

Integrating environmental and genetic effects to predict responses of tree populations to climate

TONGLI WANG,^{1,3} GREGORY A. O'NEILL,² AND SALLY N. AITKEN¹

¹Centre for Forest Conservation Genetics and Department of Forest Sciences, University of British Columbia, 3041-2424 Main Mall, Vancouver, British Columbia, Canada V6T 1Z4

²Research Branch, British Columbia Ministry of Forests and Range, Kalamalka Forestry Centre, 3401 Reservoir Road, Vernon, British Columbia, Canada V1B 2C7, and
Ecosystem Science and Management, University of Northern British Columbia, 3333 University Way, Prince George, British Columbia, Canada V2N 4Z9

Abstract. Climate is a major environmental factor affecting the phenotype of trees and is also a critical agent of natural selection that has molded among-population genetic variation. Population response functions describe the environmental effect of planting site climates on the performance of a single population, whereas transfer functions describe among-population genetic variation molded by natural selection for climate. Although these approaches are widely used to predict the responses of trees to climate change, both have limitations. We present a novel approach that integrates both genetic and environmental effects into a single “universal response function” (URF) to better predict the influence of climate on phenotypes. Using a large lodgepole pine (*Pinus contorta* Dougl. ex Loud.) field transplant experiment composed of 140 populations planted on 62 sites to demonstrate the methodology, we show that the URF makes full use of data from provenance trials to: (1) improve predictions of climate change impacts on phenotypes; (2) reduce the size and cost of future provenance trials without compromising predictive power; (3) more fully exploit existing, less comprehensive provenance tests; (4) quantify and compare environmental and genetic effects of climate on population performance; and (5) predict the performance of any population growing in any climate. Finally, we discuss how the last attribute allows the URF to be used as a mechanistic model to predict population and species ranges for the future and to guide assisted migration of seed for reforestation, restoration, or afforestation and genetic conservation in a changing climate.

Key words: adaptation; assisted migration; British Columbia, Canada; ClimateBC; climate change; provenance test; response function; sample size; species range; transfer function.

INTRODUCTION

Phenotypes are determined both by genotypes at loci that underlie quantitative traits and by the environmental conditions under which individuals live. Climate has long been recognized as playing a key role in both these determinants of phenotype in plants: in the short term, through direct effects on phenotypes via environmental effects on survival, growth, and reproduction and in the longer term, through acting as a major selection force affecting plant genotypes. Climate-related natural selection leads to local adaptation and differentiation of tree populations along climatic gradients (Epperson 2003, Howe et al. 2003). Thus, tree populations typically show moderate to strong local adaptation despite high levels of gene flow (Howe et al. 2003, Savolainen et al. 2007). Recent climate change, however, is altering environmental conditions rapidly, and locally adapted populations are likely to experience climates to which they are

not well adapted. Biological responses to climate change have already been widely reported (Chuine et al. 2004, Parmesan 2006, IPCC 2007). Strategies for management and conservation of forest genetic resources and, specifically, deployment of seed sources to long-term forest plantations in a rapidly changing climate therefore need to be modified to ensure plantations are adapted to future climates (Rehfeldt et al. 2006, Ying and Yanchuk 2006, Millar et al. 2007, Aitken et al. 2008). This requires a good understanding of both the short-term environmental (among test site) and long-term genetic (among population) effects of climate on phenotypes and the quantification of these effects to predict the impact of climate change on adaptation and growth of forest trees. Furthermore, if both environmental and genetic effects can be quantified simultaneously and compared, we will have a better understanding of plant–climate relationships and improve our ability to predict the effects of climate change on populations.

Environmental and genetic effects of climate on phenotypes have previously been used to develop population transfer functions and population response

Manuscript received 5 December 2008; revised 9 April 2009; accepted 12 May 2009. Corresponding Editor: D. McKenzie.

³ E-mail: tongli.wang@ubc.ca.

functions using provenance test data for climate change studies. Provenance tests were initially established to identify superior seed sources for reforestation (Morgenstern 1996). Such tests usually comprise many populations collected from a range of provenances and planted in common garden experiments at several or many locations. We use the term “provenance” to describe the geographic and climatic origin of individuals and “population” to describe individuals originating from that location. For climate change studies, spatial climatic variation among test sites is substituted for longer-term temporal trends.

A transfer function relates the performance of multiple populations to the climatic transfer distances between their provenances and a particular planting site (i.e., site climate minus provenance climate) (Matyas 1994, Carter 1996, Rehfeldt et al. 1999, Andalo et al. 2005). It reflects the level of local adaptation and the effect of transferring populations to different locations on phenotypes. In most studies, data from multiple test sites are pooled to generate a generalized transfer function. By pooling data from multiple sites, it is assumed that the shape and position of individual transfer functions are the same across different environments (Carter 1996, Rehfeldt et al. 1999, Andalo et al. 2005). This assumption is often not valid, as transfer effects expressed in different environments are known to differ considerably (Rehfeldt et al. 2003, Wang et al. 2006b, O'Neill et al. 2008). For example, a transfer from mean annual temperature (MAT) 10° to 8°C (i.e., -2°C transfer) may have a dramatically different effect on phenotypes than a transfer from 0° to -2°C (also a -2°C transfer).

A further complication associated with the pooling of data in generalized transfer functions is that response data must be standardized to account for large differences among sites in growth. Standardizing response data is usually accomplished by expressing population responses as a percentage of test site mean response. Standardizing the response variable precludes comparison of absolute responses in different climates, although absolute responses of tree growth to climate change can be obtained when generalized transfer functions are used in conjunction with biophysical site index models (Beaulieu and Rainville 2005).

Recently, individual transfer functions have been developed using provenance climate, instead of provenance transfer distance, as the independent variable (Wang et al. 2006b). This relationship describes the among-population genetic variation resulting from natural selection for climate at each test site. It can be used to identify the climatic conditions that have produced the most productive populations for each site. By using original values of climatic variables of each provenance instead of climatic transfer distances, individual transfer functions can better represent the among-population variation associated with climatic

variables. Therefore, a considerable proportion of the total seed transfer effect found with individual transfer functions in lodgepole pine (Wang et al. 2006b) was not revealed through general transfer functions (Rehfeldt et al. 1999). The disadvantage of individual transfer functions is that they are test site specific: an individual transfer function is valid only for the planting site for which it was developed. This limits its use in climate change studies. O'Neill et al. (2008) overcame this hurdle by associating the parameters of individual transfer functions with climatic variables of planting sites. By generalizing individual transfer functions they produced a “universal transfer function,” resolving the site-specific issue, but as the procedure is based on individual transfer functions, it does not use data from all sites simultaneously. Consequently, predictive power is compromised.

