Genet. Res., Camb. (2003), **81**, *pp.* 51–64. *With 6 figures.* © 2003 Cambridge University Press
DOI: 10.1017/S0016672302005980 Printed in the United Kingdom

Quantitative trait loci for growth trajectories in *Populus*

RONGLING WU[1]*, CHANG-XING MA[1], MARK C. K. YANG[1], MYRON CHANG[1], RAMON C. LITTELL[1], UPASANA SANTRA[1], SAMUEL S. WU[1], TONGMING YIN[2], MINREN HUANG[2], MINXIU WANG[2] AND GEORGE CASELLA[1]

[1] *Department of Statistics, University of Florida, Gainesville, FL 32611, USA*
[2] *The Key National Laboratory of Forest Genetics and Gene Engineering, Nanjing Forestry University, Nanjing, Jinagsu 210037, People's Republic of China*

(*Received 25 June 2002 and in revised form 16 September 2002*)

Summary

Growth trajectories are a biological process important to plant and animal breeding, and to evolutionary genetic studies. In this article, we report the detection of quantitative trait loci (QTLs) responsible for growth trajectories in poplars that are used as a model system for the study of forest biology. These QTLs were localized on a genetic linkage map of polymorphic markers using a statistical mapping method incorporating growth-curve models. The effects of the QTLs on growth are described as a function of age, so that age-specific changes in QTL effects can be readily projected throughout the entire growth process. The QTLs identified display increased effects on growth when trees age, yet the timing of QTL activation is earlier for stem height than diameter, which is consistent with the ecological viewpoint of canopy competition. The implications of the results for breeding and silviculture are discussed.

1. Introduction

The *Populus* genus contains ~30 species with a huge range of diversity in morphology, physiology, anatomy and reproductive behaviour (Stettler *et al*., 1996). Because of its fast growth, ease of clonal propagation, strong heterosis from interspecific hybridization and multiple industrial and bioenergetic uses, *Populus* has been widely used as a cultivated forest tree in the Northern Hemisphere. From a biological perspective, poplar is regarded as an excellent model system for fundamental research on trees owing to its favourable attributes, such as small genome size (550 Mb; $2C=$ 1·1 pg; 19 chromosomes) and amenability to genetic transformation (Stettler *et al*., 1996; Klopfenstein *et al*., 1997). More recently, genetic linkage maps based on molecular markers have been constructed in many poplars (Frewen *et al*., 2000; Wu *et al*., 2000*a*; Cervera *et al*., 2001; Yin *et al*., 2002) that will catalyse the identification of quantitative trait loci (QTLs) underlying traits of biological and economical importance.

Unlike crop annuals, poplar is long-lived and needs many years to achieve harvest ages in the field. Thus, an understanding of the genetic mechanisms underlying wood stem production in poplar should be integrated with the biological process of growth and development. In practice, knowledge of the developmental genetic basis of growth is crucial to increasing the efficiency of early selection in poplar. It has long been recognized that early selection for growth performance can accelerate the breeding process and minimize the cost of large-scale progeny tests in long-lived, large-sized forest trees (Wu *et al*., 2000*c*). The application of early selection, however, has been very limited because of our ignorance of the genetic basis of growth and development. Several authors have used molecular markers to examine the dynamic changes in the gene effect of a particular QTL over time (Verhaegen *et al*., 1997; Emebiri *et al*., 1998; Wu *et al*., 1998; Kaya *et al*., 1999). They generally claimed that, whereas some QTLs are stable, many change their action and interaction as trees age. Such a conclusion is particularly uninformative, given the much-more-complex patterns of QTL effects over time. More importantly, this conclusion might be premature because,

* Corresponding author. Tel: +1 352 392 3806. Fax: +1 352 392 8555. e-mail: rwu@stat.ufl.edu

in these studies, QTL effects on growth were estimated separately for different time points, with no consideration of possible between-age correlations.

In this article, we use a more powerful process-based statistical method to analyse QTLs that affect growth trajectories in interspecific hybrids between *Populus deltoides* and *Populus euramericana*. First, we reported the construction of genetic linkage maps for a hybrid population of *P. deltoides* × *P. euramericana* based on a pseudo-testcross strategy (Grattapaglia & Sedero, 1994) and random amplified polymorphic DNA (RAPD), restriction-fragment-length polymorphism (AFLP) and inter-simple-sequence repeats (ISSR) markers (Yin *et al*., 2002). Second, we have developed a novel statistical method using growth-model theories (Ma *et al*., 2002; Wu *et al*., 2002*a*). The basic principle of this method is to express the genotypic means of a QTL at different time points in terms of a continuous growth function with respect to time (t). Under this principle, the parameters describing the shape of growth curves, rather than the genotypic means as expected in traditional mapping strategies, are estimated within a maximum-likelihood framework. This method makes use of a nearly universal growth model in biology, the logistic or S-shaped curve, thus significantly enhancing the estimation precision of QTL parameters and the biological relevance of the QTLs detected.

2. Age-specific effect of a QTL

Growth trajectories are a complex biological process, on which the influence of a QTL could be complicated (Cheverud *et al*., 1996). However, the effect of a QTL as a function of age can be roughly illustrated by ten different patterns, using two QTL genotypes 1 and 2 as an example (Fig. 1).

(a) 'Repulsion': one QTL genotype has a larger genetic value than the other at each time point, and they never cross during the entire growth trajectory.
(b) Divergence: the two QTL genotypes differ from early growth, and their differences increase gradually with age.
(c) 'Crossover': the two QTL genotypes cross once between the initial growth and asymptotic growth.
(d) Divergence–'crossover': the two QTL genotypes differ from early growth, but cross prior to asymptotic growth.
(e) Double 'crossover': the two QTL genotypes cross twice between initial growth and asymptotic growth.
(f) Convergence: the two QTL genotypes reduce their differences with age until they merge prior to asymptotic growth.
(g) Divergence–convergence: the two QTL genotypes differ from the initial growth, but merge prior to asymptotic growth.
(h) Early 'crossover'–convergence: the two QTL genotypes cross before their inflection points and merge again prior to asymptotic growth.
(i) Middle 'crossover'–convergence: the two QTL genotypes cross at their inflection points and merge again prior to asymptotic growth.
(j) Late 'crossover'–convergence: the two QTL genotypes cross after their inflection points and merge again prior to asymptotic growth.

Quote marks around 'repulsion' and 'crossover' are to distinguish these QTL effects from the general genetic meanings of these terms. For the 'repulsion' pattern, if the two QTL genotypes have parallel genetic values during the growth trajectories, the effect of this QTL is stable over time, causing no QTL × age interaction. Yet, if nonparallel genetic values are detected between the two QTL genotypes (even if they diverge) or if any other of the patterns b–j occur then the effect of the QTL is age dependent and there is QTL × age interaction. The classification of a QTL effect into these different patterns will help us to understand the genetic basis of growth and development and to use these underlying QTL better for the genetic improvement of growth.

3. Materials and methods

(i) *Plant material*

The plant material used was derived from the hybridization of *Populus*. A *P. deltoides* clone (I-69) was used as a female parent and mated with an interspecific *P. deltoides* × *P. nigra* clone (I-45) as a male parent (Wu *et al*., 1992). Both *P. deltoides* I-69 and *P. euramericana* (*P. deltoides* × *P. nigra*) I-45 were selected at the Research Institute for Poplars in Italy during the 1950s and were introduced to China in 1972. In the spring of 1988, a total of 450 one-year-old rooted three-way-hybrid seedlings were planted at a spacing of 4 × 5 m at a forest farm near Xuchou City, Jiangsu Province, China. The total stem heights and diameters measured at the end of each of 11 growing seasons are used in this example.

