Journal of Genetics (2019) 98:39
https://doi.org/10.1007/s12041-019-1089-y
© Indian Academy of Sciences

REVIEW ARTICLE

Evolutionary dynamics models in biometrical genetics supports QTL × environment interactions

FARSHAD FATTAHI* and BARAT ALI FAKHERI

Faculty of Agriculture, Department of Plant Breeding and Biotechnology, University of Zabol, Sistan and Baluchestan 98615538, Iran
*For correspondence. E-mail: frs_fth@yahoo.com.

Received 4 May 2017; accepted 21 March 2019; published online 26 April 2019

Abstract. The process of development of quantitative trait locus (QTL) involves interactions between many factors, both environmental and genetic, in which many genes interact often in no additive pathways together and with environment. Integration of the mathematical, statistical and biological aspects of these subjects has made important and interesting results. In this review, mathematical methods offered to study the QTL × environment interactions. The topic is circumscribed, going from basic selection equations to models of evolution of QTLs. Discrete and continuous time mathematical models and subsequently, QTL modelling were introduced with and without environmental interactions. The mathematical models derived here showed that the gradients of mean fitness which have revealed in studies by many researchers had a basic role in mathematical genetics, evolutionary aspects of biometrical genetics and QTL analysis. QTL × environment interactions were studied mathematically including fitness components too. It was revealed that QTL × environment interactions in fitness could generate a balancing selection. Also, QTL analysis could be used to calculate the geometry of the phenotype landscape. In this paper, models applied in biometrical genetics corresponds to QTL analysis and matched with results from other researchers. The originality of this synthesis is the evolutionary modelling of QTL × environment interactions which can be used to investigate the extinction or stability of a population. Also to emphasize that although some scientific subjects like Brownian motion, quantum mechanics, general relativity, differential geometry, and evolutionary biometrical genetics were apparently different subjects, but the mathematical models were the backbone of these branches of science. This implies that such matters in nature have probably common and elegant basis. The perspective of the subject of this paper in future will be a new and interesting branch of interdisciplinary science.

Keywords. evolutionary dynamics; genotype × environment interactions; mathematical modelling; quantitative trait locus.

Introduction

There are two alternate but closely related ways to deal with genotype × environment interactions if the environmental variables are known, respectively for discrete time versus continuous time models. Via and Lande (1985) showed a novel insight in genotype × environment interactions based on pervious works of Falconer (1952 and 1981).

Evolutionary dynamics is the study of the mathematical procedure according to evolution of biological organisms. Also, population genetics investigate differences in the frequencies of alleles at a small number of genes. When very small effects of a large number of genes are studied, quantitative genetics is concluded (Nowak 2006). Namely, quantitative traits have polygenic control which show continuous variation within or among populations (Falconer and MacKay 1996).

Evolutionary biometrical genetics has progressed in two different ways. Some scientists hint evolutionary problems through calculating phenotypic selection and genetic variances, and covariances, i.e. G matrices. On the other hand, many studies have been performed for quantitative trait locus (QTL) cartography to investigate variation (Kelly 2009).

QTLs consist of genes containing alleles with different effects, in which some alleles are recognizable with large effects and others have small effects resulting quantitative variation (Robertson 1985). The use of genetic markers to infer genetic structure is named QTL analysis that contributes to a quantitative trait (Lynch and Walsh 1998).

In QTL mapping, the evolution of the mean phenotype expressed in different environments is called the norm of reaction. Mathematical models used in this paper describe the evolution of phenotypic response of QTLs to the varying environments, in other words, phenotypic plasticity or QTL × environment interaction (Via and Lande 1985).

The aim of this review was to study the mathematical methods which are applied in dynamical modelling of biometrical genetics and analytical biology to investigate the QTL × environment interactions. The topic is circumscribed, going from basic selection equations to models of evolution of QTLs. Discrete and continuous time mathematical models and subsequently, QTL modeling are introduced with and without environmental interactions.

Mapping population

QTL cartography consists of two parts, namely phenotyping and genotyping, and involves a number of steps. In phenotyping part, the traits are measured while in genotyping part the genotypic data based on the molecular markers are determined. Some of the steps in QTL cartography are as follows. First, two or more organisms (parents) should be different in their alleles that affect variation in a traitbut it is not necessary that the parents to be different in the mean phenotypic value of the trait, because different allelic combinations present the same phenotypic mean. Second, molecular genetic markers are used to recognize the parental lines. Thus, a polymorphic genetic map causes the two races to be determined genetically (Miles and Wayne 2008).

Further, some mating designs are used to produce mapping population in which the parents are crossed to generate heterozygous F_1 individuals and also, this population is crossed using one of the mating designs (Rodney 2001). Eventually, the F_2 population (phenotypes and genotypes) are scored and the molecular markers that are genetically related to a QTL affecting the investigated trait will segregate with trait values while unrelated molecular markers will not present significant connection with phenotype (Miles and Wayne 2008).

Molecular aspects

Progress in molecular marker methods, genomics and the relevant software (like QTL Network and QTL Cartographer) for analysis of genetic and phenotypic data have concluded in the application of QTL analysis in biosciences (Mackay 2001; Yang *et al.* 2008). Since the molecular markers are implausible to effect the investigated traits, they are applied for genotyping (Miles and Wayne 2008).

Amplified fragment length polymorphism (AFLP) and microsatellites (simple sequence repeat (SSR)) are some of the molecular genetic markers applied to produce the new linkage maps and QTL researches (Zhang *et al.* 2004).

One of the high resolution DNA markers that are used in gene mapping is single-nucleotide polymorphism (SNP) and also is called snip that is a change at a particular location in the genome in which the upper DNA molecule varies from the lower DNA molecule at a single base-pair position. SNPs molecular markers are biallelic and occur in coding and noncoding regions of genes, thus they are easily investigated (Tahira *et al.* 2009). SNPs are suggested as molecular markers utilize in QTL analysis, because they are maintained during evolution. SNP analysis has many applications in biomedical, agricultural, pharmacogenetics and forensics researches.

