ISSN 0006-2979, Biochemistry (Moscow), 2008, Vol. 73, No. 2, pp. 219-230. © Pleiades Publishing, Ltd., 2008.
Original Russian Text © K. V. Gunbin, V. V. Suslov, N. A. Kolchanov, 2008, published in Biokhimiya, 2008, Vol. 73, No. 2, pp. 270-282.

Molecular-Genetic Systems of Development: Functional Dynamics and Molecular Evolution

K. V. Gunbin*, V. V. Suslov, and N. A. Kolchanov

Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences,
pr. Lavrentieva 10, 630090 Novosibirsk, Russia; E-mail: genkvg@bionet.nsc.ru

Received August 3, 2007
Revision received September 19, 2007

Abstract—This work for the first time compares results obtained with studies of parametric robustness of the Hh- and Dpp/BMP signal cascades responsible for morphogenesis and molecular evolution of the Hh- and Dpp/BMP cascade genes. There is a link between adaptive evolution of genes and those changes in kinetic parameters of the signal cascade models, which are critical for normal functioning of these cascades. Special attention is paid to events of the molecular evolution of the Hh- and Dpp/BMP cascade genes that matched with the emergence of the major taxonomic types and classes of Bilateria.

DOI: 10.1134/S0006297908020144

Key words: Hh- and Dpp/BMP signal cascades, mathematical modeling, positive selection, Bilateria, compensatory changes

The study of molecular genetic mechanisms responsible for control of morphogenesis is an important problem of modern biology. There is significant progress not only in identification of genes controlling cell differentiation, but also in elucidation of expression patterns for many genes determining development as well as in elucidation of functions of products of these genes in invertebrates and vertebrates [1, 2]. For example, the signal transduction pathways Hh, WNT, BMP, RTK, JAK/STAT, Notch, and the nuclear receptor signal transduction pathway play an important role in morphogenesis [3]. However, analysis of large and diverse experimental data and complexity of investigated systems require use of effective information computer-aided technologies; without this approach subsequent progress in understanding of morphogenesis processes of multicellular organisms is impossible [4, 5].

A gene network, representing the functional group of coordinately expressed genes, is a basis underlying any sign [6, 7]. Any gene network contains: a gene cassette (a core) responsible for its functioning, central regulators; transcription factors responsible for coordinate regulation of the core genes (these factors interact with a corresponding binding site at the regulatory regions of these genes); receptors, which trigger functioning of the gene network in response to external signals; signaling pathways responsible for signal transduction from receptors of the gene network onto its central regulators. Some of the products of the gene network core genes also play a role of signals interacting with receptors of both their own and also other gene networks; these products may also form regulatory circuits with positive and negative feedbacks.

The promising directions of theoretical studies of gene networks include quantitative modeling of their dynamic functioning and analysis of molecular evolution of their genes. The modeling includes three steps: construction of a model, verification of its parameters, and investigation of its dynamics and robustness [8]. Study of parametric robustness allows determination of changeable parameters, which bring strong changes in gene network dynamics (parameters with hyper-response), as well as inert parameters (changes in these parameters insignificantly influence the network) [9-11]. Mutations of genes determining the hyper-response parameters may exert non-proportional influence on morphogenesis; this is important in relation to medico-biological problems (carcinogenesis, teratogenesis) and problems of the theory of evolution.

Analysis of molecular evolution of genes is based on the theory of neutral evolution [12]. Nucleotide substitu-

---

* To whom correspondence should be addressed.

tions are subdivided into synonymous (which do not result in amino acid substitution in a protein) and non-synonymous ones (the rate of their fixation should be related with functional load of amino acids in the protein). Exceeding of the fixation rate of non-synonymous substitutions over synonymous ones suggests positive selection, and in the opposite case negative selection occurs.

In this study we have compared parametric robustness of the mathematic models of Hh- and Dpp/BMP signal transduction cascades and molecular evolution of genes involved in these cascades. Genes of hyper-response identified during modeling were genes undergoing positive selection during divergence of the main types of Bilateria. Based on these data we have proposed a hypothesis on the evolutionary role of the hyper-response genes in the initiation of aromorphoses as the resources of an "internal reserve" for compensation of radical structural rearrangements of gene networks.

RESULTS

Parametric robustness of Hh cascade models. The Hh signal transduction cascade is involved in many processes related to morphogenesis and development of invertebrates and vertebrates. In insects and other arthropods (as well as in mollusks) it includes the morphogen protein Hh, its cell receptors Ptc and Smo, proteins Cos2, PKA, Slmb, Su(Fu), and Fu transducing information about the presence or absence of the morphogen Hh to the transcription factors Ci (Gli) [13-15].

In insects (Fig. 1b), in the absence of the morphogen Hh, the effect of the Smo protein on the protein complex Cos2, Fu, and Ci is blocked by the Ptc protein. Within this complex Ci undergoes ubiquitination by Slmb ubiquitin ligase followed by subsequent proteolysis. The N-terminal fragment of Ci is then transferred into the nucleus and represses the Hh cascade gene cassette. The latter also includes the *ptc* gene, encoding Ptc receptor. The interaction between Ptc receptor with the morphogen Hh prevents Smo protein inhibition. The Smo protein triggers a chain of reaction resulting in dissociation (detachment) of the Fu/Cos2/Ci complex from the cytoskeleton. Activation of Fu and Cos2 prevents proteolysis of the Ci protein. The full sized transcription factor Ci activates the Hh cascade cassette genes [13-15]. This mechanism (with minor variations) is also typical for other arthropods and mollusks.

