

Mechanistic Mapping of Ontogenetic Growth Based on Biological Principles

Min Lin,* Chang-Xing Ma,* Wei Zhao,* James M. Cheverud[†] and Rongling Wu*

**Department of Statistics, University of Florida, Gainesville, Florida 32611 USA and*

[†]Department of Anatomy and Neurobiology, Washington University School of Medicine, St.

Louis, Missouri 63110 USA

Author for correspondence:

Rongling Wu

Department of Statistics

533 McCarty Hall C

University of Florida

Gainesville, FL 32611

Phone (352)392-3806

FAX: (352)392-8555

E-mail: rwu@stat.ufl.edu

Ontogenetic growth can be described by mathematical equations constructed on the goodness of fit. Recently, the biological mechanism underlying mathematical growth equations has been explored using basic cellular properties. Here, we derive a general statistical model for understanding the genetic regulation of ontogenetic growth by integrating those biologically-proven meaningful growth equations into a quantitative trait locus (QTL) mapping framework. We can characterize the dynamic patterns of effects of QTL governing growth curves and estimate the global effect of the underlying QTL throughout the entire course of growth. The model provides the basis for deciphering genetic relationships for growth rates and the timing of life history events for any kind of organisms.

1. INTRODUCTION

According to von Bertalanffy (1957), growth in weight, length or size will occur whenever the anabolic or metabolic rate exceeds the rate of catabolism. Thus, the ratio of these two processes indicates the occurrence and change of growth. When the ratio approaches and then drops below unity, growth will decrease and eventually cease. Bertalanffy also noted that the metabolic rate of an animal (in fact, any organism, as proven by the succeeding researchers (West et al. 1997, 1999, 2001) scales as the k th power of its weight but the catabolic rate is proportional to the weight itself. Therefore, the growth rate, i.e., the difference between these two rates becomes

$$\frac{dg}{dt} = \eta g^k - \kappa g, \quad (1)$$

where g represents the growth at time t and η and κ are the constants of metabolism and catabolism, respectively. For small values of k , integration of equation (1) leads to the growth equation,

$$g = \left\{ \frac{\eta}{\kappa} - \left[\frac{\eta}{\kappa} - g_0^{1-k} \right] e^{-(1-k)\kappa t} \right\}^{\frac{1}{1-k}}, \quad (2)$$

where g_0 is the growth at $t = 0$. This growth function is sigmoidal (S-shaped), approaching asymptotically the value $(\eta/\kappa)^{1/(1-k)}$ as $t \rightarrow \infty$. Such an S-shaped pattern of growth includes an exponential growth stage, an asymptotic growth stage and the point of inflection at which these two stages are connected (Niklas 1994). At the point of inflection, an organism displays maximum growth per unit time. After the substitutions,

$$\begin{cases} a^{1-k} &= \frac{\eta}{\kappa} \\ b &= \pm \left(\frac{\eta}{\kappa} - g_0^{1-k} \right) a^{k-1} \\ r &= (1-k)\kappa \end{cases},$$

where a is the limiting growth, b is a parameter related to the growth at $t = 0$, and r is the “rate constant” that determines the spread of the curve along the time axis, Richards (1959) rewrites equation (2) as

$$g = a(1 + be^{-rt})^{1/(1-k)}. \quad (3)$$

Equation (3), referred to as the Richards growth model, has the inflection point whose coordinates are solved as $(t = \ln[\frac{b}{k-1}]/r, g = ak^{\frac{1}{1-k}})$. By letting k take different values, the Richards growth model can be reduced to three well-known growth functions,

$$g = \begin{cases} a(1 + be^{-rt}) & \text{Monomolecular curve when } k = 0 \\ ae^{-be^{-rt}} & \text{Gompertz's curve when } k = 1 \\ \frac{a}{1 + be^{-rt}} & \text{Autocatalytic or logistic curve when } k = 2. \end{cases} \quad (4)$$

The monomolecular curve actually describes the asymptotic phase of the logistic growth curve. The logistic curve is symmetric about the inflection point at which $t = (\ln b)/r$ and $g = a/2$. Yet, the Gompertz curve is not symmetric, with the inflection point occurring at $t = (\ln b)/r$ and $g = a/e$. These three curves, with different flexibilities, may best fit growth data collected from a different species, organ, or in a different environment.