Population genetics and basic selection equations

The foundation of population genetics was based on the independent inductive law of Hardy 1908 and Weinberg 1909 (Akin 1979). This law states that in a large population with random mating and in the absence of systematic factors, i.e. migration, mutation and selection, the frequency of genes will be fixed from one generation to the next generation. If the environmental conditions are unchanged, then the gene pool would be in equilibrium, i.e. the relative frequencies of genes and genotypes would be fixed, and just random changes could be occurred. The relative frequencies of genes and genotypes are correlated together which is named coadaptation, thus change in one will result change of other.

Basic selection equations were introduced by Wright (1932), Fisher (1937) and Kolmogorov *et al.* (1937). Wright, Fisher and Haldane were the founders of modern evolutionary genetics. Wright proposed the inbreeding coefficient and the ways of evaluating it in animal pedigrees. Wright and Fisher suggested the methods for calculating the gene frequencies in populations as the outcome of interactions among mutation, selection, genetic drift and migration.

On the other hand, Fisher (1937) and Kolmogorov *et al.* (1937) studied the mathematical problem of genes nondominance case in which, the viability of heterozygote genotype lies between the viabilities of homozygotes genotypes, using partial differential equations and diffusion equation. Fisher's equation (KPP equation or Fisher–KPP equation) is a semilinear reaction-diffusion equation. Fisher (1937) used Fisher–KPP equation in dynamical population genetics and investigated the spatial spread of an allele and the travelling wave solutions. Also, he studied the equilibrium states and the stability of travelling wave solutions against near and far field perturbations.

Also, Wright–Fisher genetic model including stochastic processes and Gaussian diffusion (white noise) was

studied by Norman (1975). He investigated the random fluctuations in gene frequencies under the influence of mutation, selection and random drift.

Further, Aronson and Weinberger (1978) studied the genes dominance, over-dominance and inverse over-dominance cases and they investigated the hair-trigger and threshold effects. In over-dominance case, the viability of heterozygote genotype is greater than the viabilities of homozygotes genotypes. Also in inverse over-dominance case, the viability of heterozygote genotype is lesser than the viabilities of homozygotes genotypes. Aronson and Weinberger (1978) showed that there were threshold effects in inverse over-dominance case and there were hair-trigger effects in nondominance and over-dominance cases.

In evolutionary biology, fitness (w or ω in population genetics) is the quantitative representation of selection. It describes individual reproductive success and is equal to the average contribution to the gene pool of the next generation that is made by individuals of the specified genotype or phenotype. The fitness of a genotype is presented through its phenotype, which is affected by the changing environment. The fitness of a phenotype can be different in changing selective environments (Ewens 2003; James and Kimura 1970).

Discrete time modelling of the selection dynamics

Diploid models

Suppose a diploid population in random mating with discrete nonoverlapping generations. Also, suppose that at a given gene locus, the I alleles $A_1, \ldots, A_I$ occur. Denote the frequency of allele A_i by p_i and the fitness of genotype A_iA_j by $w_{ij} = w_{ji} \rangle 0$ (Bürger 2011). Thus, the fitness of allele A_i and the mean fitness of the population are:

$$w_i = \sum_j w_{ij} p_j \quad \text{And} \quad \bar{w} = \sum_i w_i p_i = \sum_{i,j} w_{ij} p_i p_j, \tag{1}$$

respectively. The frequency p_i' of allele A_i in the next generation is

$$p_i' = p_i \frac{w_i}{\bar{w}} \quad \text{For} \quad i = 1, \ldots, I. \tag{2}$$

This equation satisfies the relation $\sum_i p_i = 1$ and presents the evolution of allele frequencies. It shows a discrete dynamical system on the simplex

$$S_I = \left\{ p = (p_1, \ldots, p_I)^T \in \mathbb{R}^I : \sum_i p_i = 1, \right.$$
$$\left. p_i \geq 0, i = 1, \ldots, I \right\}. \tag{3}$$

(Bürger 2011).

Formerly, Akin considered the unexpected occurrence of stable limit cycles in the two locus, two allele model, using Kimura's maximum principle, Sewall Wright's adaptive topography and Lyapunov function; Although his work was done on the related continuous time differential equations model (see section 'Continuous time modeling of the selection dynamics' below). The cycles resulted from the interaction between recombination and natural selection. Akin showed that the adaptive surface picture represented the effects of the forces of natural selection and recombination (Akin 1982).

Also, Akin suggested modelling genetic drift, through consideration of a Brownian motion and remarked on the cost of drift as the local optimum equilibrium point of pure selection. This valuable model in analytical biology comes from Kimura's subtle deployment of rate estimates in the neutral genes (Akin 1987).

Haploid models

If constants v_i exist such that $w_{ij} = v_i v_j$ for every i and j, then the equation 2 converts to the selection dynamics of a haploid population (Bürger 2011). Coppersmith *et al.* (1999) suggested that the time evolution of the distribution of alleles in population genetics models in haploid species with mutation (as potential energy), selection (as kinetic energy), and pleiotropy (which means that one gene affects more than one trait) could be studied with models in quantum mechanics, specifically Bose-Einstein condensation and motion of a particle in a central potential. They mapped Waxman and Peck's model which was based on the genetic model of Kimura (1965), into the mean-field theory for Bose condensation performed on the discrete-time model, and a variant of the model onto a bound-state problem in quantum theory (Waxman and Peck 1998).

