

Supplementary Material

Table S1. Genetic diversity parameters across seed orchard seedlots in the Alberta Region G1 white spruce seed orchard, using tree genomic profiles with a set of 2000 SNPs. N_e =effective population size estimated with four different methods (Ritland 1996; Waples 2006; Nomura 2008; FGRMS 2016; see Materials and Methods). Pollen contamination was estimated using SNPs and traps (see Materials and Methods). Numbers and genomic profiles corresponding to Region G1 were obtained from Galeano et al. (2021) [1].

Parameter	2003	2005	2007	2009	2018
N_e (cones)	48.6	35.3	51.4	46	18
N_e (Ritland)	166	59	158	96	12
N_e (Nomura)	58.5	19.5	27.6	23.7	16.5
N_e (Waples)	70.9	42.6	68.8	59.2	22.4
% Pollen contamination (SNPs)	51%	26%	28%	12%	18%
% Pollen contamination (traps)	100%	23%	11%	10%	16%

Table S2. List of individual trees that were unrelated to any other genotyped tree in the progeny trial and were not part of any family (unrelated to any mother from the program). These errors were diagnosed using the relationship coefficients from the original G -matrix following their expected values from the pedigree-based A -matrix (Muñoz et al. 2014). These trees were not used to obtain the corrected G -matrix (for GBLUP) or H -matrix (for ssGBLUP).

ID (progeny trial)	ID (progeny trial)
3772	7302
2918	8741
8971	422
4102	6581
7953	7359
6561	2933
9036	3308
6452	6644
4090	3805
670	3087
4640	9675
7364	9004
7351	--
Total	25

Table S3. List of individual trees that had a wrong family assignment. The diagnosis of these errors was made using the relationship coefficients from the original G -matrix following their expected values from the pedigree-based A -matrix. These trees were identified by searching across all relationships in the dataset for the mother where these individuals match the expected value, and once the new (true) mother was identified, the individuals were relabeled with the corrected pedigree. The new assignments were used to obtain the corrected G -matrix (for GBLUP), A -matrix (for ABLUP), and H -matrix (for ssGBLUP). PP=Phantom parent.

Wrong Mother	True Mother	ID (progeny trial)
XX01005	XX01251	4987
XX01251	XX01005	8272
XX01019	XX00752	7365
GoA2549	XX00754	2284
XX01850	GoA317	1637
XX00953	XX00955	9610, 509, 2085, 2492, 740, 639, 9943, 9902, 8199, 8995
XX01456	XX01126	7934, 890, 1261, 1652, 1607, 4769, 8291, 9899, 8216, 5632
XX01254	XX01252	5412, 2138, 2865, 2960, 2439, 3958, 5821, 5737, 5680, 8132

XX01427	XX01428	7087, 1340, 1298, 1615, 1573, 1473, 7474, 5756, 9861
XX01085	PP1	5459, 4655, 2129, 2448, 4914, 2304, 5774, 5743, 5672, 5642
XX01287	PP2	3347, 3726, 6680, 7049, 8213, 5239
XX01457	PP3	7109, 2174, 2126, 2451, 2403, 3961, 7452, 8219, 8198, 8977
XX01464	PP4	6259, 3775, 4130, 4929, 3941, 6638, 6560, 7357, 8157
XX01846	PP5	6261, 1311, 2941, 3300, 1602, 2319, 6636, 8258, 6526, 8981
XX01850	PP6	9552, 2981, 1276, 3260, 3979, 9585, 7393, 6510, 8972, 1639
XX01240	PP7	1351, 1562
XX01468	PP8	7064, 2968, 2110, 2467, 3273, 627, 7480, 8203, 9792
Total		111