In contrast to a transfer function, a population response function describes the relationship between the performance of a single population across test sites and the climates of those test sites. This relationship reflects the environmental effects of site climate on the performance of a particular population (i.e., genotype). Response functions have been used to assess the impact of climate change on growth and adaptation and to select populations for future climates (Rehfeldt et al. 1999, 2001, Wang et al. 2006b). Response functions provide a straightforward quantitative approach to predict the impact of climate change on a population. However, response functions are population specific; each response function is valid only for the specific population for which it is developed. This limitation restricts the ability to make predictions for other populations and exploring new seed sources for future climates is becoming particularly important. In addition, most older provenance trials were not designed for such objectives and therefore often present critical limitations. The number of populations and sites and their climatic range are often insufficient to develop accurate and reliable response functions.

The main objective of this study was to develop a novel approach that integrates both environmental effects of test site climate on phenotypes (accounted for by response functions) and among-population differentiation resulting from local adaptation to climate (accounted for by transfer functions) into a single “universal response function” (URF). This model is expected to: (1) improve the accuracy of predictions; (2) be universal, in that it can predict an adaptive response trait for any population growing in any climate; (3) reduce the necessary sample size in terms of sites and populations without reducing predictive power; and (4) facilitate comparison of the magnitude of environmental and genetic effects of climatic variables on the performance of populations. We use the Illingworth lodgepole pine (*Pinus contorta*) provenance trial, one of the most comprehensive provenance trials in the world, to test and demonstrate this model.

MATERIALS AND METHODS

Universal response function

A universal response function can be viewed as a combination of individual response functions and individual transfer functions. A typical quadratic response function can be written as

$$Y_i = b_0 + b_1X_{1i} + b_2X_{1i}^2 + e_i \quad (1)$$

where Y_i is the observation of a population at the test site i ; b 's are the intercept and regression coefficients; X_{1i} is one or more climate variables for the test site i ; and e_i is the residual.

An individual transfer function can be written as

$$Y_j = b_0 + b_1X_{2j} + b_2X_{2j}^2 + e_j \quad (2)$$

where X_{2j} is one or more climate variables for the provenance j .

Transfer functions and population response functions can be combined into a universal response function (URF), written as

$$Y_{ij} = b_0 + b_1X_{1i} + b_2X_{1i}^2 + b_3X_{2j} + b_4X_{2j}^2 + b_5X_{1i}X_{2j} + e_{ij} \quad (3)$$

where Y_{ij} is the observation of the population j at the site i ; b 's are the intercept and regression coefficients; X_{1i} and X_{2j} are as defined following Eq. 1; $X_{1i}X_{2j}$ is the interaction(s) between X_{1i} and X_{2j} ; and e_{ij} is the residual. Quadratic functions were used in this study for consistency with previous studies (Rehfeldt et al. 1999, 2001, Wang et al. 2006b).

Using climatic variables of both test sites and provenances as independent variables, both environmental effects of site climate and genetic effects (among-population variation) of provenance climate on phenotypic performance are integrated into a single model. However, among-population variation can also be considerably affected by other factors such as gene flow, demographic history, and colonization events (de Jong 1988, Epperson 2003), resulting in some spatial patterns that are independent of climate. These spatial genetic effects may be explained to some extent by geographical proximity of provenance, assuming population proximity is correlated with relatedness and shared evolutionary history, and this can also be incorporated into a universal response function. Therefore, the model can be extended to

$$Y_{ij} = b_0 + b_1X_{1i} + b_2X_{1i}^2 + b_3X_{2j} + b_4X_{2j}^2 + b_5X_{1i}X_{2j} + b_6X_{3j} + b_7X_{3j}^2 + e_{ij} \quad (4)$$

where X_{3j} is one or more geographic variables for the provenance j .

Geographic variables for test sites cannot be included in the model because the URF will be used to identify planting sites based on climatic conditions rather than geographical locations as climates change. In contrast, a

provenance can be described by either climatic or geographic variables or the combination of both. This is because among-population variation will largely remain unchanged in the near future even in a changing climate.

Plant materials

Established in 1974, the lodgepole pine (*Pinus contorta*) provenance test analyzed in this study consists of a range-wide collection of 140 populations planted at 60 test sites in British Columbia, Canada, and two sites in the southern Yukon, Canada (Illingworth 1978). The majority (107) of the populations were from subsp. *latifolia* and the rest from subsp. *contorta* and subsp. *murrayana*. An incomplete design tested 60 of the 140 populations at each site, with each population being tested at 30–40 of the sites. Within each site, a randomized complete block design was used, with one nine-tree square plot of each provenance planted in each of two blocks at 3×3 m spacing. The two Yukon sites contain six blocks, each comprising 40 provenances. In total, 69 120 trees were sown in 1971 and transplanted to the sites in the spring of 1974. Total height (HT) was measured at 60 extant sites in the fall of 1993, after 20 growing seasons in the field. Considerable differences in growth performance were observed among the populations at each planting site (see Plate 1).

Climatic and geographic variables

Values of the following 10 climatic variables were estimated for all test sites and provenances using ClimateBC version 3.2 (Wang et al. 2006a) and ClimateWNA (an extended version of ClimateBC; T. Wang, A. Hamann, and D. L. Spittlehouse, unpublished program): mean annual temperature (MAT), mean warmest month temperature (MWMT), mean coldest month temperature (MCMT), continentality (TD, the summer–winter temperature difference, MWMT – MCMT), mean annual precipitation (MAP), mean summer precipitation (MSP), growing degree-days (DD5), and frost-free period (FFP), as well as two moisture indices: log-transformed annual heat-moisture index (LAHM) and log-transformed summer heat-moisture index (LSHM). The extension “_s” or “_p” was appended to each variable to denote test sites or provenances, respectively. Details regarding the estimation of all derived climatic variables can be found in Wang et al. (2006a).

Independent variables used to develop the URF include 53 climatic and 15 geographic variables. The climatic variables included both linear and quadratic forms of each of the 10 climatic variables for both test site and provenance (40 variables), as well as 13 interactions. These interactions included an interaction between test site and provenance for each of the 10 variables at the linear level and three interactions at higher levels for MAT only (i.e., $\text{MAT}_s \times \text{MAT}_p$,

TABLE 1. Twenty-four scenarios created by various intensities of subsampling of populations, sites, or both, from the full sample provided by the Illingworth (1978) lodgepole pine provenance test used in this experiment.