Also, the discrete-time model mapped onto a continuous-time quantum mechanics problem of a particle; both the discrete and continuous time models showed qualitatively identical behaviour. Close relation between the continuous time version of haploid models of population genetics and time-dependent Schrödinger's equation was found on the long-time behaviour of haploid models (Bürger and Bomze 1996; Coppersmith *et al.* 1999). In both the discrete and continuous time models, a unique genotype could emerge only when the number of traits affected by each gene was bigger than two; this was a result similar to Waxman and Peck (1998) and Coppersmith *et al.* (1999) conclusions.

Continuous time modelling of the selection dynamics

To make the continuous time model of the selection dynamics in equation 2, it is necessary to rewrite time and fitnesses as follows

$$w_{ij} = 1 + \varepsilon m_{ij} \quad \text{And} \quad t = \left\lfloor \frac{\tau}{\varepsilon} \right\rfloor, \tag{4}$$

Here m_{ij} is the Malthusian fitness of A_iA_j and ε can be considered as the selection intensity or the generation time. If $m_i = \sum_i m_{ij}p_j$ and $\bar{m} = \sum_{i,j} m_{ij}p_ip_j$ show the marginal fitness of allele A_i and the mean fitness of population respectively, then setting $q_i = q_i(\tau) = p_i(t)$, the equation 2 is modelled such that the limit of $(q_i(\tau + \varepsilon) - q_i(\tau))/\varepsilon$ as $\varepsilon \to 0$ can be calculated. Referring to the notation p_i and t instead of q_i and τ, it is possible to obtain the weak selection approximation,

$$\dot{p}_i = p_i(m_i - \bar{m}), \quad i = 1, \ldots, I. \tag{5}$$

Which is a dynamical system on the simplex S_I and has the same equilibriums (Bürger 2011).

Consider the selection dynamics in equation 5 to provide interesting results. Thus, let the indicator variable

$$f_i(A_kA_l) = \begin{cases} 1 & \text{if } k = l = i, \\ \frac{1}{2} & \text{if } k \neq l \text{ and } k = i \text{ or } l = i, \\ 0 & \text{otherwise}, \end{cases} \tag{6}$$

where i is fixed and, k and l are independent and have probability distribution p. Thus, f_i measures the frequency of allele A_i in a given genotype (Bürger 2011) and has expectation p_i. Also, associated $I \times I$ covariance matrix is designated $\mathbf{G}_p = \left(g^{ij}\right)$ where

$$g^{ij} = Cov\left(f_i, f_j\right) = \frac{1}{2}p_i\left(\delta_{ij} - p_j\right), \tag{7}$$

and δ_{ij} is the Kronecker delta.

Therefore, the allele-frequency dynamics in equation 5 can be represented as a generalized gradient (∇) system on S_I

$$\dot{p} = \mathbf{G}_p\nabla\bar{m} = \mathbf{G}_p\left(\frac{\partial\bar{m}}{\partial p_1}, \ldots, \frac{\partial\bar{m}}{\partial p_n}\right)^T. \tag{8}$$

(Svirezhev 1972; Shahshahani 1979; Bürger 2011).

Selection produces nonrandom relationships among different loci which are named linkage disequilibria. Recombination breaks up these relationships but changes gamete frequencies. Thus, if all gamete frequencies are the products of the frequencies of the constituent alleles, therefore the population assumed to be in linkage equilibrium.

Gradient of mean fitness and differential geometry of population genetics

The differential equations which model the dynamics of QTL × environment interactions are nonlinear equations that are hard to solve and even difficult to describe the qualitative behaviour of their solutions (Akin 1979). Shahshahani first began using differential geometry to study these equations (Shahshahani 1979).

'Fisher's fundamental theorem of natural selection says that along the solutions curves of the selection differential equation, mean fitness is constantly increasing. Also, Kimura's maximum principle says that the direction of motion is the direction of greatest increase' (Ewens 2003; James and Kimura 1970).

Assuming that the diploid zygotes which make up the population are obtained by the pairing of haploid gametes and this pairing is random based on the Hardy–Weinberg law and also, using linear algebra, systems of differential equations and manifold theory, Akin suggested that the selection vector field associated with the selection differential equation, should be the gradient of mean fitness (Akin 1979).

Using the Riemannian metric which introduced by Shahshahani and Conley, Akin proved some theorems from differential geometry that were essential in mathematical biology; this was based on the linear algebra on Euclidean vector spaces and calculus on Riemannian manifolds. He applied the Shahshahani metric to population genetics and turned to evolutionary games and compared the evolutionary stable strategy (ESS) condition with other notions of stability (Akin 1990).

Riemannian metric on a manifold M associates a symmetric positive definite matrix $g(x) = (g_{ij}(x))$ to each point $x \in M$ in a smooth manner and the inner product of two vectors u and v in the tangent space T_xM at x is given by

$$\langle u, v\rangle_x = \sum_{ij} g_{ij}(x)u_iv_j = u \cdot (g(x)v)$$

where $a \cdot b$ is the Euclidean inner product of the vectors a and b. If M is an open subset of $\mathbb{R}^n$, then $T_xM = \mathbb{R}^n$. If M is the n-simplex $S_n = \left\{x \in \mathbb{R}^n | \sum_i x_i = 1\right\}$, then, for $x \in \text{int } S_n$, we have $T_xS_n = \mathbb{R}^n_0 = \left\{u \in \mathbb{R}^n | \sum_i u_i = 0\right\}$.

Shahshahani introduced a general Riemannian metric $\sum_i (|x|/x_i)dx_i \otimes dx_i$, in the positive orthant $\left\{x = (x_1, \ldots, x_n) \in \mathbb{R}^n | x_i) 0\right\}$, with $|x| = \sum_{i=1}^n x_i$. This metric which is named after him Shahshahani metric, is equivalent to the metric induced by the inner product, now called shahshahani inner product (Shahshahani 1979), on the tangent space given at the point x by $(u, v) \mapsto u \cdot (g(x)v)$ where g is the diagonal matrix with $g_{ij}(x) = |x|\delta_{ij}/x_i$. Any g-gradient is also a Shahshahani gradient.