In contrast to insects, vertebrates (fishes and mammals) have three homologs of the *Ci* gene: *Gli1*, *Gli2*, and *Gli3* (Fig. 1a) [16]. The products of all *Gli* genes (proteins Gli1, Gli2, and unprocessed Gli3) may be transcription activators. Only two Gli proteins inhibit transcription. These include modified Gli2 and Gli3, but only Gli3 undergoes proteolysis (like invertebrate Ci protein) [17-19]. In addition, vertebrates have a mechanism of control of protein Gli expression: Gli proteins activate expression of the *Gli1* and *Gli2* genes [19-21].

The nematode Hh cascade is rather specific. It is strongly rearranged and it would be more correct to define it as the Hh/Ptc cascade. In nematodes, the *Ptc* gene is similar to its homologs in other Bilateria (e.g. in mammals, birds, and insects) and was very important for develop-

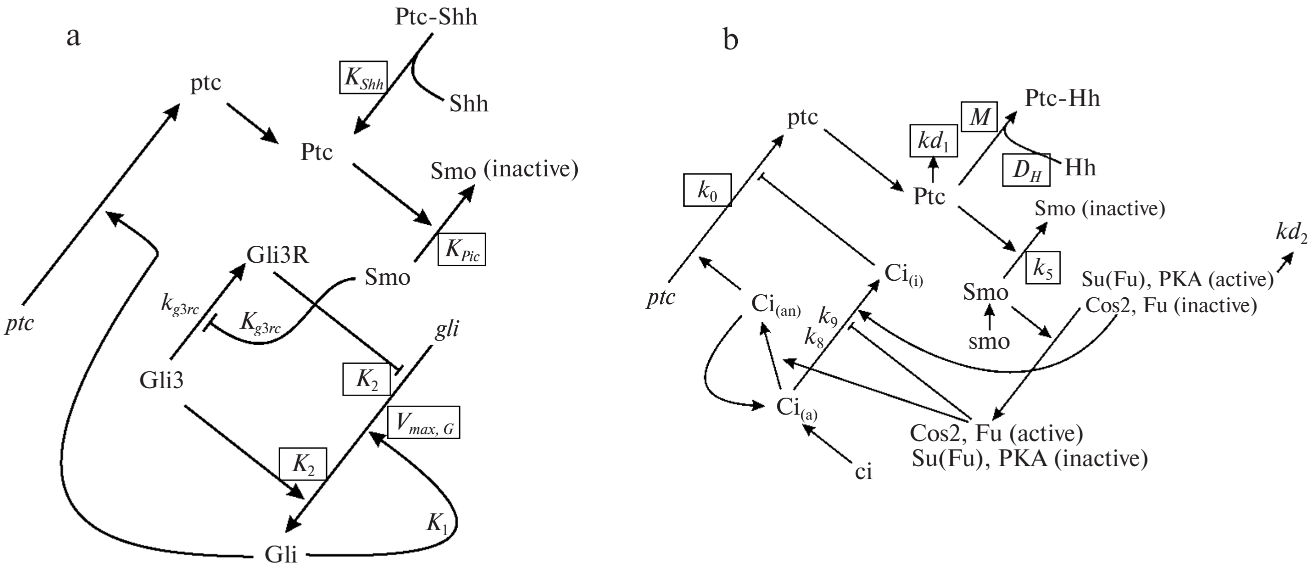

Fig. 1. Models of the Hh signal transducing cascades: a) vertebrates [11]; b) invertebrates [27]. Arrows with "sharp" and "blunt" tips designate activation and inhibition processes, respectively. Rectangles contour rate constants of simple processes, which are critical for normal dynamics of functioning of the Hh signal cascade.

ment. *TRA*, a *Ci* homolog, is also similar, but it has changed its function and regulates the development of sex cells. It should be noted that instead of the morphogen Hh, which would be lost during evolution, the nematodes have Hog-proteins encoded by a large family of genes homologous to the *Hh*-gene. Some of the Hog-proteins preserve the intein-like domain, which is characterized by sequence and structure homologous to that of Hh, whereas others lost it. In these Hog-proteins, two new domains—Wart and Ground—are revealed (giving the names of the Hog-proteins as the truncated Hedgehog), which have not been found in other Bilateria [22-24]. The rearrangement of the Hh cascade is definitely attributed to the formation of the special type of the nematode embryogenesis, when cell predetermination is related to the formation of the very early and special type of the clones of such predetermined cells [25]. This required cardinal rearrangement of the embryonic induction system [26]. Indeed, the Hh cascade represented the only truncated "medium", responsible for association with cytoskeleton, which regulates the rate of this induction. The morphogen and the transcription factor are the links responsible for joining the "medium", which have been enormously changed.

In this study, we have compared parametric robustness of the invertebrate and vertebrate Hh cascade models by means of kinetic parameters obtained earlier in experimental studies [11, 27]. The parametric robustness of the solution of a system is a degree of change of solution of the system (e.g. systems of ordinary differential equations) versus initial solution during changes of certain parameters of the system.

The model of the vertebrate Hh cascade (Fig. 1a) was developed by Lai et al. [11]. The Hh cascade of *Drosophila melanogaster* was used as the basis for the invertebrate Hh cascade model (Fig. 1b) proposed by Gunbin et al. [27]. The invertebrate Hh cascade models [27] supplement the vertebrate Hh cascade model [11] in description of signal transduction from the morphogen Hh receptors to the transcription factors Ci (Gli) (Fig. 1, a and b). This can be used to take into consideration the effect of dynamics of intracellular Hh signal transduction processes on total functioning of the Hh cascade signals.