Scenario	No. populations	No. test sites	Percentage of full data set sampled
1	14	60	10
2	21	60	15
3	28	60	20
4	35	60	25
5	49	60	35
6	70	60	50
7	105	60	75
8	140	60	100
9	140	6	10
10	140	9	15
11	140	12	20
12	140	15	25
13	140	21	35
14	140	30	50
15	140	45	75
16	140	60	100
17	14	6	1
18	21	9	2.25
19	28	12	4
20	35	15	6.25
21	49	21	12.25
22	70	30	25
23	105	45	56.25
24	140	60	100

Note: The lodgepole pine (*Pinus contorta*) provenance test analyzed in this study consists of a range-wide collection of 140 populations planted at 60 test sites in British Columbia, Canada, and two sites in the southern Yukon, Canada.

$\text{MAT}_s \times \text{MAT}_p^2$ and $\text{MAT}_s^2 \times \text{MAT}_p^2$). The 15 geographic variables for provenances included latitude (LAT), longitude (LONG), elevation (EL), and their transformations and interactions (LAT^2 , LONG^2 , EL^2 , $\text{LAT} \times \text{LONG}$, $\text{LAT} \times \text{EL}$, $\text{LONG} \times \text{EL}$, $(\text{LAT} \times \text{LONG})^2$, $(\text{LAT} \times \text{EL})^2$, $(\text{LONG} \times \text{EL})^2$, $1/\text{LAT} \times \text{LONG}$, $(1/\text{LAT} \times \text{LONG})^2$ and $1/\text{LAT} \times \text{EL}$).

In order to create maps of predicted performances of populations and to locate provenances producing fast-growing populations for each planting site across British Columbia (BC), a 400-m digital elevation model of BC was systematically subsampled by selecting every 16th point to generate a fine-scale (1.6×1.6 km) grid of 370 205 point locations. For each of these point locations ClimateBC was used to generate climate variables for three 30-year climate normal periods: 1961–1990 (reference period), 2041–2070, and 2071–2100 (future periods). These periods are hereafter referred as: 1970s, 2050s, and 2080s, respectively. We chose the A2 “enterprise” scenario implemented through the Canadian Global Circulation Model of the Canadian Centre for Climate Modeling and Analysis, scenario CGCM2 (Flato et al. 2000) to illustrate the use of the URF; for operational applications, predictions from multiple emissions scenarios and global circulation models should be evaluated to span the range of possible future conditions.

Statistical analysis

All statistical analyses were conducted using SAS (SAS Institute 1989). The mean height of trees from each population at each site was calculated, resulting in 2476 provenance-by-site means in the “full data set.” A stepwise procedure of multiple regression used the 68 climatic and geographic variables to develop a universal response function for 20-year height (HT), with significance levels for entry into and staying in the model of $\alpha = 0.01$ and 0.005, respectively. Height was selected as the response variable because it is the most heritable of growth traits for lodgepole pine (Wang et al. 1999), and population differences strongly reflect the genetic effect of climate (Rehfeldt et al. 1999, Ying and Yanchuk 2006). However, the same approach can be applied to develop URFs for other traits.

Due to the unbalanced experimental design and initial objectives of the trial, a large number of populations were tested only across a narrow range of climatic conditions; thus their performance could not be reasonably related to climatic conditions (Wang et al. 2006b). Therefore, anchor points were developed for these populations at the five coldest test sites using the same approach we used for developing conventional response functions in a previous study (Wang et al. 2006b). An individual transfer function was first developed for each of the five sites and the performance of the populations absent at these sites was then estimated using these functions.

Effects of numbers of populations and test sites on the accuracy of the URF were assessed by varying the intensity of subsampling the full data set to create 24 subsample scenarios (Table 1). For each subsample scenario, a URF was developed with the independent variables selected by the full model and used to predict each observation in the full sample. This process was bootstrapped 10 times for each scenario, and a standard error of prediction was calculated for each repetition. The mean and range of the 10 estimates of the standard error of predictions were used to assess the effects of numbers of both populations and sites on URF accuracy for each scenario. Results of modeling with the bootstrapped subset samples can also serve as validations of the model.

RESULTS

Universal response function

A universal response function for 20-year height of lodgepole pine was built through multiple regression procedures by selection of 13 of the total of 68 independent climatic and geographic variables. The selected variables include four climatic variables for site (MAT_s , MAT_s^2 , LAHM_s , LAHM_s^2), four climatic variables for provenance (MAT_p , MAT_p^2 , LAHM_p , LAHM_p^2), three interactions between site and provenance climatic variables ($\text{MAT}_s \times \text{MAT}_p$, $\text{MAT}_s^2 \times \text{MAT}_p$ and $\text{MAT}_s^2 \times \text{MAT}_p^2$), and two geographic

TABLE 2. Results of multiple regression analysis predicting lodgepole pine height from site and provenance climate and provenance geography in a universal response function.

Independent variable	Parameter estimate	Partial R^2	Model R^2	C(p)	F	P
Intercept	-20.071					
MAT_s	2.006	0.445	0.445	8400	3564.6	<0.0001
MAT_p ²	-0.039	0.195	0.640	3874	2421.0	<0.0001
MAT_s ²	-0.294	0.064	0.704	2411	950.5	<0.0001
LAHM_s ²	-2.995	0.028	0.732	1751	475.5	<0.0001
LAHM_s	15.792	0.026	0.758	1149	480.9	<0.0001
MAT_p	0.223	0.016	0.774	786	310.3	<0.0001
SWLNG_p	0.021	0.015	0.789	441	316.6	<0.0001
LAHM_p	2.061	0.006	0.795	300	134.6	<0.0001
MAT_s $\times$ MAT_p	-0.077	0.004	0.799	219	79.0	<0.0001
MAT_s ² $\times$ MAT_p	0.031	0.008	0.807	47	172.7	<0.0001
LAHM_p ²	-0.294	0.001	0.808	36	11.9	0.0006
SWLNG_p ²	-5.8E-05	0.001	0.809	22	16.0	<0.0001
MAT_s ² $\times$ MAT_p ²	-9.3E-04	0.001	0.810	14	10.8	0.0010

Notes: Abbreviations are: MAT, mean annual temperature; LAHM, log-transformed annual heat-moisture index; SWLNG, the geographic variable ($1/\text{latitude} \times \text{longitude}^2$). The extension “s” or “p” was appended to each variable to denote test sites or provenances, respectively. C(p) is Mallows’s Cp statistic.

variables for provenance (SWLNG_p, calculated as $1/\text{LAT} \times \text{LONG}^2$, and SWLNG_p²). The model explained 81% of the total variation in 20-year height (Table 2).