The metric can be applied, as Akin and Shahshahani did in their papers, to study the genetic system consisting of n types of gametes with x_i giving the number of gametes of type i (Akin 1979, 1980, 1990; Shahshahani 1979).

QTL modelling of the selection dynamics

Quantitative traits and QTLs are characters that change almost continuously. Also fitness can be considered as a quantitative trait. Quantitative traits and QTLs have genetic and environmental components and they have a

complex genetic basis because their manifestations are performed by many genes, most of them have small effects. Besides, epistasis, pleiotropy, and genotype × environment interaction are common (Bürger 2000; Mackay 2001; Barton and Keightley 2002).

'The genetic system consists of $L \geq 1$ locus. At locus n there are $I_n \geq 2$ alleles, $A_{i_n}^{(n)}(i_n = 1, \ldots, I_n)$. We use the multi-index $i = (i_1, \ldots, i_L)$ as an abbreviation for the gamete $A_{i_1}^{(1)} \ldots A_{i_L}^{(L)}$ and write $I = \Pi_n I_n$ for the number of gametes. We use the letters i, j, l for gametes and k, n for loci' (Bürger 2011).

'Let $p_i = p_i(t)$ represent the frequency of gamete i among zygotes in generation t, and $p = (p_1, \ldots, p_I)^T$ the vector of all gamete frequencies. The frequency of allele $A_{i_n}^{(n)}$ among gametes is

$$p_{i_n}^{(n)} = \sum_{i|i_n} p_i, \tag{9}$$

where the sum runs over all multi-indices i with the nth component fixed as i_n. Let w_{ij} denote the fitness of genotype ij. We designate the marginal fitness of gamete i and the mean fitness of the population by

$$w_i = w_i(p) = \sum_j w_{ij} p_j \quad \text{And}$$
$$\bar{w} = \bar{w}(p) = \sum_{i,j} w_{ij} p_i p_j, \tag{10}$$

respectively.

After selection the frequency of the genotype ij is $p_i p_j w_{ij}/\bar{w}$. The frequency of gamete i in the next generation, i.e., after recombination and reproduction, is

$$p_i' = \sum_{j,l} \frac{R_{i,jl} p_j p_l w_{jl}}{\bar{w}}, \tag{11}$$

where $R_{i,jl}$ is the probability that during gametogenesis the paternal haplotypes j and l produce a gamete i by recombination.

Let

$$\Lambda_0 = \left\{ p : p_i = p_{i_1}^{(1)} \cdot \ldots \cdot p_{i_L}^{(L)} \right\} \subseteq S_I \tag{12}$$

denote the linkage equilibrium manifold (also called the Wright manifold). If there is no position effect, i.e. if $w_{ij} = w_{ikjN,jkiN}$ for every i, j, and k, then $D_i = 0$ for every $p \in \Lambda_0$. Hence,

$$\Lambda_0 \subseteq \{ p : D = 0 \}, \tag{13}$$

where $D = (D_1, \ldots, D_I)^T$ is the vector of all linkage disequilibria. In the absence of selection equality holds in (13)' (Bürger 2011).

Evolutionary modelling of QTLs without environmental interactions

In the following models, suppose that there is no QTL × environment interaction and that the environmental component E can be treated as white noise, i.e. as a Gaussian random variable with mean zero (Bürger 2011). Thus, the phenotypic value P is

$$P = G + E, \tag{14}$$

where the genetic and environmental components, G and E, are independent. Also, suppose that G is determined additively, namely, trait effects can be related to alleles and G is the sum of the effects of the parental alleles of an individual at the considered loci. Then, the genetic variance is additive and the phenotypic variance is

$$\sigma_P^2 = \sigma_A^2 + \sigma_E^2. \tag{15}$$

Generally, the phenotypic variance can be broken up into a genetic, an environmental, and an interaction component (Bürger 2011). Also, genetic variance can be decomposed into additive, dominance, epistatic, and various interaction components. These components may contain one or several QTLs. One basic idea in biometrical genetics is the narrow sense heritability, h^2 which is equal to the ratio of additive genetic to total phenotypic variance:

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2}. \tag{16}$$

The heritability can be calculated from correlations among relatives (Bulmer 1980). 'Most traits, in most populations, show substantial heritabilities (typically between 20% and 50%), and a few patterns have been identified' (Barton and Keightley 2002). Let $\bar{P}$ and $\bar{P}_s$ represent the mean phenotype before and after selection. Since, the environment involves only white noise, then $\bar{P} = \bar{G}$. Thus, the selection response across generations is

$$\Delta \bar{P} = h^2 \left(\bar{P}_s - \bar{P} \right). \tag{17}$$

Supposing a linear parent–offspring regression, Robertson (1966, 1968) extended the breeder's equation to predict the correlated response induced by artificial and natural selection

$$\Delta \bar{G} = \frac{Cov_A(G, W)}{\bar{W}}, \tag{18}$$

(Bürger 2011). Where, $W = W(G)$ represents the fitness of individuals with genotype value G, also, $Cov_A(G, W)$ is the additive genetic covariance of G and W. This is the secondary theorem of natural selection.

Nagylaki (1993) extended this result and derived an asymptotic version from first principles and in great generality; he proved

$$\Delta \bar{G} = \frac{Cov_A(G, W)}{\bar{W}} + O\left(s^2\right), \tag{19}$$

If selection is weak relative to recombination ($s \ll c_{\min}$).

Besides, if there is epistasis and either selection is strong or linkage disequilibria are large, changes in the linkage disequilibria can produce big changes in $\bar{G}$ even if $\sigma_A^2 = 0$ (Gimelfarb 1989; Nagylaki 1993).