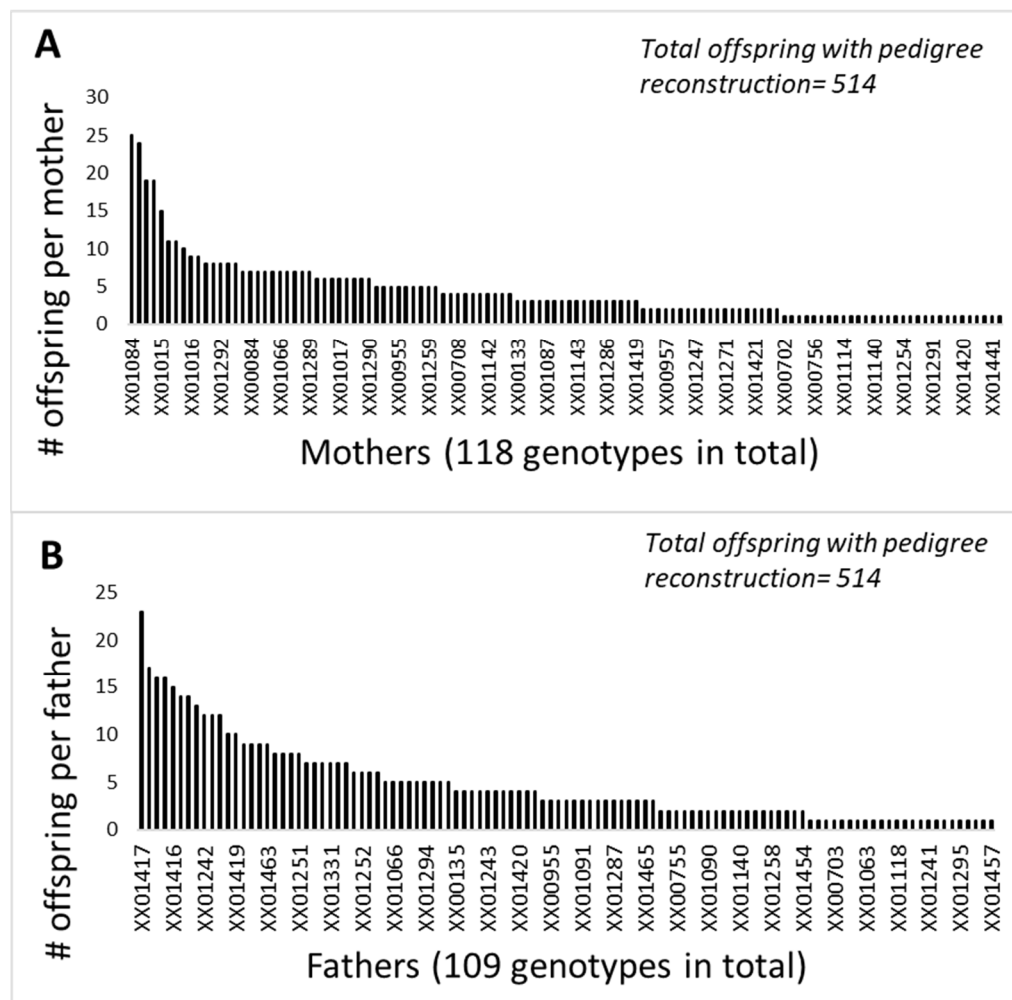


Figure S1. Maternal and paternal contributions in the Alberta Region I white spruce seed orchard based on six years of assessment. Diagrams were based on 514 seedlings and 166 parents for the years 2007, 2009, 2010, 2011, 2013, 2015. **(A)** Number of offspring per mother. **(B)** Number of offspring per father. Parents were inferred by the software CERVUS.