(ii) *Linkage maps*

A subset of genotypes (90) randomly selected from the 450 hybrids were used to construct a genetic linkage map by AFLP and ISSR (Yin *et al*., 2002). Of a total of 839 polymorphic markers identified, 560 (67%) were the testcross markers heterozygous in one parent but null in the other (segregating 1 : 1), 206 (25%) were the intercross dominant markers heterozygous in both parents (segregating 3 : 1) and the remaining 73 (9%) were the co-dominant markers heterozygous in both parents (segregating 1 : 2 : 1). A mixed set of the testcross markers and co-dominant markers were used

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

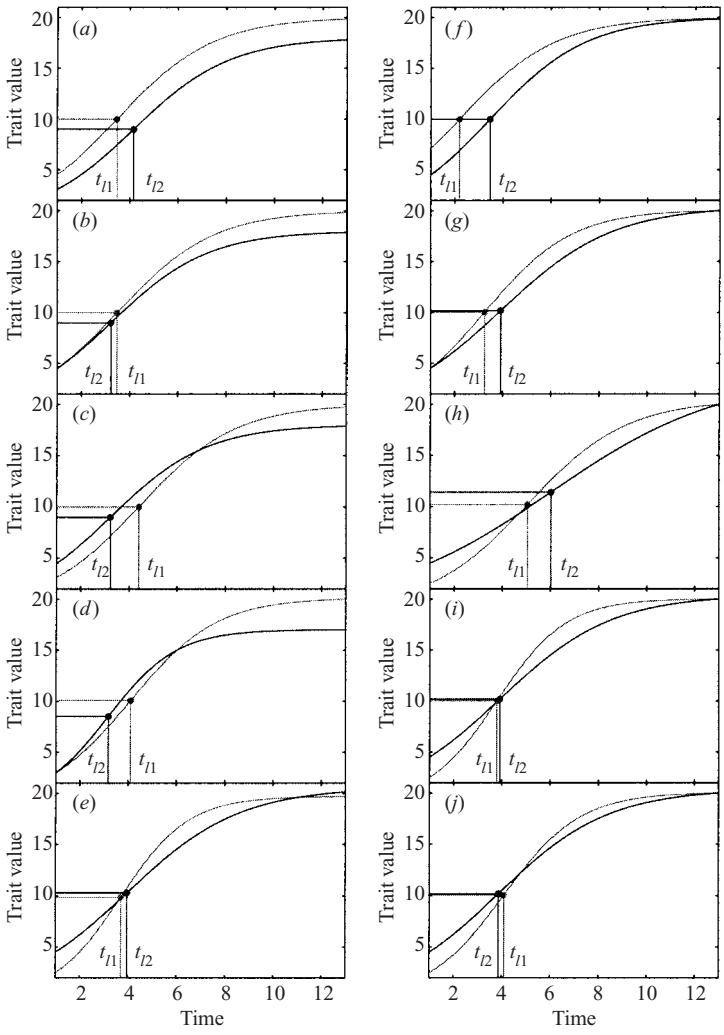

Fig. 1. Diagrams depicting different dynamic patterns of a QTL affecting growth trajectories. In each graph, the curves represent two QTL genotypes, and the difference between the two curves depicts the magnitude of the QTL effect. The growth and time at the reflect point (at which a tree has a maximum rate of growth) for a different QTL genotype are indicated by coordinates. (*a*) 'Repulsive' pattern. (*b*) Divergence pattern. (*c*) 'Crossover' pattern. (*d*) Divergence–'crossover' pattern. (*e*) Double 'crossover' pattern. (*f*) Convergence pattern. (*g*) Divergence–convergence pattern. (*h*) Early 'crossover'–convergence pattern. (*i*) Middle 'crossover'–convergence pattern. (*j*) Late 'crossover'–convergence pattern.

to construct two linkage maps based on a pseudo-testcross strategy (Grattapaglia & Sedero, 1994), one for the *P. deltoides* genome (D) and the other for the *P. euramericana* genome (E). The intercross dominant markers were not included in map construction because they do not provide accurate estimates of the recombination fractions. The two maps constructed showed nearly complete coverage of the genome, spanning 3801 cM and 3452 cM, respectively. The availability of co-dominant markers as orthologous genes allowed a direct comparison of the rates of meiotic recombination in the two different parent species. In this study, five of the 19 largest linkage groups for each parental map, which approximately represent 19 pairs of chromosomes, are used for QTL identification. Of these, linkage groups D1/E1 to D11/E11 have orthologous markers to align the two parental maps, whereas linkage groups D12–D19 and E12–E19 are separated for each parent (Yin *et al.*, 2002).

(iii) *Statistical analysis*

The method of Ma *et al.* (2002) for mapping age-specific QTLs on growth was developed for the simplest backcross design, initiated with two inbred lines. This method does not directly estimate the expected means of different QTL genotypes at different time points but fits these means using growth curves as a function of time. Thus, the estimation of QTL effects on growth is equivalent to the estimation of model parameters describing the shape of growth curves. Ma *et al.* (2002) based their QTL analysis on a logistic or S-shaped curve, which is an almost-universal rule for capturing age-specific changes in all organisms

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

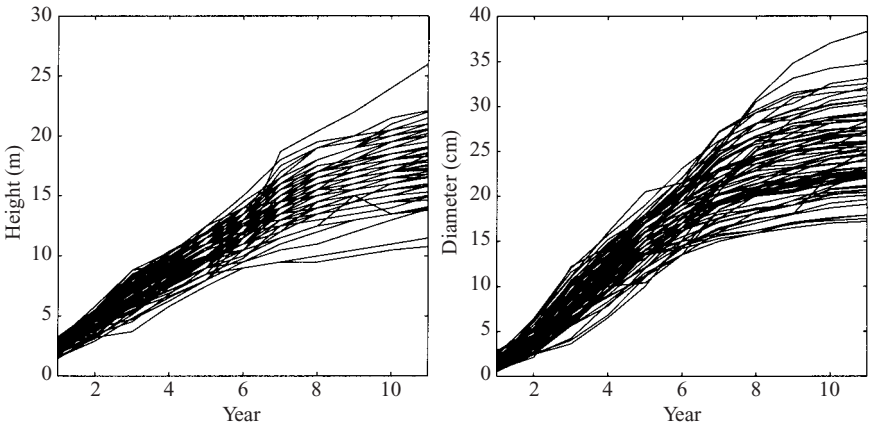

Fig. 2. Plots of stem height and diameter growth against age for 90 genotypes used to construct linkage maps in poplar hybrids (Yin *et al.*, 2002).

(Pearl, 1925; Niklas, 1994; West *et al.*, 2001). When growth (g) is plotted against time, it typically follows a logistic growth curve

$$g(t)=\frac{a}{1+be^{-rt}}, \tag{1}$$

where a is the asymptotic or limit value of g when $t\to\infty$, $g(0)=a/(1+b)$ is the initial value of g when $t=0$, and r is the relative rate of growth (von Bertalanffy, 1957). The overall form of the curve is determined by different combinations of parameters a, b and r.

The logistic growth curve described in eqn 1 can be used to determine the coordinates of the most biologically important point in the entire growth trajectory, the inflection point, at which a maximum growth rate occurs. The time (t_I) and growth ($g(t_I)$) at the inflection point for a QTL genotype can be derived as

$$\begin{cases} t_I=\dfrac{\log b}{r} \\ g(t_I)=\dfrac{a}{2}. \end{cases} \tag{2}$$

The time at the inflection point, together with the initial growth ($g(0)$) and asymptotic growth (a) determine exclusively the difference between any two growth curves. That is, any two curves will be indistinguishable if they have the same values for these three variables.