Among the 13 independent variables in the model, site MAT (MAT_s and MAT_s²), provenance MAT (MAT_p and MAT_p²), and their three interactions accounted for the vast proportion of the variation (51%, 21%, and 1%, respectively, of the sums of the partial R^2 values from their linear and quadratic forms in Table 2). In order to illustrate the universal response function graphically, only two independent variables and their transformations and interactions can be displayed in a 3-D graph. Thus, a simplified function was developed

using only MAT_s and MAT_p and their transformations and interactions listed in Table 2 to predict HT (Fig. 1). The response surface shown in Fig. 1 describes the majority of phenotypic variation among populations in growth response in this study (cumulative $R^2 = 0.73$) and thus generally represents the trends in the data well, although a small group of outliers can be observed in the upper right portion of the distribution.

Effects of sample size on the URF

The number of populations and test sites showed strong positive effects on the accuracy of URF predictions (Fig. 2). However, increasing the number of populations beyond 20 (14% of the full sample) when

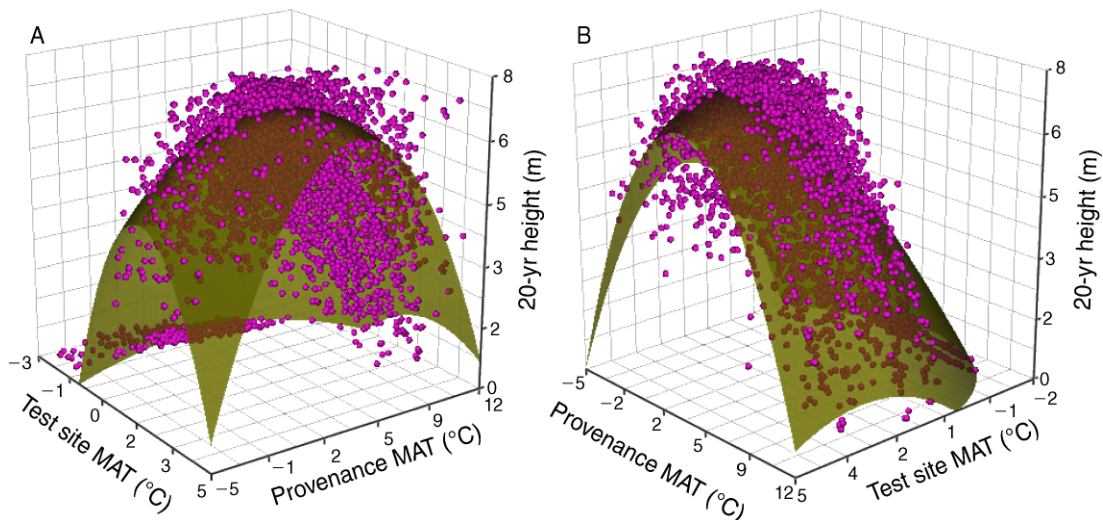


FIG. 1. Observed heights of populations in lodgepole pine (*Pinus contorta*) provenance tests and modeled universal response function in which height is predicted as a joint function of site and provenance mean annual temperature (MAT). The fits of the data points to the model surface are shown in two rotations (A and B, where B is rotated 90° from A). The lodgepole pine provenance test analyzed in this study consists of a range-wide collection of 140 populations planted at 60 test sites in British Columbia, Canada, and two sites in the southern Yukon, Canada.

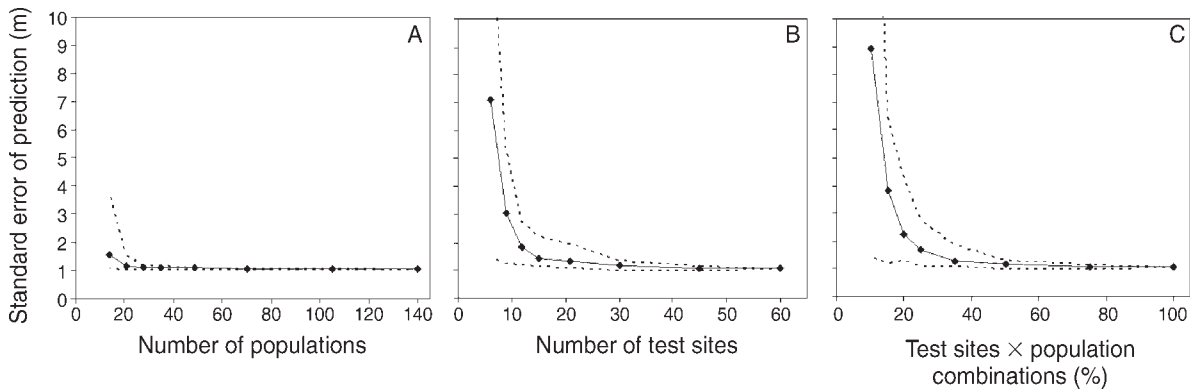


FIG. 2. Effects of number of (A) provenances, (B) test sites, and (C) both sites and provenances on standard error of predictions from the universal response function for lodgepole pine 20-yr height. The solid line is the average standard error of predictions, and dashed lines are the minimum and maximum among the 10 repeated assessments.

all test sites were included or increasing the number of sites beyond 15 (25%) when all provenances were used did not improve URF accuracy substantially. Scenarios in which numbers of both populations and sites increased showed that URF accuracy did not improve substantially when numbers of populations and sites exceeded 49 and 21 (i.e., 35% of sites and 35% of populations [i.e., 12% of the full sample]), respectively.

To assess the validity of the model, we examined the range of standard errors of prediction for each scenario. The range in standard errors of predictions among the 10 repeated predictions for each sampling scenario also declined sharply as numbers of populations and sites increased and approached the asymptote at sample sizes of 28 populations (20% of full sample), 30 sites (50% of full sample), and a combination of 70 populations and 30 sites (25% of full sample), suggesting that the model is highly consistent among the bootstrapped subsamples and not sensitive to a few particular data points.