We have demonstrated that growth laws can be incorporated into a mixture model composed of different components, by which we can map quantitative trait loci (QTL) responsible

for the differentiation in growth trajectories (Wu et al. 2002; Ma et al. 2002). Here, we generalized the growth law-based mapping model to show its universal properties in a different biological model system – mouse and its biological relevance in asking, disseminating and answering questions at the interface among genetics, development and evolution. Statistically, we used the Legendre polynomial approach, as advocated by Kirkpatrick and Heckman (1989) and Kirkpatrick et al. (1990, 1994), for modelling the structure of the residual covariance matrix in QTL mapping.

2. MAPPING MODEL

The basic statistical model for mapping QTL is a mixture model, in which each observation y is assumed to have arisen from one of k groups of QTL genotypes, each group being modelled by a density from the parametric family f . The population density function of y is

$$p(y|\pi, \phi, \eta) = \pi_1 f(y; \phi_1, \eta) + \dots + \pi_k f(y; \phi_k, \eta)$$

where $\pi = (\pi_1, \dots, \pi_k)$ are the mixture proportions which are constrained to be non-negative and sum to unity; $\phi = (\phi_1, \dots, \phi_k)$ are the genotype-specific parameters, with ϕ_j being specific to genotype group j ; and η is a parameter which is common to all genotype groups. The mixture proportions denoted as the frequencies of QTL genotypes depend on the marker genotypes of two flanking markers bracketing the QTL. The normal density functions associated with different QTL genotypes are expressed in terms of the expected value of each genotype.

For QTL mapping of ontogenetic growth, we measure multivariate phenotypes for each individual (denoted by a vector \mathbf{y}) on a single trait measured at many different time points (say τ). Thus, the phenotypes of each QTL genotype group follow a multivariate normal density,

$$f_j(\mathbf{y}) = \frac{1}{(2\pi)^{m/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[-(\mathbf{y} - \mathbf{u}_j)^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{u}_j) / 2 \right],$$

where $\mathbf{u}_j = (u_{j1}, \dots, u_{j\tau})$ is the vector of the expected genotypic values of QTL genotype j for all τ measurements and Σ is the $(\tau \times \tau)$ residual variance-covariance matrix of the variables. Indeed, \mathbf{u}_j can be modelled by a time-dependent mathematical function, e.g, equation (3). Thus, estimating the genotypic values (and therefore additive, dominant and epistatic genetic effects) is equivalent to estimating the mathematical parameters that specify the shape of growth curves. In modelling functional mapping, we usually estimate the structured form of the variance-covariance matrix to increase the model's power rather than estimate every element in the matrix. Wu et al. (2002) and Ma et al. (2002) used the first-order autoregressive [AR(1)] model to structure the covariance matrix, but this assumes that the variance is constant over time and the correlation decay in a proportion with time interval. Many other approaches have been proposed to structure the matrix (Kirkpatrick et al. 1994; Pletcher et al. 1999; Jaffrezix and Pletcher 2000). In this article, we introduce Kirkpatrick and Heckman's (1989) Legendre polynomial approach to model the structure of the residual covariance matrix Σ .

3. RESULTS

Our newly-developed model is used to map age-dependent QTL in a model system – mouse. Cheverud et al. (1996) constructed a linkage map based on 75 microsatellite markers in 535 F₂ progeny population derived from two strains, the Large (LG/J) and Small (SM/J). Initiated with these two same strains, a second F₂ population of 502 mice was generated, in which Vaughn et al. (1999) constructed a linkage map by increasing microsatellite markers to 96. In both experiments, the F₂ progeny was weighted at 10 weekly intervals starting at age 7 days. The raw weights were corrected for the effects of each covariate due to dam, litter size at birth, parity and sex. Figure 1 demonstrates statistical significance for the sigmoidal shape of body weight, as a function of age during the first 10 weeks of growth in the growth chamber, for one mouse (#44), randomly drawn from the first F₂ population (Cheverud et

al. 1996).