According to the method of Ma *et al.* (2002), unknown parameters fitting the growth curves of two QTL genotypes (j) include $\Omega=(a_j, b_j, r_j, \theta, \rho, \sigma^2)^T$, where θ specifies the position of the QTL and ρ and σ^2 are related to the residual (co)variances among different ages, whose structure is specified by the AR(1) model (Davidian & Giltinan, 1995) expressed as

$$\Sigma=\sigma^2\begin{bmatrix} 1 & \rho & \cdots & \rho^{m-1} \\ \rho & 1 & \cdots & \rho^{m-2} \\ \cdots & \cdots & \cdots & \cdots \\ \rho^{m-1} & \rho^{m-2} & \cdots & 1 \end{bmatrix}. \tag{3}$$

Wu *et al.* (2002*a*) discussed the advantages and disadvantages of this model when used to specify the structure of a age-dependent residual (co)variance matrix. The genetic effects of the QTL on growth trajectories are described by the growth-curve parameters contained in Ω. The crucial value for claiming the existence of growth QTL can be determined from the Bonferroni argument for the sparse-map case (Lander & Bostein, 1989) or by permutation tests proposed by Doerge & Churchill (1996). In this study, permutation tests repeated 100 times for each linkage group are used to determine these crucial thresholds at significance levels $\alpha=0.05$.

Our plant material in this study is a full-sib family derived from two outcrossing parents. The genetic maps for this full-sib family were constructed based on a two-way pseudotest backcross in which one parent is heterozygous and the other is homozygous. For the heterozygous parent, two markers and their bracketed QTL can be arranged in different linkage phases (coupling vs repulsion). Hence, for the pseudo-test-backcross strategy, the determination of a more-likely linkage phase across the markers and QTLs is a prerequisite for estimating the position and effects of a putative QTL responsible for growth differentiation. In the appendix, we provide a modified statistical procedure for estimating a more-likely linkage phase and QTL positions and effects.

4. Results

(i) *QTL detection*

By plotting total growth against year, we found that each of the 90 mapped genotypes followed a different S-shaped growth curve (Fig. 2; $R^2=0{\cdot}90$–$0{\cdot}98$). The statistical model built on the S-shaped growth curve that is regarded as nearly universal in biology (West *et al.*, 2001) is used to map QTLs responsible for the difference in growth trajectories between these

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

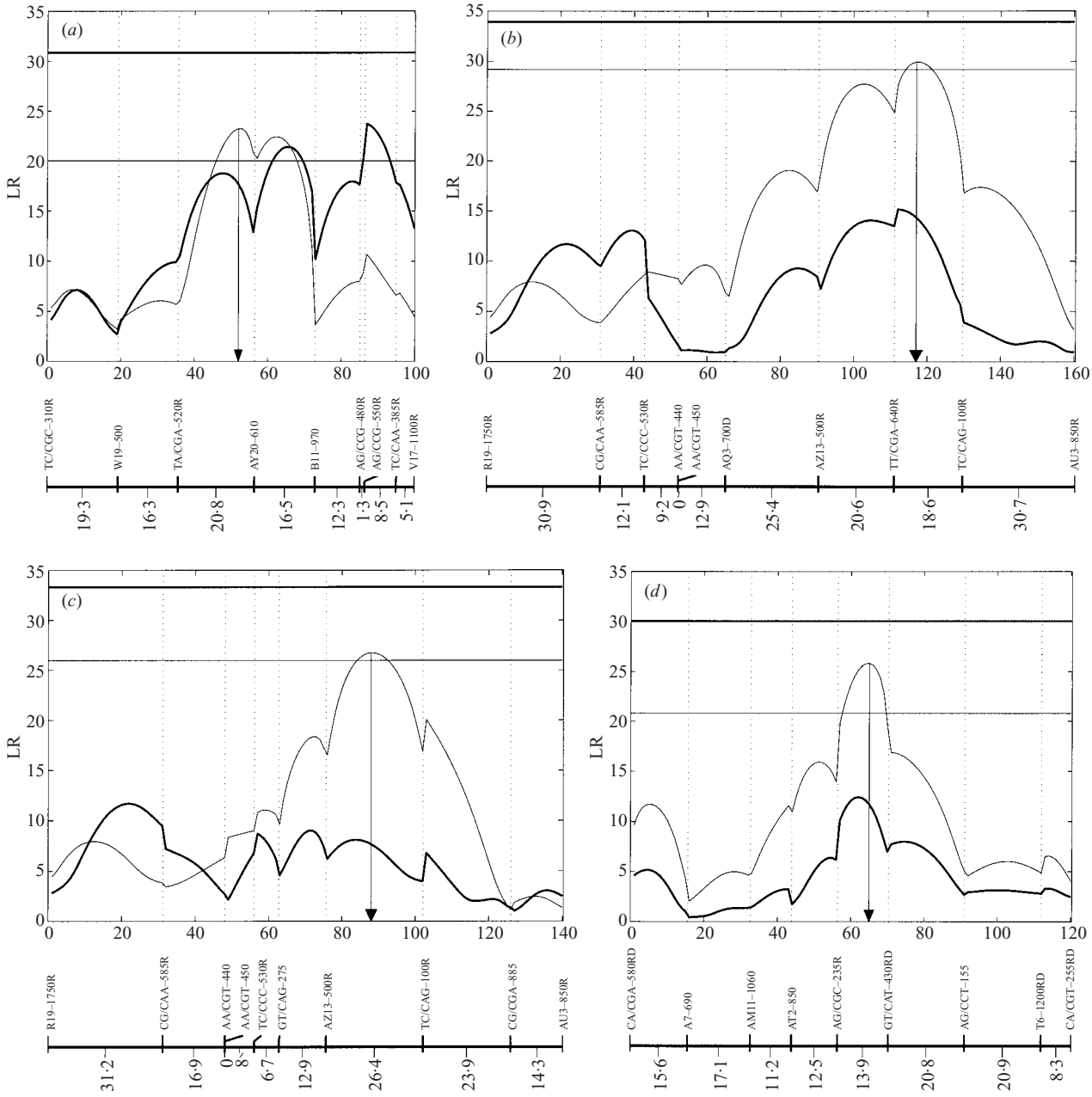

Fig. 3. The profile of the log-likelihood ratios (thin curve) between the full model and reduced (no QTL) model for height growth trajectories across different linkage groups. The genomic positions corresponding to the peaks of the curves are the MLEs of the QTL localization. For comparison, the profiles for diameter growth are also given (thick curve). The two horizontal lines indicate the threshold values for height (thin) and diameter (thick). The positions of markers on linkage groups and their names (Yin *et al.*, 2002) are given beneath. (*a*) Linkage group E1-1. (*b*) Linkage group D7. (*c*) Linkage group E7. (*d*) Linkage group E17.

genotypes. When QTL mapping is based on a pseudo-test-backcross population, as in this study, we should first determine a more-likely linkage phase between the QTL hypothesized and its flanking markers. A procedure is described in the appendix, using the logistic-mixture mapping model originally proposed by Wu *et al.* (2002*a*) and Ma *et al.* (2002) to map growth QTLs, for the determination of a more-likely linkage phase based on four different criteria for making a constraint on genotypic differences in growth trajectories between two groups of QTL genotypes in the pseudo-test backcross.