Environmental vs. genetic effects of climate on phenotype

The universal response function integrates both environmental and genetic effects of climate by describing phenotype as a function simultaneously of the climate of the site and the provenance. Mean annual temperature (MAT) is used for both test site (MAT_s) and provenance (MAT_p) and accounts for a major proportion of the total variation (73%); thus its environmental and genetic effects can be quantified and compared. With the simplified function shown in Fig. 1, the rate of change in HT along each independent axis represents the environmental and genetic effects of MAT and can be calculated by taking the first-order partial derivatives of the URF with respect to each of the two variables (MAT_s and MAT_p). However, for better quantifications of these effects, the calculations were based on the full model with the interactions removed. The model was rebuilt with the variables listed in Table 2 except the interaction terms. Their interactions were ignored in this case because of the small

amount of variation they explained (1%) and because we are only interested in determining the environmental and genetic effects of MAT in general across sites and across provenances. From the rebuilt model the environmental effect can be estimated as

$$\frac{\partial y}{\partial \text{MAT}_s} = 1.651 - (2 \times 0.185 \times \text{MAT}_s)$$

and the genetic effect as

$$\frac{\partial y}{\partial \text{MAT}_p} = 0.260 - (2 \times 0.0440 \times \text{MAT}_p).$$

Both environmental and genetic effects, represented by the partial derivatives above, were linearly associated with MAT, although the environmental effect on phenotype was substantially stronger than the genetic effect (Fig. 3). Their values are positive when MAT is lower than 3°C and 4.5°C for genetic and environmental

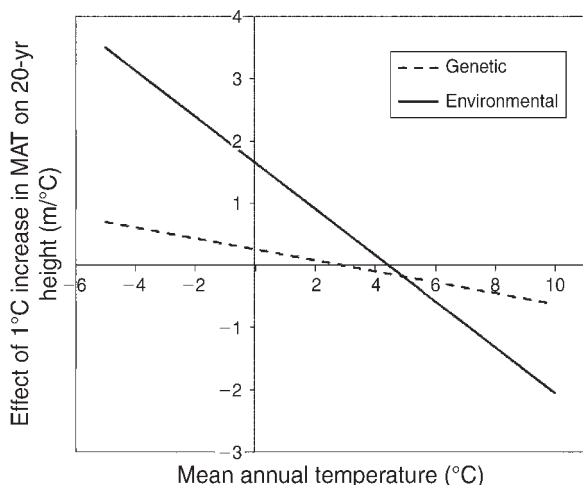


FIG. 3. Environmental vs. genetic effects of mean annual temperature (MAT) on lodgepole pine 20-yr height. Effects are expressed by the rate of change in 20-yr height with an increase of MAT by 1°C at a given MAT (x-axis).

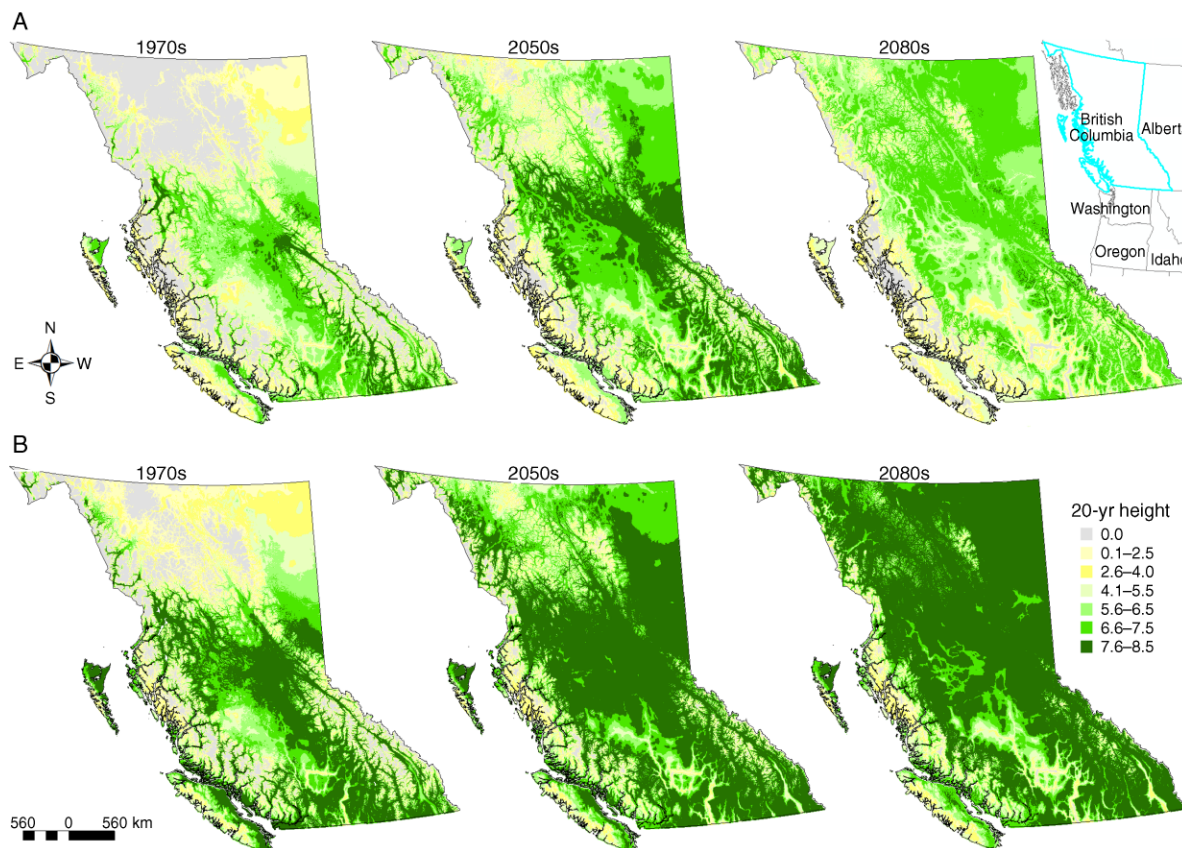


FIG. 4. Universal response function predictions of height at age 20 yr in 1970s, 2050s, and 2080s (A) for lodgepole pine populations growing in their native provenances and (B) when each location in the province is planted with the lodgepole pine provenance predicted to be the fastest growing at that location and at that date. Future climates are based on the A2 “enterprise” scenario implemented through the Canadian Global Circulation Model of the Canadian Centre for Climate Modeling and Analysis (CGCM2).

effects, respectively, and become negative at MATs greater than these values. For example, in an environment where $\text{MAT} = 0^{\circ}\text{C}$, an increase of MAT by 1°C at the test site would cause an increase in 20-year height of 1.65 m, whereas a population from $\text{MAT} = 1^{\circ}\text{C}$ would be 0.26 m taller than a population from $\text{MAT} = 0^{\circ}\text{C}$. In terms of impacts of climate change on existing forests, these findings imply that at sites colder than $\text{MAT} = 4.5^{\circ}\text{C}$, lodgepole pine growth will be accelerated due to climate warming, whereas the opposite response is expected at sites warmer than 4.5°C MAT.