By pooling the two F_2 populations together, a new linkage map was constructed to cover 19 mouse chromosomes. Our model identified several significant QTL located on different chromosomes for the growth process of body weight in the pooled F_2 population. Figure 1 illustrates the profile of the log-likelihood ratio (LR) test statistics for claiming the existence of a growth QTL on mouse chromosome 1. This QTL located between markers D1Mit14 and D1Mit17, as indicated by a peak of the LR profile, exerts an age-specific effect on growth. The maximal LR value is 248, beyond the $\alpha = 0.001$ critical threshold determined from permutation tests.

The three growth curves estimated each corresponding to one of the three QTL genotypes, QQ , Qq and qq , in the F_2 population, are compared in Figure 3A. Statistical tests suggest that the three curves are not significantly different before ages 4 weeks. Increased differentiation in growth occurs after age 5 – 6 weeks (results not shown). This indicates that the detected QTL triggers varying impacts on growth depending on stages of mouse development. The overlapping of the two homozygotes (QQ and qq) implies that there is no additive effect on growth trajectories, whereas the divergence of the heterozygote (Qq) from the two homozygotes implies that the dominant effect on body weight growth is increased with age (Fig. 3B).

We further analyzed the difference in growth rate among the three genotypes at the QTL detected on mouse chromosome 1. The inflection point, at which growth rate is maximal ($d^2g(t)/dt^2 = 0$; see equation 2), occurs about 3 – 4 days earlier for the heterozygote Qq than the homozygotes QQ and qq (Fig. 3B). It is possible that this difference in development causes the homozygote Qq to reach the asymptotic growth earlier than the other genotypes. With the three different growth curves, each corresponding to a QTL genotype, we can investigate possible pleiotropic effects of this growth QTL on many different developmental

events, such as the timing of sexually maturity and reproductive fitness, or biomedically important traits, such as metabolic rate and fatness. We can therefore integrate growth and development, which are historically regarded as two different biological problems, into a comprehensive framework under which their common or unique underlying genetic machineries are identified.

DISCUSSION

A number of statistical methods have been proposed to map quantitative trait loci (QTL) underlying complex phenotypes primarily based on the goodness of fit to observational data rather than on any biological mechanism. We have framed a new statistical strategy for QTL mapping through specific incorporation of biological laws behind the phenotypic expression of complex traits. The new strategy, termed *functional mapping* (Ma et al. 2002), displays great potential in improving the precision, power and resolution of QTL mapping in any kind of organism. Functional mapping grounds theoretical genetic models in integrated developmental networks or processes and, consequently, has a direct impact on the interface of development and evolution. With this new strategy, we are in an excellent position to address two major challenging questions in contemporary developmental biology regarding the existence of particular regulatory genes guiding growth differentiation throughout an entire biological process and the alteration of the genetic architecture of a complex trait over developmental times.

This work is supported by an Outstanding Young Investigator Award of the National Natural Science Foundation of China (30128017), a University of Florida Research Opportunity Fund (02050259) and a University of South Florida Biodefense grant (7222061-12) to R. W. The publication of this manuscript is approved as Journal Series R-08963 by the Florida Agricultural Experiment Station.