Here, we use the constraint on asymptotic growth, for which four significant QTLs are identified for stem-height growth and two significant QTLs are identified for stem-diameter growth during the first 11 years in the interspecific hybrids of poplar. The four height QTLs and two diameter QTLs detected are located on different linkage groups (Figs 3, 4). For discrimination, a QTLs guiding stem-height growth are defined as

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

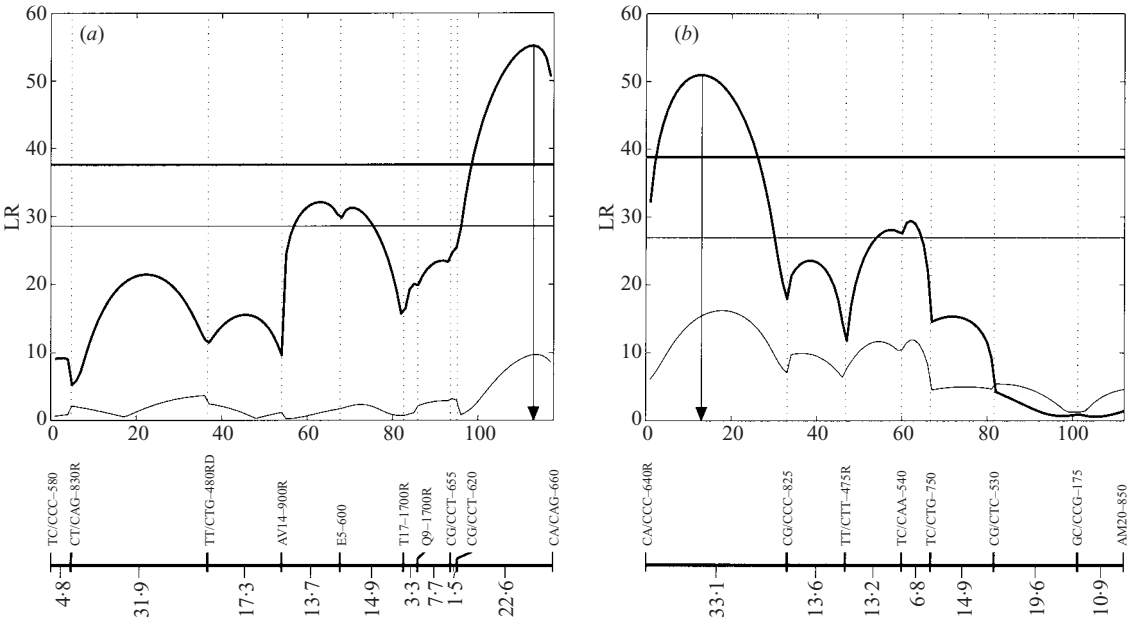

Fig. 4. The profile of the log-likelihood ratios (thick curve) between the full model and reduced (no QTL) model for diameter growth trajectories across different linkage groups. The genomic positions corresponding to the peaks of the curves are the MLEs of the QTL localization. For comparison, the profiles for height growth are also given (thin curve). The two horizontal lines indicate the threshold values for height (thin) and diameter (thick). The positions of markers on linkage groups and their names (Yin *et al.*, 2002) are given beneath. (*a*) Linkage group D9. (*b*) Linkage group D10.

follows. Height QTL 1 is located between markers TA/CGA-520R and AY20-610 on linkage group E1-2 (Fig. 3*a*). Height QTL 2 is located between markers TT/CGA-640R and TC/CAG-100R on linkage group D7 (Fig. 3*b*). Height QTL 3 is located between markers AZ13-500R and TC/CAG-100R on linkage group E7 (Fig. 3*c*). Height QTL 4 is located between markers AG/CGC-235R and GT/CAT-430RD on linkage group E17 (Fig. 3*d*). The QTLs for diameter growth are defined as follows. Diameter QTL 1 is located between markers CG/CCT-620 and CA/CAG-660 on linkage group 9D (Fig. 4*a*). Diameter QTL 2 is located between markers CA/CCC-640R and CG/CCC-825 on linkage group D10 (Fig. 4*b*).

Height QTLs 2 and 3 are probably present orthologous, because these two QTLs are associated with intercross markers AZ13-500R and TC/CAG-100R on homologous linkage groups 7D and 7E (Yin *et al.*, 2002). It seems that this orthologous QTL has a consistent map distance ($\sim$13 cM) from TC/CAG-100R, although the two intercross markers are not equally distanced between the two parents. The height QTLs were found to differ from the diameter QTLs in two respects: their heterozygosity in the parental species; and the sizes of QTL effects. It is consistent that favourable alleles for the height QTLs are contributed from the *P. euramericana* parent and favourable alleles for the diameter QTLs are from the *P. deltoides* parent. Although more QTLs are detected for height growth than diameter growth, the effects of the diameter QTLs tend to be stronger than those of the height QTLs, as shown by different test statistics at the peaks of log-likelihood ratio curves (Figs 3, 4). It is surprising that we did not find any co-location between QTLs for height and diameter growth. There is probably a pleiotropic QTL affecting both of these traits, given their strong genetic correlation (Wu & Stettler, 1994, 1996). These QTL might be detected when the sample size and/or the marker density are increased over the current level.

For linkage group E1-2, the curve of the log-likelihood ratio across the map length displays more than one peak for stem height above the threshold at significance level $\alpha = 0{\cdot}01$. In general, unless different peaks on the same group are distant, they cannot be interpreted as definite evidence of multiple QTLs located at the intervals corresponding to the peaks. This is because false 'ghost' peaks can arise if other QTLs are linked to the intervals of interest (Martinez & Curnow, 1992). Multiple QTLs can be detected more accurately when our mapping strategy is combined with composite interval mapping (Jansen & Stam, 1994).

(ii) *The dynamic pattern of QTL effects*

The growth curves of height and diameter are drawn using the estimates of logistic parameters for two genotypes at a QTL detected on the linkage groups (Figs 5, 6). As mentioned above, by comparing these curves, we can ask many biologically relevant questions, such as how the QTL affects the entire growth