Applications of URF

To predict the performance of a given population at any planting site (i.e., in any climate), the URF (Table 2) can be simplified into a population response function as the variables for the provenance become constants. Similarly, to predict the performance of any population (i.e., a population from any climate) at a given planting site, the URF (Table 2) can be simplified into an individual transfer function as the variables for the planting site become constants.

The universal response function can also be used to provide spatially explicit estimates of growth of local and optimum (i.e., fastest growing) populations for current and future climates. To estimate the growth of local provenances in current climate, values of planting site climate variables are set equal to that of the provenance climate variables in the URF (i.e., $\text{MAT}_p = \text{MAT}_s$ and $\text{LAHM}_p = \text{LAHM}_s$). To predict growth of local populations under future climates, values of site current climatic variables are replaced by site future climate values. To illustrate this application of the URF, height growth of local lodgepole pine populations at age 20 was estimated for 370 205 grid points in BC for the reference period (1970s) and two future periods (2050s and 2080s) based on CGCM2 A2 (Fig. 4A).

To estimate the growth rate of the optimal population for a planting site, we first needed to identify the source of the optimal population. This was achieved by (1) taking first-order partial derivatives of the URF (Table 2) with respect to MAT_p , LAHM_p , and SWLNG_p , respectively; (2) setting these partial derivatives to zero; and (3) solving the equations to obtain the values of the

three provenance climate variables expected to provide greatest growth at the site. After applying these optimum values to the URF, height at age 20 for each of the planting sites (grid point locations) with the optimal population (largest height) were estimated for a base case scenario (1970s) and two future periods (2050s and 2080s; Fig. 4B).

DISCUSSION

Quantitative approaches for predicting population responses to a changing climate using population transfer functions (Matyas 1994) and population response functions (Rehfeldt et al. 1999) continue to improve. We believe that the development of the universal response function, which integrates both genetic and environmental effects of climate on phenotypes into a single model, is another significant step in this field. With the growing concern about the impact of climate change on the productivity of tree populations, URFs can play an important role in understanding the environmental and genetic effects of climate on the performance of populations and can assist in developing seed deployment strategies that help ensure plantations are optimally adapted to future climates (Aitken et al. 2008).

Accuracy of URF predictions

In a previous study, we used among-population variation in individual transfer functions to produce “anchor points” to strengthen response functions (Wang et al. 2006b). In the present study, among-population variation that can be attributed to climatic and geographic variables is fully integrated into the universal response function, explaining the vast majority (81%) of the total phenotypic variation in a provenance trial. In contrast, using climatic transfer distances ($\text{MAT}_s - \text{MAT}_p$ and $(\text{MAT}_s - \text{MAT}_p)^2$) as independent variables to develop a general transfer function explained only 41% of the total variation. The accuracy of general transfer functions developed for other species is also generally low. For example, they explained on average ~30% of the total variation in height for 11 species in North America (Carter 1996, Andalo et al. 2005). Individual transfer functions and conventional response functions (with anchor points) explained ~55% and 70% of total height variation, respectively (Wang et al. 2006b), substantially less than the variation explained by the URF.

The use of the anchor points increased R^2 by just 0.04, which is much less than that in developing population response functions (+0.15) in our early study (Wang et al. 2006b). This is because the anchor points were generated using individual transfer functions based on genetic (among-population) effects of climate, and these effects have already been integrated into the URF in this study. However, as the genetic effect accounted for by the URF is global, we believe that the anchor points generated from local modeling for the five critical sites at

the extremes would further increase the accuracy of the URF.

In comparison with conventional population-specific response functions, a URF can be regarded as an average over many population response functions weighted by provenance climatic and geographic variables. Similarly, it can also be viewed as an average over many site-specific individual transfer functions adjusted by test site climatic variables. Therefore, random error of a URF would be much smaller than on a conventional response function or an individual transfer function. Random error could result from population sampling effects, an imperfect climate model, site-to-site edaphic (non-climatic environmental) variation (O'Neill et al. 2007), or biotic stresses (i.e., insect, disease, or animal damage). On the other hand, the averaging effect may also lead to either under- or overestimating the performance of some individual populations because the patterns of among-population variation can be complex, and some among-population variation cannot be explained by climatic and geographic variables alone. For example, there is a small group of observations with performances higher than predicted shown in the upper right corner of the distribution in Fig. 1A. Although the deviation of these outliers on average was much smaller in the full model (+0.57 m; Table 2) than in the simplified model (+1.02 m; Fig. 1), it is still considerable. These outlier populations from southeast BC are predicted to be the fastest growing across a wide range of locations in current and future climates (Wang et al. 2006b) and also have higher growth than others across a range of temperature regimes in growth chamber experiments (P. Smets and S. Aitken, *unpublished manuscript*). The formation of this population “hot spot” of growth potential may be related to migrational history and interspecific competition (T. Wang, A. Hamann, and S. Aitken, *unpublished manuscript*).

Contributions of independent variables

Variables selected for the final URF model include two climatic variables (MAT and LAHM) for both test sites and provenances and one geographic variable (SWLNG) for provenance. Mean annual temperature of site and provenance together explain >70% of the total phenotypic variation observed in 20-year height. This is expected as MAT is a major driver of variation in growth rates of lodgepole pine populations among test sites (Rehfeldt et al. 1999, Wang et al. 2006b). It is also the most important factor determining among-population variation in individual transfer functions (Wang et al. 2006b). The combination of MAT and AHM was chosen to define the climatic environments in developing population response functions in our previous study (Wang et al. 2006b) and was also selected in the stepwise multiple regression process used here.

As both MAT_s and MAT_p are in the model, the effects of their linear combinations (i.e., $\text{MAT}_s \pm \text{MAT}_p$), particularly the transfer distance ($\text{MAT}_s -$



PLATE 1. Photos taken from one of the 62 test sites in the lodgepole provenance test in June, 2005. This site is located in Smithers, northwestern British Columbia (55.32° N, 127.27° W and elevation 640 m). The photo on the left shows the performance of trees of an adapted population, while the photo on the right shows the ones from a maladapted population. Photo credits: T. Wang.

MAT_p), are accounted for by the model. This is also true for LAHM_s and LAHM_p. Therefore, although climate transfer distances are not specifically calculated, in contrast to general transfer functions, their effects are accounted for by the URF.

