cutting-edge technologies used in this work appear promising to help advance Alberta's tree breeding programs.

## Methodology

The following plant material was genotyped using needle tissue: (1) a random sample of 105 open-pollinated bulk seeds from five different commercial seedlots collected from the G1 orchard (525 seedlings), (2) progeny from the top 10 families at two different G365 progeny trial sites (328 trees), and (3) the founder population (including all original parent tree selections) at the G218 clone bank (151 trees). An Infinium iSelect SNP array with 2,000 SNPs was used to reconstruct the pedigree, estimate effective population size ($N_e$) and calculate pollen contamination levels for each population. Pollen contamination levels were compared using two methods (SNPs vs pollen traps) and with wind direction during the days with the highest rate of pollen flow and conelet receptivity. Weather data for 2003-2018 and for 2019-2020 were obtained from www.climate.weather.gc.ca and www.acis.alberta.ca, respectively.

## Results

Backward selection was used to rogue 65% of the genotypes in the G1 white spruce orchard population which resulted in doubling the gain in height (from 2.5% to 5%), while increasing the coancestry (relatedness) 8-fold and reducing the $N_e$ 8-fold in the 2018 seedlot (Table 1, purple squares). These results suggest that fewer trees and less pollen diversity remained in the orchard. Pedigree reconstruction showed unequal parental contributions across years with pollen contamination levels ranging between 12-51% (average 27%) among seedlots (Table 1), and 7-68% (average 30%) among individual genotypes within a seedlot (Fig. 1).

**Table 1.** Summary table with coancestry coefficient ($\Theta$), parental contributions, effective population size (*Ne*), and pollen contamination levels across populations in the G1 white spruce seed orchard, using tree genomic profiles with a sub-set of 2,000 SNPs. *N* is the actual population size. Wind direction in degrees and number of days with a west wind direction out of 15 days of assessment for each year is also shown.

| Parameter | Population | | | | | | |
|---|---|---|---|---|---|---|---|
| | Founders | 2003 seedlot | 2005 seedlot | 2007 seedlot | 2009 seedlot | 2018 seedlot | Progeny trials |
| Coancestry ($\theta$) | 0.001 | 0.003 | 0.008 | 0.003 | 0.005 | 0.04 | 0.016 |
| Parental contribution (both sexes) | n/a | 28% | 25% | 35% | 40% | 57% | n/a |
| *Ne* (using number of cones*) | n/a | 48.6 | 35.3 | 51.4 | 46 | 18 | n/a |
| ***Ne* (using SNPs)** | **500** | **166** | **59** | **158** | **96** | **12** | **31** |
| Pollen contamination (using traps*) | n/a | 100% | 23% | 11% | 10% | n/a | n/a |
| **Pollen contamination (using SNPs)** | n/a | **51%** | **26%** | **28%** | **12%** | **18%** | n/a |
| Wind direction (average) | n/a | 236° | 201° | 223° | 191° | 201° | n/a |
| No. days with west wind direction | n/a | 11 | 6 | 7 | 8 | 6 | n/a |

*These values were calculated following FGRMS (2016).

**Relatedness levels for progeny trials were calculated among families.

**Calculated using the weather database from 'Grande Prairie A' station (https://climate.weather.gc.ca/)

The two methods used to calculate the effective population size showed a Pearson's correlation of 0.98 (Fig. 2a). Levels of pollen contamination showed a Pearson's correlation of 0.92 with wind direction, likely from a pollen source 1 km away from the orchard under study (Fig. 2b).