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

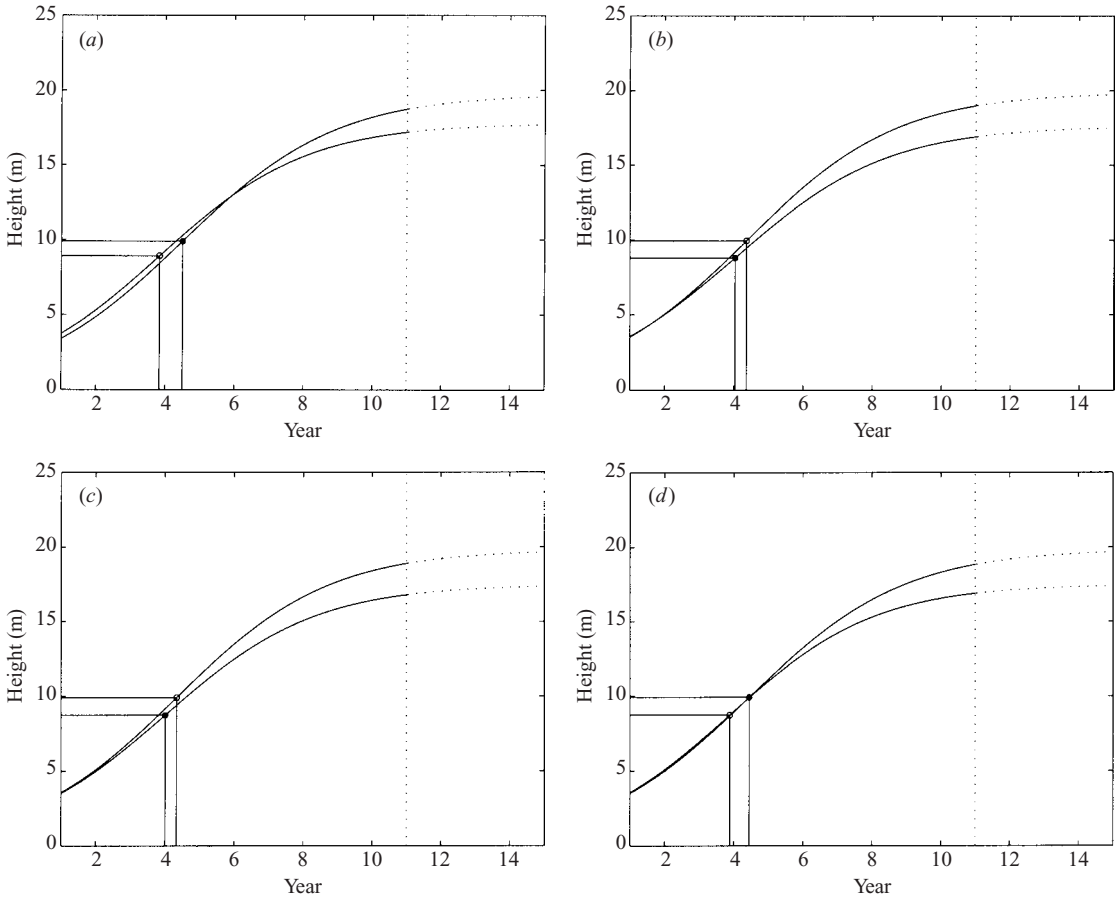

Fig. 5. Two height-growth curves, each presenting a group of genotypes at a QTL detected on different linkage groups. The growth and time at the inflection point are indicated for each QTL genotype. The differentiation patterns of growth curves beyond the maximum observed age (11) affected by the QTL are represented by the dotted lines. (*a*) Linkage group E1-1. (*b*) Linkage group D7. (*c*) Linkage group E7. (*d*) Linkage group E17.

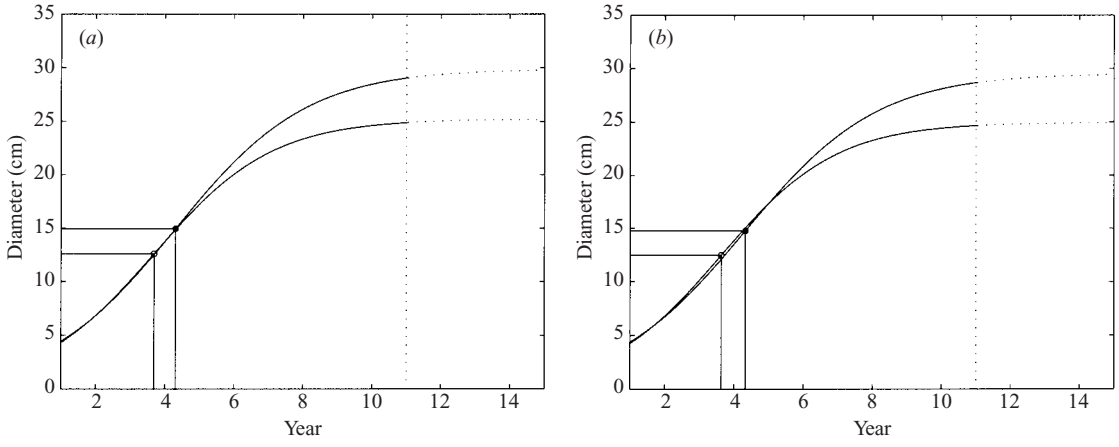

Fig. 6. Two diameter-growth curves, each presenting a group of genotypes at a QTL detected on different linkage groups. The growth and time at the inflection point are indicated for each QTL genotype. The differentiation patterns of growth curves beyond the maximum observed age (11) affected by the QTL are represented by the dotted lines. (*a*) Linkage group D9. (*b*) Linkage group D10.

process (local vs global effect) and when a QTL starts to affect growth trajectories. In our example, some of the height QTLs are active throughout the entire curve and are thus more globally expressed (Fig. 5*a*), whereas the diameter QTLs are more locally expressed because the two QTL genotypes differentiate only after age 5–6 years (Fig. 6).

If two growth curves affected by a QTL have different ages and/or growth at the inflection point, this indicates that the inflection point is under the

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

determination of the QTL. Of the QTLs detected, some have stronger control over the inflection point than others (Figs 5, 6), although all are statistically significant. Because the inflection point represents the point at which a tree has a maximum rate of growth, the genetic control of the inflection point suggests that the growth trajectory can be genetically modified to increase a tree's capacity to acquire spatial resources.

Different patterns of the QTL effect between height and diameter are observed when a QTL starts to exert an effect. Both height and diameter QTLs exert an increased influence on growth with age, but the timing of gene activation is generally earlier for the height QTLs (Fig. 5) than the diameter QTLs (Fig. 6). More specifically, the height QTLs affect growth differentiation in height from 1–5 years of age (before or around the inflection point), whereas the diameter QTLs are effective after 5–6 years of age (after the inflection point). The QTL on linkage group E1-2 displays an effect on height growth with different directions between younger and older ages (Fig. 5*a*).

5. Discussion

The change in an individual's phenotype with age is called its growth trajectory. This is a complex trait that, in addition to genetic and environmental factors, can also depend on a number of physiological and developmental control mechanisms that guide growth process in an integrated manner (Cheverud *et al*., 1983; Atchley, 1984). Both theoretical predictions and empirical studies suggest that genetic factors or QTLs play a central role in the differentiation of growth trajectory and can be used to alter growth patterns through artificial selection (Kirkpatrick & Heckman, 1989; Kirkpatrick *et al*., 1990, 1994; Atchley & Zhu, 1997; Pletcher & Geyer, 1999; Vaughn *et al*., 1999). Previous genetic studies, aimed at identifying QTLs for growth using molecular markers, have not considered the physiological control mechanisms of growth (Cheverud *et al*., 1996; Verhaegen *et al*., 1997; Emebiri *et al*., 1998; Wu *et al*., 1998; Kaya *et al*., 1999). Yan *et al*. (1998) and Wu *et al*. (2002*b*) used a conditional developmental quantitative genetic model developed by Zhu (1995) to detect QTLs expressed at different growth stages in plants. These studies base QTL mapping on a finite set of discrete time points and have two limitations. First, from a biological perspective, QTLs detected separately at individual time points are expressed at one time point and cannot reflect a dynamic nature of gene effect as a function of the previous physiological status of a plant and, therefore, as a function of time. Second, from an analytical perspective, the identification of QTLs for separate time points has less accuracy and precision of parameter estimates than when all time points are considered simultaneously, as demonstrated in multi-trait-mapping studies (Knott & Haley, 2000; Korol *et al*., 2001).

To circumvent the biological limit, we use an innovative statistical model to map QTLs affecting growth trajectories. This model (called the Wu–Ma model thereafter), originally produced by Wu *et al*. (2002*a*) and Ma *et al*. (2002), incorporates growth laws into a mapping strategy and has proved to have advantages for detecting biologically relevant QTLs for growth. Using the Wu–Ma model, we can test how a QTL globally guides growth trajectories and when the QTL starts or stops its effect during growth process. Also, some developmentally or ecologically important questions can be addressed; for example, how important the QTL is in determining a specific developmental event, such as the occurrence of the inflection point associated with a plant's fate in a competitive environment and, therefore, its fitness.

In this study, we use the Wu–Ma model to map the QTLs governing the differences in the growth trajectories of an interspecific poplar hybrid population between two commercially important poplar species, *P. deltoides* and *P. euramericana*. Our study is based on high-density linkage maps constructed for this population using polymorphic molecular markers (Yin *et al*., 2002). Several previous investigations in poplars have shown that there is a significant genetic variability between different stages of stem growth before and after canopy closure (Wu & Stettler, 1994, 1996, 1997). In our current analysis, this information is deciphered more precisely by modelling the genetic variation as a function of time. For both stem-height and stem-diameter growth, resultant genetic effects of the QTLs detected tend to increase as trees age (Fig. 6), conforming to our earlier results from quantitative genetic analysis, in which broad-sense heritabilities increase with age (Wu & Stettler, 1996; Wu *et al*., 1998). However, these two traits display different age-dependent patterns in gene effect. The effects of the QTL on height growth seems to occur earlier on diameter growth. Also, height is predicted to tend to have reduced differentiation after year 11, whereas diameter maintains a high degree of differentiation after this time point (Fig. 6). These differences in genetic control mechanism between height and diameter are consistent with ecological viewpoints of allometric scaling (Wu & Stettler, 1996). When canopies close, strong competitive interactions occur among poplar trees in the stand. Thus, to maintain vigorous growth, larger individuals must allocate more stem biomass to radial growth than to height growth, whereas smaller individuals tend to emphasize height growth at the cost of radial growth to gain access to light. Such changes in biomass allocation during canopy closure result in reduced differences in height but increased differences in diameter. It is not surprising that genes are responsible for this transition between growth phases.