Some interactions between test site and provenance climate variables (one type of genotype $\times$ environment interaction) were also statistically significant. However, the total amount of variation explained by these interactions was small ($\sim 1\%$). This is because: (1) populations from an area in southeast BC outperformed other populations across a wide range of climate; (2) the superiority of these populations was particularly large in productive areas, where variation among populations was substantially greater than that in nonproductive areas and contributed the majority of the total variation, thus driving the shape of the URF; and (3) although these populations were not the fastest growing in the northern sites, they were still above average. Although their contributions to the model were limited, they played an important role in changing the shape of the response surface (Fig. 1) from a peak (a single optimum provenance for all sites) to a ridge (optimum provenance climate tracks site climate).

Although 15 geographic variables for provenance were tested in the analysis, SWLNG is the only one that contributed significantly to the model. This is an integrated variable comprising latitude and longitude and thus might explain some phenotypic variation associated with a geographical pattern following such a combination of latitude and longitude. It may reflect the major mountain ranges throughout much of the area sampled running from southeast to northwest or the Pacific Coast, which follows the same approximate orientation. Differences in latitude between provenance and test site were also tested to account for the effect of photoperiod, but no significant effect was detected.

Effects of numbers of populations and sites

The high cost of establishing, maintaining, and measuring provenance tests limits the number of populations and test sites, but predictive model strength and accuracy can be compromised with these limitations. Our analyses show that in the case of our experiment, both population and test site sample size could be reduced by $\sim 65\%$ without affecting the prediction accuracy of the URF. If the sampling design used in the present test is representative of conventional provenance tests, it may be possible to reduce provenance and test sample size in future experiments without significantly forfeiting model accuracy. Furthermore, if the climatic distribution of test sites and provenances along climatic gradients in future tests were more uniform, it might be possible to further reduce sample size as indicated by the relatively small minimum standard errors for small sample sizes in Fig. 2. A geographic information system in combination with high-resolution climate models could be employed to develop a strategic sampling design to minimize the numbers of sites and populations required to produce an accurate URF. Similarly, the URF makes it possible to utilize incomplete or less comprehensive provenance tests for climate change studies that would be under-designed for establishing transfer or response functions.

Environmental vs. genetic effects of climate

It is of great interest to evolutionary biologists and ecologists to separate and understand the environmental and genetic effects of climate on plant performances. In the present study we were able to identify and quantify the effects of MAT, the climatic variable that explains the majority of population and phenotypic variation for lodgepole pine height growth, using the universal response function. We found that the environmental effect was much greater (either positive or negative; Fig.

3) than the genetic effect on height growth as expected. This is because a genotypic response to selection occurs over generations and the magnitude of the response is usually only a fraction of the total phenotypic response as a function of the heritability (Falconer and Mackay 1996).

Both environmental and genetic effects are relatively small in environments where the MAT approaches the optimum for this species' growth on average (i.e., $\sim 4^{\circ}\text{C}$). This is also the environment in which the central populations are located and should be the geographic area in which populations experience the least migrational load from gene flow (Garcia-Ramos and Kirkpatrick 1997). These populations may have had the longest period of evolution to adapt to their local environments. In contrast, populations away from the central distribution range are more likely to experience environmental stress and to be exposed to natural selection pressure. Thus, the magnitude of genetic to environmental effects is probably associated with population adaptational lag (Aitken et al. 2008).

Applications of the URF in adapting to climate change

An important application of the URF is to develop a system of assisted migration (Rehfeldt et al. 2001, Hulme 2005, Millar et al. 2007). Planting seedlings adapted to predicted future climates has been proposed as a key forest management strategy to mitigate negative impacts of climate change and in some areas, to capitalize on potential opportunities for increased growth associated with warmer climates (Ledig and Kitzmiller 1992, Rehfeldt et al. 2003, Rehfeldt 2004, Sonesson 2004). However, modeling exercises are required to identify the population that will be most productive over several or many decades of the next rotation for each location. The URF is ideally suited to this challenge because of its ability to estimate the growth of any population growing in any climate. It can be used to identify optimal populations and predict their growth potentials for a given planting site for a given future climate.

Predicted heights at age 20 for optimum lodgepole pine populations in southeast BC for the 2080s were greater than predictions obtained using conventional growth response functions in our previous study (Wang et al. 2006b). This is because we developed response functions only for subsp. *latifolia* in the previous study. In the present study, we also included populations from subsp. *contorta* and subsp. *murrayana*. Some populations from subsp. *contorta* with warm provenance climates are predicted to perform well in warm interior sites, and these apparently contribute to the higher growth potential predictions in this area.

The URF can also be used to predict species range shifts. The predicted potential distribution map of 20-year height for current climate shown in Fig. 4A is comparable to the current species range (Hamann and Wang 2006). Thus, the predicted potential distribution

for the 2050s and the 2080s may well represent the distribution of appropriate habitat as described by climate for this species in those periods, at least for planted trees (Aitken et al. 2008). As the URF is mechanistically based on observations of growth responses and accounts for among-population variation, it may produce more accurate predictions of future species distributions than the climatic envelope approach (Hamann and Wang 2006, Rehfeldt et al. 2006), which utilizes the climatic breadth of niches from natural species ranges and does not consider population-specific niches or the climatic niches of planted (rather than naturally regenerated) trees. The URF predictions for species range shifts may therefore overcome a critical shortfall of commonly used species range prediction models (reviewed by Aitken et al. 2008).

Because of the universal features of the URF, it is much easier to be integrated into individual-tree stand growth models, such as the widely used growth and yield model TASS (tree and stand simulator; Mitchell and Bloomberg 1986) than individual transfer or response functions. By using growth trajectories from the URF, the output of such models can be used to predict the effects of climate change. It is also possible to integrate the URF into a high-resolution climate model such as ClimateBC. In this application, users could simply input coordinates of provenance and planting site, and the integrated model could predict the relevant climatic variables and the growth or productivity for the site.

The need to predict the growth response of trees to new climate conditions is ever growing. The URF approach for predicting such responses will increase the functionality and accuracy of climate change predictions from existing trials and increase the efficiency of new experiment designs. These applications should assist forest managers in planning reforestation for uncertain future conditions.

ACKNOWLEDGMENTS

We greatly appreciate A. D. Yanchuk at the Research Branch, BC Ministry of Forests and Range, for his review of an earlier version of the paper. We also thank two anonymous reviewers for their critical reviewing and helpful comments. Funding for this study was provided by the Forest Investment Account through both the BC Forest Science Program and the Forest Genetics Council of British Columbia.