We have compared data on the parametric robustness and mutation portraits of both Hh cascade models. (The mutation portrait of the gene network is a set of its steady states and dynamic characteristics obtained during variation within given limits of rates for each elementary process in the gene network. Limits of variations at which the system functions in normal and aberrant modes are determined using experimentally obtained data on various mutations [28].) We found the set of hypersensitive parameters; even small changes in these dynamic parameters sharply change dynamics of the Hh cascade functioning (Table 1). Such parameters characterize: interaction between Ptc and Smo proteins, affinity of transcription factors Ci/Gli to gene enhancers, formation/distribution of the morphogen Hh and its interaction with Ptc receptor. Linking dynamic parameters of these models with functions of Hh cascade proteins, we have determined hyper-response genes, which may sharply change dynamics of Hh cascade functioning in dependence on certain mutations (Table 1).

Parametric robustness of models of the Dpp/BMP cascade. Dpp protein (and its homologs BMP2/BMP4 in mammals) is involved in determination of body axes and in organogenesis. In invertebrate and vertebrate tissues, specific Dpp paralogs increasing its activity in certain tissues have been identified; these include Scw and Gbb in Drosophila and their homologs BMP5/BMP6/BMP7 in invertebrates [29]. The Dpp/BMP cascade [30-32] includes Dpp(BMP) morphogens, their receptors Tkv and Sax (type I receptors), Put (type II receptors), and transcription factors Mad (R-Smad) and Med (Co-Smad) regulating a gene cassette. Normal functioning of the Dpp/BMP cascade requires increased expression of receptors in the regions of high concentration of the Dpp(BMP) morphogens. This involves SBP proteins (particularly Cv-2 protein), which bind the Dpp(BMP) morphogens on the cell surface [33, 34]. The monomeric form of the Dpp(BMP) morphogen is inactive. Dpp(BMP) dimers activate Put receptor, which forms heterodimer with Tkv receptor and phosphorylates membrane proteins Tkv and Sax, and also cytoplasmic protein Mad (R-Smad) [35, 36]. Mad binds to Med (Co-Smad) and forms a transcription complex regulating expression of a gene cassette of the Dpp/BMP cascade [30-32] (Fig. 2). The composition of Dpp/BMP and its functions in nematodes are comparable with those of other Bilateria [37, 38].

Umulis et al. theoretically analyzed dynamics of general functioning of the Dpp/BMP cascade without detailed description of BMP dimer formation [39]. Shimmi et al. theoretically analyzed the process of BMP dimer formation [40]. Values of kinetic parameters and limits of their variations in norm and in mutations were set on the basis of experimental data [39, 40].

We have compared parametric robustness of these mutually supplemented models (Fig. 2) and their mutation portrait. Linking the hypersensitive parameters (Table 2) with functions of Dpp/BMP cascade proteins, we have determined the hyper-response genes (Table 2). Changes in interactions of Tkv, Sax, and Put protein, affinity of Mad/Med transcription factors to gene cassette enhancers, interaction of Dpp(BMP) dimers, receptors (co-receptors) Tkv and SBP, and formation of various types of BMP dimers may drastically influence dynamics of the Dpp/BMP cascade.

Molecular evolution of genes encoding proteins of Hh and Dpp/BMP cascade signals. Gene evolution modes were analyzed in two steps [41, 42]. The first step included search of protein sites for which there was positive selection at the codon level in corresponding gene regions. During the second step we searched for branches

Table 1. Hypersensitive parameters of the Hh signal cascade models for vertebrates [11] and invertebrates [27]

Parameters of models	Parameter description	Proteins or genes determining parameter
	Vertebrate Hh cascade model [11]	
$v_{max,G}$	maximal rate of transcription of *Gli* genes	*Gli*, Gli3
K_2	dissociation constant for protein Gli3 and promoter of *ptc* gene	Gli3
K_{Ptc}	half-maximal concentration of Ptc protein required for inhibition of Smo protein	Ptc, Smo
K_{Shh}	dissociation constant for Ptc–Hh complex	Ptc, Hh
	Invertebrate Hh cascade model [27], dynamic variables of simple biochemical processes	
M	rate constant of protein Ptc inactivation in complex Ptc–Hh	Ptc, Hh
D_H	diffusion coefficient of Hh morphogen	Hh, enzymes of Hh modification
kd_1	degradation constant for Ptc protein	Ptc, enzymes of Ptc degradation
k_5	half-maximal concentration of Ptc protein required for Smo inhibition	Ptc, Smo
	Invertebrate Hh cascade model [27], dynamic variables of generalized biochemical processes*	
k_3	rate constant for degradation reaction of *ptc* gene mRNA	*ptc*, enzymes of active or passive degradation of *ptc* gene mRNA
k_{3a}	rate constant of synthesis of membrane form of Ptc protein	Ptc, enzymes of synthesis of membrane form of Ptc protein
I	basal expression of *ptc* gene	*ptc*, enzymes of basal expression of *ptc*
kd_3	degradation constant for protein complex mediating Hh signal transduction into cell nucleus	Fu, Su(Fu), Cos2, enzymes of active or passive degradation

* These constants describe poorly investigated molecular biological processes and so it is difficult to link changes in these dynamic variables with changes in certain proteins.

of the phylogenic tree with positive selection. Gene regions in which positive selection was identified at the first step were analyzed.

Search for gene regions involved in adaptive evolution was carried out by means of alignment of nucleotide sequences without gaps; they were built by alignment of amino acid sequences. We used standard programs of molecular evolution analysis [41, 42]. Positive selection was considered as valid when it was confirmed by most programs.