The QTL for growth identified from the Wu–Ma model, although they need to be confirmed with more trees and markers, might have immediate implications for early selection in poplar. Based on the growth curves of different genotypes at an identified QTL, one can predict how the QTL affects stem wood growth when trees are older. Although some of the QTLs, especially those for diameter growth, might have minimal effects on growth at earlier stages, they can still be incorporated in a marker-assisted selection strategy for superior growth at harvesting ages (Wu *et al*., 2000*b*), because their increasingly growing effects are predicted. In applied silviculture, growth curves of a QTL offer important information for better management of poplar trees. For a given QTL genotype, we can identify the age of its inflection point at which this genotype has a maximum rate of growth. Thus, from growth curves, we can better determine the optimal ages of trees, at which fertilization and irrigation might result in maximum stem wood growth and production.

In our theoretical analyses of the Wu–Ma model, we claimed that the implementation of growth laws in a QTL mapping strategy has provided a more powerful way of studying the genetic architecture of quantitative traits and shedding light on the interplay of genetics, development, physiology and evolution (Rice, 1997). By analysing real-world data from poplar hybrids, we further claim that results from the Wu–Ma model have great potential to advance theoretical genetics at the frontiers of applications to breeding and silviculture.

We are grateful to Bill Hill and two anonymous referees for their constructive comments on this manuscript. This work is partially supported by grant from the National Science Foundation to G. C. (DMS9971586), an Outstanding Young Investigators Award of the National Natural Science Foundation of China to R. W. (30128017) and the University of Florida Research Opportunity Fund to R. W. (02050259). The publication of this manuscript is approved as journal series R-08695 by the Florida Agricultural Experiment Station.

Appendix

In what follows, we modify the growth model of Ma *et al*. (2002) to suit a pseudo-test backcross derived from two outcrossing parents.

(i) *Linkage phases*

For a pseudo-test backcross, the heterozygous parent might have different linkage phases or nonallelic assignments between the two flanking markers (M_1 and M_2) and QTL (Q). The alleles of the markers and QTLs are denoted by M vs m and Q vs q, respectively. There are a total of four possible nonallelic assignments. (1) The nonalleles M_1 and M_2 of the two markers are in a coupling phase, both of which are in a coupling phase with QTL allele Q, expressed as M_1QM_2/m_1qm_2. (2) The nonalleles M_1 and M_2 of the two markers are in a coupling phase, both of which are in a repulsion phase with QTL allele Q, expressed as M_1qM_2/m_1Qm_2. (3) The nonalleles M_1 and M_2 of the two markers are in a repulsion phase, of which M_1 is in a coupling phase with QTL allele Q, expressed as M_1Qm_2/m_1qM_2. (4) The nonalleles M_1 and M_2 of the two markers are in a repulsion phase, of which M_1 is in a repulsion phase with QTL allele Q, expressed as M_1qm_2/m_1QM_2.

In practice, the parental linkage phase between the two markers can be known prior to QTL identification (Grattapaglia & Sedero, 1994). Thus, a more likely linkage phase of the two markers relative to the QTL is just selected from cases 1 and 2 or cases 3 and 4 by comparing their respective likelihoods. Assuming no double recombination, we derive the conditional probabilities of QTL genotypes upon marker genotypes under each of the four nonallelic assignments for the pseudo-test backcross design (Table 1).

(ii) *Characterizing a most likely linkage phase and QTL effect pattern*

The computational algorithm for a pseudo-test backcross should consider two alternative nonallelic assignments of a QTL (1 and 2 or 3 and 4) relative to the two flanking markers of known linkage phase (Table 1). For each of these two pairs of assignments, however, the conditional probabilities for one QTL genotype under one assignment have the same effects as those for the alternative QTL genotype under the other assignment (Table 1). For this reason, two different logistic curves, each described by a set of parameters $\pi=(a_j, b_j, r_j)^T$ ($j=1$ for Qq and 2 for qq) cannot be used to characterize the identities of the two QTL genotypes when a pair of assignments, A1 and A2 or A3 and A4, are compared. Because the likelihood values calculated are not different between the two alternative assignments, there is no way to make a choice for a more likely assignment. The alternative nonallelic assignments can be distinguished by posing some constraints on the difference in genetic values between the two QTL genotypes. The difference between the QTL genotypes can be specified by assuming that one QTL allele is favourable to the other. But the allelic relationship of a QTL may change from one time point to the next in growth

Table 1. *Conditional probability of a QTL genotype upon the marker genotypes for* M_1 *and* M_2 *under four different nonallelic assignments* (*1*−*4*)

Marker genotype	1		2		3		4	
	Qq	*qq*	*Qq*	*qq*	*Qq*	*qq*	*Qq*	*qq*
$M_1m_1M_2m_2$	1	0	0	1	$1-\theta$	θ	θ	$1-\theta$
$M_1m_1m_2m_2$	$1-\theta$	θ	θ	$1-\theta$	1	0	0	1
$m_1m_1M_2m_2$	θ	$1-\theta$	$1-\theta$	θ	0	1	1	0
$m_1m_1m_2m_2$	0	1	1	0	θ	$1-\theta$	$1-\theta$	θ

$\theta = r_1/r$, where r is the recombination fraction between two markers M_1 and M_2 and r_1 is the recombination fraction between QTL and marker M_1. Double crossover is ignored.

trajectories, which is different from an ordinary single-trait analysis. In fact, the allelic relationship can be assumed to be based on four different criteria: (1) final (asymptotic) growth; (2) time at the inflection point; (3) growth at a time point of typical interest (e.g. harvesting age); and (4) growth during a particular time interval. Without loss of generality, we assume that allele Q is favourable to q for each of these criteria. The constraints on genotypic differences in growth trajectories are given for each criterion as follows.

(a) *Asymptotic growth*

Asymptotic growth represents the limit of growth. For an individual, greater asymptotic growth helps to maintain its space advantage for defending against the neighbours in nature, or to yield high final food or fibre products in agriculture and forestry. For QTL-effect patterns a–e (Fig. 1), we can simply pose the constraint

$$a_1 > a_2 \tag{A1}$$

because QTL genotype Qq, carrying a favourable allele, is supposed to have greater asymptotic growth than genotype qq, carrying no favourable allele. To incorporate the inequality constraint of A1, we introduce a new parameter x

$$x^2 = a_1 - a_2 > 0,$$

so

$$a_1 = x^2 + a_2. \tag{A2}$$

With this substitution, the vector for the unknown parameters to be estimated is changed from $\Omega = (a_j, b_j, r_j, \theta, \rho, \sigma^2)^T$ to $\Omega^* = (x, a_2, b_j, r_j, \theta, \rho, \sigma^2)^T$. The new vector Ω^* can be estimated using the Expectation (E)–Maximization (M) algorithm described by Ma *et al.* (2002).

For QTL-effect patterns f–j (Fig. 1), in which the two QTL genotypes have the same asymptotic growth $a_1 = a_2$, different constraints are needed to specify the identities of the two genotypes. Generally, it can be assumed that the QTL genotype with greater late growth (N.B. not asymptotic growth) contains a favourable allele. For QTL-effect pattern f (convergence), the constraints that can specify the identities of the QTL genotypes are given in Table A1, in which the QTL genotype containing a favourable allele also has greater initial growth and achieves the time at the inflection point earlier than the alternative genotype containing no favourable allele given this particular QTL-effect pattern. Referring to Fig. 1, similar sets of restrictions can be given for QTL-effect patterns g (divergence–convergence), h (early 'crossover'–convergence), i (middle 'crossover'–convergence) and j (late 'crossover'–convergence). These are all given in Table A1.