References

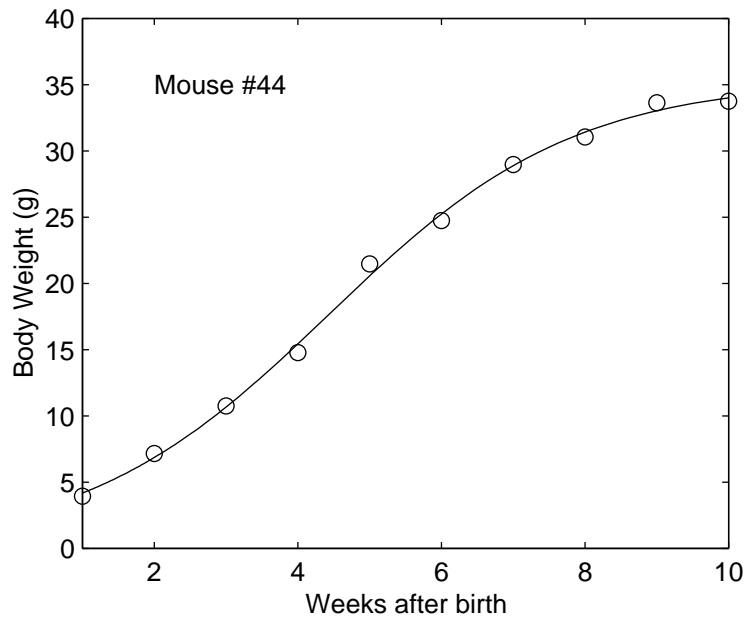
- Bertalanffy, V. L., 1957 Quantitative laws for metabolism and growth. *Quart. Rev. Biol.* 32: 217-231.
- Cheverud, J. M., E. J. Routman, F. A. M. Duarte, B. van Swinderen, K. Cothran and C. Perel, 1996 Quantitative trait loci for murine growth. *Genetics* 142: 1305-1319.
- Churchhill, G. A., and R. W. Doerge, 1994, Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963-971.
- Jaffrezix, F., and S. D. Pletcher, 2000 Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics* 156: 913-922.
- Kirkpatrick, M., and N. Heckman, 1989 A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.* 27: 429-450.
- Kirkpatrick, M., W. G. Hill and R. Thompson, 1994 Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genet. Res.* 64: 57-69.
- Kirkpatrick, M., D. Lofsvold and M. Bulmer, 1990 Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124: 979-993.
- Lander, E. S., and D. Botstein, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185-199.
- Ma, C. X., G. Casella, and R. L. Wu, 2002 Functional mapping of quantitative trait loci underlying the character process: A theoretical framework. *Genetics* 161: 1751-1762.
- Niklas, K. L. (1994) *Plant Allometry: The Scaling of Form and Process*. University of Chicago, Chicago, IL.
- Pletcher, S. D., and C. J. Geyer, 1999 The genetic analysis of age-dependent traits: Modeling

- the character process. *Genetics* 153: 825-835.
- Richards, F. J., 1959 A flexible growth function for empirical use. *J. Exp. Bot.* 10: 290-300.
- Vaughn, T. T., L. S. Pletscher, A. Peripato, K. King-Ellison, E. Adams, C. Erikson and J. M. Cheverud, 1999 Mapping quantitative trait loci for murine growth – A closer look at genetic architecture. *Genet. Res.* 74: 313-22.
- West, G. B., J. H. Brown and B. J. Enquist, 1997 A general model for the origin of allometric scaling laws in biology. *Science* 276: 122-126.
- West, G. B., J. H. Brown, and B. J. Enquist, 1999 The fourth dimension of life: Fractal geometry and allometric scaling of organisms. *Science* 284: 1677-1679.
- West, G. B., J. H. Brown and B. J. Enquist, 2001 A general model for ontogenetic growth. *Nature* 413: 628-631.
- Wu, R. L., C. X. Ma, M. Chang, R. C. Littell, S. S. Wu, M. Huang, M. Wang and G. Casella, 2002 A logistic mixture model for detecting major genes governing growth trajectories. *Genet. Res.* 79: 235-245.