In breeding programs and in nature, selection acts on many traits (Bürger 2011). To study the evolution of multivariate phenotypes, Lande (1979) developed a theory for the evolutionary dynamics. The phenotype of an individual is characterized by a vector of measurements of K QTLs, $P = (P_1, \ldots, P_K)^T$ and assumed to be specified by an additive genetic component G which may contain one or more QTLs, and an environmental component E such that $P = G + E$, where the mean of E disappears. Therefore, the mean vectors satisfy $\bar{P} = \bar{G}$. Lande's central assumption is that the distributions of G and of E are both multivariate normal and independent. If the corresponding covariance matrices are represented by **P**, **G**, and **E**, then $\mathbf{P} = \mathbf{G} + \mathbf{E}$ (Bürger 2011). If the fitness of individuals with phenotype P is $W(P)$ and $f(P)$ denotes the Gaussian density of phenotypes, then the mean fitness of the population is $\bar{W} = \int f(P)W(P)dP$, which is a function of $\bar{P}$ and **P**.

Lande proved that the change of the mean phenotype between generations is

$$\Delta\bar{P} = G\nabla ln\bar{W} = G\left(\frac{\partial \ln \bar{W}}{\partial \bar{P}_1}, \ldots, \frac{\partial \ln \bar{W}}{\partial \bar{P}_K}\right)^T, \tag{20}$$

where $\nabla \ln \bar{W}$ is named the selection gradient (Bürger 2011). In the univariate model, this can be proved to be equivalent to Robertson's equation (18). Also the similarity with the Svirezhev–Shahshahani gradient (8) is considerable. Besides, it is clear that (20) implies

$$\Delta \ln \bar{W} \geq 0, \tag{21}$$

Thus, Lande's theory combined QTL analysis into evolutionary genetics.

These results are based on the assumption that allelic effects contribute additively, namely without dominance and epistasis to the trait. But, because the fitness function is nonlinear, thus the dominance and epistasis in fitness are included and they are not necessarily weak (Bürger 2011).

QTL mapping and G-Matrix

The value of phenotype can be separated into QTL, genomic context, and environmental values. In other words, P = (**G**-Matrix) + **E**, where **P** is the vector of phenotypic values, **G**-Matrix equals to $(\mathbf{g} + \hat{\mathbf{g}})$, and **g** is the vector of QTL values, $\hat{\mathbf{g}}$ is the vector of genomic context values, and E is the vector of environmental residuals (Kelly 2009). Goldgar (1990) mentioned that, **g** can include the whole of genome as well as all of the chromosomes. Also, **g** and $\hat{\mathbf{g}}$ have their own variance and covariance matrices, which can be separated into dominance and additive parts (Kelly 2009). See section 'Evolutionary modelling of QTL × environment interactions'.

QTL mapping segregates the variation of trait into genomic values (Tanksley 1993), that results the genetic structure of quantitative variation (Mackay 2004).

Allele frequencies of QTL associates the **G**-matrices to QTL mapping trials in which molecular marker-trait relationships are applied to compute population allele frequencies. QTL mapping is utilized to determine the genes and the nucleotides, i.e. quantitative trait nucleotide (QTN) that applies to variation in quantitative traits (Mackay 2001).

If a QTL is under selection, then the changing of allele frequency and average effect of the selected QTL could be a portion to the variance of the QTL (Griffing 1960). Besides, in QTL analysis, the direction and quantity of variation in the variance of the QTL pertains to the higher moments of the distribution (Kelly 2009).

In QTL mapping, genotyping estimates the allele frequencies of QTL within each population; thus, the simultaneous separation of QTL variation with phenotypic variation results the allelic effects to be estimated (Kelly 2009).

Evolutionary modelling of QTL × environment interactions

The most frequently investigated subject in QTL analysis has been the balance between stabilizing selection and mutation (Gillespie and Turelli 1989). QTL × environment interactions (phenotypic plasticity) are observed in QTLs including fitness components which provide an explanation for quantitative variation. Random changes in temperature, climate, entering of pests or pesticides into the new habitat, fluctuations in the life cycles of prey predators or parasites, environmental pollutions and spread of a predator or pest are some kinds of environmental variations (Bürger and Krall 2004).

Study of QTL × environment interactions is important not only to understand how the genes interact with the environment, but also to correctly document the relative effect of QTL (Bürger and Krall 2004). Environmental tolerance curves relating fitness to abiotic environmental variables such as temperature (Angilletta 2009) or salinity (Browne and Wanigasekera 2000) have been investigated for some kinds of organisms like reptiles, amphibians, insects and plants (Angilletta 2009).

In QTLs, polymorphism is conserved by balancing selection through some kinds of rules like QTL × environment interactions, environmental fluctuations, pleiotropy and over dominance.

Falconer (1952) showed that a QTL expressed in two environments can be considered as two QTLs which are correlated together. Thus, the expression of a QTL in an environment is called a QTL state. For instance, body weight or grain yield in two environments are considered as two correlated QTL states (Falconer 1952; Via and Lande 1985). Genetic correlation across environments represents that the same alleles affect the QTL states

in the environments (Via and Lande 1985). If genetic correlation among QTLs which are expressed in environments is less than one hundred per cent, it is considered as QTL × environment interaction (Robertson 1959). If QTL states in different environments are not to be assessed on the same individual, then the phenotypic covariance is undetermined (Falconer 1981).

The matrix of additive genetic variances and covariances of QTL states is presented as **G**, where the elements of G_{ii} are the additive genetic variances of the QTL states in the ith environment. Also, G_{ij} is the additive genetic covariance of QTL states expressed in the ith and jth environments (Lande 1976, 1980).

Thus, the random mating is supposed to occur in the population and individuals diffuse into the habitats randomly in each generation; also, selection acts on the population before diffusion and mating (Via and Lande 1985).