Search for branches of the phylogenic tree (where positive selection was carried out) included paired comparison of sequences. Reconstruction of ancestor nucleotide sequences was not employed due to large evolutionary distances and low homology. In addition, use of reconstruction of ancestor nucleotide sequences for the Hh cascade genes would have doubtful validity due to the abovementioned features of the Hh/Ptc cascade structure in nematodes. Although nematodes are not an ancestral group for Bilateria, they diverged from the common trunk a long time ago and therefore the presence of some features typical for the nematode Hh/Ptc cascade could not be ruled out in common Bilateria ancestors. Several methods were used, and positive selection was considered as valid when most methods indicated statistically significant presence of the positive selection [42]. Statistical significance under conditions of large evolutionary distances was additionally evaluated by the computer test based on modeling of protein evolution [42]. Simulated genes were built up by means of random reverse translation from amino acids into codons of the modeled proteins. In the case of suggested absence of positive selection, paired analysis of simulated genes employed the same methods as in the case of real genes. For each pair of simulated

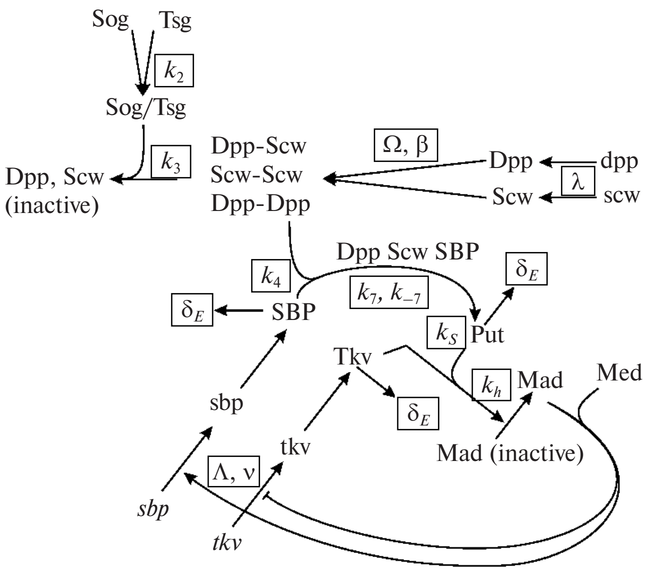

Fig. 2. Models of the Dpp/BMP signal transduction cascade [39, 40]. The model [40] describes the process of BMP dimer formation (upper part of this figure); the model [39] describes the remaining part of the Dpp/BMP signal cascades (lower part of this figure). Other designations are the same as in Fig. 1.

genes we evaluated limits of statistical intervals of degree of positive selection (for various methods). The resulting limits of intervals were compared with degree of positive selection [42] calculated for real genes. If degree of positive selection calculated for real genes exceeded upper limits of statistical intervals, positive selection was identified for the tree branch. In this analysis all compared pairs of genes were taken into consideration.

Results on identification of protein domains subjected to positive selection. Analysis has shown that the pressure of positive selection on proteins and/or their domains generally depends on their role in functioning of the Hh and Dpp/BMP cascades (Table 3). For example, change in the intein domain of the Hh morphogen influences the D_{H} parameters (diffusion coefficient of the Hh morphogen) [22] and K_{Shh} (dissociation constant for the Ptc–Hh complex) [11, 41, 42] (Table 1). In the *Ptc* gene, positive selection has been identified in sites encoding protein domains that bind the Hh morphogen [41, 42]. They determine parameter M, influencing changes in dynamics of system functioning [27], and can also influence kd_1, k_{3a} [27], and K_{Shh} [11] (Table 1).

Changes in proteins Ptc and Smo as well as in proteins mediating their interaction influence k_5, kd_3 [27], and K_{Ptc} [11] (Table 1). Change in k_5 [27] may be also determined by positive selection in the *Smo* gene [41, 42], in regions encoding domain responsible for binding to a high molecular weight complex Fu/Cos2/Ci and G-protein binding domain [42].

In the *Ci* gene and its *Gli* homologs, positive selection has been found for regions encoding domains responsible for positioning of Ci(Gli) proteins in cytoplasm and their binding to transcription cofactors [41, 42]. The rate of *Ptc* gene transcription is regulated by transcription factors Ci(Gli), and so changes in k_0 [27], K_2, and $v_{max,G}$ [11], which are correlated with affinity of these factors to gene enhancers (including *Ptc* gene), strongly influenced dynamic changes of the Hh signal cascade functioning (Table 1).

In a Dpp(BMP) morphogen protein, positive selection was found in the domain responsible for protein dimerization, and in BMP morphogen receptors positive selection was detected in domains responsible for binding of morphogen molecules [41]. Such changes in morphogens and their receptors may significantly influence ratios of concentrations of various BMP dimers and their activity (parameters Ω, β, and λ [40], parameters k_4, k_2, k_3 [39]) (Table 2).

In proteins Tkv and Put (morphogen Dpp(BMP) receptors), positive selection was found not only in the extracellular domain binding Dpp(BMP) dimers, but also to intracellular C-terminus [41] modulating the rate of signal transduction in the Dpp cascade due to binding with protein inhibitors (I-smads, Dad) and activators (R-smad, SARA) [43-46]. Evolution of extracellular (receptor) and intracellular (effector) domain of Dpp(BMP) protein receptors [39] influences k_7, k_{-7}, k_{-5}, K_h, δ_{E}, the dynamic parameters reflecting efficiency of formation of active protein complexes that can trigger the Dpp(BMP) cascade signal transduction inside the cell (Table 2). It is important that such characteristic positive selection of Dpp(BMP) receptors is predictable if one takes into consideration the huge diversity of paralogic groups of Dpp(BMP) morphogens and their adaptive evolution.

Positive selection in Mad and Med proteins influences dynamic parameters of effective Dpp/BMP cascade functioning k_4, Λ, K_h, and ν (Table 2) [39] and involves a region responsible for efficiency of regulation of gene expression, the linker region, containing sites for nuclear export and sites for binding with transcription cofactors [41].