To incorporate these inequality restrictions for each pattern f–j, we introduce three new parameters x, y and z, and use them to substitute for a_1, b_1 and r_1 in the logistic mixture model of eqn 1. These substitutions are given in Table A1.

Similarly, with each of these substitutions (Table A1), we can obtain the MLEs of the corresponding new vector Ω^* for each QTL-effect patterns f–j. For all possible QTL-effect patterns listed in Fig. 1, we have a total of six substitutions (Eqn A2; Table A1) describing their properties. Under each of these substitutions, there are two alternative nonallelic assignments (1 and 2 for the coupling markers, or 3 and 4 for the repulsion markers). To determine a more-likely nonallelic assignment (specified by the conditional probabilities; Table 1) and a

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

Table A1. *Constraints posed on genotypic differences in growth at* t = 0 *and time at the inflection point and resulting substitutions for growth curve parameters in different QTL effect patterns* (*f, g, h, i and j*)

Pattern	Constraint	Substitution
f	$\begin{cases} g_1(0) > g_2(0) \\ t_{I1} < t_{I2} \end{cases}$	$\begin{cases} x = a_1 = a_2 \\ b_1 = \dfrac{(x-y^2)(1+b_2)-x}{y_2(1+b_2)+x} \\ r_1 = \dfrac{r_2 \log b_1}{\log b_2 - r_2 z^2} \end{cases}$
g	$\begin{cases} g_1(0) = g_2(0) \\ t_{I1} < t_{I2} \end{cases}$	$\begin{cases} x = a_1 = a_2 \\ y = b_1 = b_2 \\ r_1 = \dfrac{r_2 \log y}{\log y - r_2 z^2} \end{cases}$
h	$\begin{cases} g_1(0) < g_2(0) \\ t_{I1} < t_{I2} \end{cases}$	$\begin{cases} x = a_1 = a_2 \\ b_1 = \dfrac{(x+y^2)(1+b_2)-x}{x-y^2(1+b_2)} \\ r_1 = \dfrac{r_2 \log b_1}{\log b_2 - r_2 z^2} \end{cases}$
i	$\begin{cases} g_1(0) < g_2(0) \\ t_{I1} = t_{I2} \end{cases}$	$\begin{cases} x = a_1 = a_2 \\ b_1 = \dfrac{(x+y^2)(1+b_2)-x}{x-y_2(1+b_2)} \\ r_1 = \dfrac{r_2 \log b_1}{\log b_2} \end{cases}$
j	$\begin{cases} g_1(0) < g_2(0) \\ t_{I1} > t_{I2} \end{cases}$	$\begin{cases} x = a_1 = a_2 \\ b_1 = \dfrac{(x+y^2)(1+b_2)-x}{x-y_2(1+b_2)} \\ r_1 = \dfrac{r_2 \log b_1}{\log b_2 - r_2 z^2} \end{cases}$

most-likely QTL-effect pattern (specified by the substitutions above), we obtain the maximum likelihood estimates (MLEs) of the unknown parameters using the EM algorithm and compare the likelihood values for all 12 possible combinations between the assignments and substitutions. Finally, a most-likely combination is chosen based on the maximum of the likelihood values.

(b) *Time to the inflection point*

As illustrated in Niklas (1994), the growth rate at time (t_I) to the inflection point is maximum. An individual achieving t_I earlier is often thought to have evolutionary advantage in competing for space resources. Similarly, selection for the genotypes that achieve maximum growth rate more rapidly might be interesting to plant breeders. As a result, it is useful to describe QTL-effect patterns in terms of the time to the inflection point.

If the QTL genotype carrying a favourable allele achieves the inflection point more rapidly than the alternative genotype with no favourable allele, the constraint on t_I can be written as

$$t_{I1} < t_{I2}. \tag{A3}$$

If the two QTL genotypes achieve the inflection point at the same time, we assume that the QTL genotype of greater asymptotic growth carries a favourable allele, thus having the constraints

$$\begin{cases} t_{I1} = t_{I2} \\ a_1 > a_2. \end{cases} \tag{A4}$$

If the two QTL genotypes have the same t_I and the same asymptotic growth, as shown by QTL-effect pattern i (middle 'crossover'–convergence), then the QTL genotype with the greater growth potential after t_I is assumed to

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

carry a favourable allele. In this case, the constraints are given as

$$\begin{cases} t_{I1} = t_{I2} \\ a_1 = a_2 \\ G_1[0, t_o] < G_2[0, t_o], \end{cases} \tag{A5}$$

where the last constraint is imposed on the areas of the two growth curves from time $t=0$ to the time at which the curves cross, $t_o = (\log b_1 - \log b_2)/(r_1 - r_2)$, and the area of a QTL genotype between times 0 and t_o is calculated as

$$G_j[0, t_o] = \frac{a_j}{r_j}[\ln(b_j + e^{r_j t_o}) - \ln(b_j + 1)].$$

The three constraints expressed by eqns A3–A5 cover all possible QTL-effect patterns in Fig. 1. These constraints can be removed by introducing new parameters, which are estimated using the EM algorithm. Similarly, a more-likely nonallelic assignment and a most-likely QTL effect pattern described in terms of t_I can be obtained by comparing the likelihood values of all possible combinations.

(c) *Growth at a time point of typical interest*

In practice, growth at a particular time point might interest biologists or breeders. For example, many man-made forests of loblolly pine used for the pulp and paper industry are harvested at age 15 years in the south-eastern United States. Thus, final growth at this age is an important selection criterion. The patterns of QTL effect can be described in terms of the biomass growth of loblolly pine at age 15 years. The constraint for growth at a single time point is given as

$$g_1(t_h) > g_2(t_h), \tag{A6}$$

if we assume that QTL genotype Qq carries a favourable allele for the growth at time t_h.

(d) *Growth at a particular time interval*

We propose to use the area between two growth curves as a criterion to assess the contribution of a QTL to growth. The area at a time interval $[t_1\ t_2]$ is expressed as the integral of the differences of the two growth curves.

$$\begin{aligned} \Delta G(t_1, t_2) &= \int_{t_1}^{t_2} \left[\frac{a_1}{1 + b_1 e^{r_1 t}} - \frac{a_2}{1 + b_2 e^{r_2 t}}\right] \mathrm{d}t \\ &= \frac{a_1}{r_1}[\ln(b_1 + e^{r_1 t_2}) - \ln(b_1 + e^{r_1 t_1})] - \frac{a_2}{r_2}[\ln(b_2 + e^{r_2 t_2}) - \ln(b_2 + e^{r_2 t_1})]. \end{aligned}$$

If the time interval includes the point at which the two QTL genotypes cross (Fig. 1), the difference between the areas between the two genotypes has a different direction. In this case, the area can be calculated in two parts, one from $[t_1\ t_o]$ and the second from $[t_o\ t_2]$, where t_o denotes the crossing point. Depending upon the actual problem, researchers might select one of the two parts as a criterion for determining the allelic relationship of a putative QTL.

The four different criteria proposed for modelling the allelic relationship of a QTL might have different implications. The asymptotic growth and the time of achieving the inflection point are thought of a global search strategy for a QTL of any possible effect pattern. Thus, using these two criteria, all possible QTL-effect patterns can be identified. By contrast, the constraints on growth at a particular time point or interval permit us to detect the QTL of a particular effect pattern. For example, if stem growth at age 15 years is selected in a loblolly-pine breeding program, only those QTLs exerting an effect on growth difference at this age will be detected. Other QTLs responsible for growth at different ages, in which a breeder might not be interested, will not be detected.