Dynamical equations for QTL evolution under mild and severe selections in environments:

Let α_i be the portion of the population immigrating the ith habitat where $\sum \alpha_i = 1$, $\bar{\eta}_i$ be the average value of the QTL state expressed in the ith environment, C_{ii}^{-1} be the inverse of the phenotypic variance in the ith environment, and ζ_i be the difference between the average phenotype after and before selection in the ith environment. Then, the dynamical equation for mild selection in n environments is

$$\begin{pmatrix} \Delta\bar{\eta}_1 \\ \vdots \\ \Delta\bar{\eta}_n \end{pmatrix} = \begin{pmatrix} G_{11} \ldots G_{1n} \\ \vdots \quad \ddots \vdots \\ G_{n1} \cdots G_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 C_{11}^{-1} & \zeta_1 \\ \vdots & \vdots \\ \alpha_n C_{nn}^{-1} & \zeta_n \end{pmatrix}$$

and in special case for two environments we have

$$\begin{pmatrix} \Delta\bar{\eta}_1 \\ \Delta\bar{\eta}_2 \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} \alpha_1 C_{11}^{-1} & \zeta_1 \\ \alpha_2 C_{22}^{-1} & \zeta_2 \end{pmatrix}. \tag{22}$$

Thus,

$$\begin{aligned} \Delta\bar{\eta}_1 &= \alpha G_{11} C_{11}^{-1} \zeta_1 + (1-\alpha)\, G_{12} C_{22}^{-1} \zeta_2 \\ \Delta\bar{\eta}_2 &= \alpha G_{21} C_{11}^{-1} \zeta_1 + (1-\alpha)\, G_{22} C_{22}^{-1} \zeta_2 \end{aligned} \tag{23}$$

where $\alpha = \alpha_1$ (Via and Lande 1985). Equations (23) shows that the evolution of each of the QTL states covers the response to selection in the environment in which the QTL state is expressed and also, the related response to selection on the state which is expressed in the other environments. Hence, if two QTL states are selected together to increase environments, namely $\zeta_1 > 0$ and $\zeta_2 > 0, \alpha = 0.5$, then a negative or small genetic covariance in the phenotype across environment, i.e. G_{12} will decrease the rate of evolution. Also, evolution under disruptive selection, i.e. $\zeta_1 > 0$ and $\zeta_2 < 0$ will be delayed by positive genetic covariance (Via and Lande 1985). If in each environment the QTLs are covered, then selection on other QTLs in the same environment can also result to change the average phenotype.

Let the fitness of an individual of phenotype η in environment i be represented as $W_i(\eta_i)$. If a normal distribution of phenotypes, $p_i(\eta_i)$ exists and average fitness be defined as $\bar{W}_i$ then, the selection differential in the ith environment is

$$\zeta_i = \frac{\left[\int \eta_i p_i(\eta_i)\, W_i(\eta_i)\, d\eta_i\right]}{\bar{W}_i - \bar{\eta}_i}. \tag{24}$$

(Via and Lande 1985).

Denoting the change in average fitness with an increasing change in $\bar{\eta}_i$ and also referring to (24), the observed selection differential can be rearranged in terms of the gradient of the logarithm of average fitness in the ith environment

$$C_{ii}^{-1} \zeta_l = \nabla_i \ln \bar{W}_i, \tag{25}$$

(Lande 1979). Where the gradient operator $\nabla_i = \delta/\delta\bar{\eta}_i$ working on $\ln \bar{W}_i$ shows the effect of selection in the ith environment on the QTL state. The selective force $\nabla_i \ln \bar{W}_i$ is equitable to the partial regression coefficient of individual relative fitness in the ith environment $\left(W_i/\bar{W}_i\right)$ on η_i (Lande and Arnold 1983). The dynamic equations for mild selection in two environments can be rearranged as follows, if the equation (25) substitutes into (22)

$$\begin{pmatrix} \Delta\bar{\eta}_1 \\ \Delta\bar{\eta}_2 \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \cdot \begin{pmatrix} \nabla_1 \\ \nabla_2 \end{pmatrix} \ln\left[\bar{W}_1^{\alpha} \bar{W}_2^{(1-\alpha)}\right]. \tag{26}$$

The function $\bar{W} = \bar{W}_1^{\alpha} \bar{W}_2^{1-\alpha}$ represents the joint average fitness under mild selection as the weighted geometric mean of the average fitness in the two environments, which shows an adaptive topography for evolution in two environments (Via and Lande 1985).

The evolutionary dynamics of average fitness is estimated through by expanding $\bar{W}$ in a Taylor series around an arbitrary point. The higher order terms can be relinquished if mild selection occurs, thus

$$\begin{aligned} \Delta \ln\left[\bar{W}_1^{\alpha} \bar{W}_2^{(1-\alpha)}\right] &= \left(\Delta\bar{\eta}_1, \Delta\bar{\eta}_2\right) . \nabla \ln\left[\bar{W}_1^{\alpha} \bar{W}_2^{(1-\alpha)}\right] \\ &= \left(\nabla \ln \bar{W}\right)^T G \nabla \ln \bar{W} \geq 0, \end{aligned} \tag{27}$$

where $\nabla^T = (\nabla_1, \nabla_2)$, and T represents matrix transposition (Via and Lande 1985).