Positive selection in proteins playing “universal” functions (PKA, Slmb, and Exd) has not been found (Table 3) [41, 42]. Interestingly, analysis of parametric robustness of the mathematic model of the Hh signal cascades also did not find sensitivity of the Hh cascade system to changes in parameters of “universal” cell functions [27].

Comparison of modes of evolution of genes involved into the Hh and Dpp(BMP) signal cascades with their functions and types of response. So, based on analysis of parametric robustness of the mathematic models of the Hh and Dpp/BMP signal cascades, hyper-response and inert genes have been identified. Based on analysis of molecular evolution, genes accumulating non-synonymous mutations have been identified [41, 42]. Available nematode genes involved in the Hh/Ptc cascade and homologous to both hyper-response genes and inert genes have also been taken into analysis. Table 4 shows compar-

Table 2. Hypersensitive parameters of the Dpp/BMP signal cascade models [39, 40]

Parameters of models	Parameter description	Proteins or genes determining parameter
	Model of BMP dimer formation [40]	
Ω	ratio of rates of formation of BMP heterodimers to BMP homodimers	Dpp, Scw, enzymes of processing of Dpp and Scw proteins
β	ratio of rates of formation of various BMP monomers (Dpp and Scw)	*dpp*, *scw*
λ	ratio of rates of formation of Scw monomer mutant to normal Scw	*Scw*
	Dpp/BMP cascade model [39], dynamic variables of simple biochemical processes	
k_4	rate constant for SBP/BMP dimer complex formation	SBP, BMP dimers
k_7	rate constant for formation of SBP/BMP dimer/Tkv complex	SBP, BMP dimers, Tkv, Put
k_{-7}	rate constant for dissociation of complex SBP/BMP dimer/Tkv	SBP, BMP dimers, Tkv, Put
k_{-5}	rate constant for dissociation of complex BMP dimer/Tkv	BMP dimers, Tkv, Put
k_2	rate constant for formation of soluble extracellular inhibitor of BMP dimer	Sog, Tsg
k_3	rate constant for BMP dimer binding to its soluble inhibitor	Sog, Tsg, BMP dimers
	Dpp/BMP cascade model [39], dynamic variables of generalized biochemical processes	
Λ	maximal rate of co-receptor SBP formation	SBP expression
δ_E	constant of membrane protein degradation	Tkv, Put, SBP, enzymes of Tkv, Put, and SBP degradation
K_h	half-maximal concentration of complex BMP/Tkv required for SBP production	*SBP*, Tkv, Mad, Med, Mad antagonists
ν	efficiency of production of SBP activated by complex BMP/Tkv	*SBP*, Tkv, Mad, Med, Mad antagonists

ison of evolution modes of genes of the Hh and Dpp/BMP cascades with their functions and type of response of the Hh and Dpp/BMP cascades models. Analysis of parametric robustness of the mathematic models of the Hh and Dpp/BMP cascades has shown that nine of 12 investigated genes may be referred to the genes of hyper-response. Interestingly, they are referred to the class of developmental genes. In eight of nine hyper-response genes *Hh*, *Smo*, *Ci*(*Gli*), *Dpp*, *Tkv*, *Put*, *Mad*, and *Med*, positive selection was related to divergence and formation of large taxons of Bilateria: Ecdysozoa (insects and crustaceous), Lophotrochozoa (mollusks, annelids), and Deuterostomia (chordates, echinoderms) [41, 42] (Fig. 3). One may conclude that the main periods of positive selection during evolution of the Hh and Dpp/BMP signal cascade genes corresponded to the step of divergence of the main large Bilateria taxons. According to paleontology data, this divergence step took place at the Vendian/Cambrian border. Interestingly, divergence of these genes in nematodes and other Bilateria (except Smo lacking homologs of nematodes [22-24]) also accompanied positive selection.

For three genes (*PKA*, *Exd*, *Slmb*), which are not hyper-response genes, the mode of positive selection has not been recognized in any branches [41, 42]. Positive selection has not also been found in the case of the *Ptc* hyper-response gene. In this connection, it should be noted that *Ptc* was the only conservative gene in the nematode Hh/Ptc cascade and in the Hh cascade of other Bilateria. It is possible that adaptive evolution is suppressed by the stabilizing selection due to the important role that it plays in cytoplasmic signal transduction.

DISCUSSION

Genetic determination of morphological signs is complex. A phenotypic sign represents a pooled result of functioning of all components of the gene network.

Table 3. Domain composition of investigated proteins and their functions [41, 42]