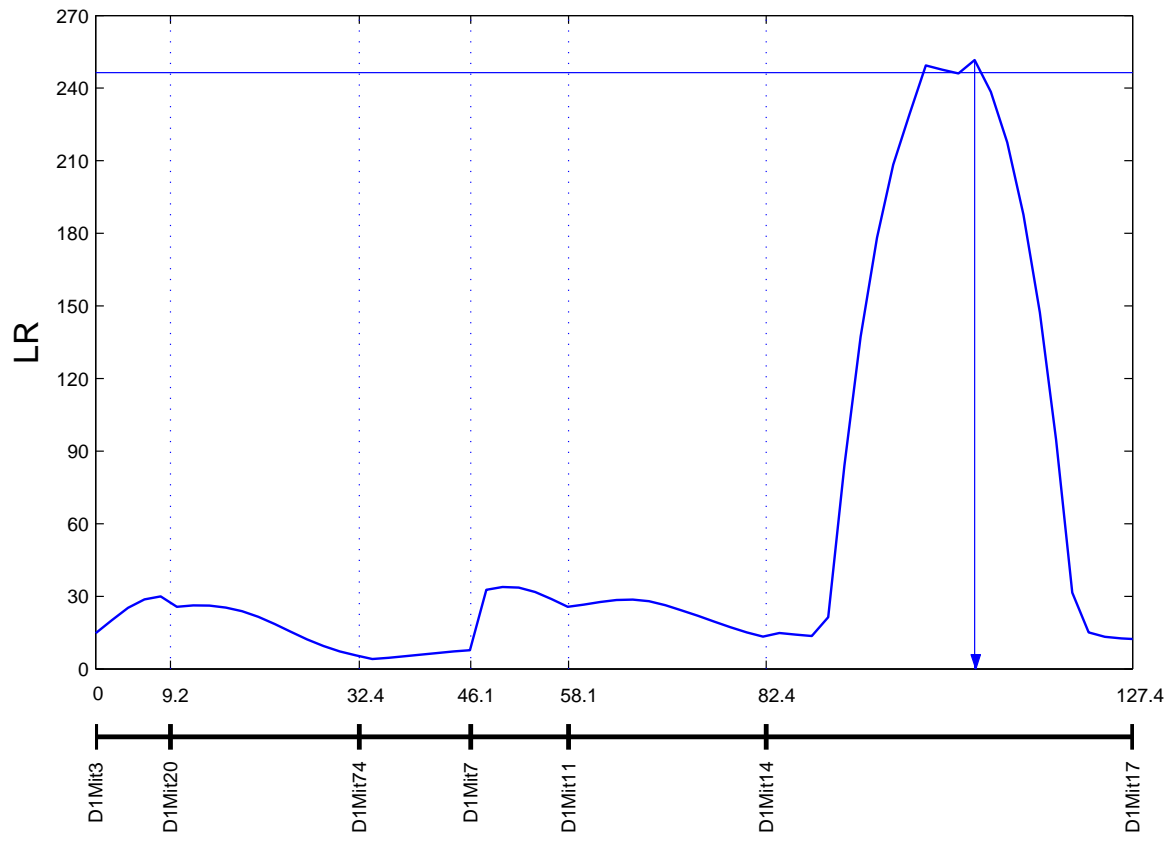
Figure 1.—One typical example of fit to growth curve using logistic equation from an F_2 population of two mouse inbred strains. The data from Cheverud et al. (1996).

Figure 2.—The profile of the log-likelihood ratios (LR) between the full model (there is a QTL) and reduced (there is no QTL) model for body weight growth trajectories across chromosome 1 for the mouse map. The genomic position corresponding to the peak of the curve is the maximum likelihood estimate of the QTL localization. The vertical broken lines indicate the positions of markers on this chromosome shown beneath. The map distances (in centiMorgan) between two markers are calculated using the Haldane mapping function. The thresholds for proclaiming the genome-wide existence of a QTL are obtained from permutation tests (Churchill and Doerge 1994). The distribution of each of the LR values over 1000 simulation replicates can be approximated by a χ^2 distribution. The 99.9th percentile of the distribution of the maximum is used as an empirical critical value to declare genome-wide existence of a QTL at $\alpha = 0.001$. This percentile is indicated at a horizontal line. The data pooled from Cheverud et al. (1996) and Vaughn et al. (1999).