References

Atchley, W. R. (1984). Ontogeny, timing of development and genetic variance–covariance structure. *American Naturalist* **123**, 519–540.

Atchley, W. R. & Zhu, J. (1997). Developmental quantitative genetics, conditional epigenetic variability and growth in mice. *Genetics* **147**, 765–776.

Cervera, M. T., Storme, V., Ivens, B., Gusmao, J., Liu, B. H., Hostyn, V., Van Slycken, J., Van Montagu, M. &

https://doi.org/10.1017/S0016672302005980 Published online by Cambridge University Press

Boerjan, W. (2001). Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics* **158**, 787–809.

Cheverud, J. M., Rutledge, J. J. & Atchley, W. R. (1983). Quantitative genetics of development – genetic correlations among age-specific trait values and the evolution of ontogeny. *Evolution* **37**, 895–905.

Cheverud, J. M., Routman, E. J., Duarte, F. A. M., van Swinderen, B., Cothran, K. & Perel, C. (1996). Quantitative trait loci for murine growth. *Genetics* **142**, 1305–1319.

Davidian, M. & Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data*. London: Chapman and Hall.

Doerge, R. W. & Churchill, G. A. (1996). Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285–294.

Emebiri, L. C., Devey, M. E., Matheson, A. C. & Slee, M. U. (1998). Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theoretical and Applied Genetics* **97**, 1053–1061.

Frewen, B. E., Chen, T. H. H., Howe, G. T., Davis, J., Rohde, A., Boerjan, W. & Bradshaw, H. D. (2000). Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics* **154**, 837–845.

Grattapaglia, D. & Sedero, R. (1994). Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* **137**, 1121–1137.

Jansen, R. C. & Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**, 1447–1455.

Kaya, Z., Sewell, M. M. & Neale, D. B. (1999). Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics* **98**, 586–592.

Kirkpatrick, M. & Heckman, N. (1989). A quantitative genetic model for growth, shape, reaction norms and other infinite-dimensional characters. *Journal of Mathematical Biology* **27**, 429–450.

Kirkpatrick, M., Hill, W. G. & Thompson, R. (1994). Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genetical Research* **64**, 57–69.

Kirkpatrick, M., Lofsvold, D. & Bulmer, M. (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* **124**, 979–993.

Klopfenstein, N. B., Chun, Y. W., Kim, M.-S. & Ahuja, M. R. (1997). Micropropagation, Genetic Engineering and Molecular Biology of *Populus*. General Technical Report RMGTR–297. Fort Collins, CO: Rocky Mountain Forest and Range Experiment Station.

Knott, S. A. & Haley, C. S. (2000). Multitrait least squares for quantitative trait loci detection. *Genetics* **156**, 899–911.

Korol, A. B., Ronin, Y. I., Itskovich, A. M., Peng, J. & Nevo, E. (2001). Enhanced efficiency of quantitative trait loci mapping analysis based on multivariate complexes of quantitative traits. *Genetics* **157**, 1789–1803.

Lander, E. S. & Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199.

Ma, C.-X., Casella, G. & Wu, R. L. (2002). Functional mapping of quantitative trait loci underlying the character process: a theoretical framework. *Genetics* **161**, 1751–1762.

Martinez, O. & Curnow, R. N. (1992). Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theoretical and Applied Genetics* **85**, 480–488.

Niklas, K. L. (1994). *Plant Allometry: The Scaling of Form and Process*. Chicago, IL: University of Chicago.

Pearl, R. (1925). *The Biology of Population Growth*. New York: Knopf.

Pletcher, S. D. & Geyer, C. J. (1999). The genetic analysis of age-dependent traits: modeling the character process. *Genetics* **153**, 825–835.

Rice, S. H. (1997). The analysis of ontogenetic trajectories: when a change in size or shape is not heterochrony. *Proceedings of the National Academy of Sciences of the USA* **94**, 907–912.

Stettler, R. F., Bradshaw, H. D. Jr, Heilman, P. E. & Hinckley, T. M. (1996). *Biology of Populus and its Implications for Management and Conservation*. Part I. Ottawa, Canada: CRC Research Press.

Vaughn, T. T., Pletscher, L. S., Peripato, A., King-Ellison, K., Adams, E., Erikson, C. & Cheverud, J. M. (1999). Mapping quantitative trait loci for murine growth: a closer look at genetic architecture. *Genetical Research* **74**, 313–322.

Verhaegen, D., Plomion, C., Gion, J. M., Poitel, M., Costa, P. & Kremer, A. (1997). Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers. 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theoretical and Applied Genetics* **95**, 597–608.

von Bertalanffy, L. (1957). Quantitative laws in metabolism and growth. *Quarterly Review of Biology* **32**, 217–231.

West, G. B., Brown, J. H. & Enquist, B. J. (2001). A general model for ontogenetic growth. *Nature* **413**, 628–631.

Wu, R. & Stettler, R. F. (1994). Quantitative genetics of growth and development in *Populus*. I. A three-generation comparison of tree architecture during the first two years of growth. *Theoretical and Applied Genetics* **89**, 1046–1054.

Wu, R. & Stettler, R. F. (1996). The genetic dissection of juvenile canopy structure and function in a three-generation pedigree of *Populus*. *Trees – Structure and Function* **11**, 99–108.

Wu, R. & Stettler, R. F. (1997). Quantitative genetics of growth and development in *Populus*. II. The partitioning of genotype × environment interactions in growth. *Heredity* **78**, 124–134.

Wu, R. L., Wang, M. X. & Huang, M. R. (1992). Quantitative genetics of yield breeding for *Populus* short rotation culture. I. Dynamics of genetic control and selection models of yield traits. *Canadian Journal of Forest Research* **22**, 175–182.

Wu, R., Bradshaw, H. D. & Stettler, R. F. (1998). Developmental quantitative genetics of growth in *Populus*. *Theoretical and Applied Genetics* **97**, 1110–1119.

Wu, R. L., Han, Y. F., Hu, J. J., Li, L., Li, M. L. & Zeng, Z.-B. (2000*a*). An integrated genetic map of *Populus* based on amplified fragment length polymorphisms. *Theoretical and Applied Genetics* **100**, 1249–1256.

Wu, R. L., Yin, D. M., Huang, M. R. & Wang, M. X. (2000*b*). The application of marker assisted selection to forest tree breeding. *Scientia Silvae Sinicae* **36**, 103–113.

Wu, R. L., Zeng, Z. B., McKeand, S. E. & O'Malley, D. M. (2000*c*). The case for molecular mapping in forest tree breeding. *Plant Breeding Reviews* **19**, 41–67.

Wu, R. L., Ma, C.-X., Chang, M., Littell, R. C., Wu, S. S., Yin, T. M., Huang, M. R., Wang, M. X. & Casella, G. (2002*a*). A logistic mixture model for characterizing genetic determinants causing differentiation in growth trajectories. *Genetical Research* **79**, 235–245.

Wu, R. L., Ma, C. X., Zhu, J. & Casella, G. (2002*b*). Mapping epigenetic quantitative trait loci (QTL) altering a developmental trajectory. *Genome* **45**, 28–33.

Yan, J. Q., Zhu, J., He, C. X., Benmoussa, M. & Wu, P. (1998). Molecular dissection of developmental behavior of plant height in rice (*Oryza sativa* L.). *Genetics* **150**, 1257–1265.

Yin, T. M., Zhang, X. Y., Huang, M. R., Wang, M. X., Zhuge, Q., Zhu, L. H. & Wu, R. L. (2002). Molecular linkage maps of the *Populus* genome. *Genome* **45**, 541–555.

Zhu, J. (1995). Analysis of conditional genetic effects and variance components in developmental genetics. *Genetics* **141**, 1633–1639.