Protein	Protein domains and their functions	Positive selection
Hh	N-terminal; binding with Ptc receptor	–
	intein; self-splicing	+
Dpp/BMP	N-terminal; Dpp(BMP) dimerization	+
	C-terminal, TGF-β domain; binding to Tkv, Put, and Sax receptors	–
Ptc	12 transmembrane domains	
	sterol-responsive domain	–
	exporter protein domain	–
	extracellular domain located between 1st and 2nd transmembrane domains; contact with the Hh morphogen	+
	extracellular domain located between 7th and 8th transmembrane domains; contact with the Hh morphogen	+
Smo	Seven transmembrane domains	
	C-terminal; binding with high molecular weight complex Fu/Cos2/Ci	+
	G-protein binding domain	+
	N-terminal, extracellular; possibly binding with Ptc receptor	–
Tkv (type I)	One transmembrane domain	
	serine/threonine kinase domain; catalysis of protein phosphorylation reaction, binding with proteins inhibiting (I-smads, DAD) and activating (R-smad, SARA) catalytic activity of Dpp(BMP) morphogen receptors	+ (domain C-terminus)
	type I activin domain; binding with Dpp(BMP) morphogens	+
Put (type II)	One transmembrane domain	
	G-protein binding domain; catalysis of protein phosphorylation reaction, binding to proteins involved in modulation of Dpp/BMP signal cascade	+ (domain C-terminus)
	type II activin domain; binding to Dpp(BMP) morphogen	+
Ci	zinc finger domain; DNA binding	–
	domains responsible for cytoplasmic localization of protein (positioned from C- and N-end of zinc finger domain)	+
	transcription cofactor binding domain	+
Mad	MH1-domain; DNA binding	–
	MH2-domain; interaction with receptors, transactivation, dimerization	–
	linker domain; contains phosphorylation sites, sites for interaction with ubiquitin ligases, sites of nuclear export, sites for transcription cofactor binding	+
Med	MH1-domain; DNA binding	–
	MH2-domain; interaction with receptors, transactivation, dimerization	–
	linker domain; functions are similar to Mad	+
PKA	catalytic domain of protein phosphorylation	–
	C-terminal flanking domain; unknown function	–
	N-terminal flanking domain; unknown function	+
Slmb	WD-domain; protein–protein interactions	–
	F-box; protein–protein interactions	–
Exd (PBX)	homeodomain; DNA binding	–
	PBC domain; heterodimerization with homeobox proteins, protein–protein interactions	–

Table 4. Relationship between positive selection of genes with formation of large animal taxons and with changeable kinetic parameters influencing dynamics of the Hh and Dpp(BMP) signal cascades

Name of protein (gene)	Kinetic parameters of models corresponding to functioning of gene product	Type of network response to changes in kinetic parameters (+, hyper-response; –, inert)	Functional group of proteins	Positive selection of genes related to formation of large animal taxons [41, 42]
Hh	D_H**; $KShh$*	+ **, *	development	+
Dpp/BMP	k_4, k_2, k_3***; Ω, β, λ****	+ ***, ****	development	+
Ptc	$kd_1, M, k_5, kd_3, k_{3a}$**; K_{Ptc}*	+ **, *	development	–
Smo	kd_3, k_{3a}, k_5**; K_{Ptc}*	+ **, *	development	+
Put	$k_7, k_{-7}, k_{-5}, K_h, \delta_E$	+ ***	development	+
Tkv	$k_7, k_{-7}, k_{-5}, K_h, \delta_E$	+ ***	development	+
PKA	k_8, k_9, kd_2**; $k_{g3rc}, \mathrm{K}_{g3rc}$*	– **, *	housekeeping	–
Slmb	k_8, k_9, kd_2**; $k_{g3rc}, \mathrm{K}_{g3rc}$*	– **, *	housekeeping	–
Ci	k_0**; $v_{max,G}, K_2$*	+ **, *	development	+
Mad	k_4, Λ, K_h, ν	+ ***	development	+
Med	k_4, Λ, K_h, ν	+ ***	development	+
Exd	no data	no data	housekeeping	–

* Results of analysis of parametric robustness of the vertebrate Hh cascade model [11].
** Same for invertebrate Hh cascade model [27].
*** Same for Dpp/BMP cascade model [39].
**** Same for model of morphogen Dpp(BMP) dimerization [40].

Contribution of an individual gene into the phenotypic sign may vary from contribution of a major gene (or genes) to contribution of a modulator gene [47, 48]. Signs may be arbitrarily subdivided into "biochemical" and "morpho-physiological". In the biochemical signs contribution of the gene network core prevails and so it is easy to detect major genes by protein functions. For example, function of protamine proteins consists in compact package of DNA in spermatozoa head, requires basic nature of these proteins, and therefore protamines are characterized by high proportion of basic amino acids, particularly arginine. Mutation eliminating a basic amino acid will increase probability of fixation of the mutation compensating this effect (so-called co-adaptive substitutions [49]). In the result, strong positive selection [50] is not related to change in protamine functions. (Protamines play an important role in fusion of male and female pronuclei during fertilization and motive selection, fixed by Kimura criterion, probably plays a role in two opposite evolutionary tendencies: it forms a reproductive barrier by changing protein structure [51] and also controls fluctuations in common parameters of this protein by means of co-adaptive substitutions [52]. However, the Kimura criterion cannot discriminate these tendencies.) However, it leads to prevailing of non-synonymous substitutions and to change of DNA sequence [52]. Thus, one should be very careful in interpretation of the Kimura positive selection as a motive one.

Study of evolution of biochemical signs revealed a significant role of selection by the gene network core. For example, sperm lysine genes are subjected to positive selection in 25 *Haliotis* mollusk species [53]. The sperm lysine protein is responsible for tissue-specific fusion (acrosomal reaction) of spermatozoa and oocyte. In the *Drosophila* genus, adaptive evolution involves genes encoding sperm proteins semenogelins [54] and genes encoding female genital tract proteins [55]. In humanoid primates and rodents, rapid evolution of *ZP2* and *ZP3* genes has been detected [56, 57]; protein products of these genes determine formation of oocyte *zona pellucida*. The biological sense of positive selection consisted in formation of the reproductive barrier. Biochemical signs also include rapidly evolving genes of the immune system, genes of histocompatibility complex, and genes encoding enzymes [58-60].