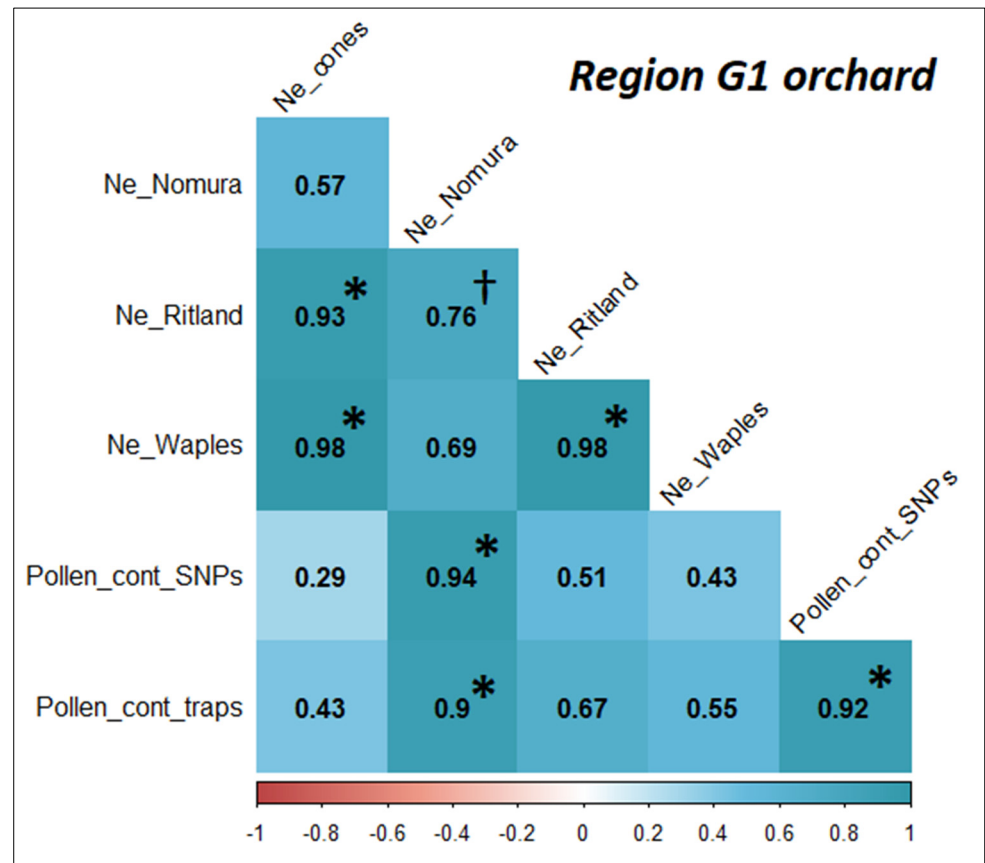


Figure S2. Pearson's correlation matrices for the different pollen contamination assessments and effective population size (N_e) methods in the Alberta Region G1 white spruce seed orchard. Pollen contamination was estimated using traps (monitors) and using SNPs (genomic profiles) (see Materials and Methods). The N_e methods are Nomura, Ritland, Waples, and cones (Ritland 1996; Waples 2006; Nomura 2008; FGRMS 2016; see Materials and Methods). Asterisk (*) indicates statistically significant values at $p < 0.01$ and cross (†) indicates statistically significant values at $p < 0.05$. Numbers and genomic profiles corresponding to Region G1 were obtained from Galeano et al. (2021) [1].