The dynamical equation for phenotypic evolution under severe selection for two environments can be shown as

$$\begin{pmatrix} \Delta\bar{\eta}_1 \\ \Delta\bar{\eta}_2 \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \cdot \begin{pmatrix} \nabla_1 \\ \nabla_2 \end{pmatrix} \ln\left[\alpha \bar{W}_1 + (1-\alpha)\, \bar{W}_2\right] \tag{28}$$

where $\bar{W} = \alpha \bar{W}_1 + (1-\alpha)\, \bar{W}_2$ and the relative average fitness of selected individuals is $\left(\bar{W}_i/\bar{W}\right)$. Since $\delta \bar{W}_2/\delta \bar{\eta}_1 = 0$, then $\bar{W}_2$ is not a function of $\bar{\eta}_1$ and we have

$$\left(\alpha \frac{\bar{W}_1}{\bar{W}}\right) \nabla_1 \ln \bar{W}_1 = \left(\frac{\alpha}{\bar{W}}\right)\left(\frac{\delta \bar{W}_1}{\delta \bar{\eta}_1}\right) = \left(\frac{1}{\bar{W}}\right)\left[\frac{\alpha \delta \bar{W}_1}{\delta \bar{\eta}_1} + \frac{(1-\alpha)\, \delta \bar{W}_2}{\delta \bar{\eta}_1}\right] = \nabla_1 \ln \bar{W} \tag{29}$$

(Via and Lande 1985).

Conclusions

Evolutionary and biometrical genetics are scientific achievements. Integrating these mathematical, statistical and biological branches of science has made important results to this time. The development of most phenotypic traits involves complex interactions between many underlying factors, both genetic and environmental. During development of a phenotypic trait, gene products interact in highly no additive ways with one another and with environmental factors (Rice 2002).

In previous two decades, many researches have investigated about QTL analysis and its basic role in the evolution of quantitative traits. Also, progress in mathematical and molecular techniques has answered a lot of questions about QTL analysis. Here, QTL analysis can be used to estimate the local geometry of the phenotype landscape, a plot of a phenotypic trait as a function of the underlying genetic and environmental factors that contribute to its development (Rice 2002).

The mathematical models derived here showed that the gradients of mean fitness which have revealed in studies made by Akin (1979, 1990), Shahshahani (1979), Lande (1979) and Bürger (2011), had a basic role in mathematical genetics and evolutionary aspects of biometrical genetics. Also, it was marvelous that Riemannian geometry which had fundamental role in Einstein's general relativity had great applications in quantitative and population genetics; these were perfectly shown in Akin (1979, 1980, 1982, 1987, 1990) scientific articles.

In this review, mathematical methods have been applied in dynamical modelling of biometrical genetics and analytical biology, offered to study the QTL × environment interactions. The topic is circumscribed, going from basic selection equations to models of evolution of QTLs. Discrete and continuous mathematical models and subsequently, QTL modelling were introduced with and without environmental interactions.

Supports for subjects presented in this paper were models applied in biometrical genetics which corresponded to QTL analysis and matched with results from Bürger (2011), Kelly (2009), Bürger and Krall (2004), Lande (1979, 2014).

One aspect of this paper is documentation to clear that evolutionary feature of biometrical genetics support QTL × environment interactions. Based on this review and other studies, QTL × environment interactions are important to investigate both genes × environment interactions and relative effects of QTLs.

The originality of this synthesis is both the evolutionary modelling of quantitative aspects of QTL × environment interactions which can be used to investigate the extinction or stability of a population, and to emphasize that although some scientific subjects like Brownian motion, quantum mechanics, general relativity, differential geometry, and evolutionary biometrical genetics are apparently separate and different subjects, but the mathematical models are the backbone of these sciences which implies that such matters in nature have probably common foundations.

Since there have been many studies in mathematics and statistics related to genetics in the last decades, then a new and interesting branch of interdisciplinary science in future will be the perspective of the subject of this article.

Acknowledgements

We appreciate Prof. Ethan Akin, Department of Mathematics, The City College of New York, USA, for his helpful suggestions on earlier versions of this paper and Prof. Russell Scott Lande, Centre for Biodiversity Dynamics, Norwegian University of Science and Technology for his beneficial assistance. We also thank Prof. Reinhard Bürger, Faculty of Mathematics, University of Vienna, Austria.

References

Akin E. 1979 *The geometry of population genetics. Lecture notes in biomathematics*, vol. 31, pp. 1–211. Springer.

Akin E. 1980 Domination or equilibrium. *Math. Biosci.* **50**, 239–250.

Akin E. 1982 Cycling in simple genetic systems. *J. Math. Biol.* **13**, 305–324.

Akin E. 1987 *Cycling in simple genetic systems: II. The symmetric cases. Lecture notes in economics and mathematical systems*, vol. 287, pp. 139–153. Springer.

Akin E. 1990 *The differential geometry of population genetics and evolutionary games. Mathematical and Statistical developments of evolutionary theory*, pp. 1–93. Kluwer Academic Publishers, Dordrecht.

Angilletta M. 2009 *Thermal adaptation: a theoretical and empirical synthesis*. Oxford University Press, Oxford.

Aronson D. G. and Weinberger H. F. 1978 Multidimensional nonlinear diffusion arising in population genetics. *Adv. Math.* **30**, 33–76.

Barton N. H. and Keightley P. D. 2002 Understanding quantitative genetic variation. *Nat. Rev. Genet.* **3**, 11–21.

Browne R. A. and Wanigasekera G. 2000 Combined effects of salinity and temperature on survival and reproduction of five species of Artemia. *J. Exp. Mar. Biol. Ecol.* **244**, 29–44.