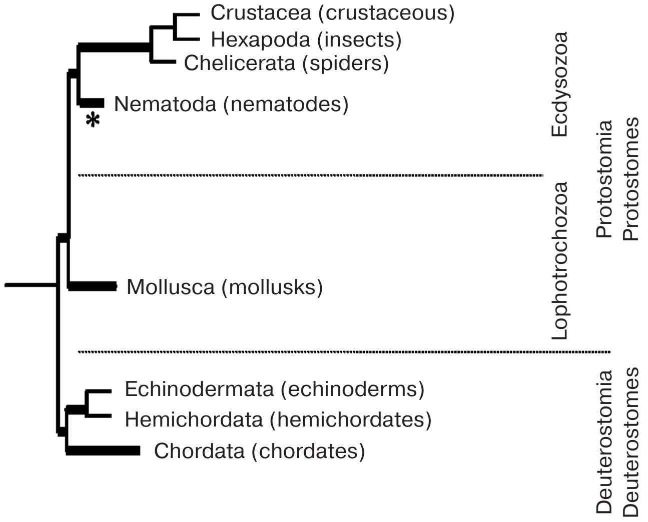

Fig. 3. Projection of events of positive selection of genes of the Hh and Dpp(BMP) signal cascades on the phylogenic tree of types and supraclasses of investigated multicellular animals [41, 42]. Bold lines designate events of detected positive selection in at least one gene of Hh or Dpp(BMP) signal cascade. The asterisk shows events of positive selection of nematode genes.

A role of mutations involving a regulatory component of the gene networks is especially important in evolution of morphophysiological signs, the signs related to shape and size of organs. Formation of each organ requires complex coordinated expression of genes hierarchically organized into the gene network, sometimes containing many hundreds of genes. Decoding of gene networks of Bilateria embryogenesis revealed their block-module nature [41]. Blocks of gene networks form hierarchical structure, in which order and time of module inclusion correspond to formation of compartments from embryonic cells. A link between blocks of the gene networks involves signal molecules, providing distant (morphogen gradient) and local (intercellular contacts) interactions, which are then transmitted into the cell nucleus via the gene networks of signal pathways [61].

Accelerated evolution of G-protein coupled receptors in humanoid primates, motive selection in genes involved in regulation of mammalian transcription [59, 60] and evolution of hormones (particularly, mammalian prolactin [62]) and their receptors are definitely related to evolution of morphophysiological signs. For example, in rodents adaptive evolution of the insulin gene found in caviomorphs can also be related with their large (within the rodent's scale) sizes [63]. Adaptive evolution of paralogs of the growth hormone gene in teleost fishes [64] is attributed to characteristic features in growth and development of these fishes.

In contrast to the biochemical signs, molecular mechanisms of adaptive evolution of the morphophysiological signs are much less studied. Due to experimental difficulties, researchers usually state the fact of positive selection by certain genes and do not investigate the mechanism of its effect on the gene network functioning and expression of the sign. For example, evolution of microcephalin [65] and ASPM [66] genes is attributed to brain development in the human line, because nonsense mutations in these genes cause strong reduction (up to 70%) of human brain size. It should be noted that even in modern man this parameter is subjected to strong variations; for example, masses of brain of O. Cromwell, I. S. Turgenev, and A. France were 2300, 2012, and 1017 g, respectively [67], and in the light of paleontologic data (e.g. discovery of the dwarf form, *Homo florensis* [68]) the range of changeability in the line of human ancestors *Homo ergaster* was even wider [69].

Within the theory of the dynamic systems, there are two modes of formation of new structures, force and parametric ones. In the first case changes in the system definitely reflects the influence (mutation) and consequently may be directly tested by selection. This corresponds to evolution of biochemical signs. In the second case change of parameters results in unstable state of the system; this influences probability of realization of a certain variant of its functioning. During selection, several variants of stabilization of appearing instability are possible [70, 71]. An initial change is not basically tested and it does not determine subsequent trend of evolution. For example, in the simplest circuit with negative feedback regulating protein concentration the parameter "deviation of concentration from normal" is monitored by a regulatory link of the negative feedback and compensates it by changing the rate of protein biosynthesis (the effector link of negative feedback). The circuit is inert to reasons causing deviations from the norm. These may include any mutations. Thus, negative feedback narrows the reaction norm and converts into a neutral one any change in the lower hierarchical levels by turning out selection. It has been theoretically demonstrated: the stronger the negative of feedback the higher this effect is and the lower the phenotypic changeability is in a population [72]. The stabilizing selection favors taxons with negative feedback mainly by fixing negative feedback of higher hierarchical level. This increases the number of regulatory levels, decreases the range of modification changeability; interestingly, non-synonymous mutations accumulated at lower hierarchical levels may evolve in the neutral mode until the regulatory contour exists. On the contrary, collapse of negative feedback during evolution or formation of positive feedback promotes manifestations of all earlier fixed neutral mutations (hyper-manifestation of variability) [7, 72].

Mathematical modeling of dynamics of real gene network functioning allows avoiding experimental difficulties and monitoring how mutations of certain genes would influence functioning of the whole gene network. The pattern of molecular evolution of such well-investigated gene networks as the Hh and Dpp/BMP signal cascades networks, which we have obtained in this study, is consistent with the theory of stabilizing selection [73-76]. This theory based on morphology and paleontology data indicates that complex signs (e.g. tetrapods jaws apparatus [75, 76]) are organized in a way always containing inert links and weak link: essential changes in the inert links insignificantly alter the system, whereas small changes in the weak link may changes functioning of the whole system. Such hyper-response may significantly extend reaction norm. Extension of the reaction norm means attenuation, destabilization of the formed correlation systems of parts of a developing organism. The creative role of the stabilizing selection consists in canalization of some of such labile types of ontogenetic reactions and this forms new directions for taxon evolution. Mutations in a “weak link” may change status of inert links by decreasing their stability resources or triggering their evolution for compensation of the destabilizing effect of the “weak link”. Thus, the evolutionary concept based on morphological data is applicable for molecular evolution!