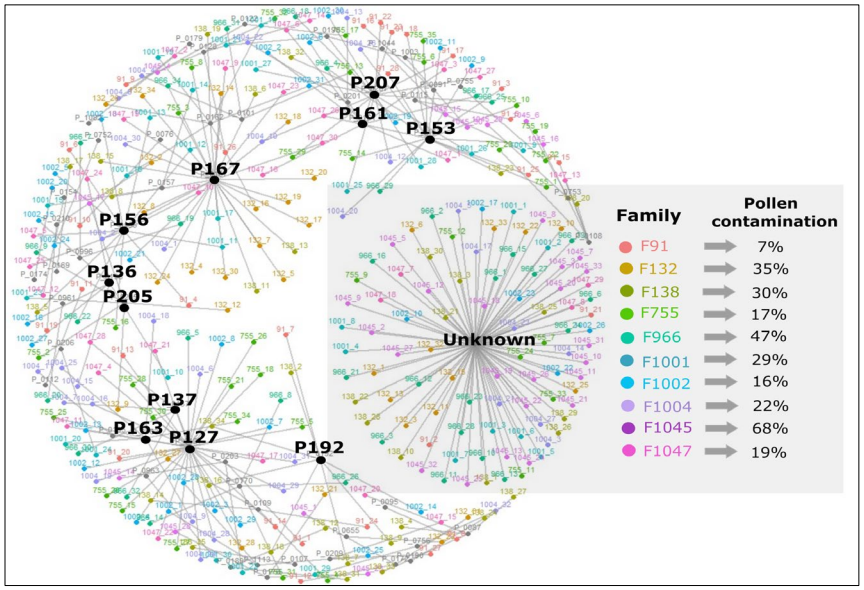


**Figure 1.** Pedigree reconstruction of 10 families (328 trees) from a 15-year-old progeny trial (two sites) associated with the Region G1 white spruce tree improvement program. Known mothers (coloured dots), and fathers (P) could be assigned from the breeding program using genomic profiles. Black dots represent the 11 fathers with the highest contributions. The diagram inside the grey square represents the offspring pollinated by fathers from outside the seed orchard, and the table details the percent pollen contamination identified in each family (F). 'Unknown' indicates all unknown contaminating pollen sources combined.

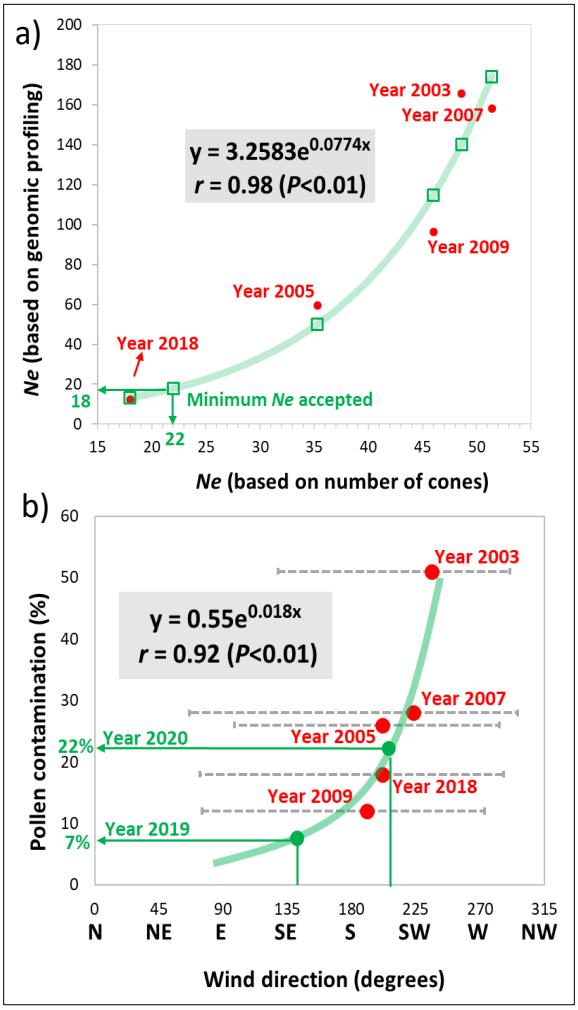


**Figure 2.** Correlation plots (Pearson's) showing exponential trendlines (green lines), equations, R-squared values and *P* values (gray squares) for different traits. (a) Effective population size ($N_e$) calculated using the number of cones per genotype versus genomic profiles; (b) Pollen contamination levels (%) estimated using SNP profiling versus wind direction average between 15-31 May of each year. Red dots show data used for fitting the equations. Green dots and squares show predicted values using the equations. Grey dotted lines indicate all values, including maximum and minimum wind speed values. The "Minimum $N_e$ accepted" corresponds to the value required in the Alberta forestry policy (FGRMS, 2016), which is $N_e$=18.

## Conclusion

The achievement of 5% genetic gain in height at rotation through eliminating two thirds of the orchard, generated a loss in genetic diversity as determined by the reduction in effective population size in the seedlot following roguing. Use of genomic profiles revealed the considerable impact of roguing on genetic diversity, and pedigree reconstruction of full-sib families showed the unanticipated impact of pollen contamination from a previously unconsidered source. This work highlights the benefits of incorporating genomic tools to assess genetic diversity, estimate pollen contamination levels and reconstruct pedigrees to advance the breeding cycles as quickly and efficiently as possible to realize the economic benefits. We recommend that these technologies be considered for inclusion in the next revision of the Alberta Forest Genetic Resource Management and Conservation Standards.

***Note: This work is published in Nature Scientific Reports (2021) 11:4990***
(https://www.nature.com/articles/s41598-021-84566-2)

*Please contact Barb Thomas (bthomas@ualberta.ca), or Esteban Galeano (galeanog@ualberta.ca) for more information.*

NSERC CRSNG

TREE IMPROVEMENT LAB

UNIVERSITY OF ALBERTA
QUAECUMQUE VERA