Bulmer M. G. 1980 *The mathematical theory of quantitative genetics*. Clarendon Press, Oxford.
Bürger R. and Bomze I. M. 1996 Stationary distributions under mutation-selection balance: structure and properties. *Adv. Appl. Prob.* **28**, 227–251.
Bürger R. 2000 *The mathematical theory of selection, recombination, and mutation*. Wiley, Chichester.
Bürger R. and Krall C. 2004 Quantitative-genetic models and changing environments. In *Evolutionary conservation biology* (ed. R. Ferrière, U. Dieckmann and D. Couvet), pp. 171–187. Cambridge University Press, Cambridge.
Bürger R. 2011 *Some mathematical models in evolutionary genetics*, pp. 67–89. Springer, Berlin.
Coppersmith S. N., Blank R. D. and Kadanoff L. P. 1999 Analysis of a population genetics model with mutation, selection, and pleiotropy. *J. Stat. Phys.* **97**, 429–457.
Erickson D., Fenster C. B., Stenoien H. K. and Price D. 2004 Quantitative trait locus analyses and the study of evolutionary process. *Mol. Ecol.* **13**, 2505–2522.
Ewens W. J. 2003 *Mathematical population genetics*, 2nd edition. Springer, New York.
Falconer D. S. 1952 The problem of environment and selection. *Am. Nat.* **86**, 293–298.
Falconer D. S. 1981 *Introduction to quantitative genetics*, 2nd edition. Longman, N.Y.
Falconer D. S. and MacKay T. F. C. 1996 *Introduction to quantitative genetics*, 4th edition. Longman, Essex.
Fisher R. A. 1937 The advance of advantageous genes. *Ann. Eugen.* **7**, 355–369.
Gillespie J. H. and Turelli M. 1989 Genotype-environment interactions and the maintenance of polygenic variation. *Genetics* **141**, 129–138.
Gimelfarb A. 1989 Genotypic variation for a quantitative character maintained under stabilizing selection without mutations: epistasis. *Genetics* **123**, 217–227.
Goldgar D. 1990 Multipoint analysis of human quantitative genetic-variation. *Am. J. Hum. Genet.* **47**, 957–967.
Griffing B. 1960 Theoretical consequences of truncation selection based on the individual phenotype. *Aust. J. Biol. Sci.* **13**, 309–343.
James F. C. and Kimura M. 1970 *An introduction to population genetics theory*. Blackburn Press, New Jersey.
Kelly J. K. 2009 Connecting QTLs to the G-matrix of evolutionary quantitative genetics. *Evolution* **63**, 813–825.
Kimura M. 1965 A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. USA* **54**, 731–736.
Kolmogorov A., Petrovsky I. and Piscounov N. 1937 Étude de l'équation de la diffusion avec croissance de la quantité de matiére et son application a un probléme biologique. *Bull. Univ. Mosc. Ser. Int. Sect. A.* **1**, 1–25.
Lande R. 1976 The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**, 221–235.
Lande R. 1979 Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**, 402–416.
Lande R. 1980 The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**, 203–215.
Lande R. and Arnold S. J. 1983 The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226.
Lande R. 2014 Evolution of phenotypic plasticity and environmental tolerance of a labile quantitative character in a fluctuating environment. *J. Evol. Biol.* **27**, 866–875.
Lynch M. and Walsh B. 1998 *Genetics and analysis of quantitative characters*. Sinauer associates, Sunderland.
Mackay T. F. C. 2001 The genetic architecture of quantitative traits. *Annu. Rev. Genet.* **35**, 303–339.
Mackay T. F. C. 2004 Genetic dissection of quantitative traits. In *The evolution of population biology* (ed. R. S. Singh and M. K. Uyenoyama), pp. 51–73. Cambridge University Press, Cambridge.
Miles C. and Wayne M. 2008 Quantitative trait locus (QTL) analysis. *Nat. Edu.* 1. (https://www.nature.com/scitable/topicpage/quantitative-trait-locus-qtl-analysis-53904).
Nagylaki T. 1993 The evolution of multilocus systems under weak selection. *Genetics* **134**, 627–647.
Norman M. F. 1975 Approximation of stochastic processes by Gaussian diffusion and applications to Wright-Fisher genetic models, SIAM. *J. Appl. Math.* **29**, 225–242.
Nowak M. A. 2006 *Evolutionary dynamics: exploring the equations of life*. Harvard University Press, Cambridge.
Rice H. S. 2002 A general population genetic theory for the evolution of developmental interactions. *Proc. Natl. Acad. Sci. USA* **99**, 15518–15523.
Robertson A. 1959 The sampling variance of the genetic correlation coefficient. *Biometrics* **15**, 469–485.
Robertson A. 1966 A mathematical model of the culling process in dairy cattle. *Anim. Prod.* **8**, 95–108.
Robertson A. 1968 The spectrum of genetic variation. In *Population biology and evolution* (ed. R. C. Lewontin), pp. 5–16. University Press, Syracuse.
Robertson D. S. 1985 A possible technique for isolating genic DNA for quantitative traits in plants. *J. Theor. Biol.* **117**, 1–10.
Rodney M. 2001 Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* **2**, 370–381.
Shahshahani S. 1979 A new mathematical framework for the study of linkage and selection. *Mem. Am. Math. Soc*. **17**, ix+34. https://doi.org/10.1090/memo/0211.
Svirezhev Yu. M. 1972 Optimality principles in population genetics. In *Studies in theoretical genetics*, pp. 86–102. Institute of Cytology and Genetics, Novisibirsk (In Russian).
Tahira T., Kukita Y., Higasa K., Okazaki Y., Yoshinaga A. and Hayashi K. 2009 Estimation of SNP allele frequencies by SSCP analysis of pooled DNA. *Methods Mol. Biol.* **578**, 193–207.
Tanksley S. D. 1993 Mapping polygenes. *Annu. Rev. Genet.* **27**, 205–233.
Via S. and Lande R. 1985 Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **39**, 505–522.
Waxman D. and Peck J. R. 1998 Pleiotropy and the preservation of perfection. *Science* **279**, 1210–1213.
Wright S. 1932 The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. 6th Int. Cong. Genet.* **1**, 356–366.
Yang J., Hu C., Hu H., Yu R., Xia Z., Ye X. *et al*. 2008 QTL Network: mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics* **24**, 721–723.
Zhang W. K., Wang Y. J., Luo G. Z., Zhang J. S., He C. Y. and Chen S. Y. 2004 QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. *Theor. Appl. Genet.* **108**, 1131–1139.

Corresponding editor: H. A. Ranganath