LITERATURE CITED

- Aitken, S. N., S. Yeaman, J. A. Holliday, T. Wang, and S. Curtis-McLane. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* 1:95–111.
- Andalo, C., J. Beaulieu, and J. Bousquet. 2005. The impact of climate change on growth of local white spruce populations in Quebec, Canada. *Forest Ecology and Management* 205: 169–182.
- Beaulieu, J., and A. Rainville. 2005. Adaptation to climate change: Genetic variation is both a short- and a long-term solution. *Forestry Chronicle* 81:704–709.
- Carter, K. K. 1996. Provenance tests as indicators of growth response to climate change in 10 north temperate tree species. *Canadian Journal of Forest Research* 26:1089–1095.

- Chuine, I., P. Yiou, N. Viovy, B. Seguin, V. Daux, and E. L. Ladurie. 2004. Historical phenology: grape ripening as a past climate indicator. *Nature* 432:289–290.
- de Jong, G. 1988. Population genetics and evolution. Springer-Verlag, New York, New York, USA.
- Epperson, B. K. 2003. Geographical genetics. Monographs in population biology 38. Princeton University Press, Princeton, New Jersey, USA.
- Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman, Harlow, Essex, UK.
- Flato, G. M., G. J. Boer, W. G. Lee, N. A. McFarlane, D. Ramsden, M. C. Reader, and A. J. Weaver. 2000. The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate. *Climate Dynamics* 16:451–467.
- García-Ramos, G., and M. Kirkpatrick. 1997. Genetic models of adaptation and gene flow in peripheral populations. *Evolution* 51:21–28.
- Hamann, A., and T. Wang. 2006. Potential effects of climate change on ecosystem and tree species distribution in British Columbia. *Ecology* 87:2773–2786.
- Howe, G. T., S. N. Aitken, D. B. Neale, K. D. Jermstad, N. C. Wheeler, and T. H. H. Chen. 2003. From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany* 81:1247–1266.
- Hulme, P. E. 2005. Adapting to climate change: Is there scope for ecological management in the face of a global threat? *Journal of Applied Ecology* 42:784–794.
- Illingworth, K. 1978. Study of lodgepole pine genotype–environment interaction in B.C. Pages 151–158 in *Proceedings of the International Union of Forestry Research Organizations (IUFRO) joint meeting of working parties: Douglas-fir provenances, lodgepole pine provenances, Sitka spruce provenances and Abies provenances*. British Columbia Ministry of Forests, Information Services Branch, Victoria, British Columbia, Canada.
- IPCC [Intergovernmental Panel on Climate Change]. 2007. Climate Change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.
- Ledig, F. T., and J. H. Kitzmiller. 1992. Genetic strategies for reforestation in the face of global climate change. *Forest Ecology and Management* 50:153–169.
- Matyas, C. 1994. Modeling climate change effects with provenance test data. *Tree Physiology* 14:797–804.
- Millar, C. I., N. L. Stephenson, and S. L. Stephens. 2007. Climate change and forests of the future: managing in the face of uncertainty. *Ecological Applications* 17:2145–2151.
- Mitchell, K. J., and W. J. Bloomberg. 1986. Expanding concepts of growth and yield modeling to disease impacts and forest products. Pages 3–18 in *Second Canadian Forest Service Modelling Workshop*. Canadian Forest Service, Victoria, British Columbia, Canada.
- Morgenstern, E. K. 1996. Geographic variation in forest trees: genetic basis and application of knowledge in silviculture. University of British Columbia Press, Vancouver, British Columbia, Canada.
- O'Neill, G. A., N. Gordon, T. Wang, and P. K. Ott. 2007. Growth response functions improved by accounting for non-climatic site effects. *Canadian Journal of Forest Research* 37:2724–2730.
- O'Neill, G., A. Hamann, and T. Wang. 2008. Accounting for population variation improves estimates of climate change impacts on species' growth and distribution. *Journal of Applied Ecology* 45:1040–1049.
- Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics* 37:637–669.
- Rehfeldt, G. E. 2004. Interspecific and intraspecific variation in *Picea engelmannii* and its congeneric cohorts: biosystematics, genecology and climate change. General Technical Report RMRS-GTR-134. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, USA.
- Rehfeldt, G. E., N. L. Crookston, M. V. Warwell, and J. S. Evans. 2006. Empirical analyses of plant–climate relationships for the western United States. *International Journal of Plant Sciences* 167:1123–1150.
- Rehfeldt, G. E., N. M. Tchepakova, L. I. Milyutin, E. I. Parfenova, W. R. Wykoff, and N. A. Kouzmina. 2003. Assessing population responses to climate in *Pinus sylvestris* and *Larix* spp. of Eurasia with climate transfer models. *Eurasian Journal of Forest Research* 6:83–98.
- Rehfeldt, G. E., W. R. Wykoff, and C. C. Ying. 2001. Physiologic plasticity, evolution, and impacts of a changing climate on *Pinus contorta*. *Climatic Change* 50:355–376.
- Rehfeldt, G. E., C. C. Ying, D. L. Spittlehouse, and D. A. Hamilton. 1999. Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecological Monographs* 69:375–407.
- SAS Institute. 1989. SAS/STAT(R) user's guide. Version 6. Fourth Edition. SAS Institute, Cary, North Carolina, USA.
- Savolainen, O., T. Pyhäjärvi, and T. Knürr. 2007. Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics* 38:595–619.
- Sonesson, J. 2004. Climate change and forestry in Sweden. *Kungliga Skogs-och Lantbruksakademiens* 143:1–40.
- Wang, T., S. N. Aitken, P. Rozenberg, and M. R. Carlson. 1999. Selection for height growth and Pilodyn pin penetration in lodgepole pine: effects on growth traits, wood properties, and their relationships. *Canadian Journal of Forest Research* 29:434–445.
- Wang, T., A. Hamann, D. L. Spittlehouse, and S. N. Aitken. 2006a. Development of scale-free climate data for western Canada for use in resource management. *International Journal of Climatology* 26:383–397.
- Wang, T., A. Hamann, A. Yanchuk, G. A. O'Neill, and S. N. Aitken. 2006b. Use of response functions in selecting lodgepole pine populations for future climate. *Global Change Biology* 12:2404–2416.
- Ying, C. C., and A. D. Yanchuk. 2006. The development of British Columbia's tree seed transfer guidelines: purpose, concept, methodology, and implementation. *Forest Ecology and Management* 227:1–13.