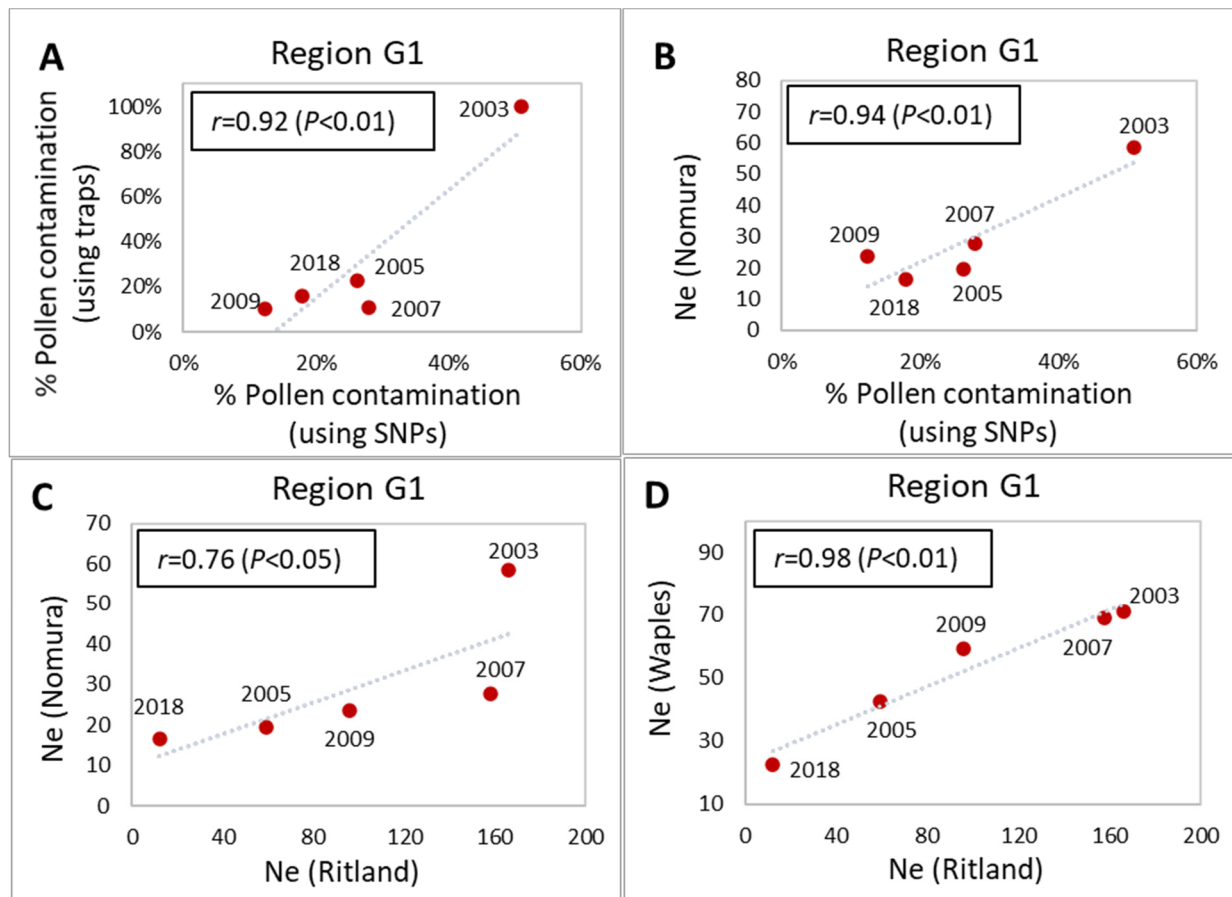


Figure S3. Scatterplots showing linear trendlines, Pearson's correlation, and p values for pollen contamination assessments and N_e using genomic profiles in the Alberta Region G1 white spruce seed orchard. (A) Pollen contamination using SNPs vs. using traps in Region G1, (B) Pollen contamination using SNPs vs. N_e (Nomura) in Region G1, (C) N_e (Ritland) vs. N_e (Nomura) in Region G1, (D) N_e (Ritland) vs. N_e (Waples) in Region G1 (Ritland 1996; Waples 2006; Nomura 2008; see Materials and Methods). The different seedlot years (red dots) are denoted inside the graphs for Region G1. Numbers and genomic profiles corresponding to Region G1 were obtained from Galeano et al. (2021) [1].

References

1. Galeano, E.; Bousquet, J.; Thomas, B.R. SNP-based Analysis Reveals Unexpected Features of Genetic Diversity, Parental Contributions and Pollen Contamination in a White Spruce Breeding Program. *Sci. Rep.* **2021**, *11*, 4990. <https://doi.org/10.1038/s41598-021-84566-2>.