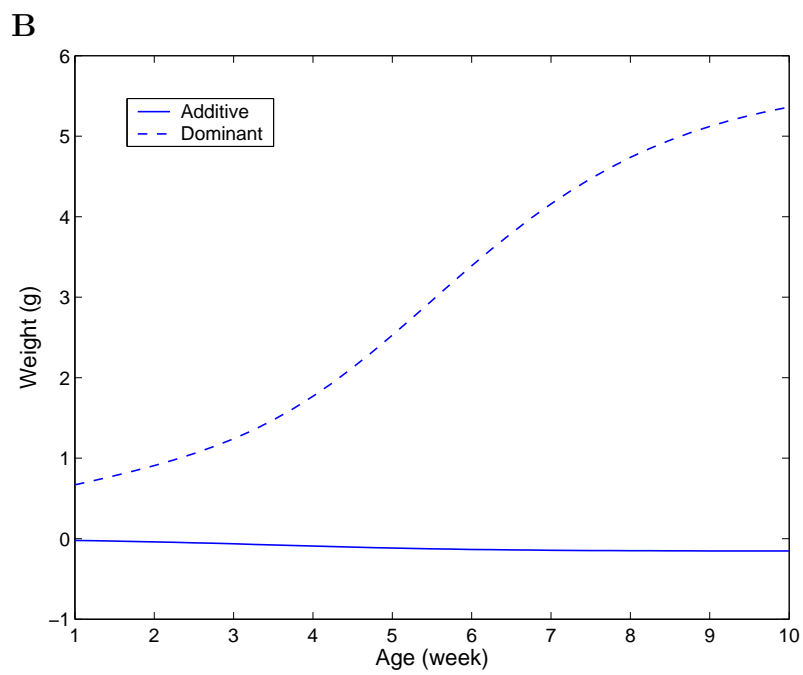
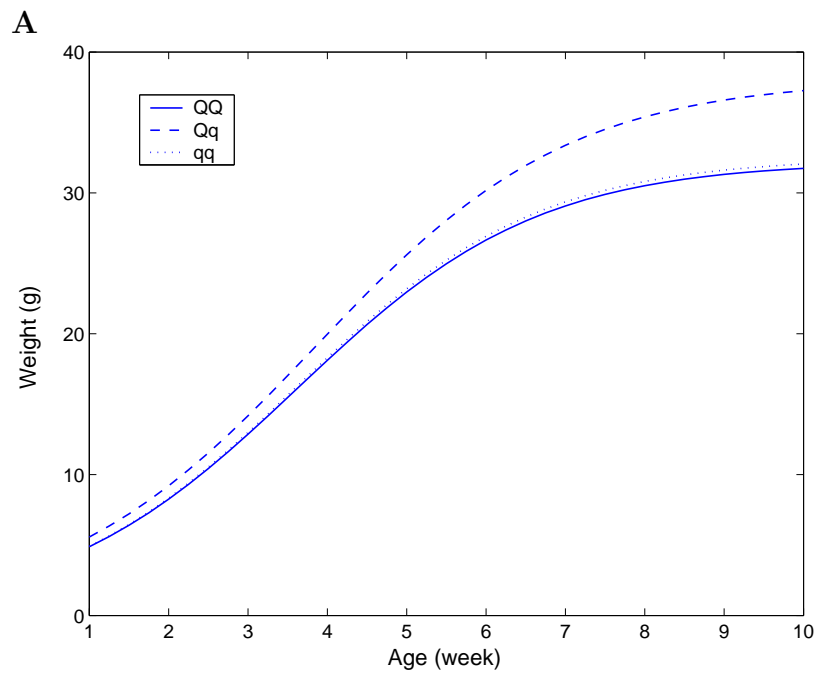
Figure 3.—**(A)**: Three growth curves each presenting a group of genotypes at the QTL detected on chromosome 1 in the F_2 population of mice. **(B)**: The dynamics of the additive and dominant effects on body mass growth.



(Figure 1)



(Figure 2)



(Figure 3)