Identification of the hyper-response genes and inert genes give a new look at the definition of neutral mutations and interrelationship between positive selection at the gene level and evolution at the phenotypic sign of certain function of the gene network. The organism represents a total sum of hierarchically organized regulatory circuits of gene networks. Evolution of genes included in a regulatory circuit should occur in a punctuational manner: initially there is long-term accumulation of mutations, which are then tested altogether by selection during a short period of hyper-manifestations [7, 41]. Evolution of inert genes should proceed in the same way: non-synonymous mutations in these genes insignificantly influence parameters of gene network functioning. Thus, any mutation is neutral until it confers no changes in regulation of the gene network. This conclusion is supported by detection of non-null relative harmfulness of synonymous substitutions based on data on SNPs in the human genome (provided that the harmfulness is defined as real under presentation of certain classes of mutations compared with theoretically possible and that the harmfulness of nonsense mutations is defined as 100%). Some authors [77] have explained this by the effect of synonymous substitutions on splicing sites playing an important role in eukaryotic genome (this is consistent with data obtained from drosophila [78]). Genes that are characterized by mutations weakly influencing the gene network functioning will rapidly evolve in the quasineutral mode. Comparison of genomes of chimpanzee and man has shown that evolution of functionally important genes (including *FOXP2*, *BRCA1* [60], *MYH16* [79] genes, which are related to anthropogenesis) is better explained by the hypothesis of long term attenuation of selection followed by subsequent short burst of adaptive evolution rather than by the hypothesis of strong positive selection.

However, accumulation of quasineutral mutations in a gene is not endless. Any regulatory circuit, any protein structures have certain “factor of safety” [7, 41], and when it is exhausted, all mutations will be tested in the “wholesale” mode (in randomly selected ensembles). We suggest defining such mode as “adaptive optimization” in contrast to classic scheme of the motive selection, when mutations are tested right after their appearance (this is typical for biochemical signs). Thus, evolution of morphophysiological signs should be jerky. Long expectation of mutations in the hyper-response genes occurs simultaneously with inert gene erosion, which occurs in the quasineutral mode. In this case, mutations in hyper-response genes are accompanied by subsequent compensatory evolution followed by divergence. In each of the diverged phyla, stabilizing selection optimizes “its own” parameter. If such mutation does not occur, inert genes accumulated enough mutations evolve towards hyper-response genes and participate in formation of the next hot point of evolution. Such hypothesis is consistent with uneven rates of evolution [80-83] recognized by paleontologists.

The authors are grateful to Dr. D. A. Afonnikov for discussion of the results.

This work was supported by the Russian Foundation for Basic Research (grants 05-04-49283, No. 10104-34/P-18/155-270/1105-06-001/28/2006-1, 05-07-98012-r_Ob, 03-04-48506-a), State contract (02.467.11.1005), Federal Scientific Technical Program “Studies and developments on priority directions of the development of science and technology”, Interdisciplinary Integration Project on Basic Research of the Siberian Branch of the Russian Academy of Sciences “Gene networks: theoretic analysis, computer modeling and experimental construction” No. 119, and program of Presidium of the Russian Academy of Sciences (project “Origin and evolution of biosphere”).

REFERENCES

1. Held, L. I. (2002) *Imaginal Discs: the Genetic and Cellular Logic of Pattern Formation*, Cambridge University Press, Cambridge.
2. Rossant, J., and Tam, P. (2002) *Mouse Development: Patterning*, *Morphogenesis*, *and Organogenesis*, Academic Press, San Diego.
3. Pires-daSilva, A., and Sommer, R. J. (2003) *Nat. Rev. Genet.*, **4**, 39-49.
4. Eppig, J. T., Bult, C. J., Kadin, J. A., et al. (2005) *Nucleic Acids Res.*, **33**, D471-D475.
5. Grumbling, G., and Strelets, V. (2006) *Nucleic Acids Res.*, **34**, D484-D488.
6. Kolchanov, N. A., Nedosekina, E. A., Ananko, E. A., et al. (2002) *In Silico Biol.*, **2**, 0009.
7. Suslov, V. V., Kolchanov, N. A., and Sergeev, M. G. (2006) in *Biodiversity of Dynamics of Ecosystems* (Shumnoi, V. K., Shokin, Yu. I., Kolchanov, N. A., and Fedotov, A. M.,

eds.) [in Russian], SB RAS Press, Novosibirsk, pp. 317-344.

8. De Jong, H. (2002) *J. Comput. Biol.*, **9**, 67-103.
9. Dillon, R., Gadgil, C., and Othmer, H. G. (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 10152-10157.
10. Eldar, A., Dorfman, R., Weiss, D., et al. (2002) *Nature*, **419**, 304-308.
11. Lai, K., Robertson, M. J., and Schaffer, D. V. (2004) *Biophys. J.*, **86**, 2748-2757.
12. Kimura, M. (1985) *Molecular Evolution: a Theory of Neutrality* [Russian translation], Mir, Moscow.
13. Lum, L., and Beachy, P. A. (2004) *Science*, **304**, 1755-1759.
14. Ingham, P. W., and McMahon, A. P. (2001) *Genes Dev.*, **15**, 3059-3087.
15. Nybakken, K., and Perrimon, N. (2002) *Curr. Opin. Genet. Dev.*, **12**, 503-511.
16. Huangfu, D., and Anderson, K. V. (2006) *Development*, **133**, 3-14.
17. Muller, B., and Basler, K. (2000) *Development*, **127**, 2999-3007.
18. Aza-Blanc, P., Lin, H. Y., Ruiz i Altaba, A., and Kornberg, T. B. (2000) *Development*, **127**, 4293-4301.
19. Tyurina, O. V., Guner, B., Popova, E., et al. (2005) *Dev. Biol.*, **277**, 537-556.
20. Dai, P., Akimaru, H., Tanaka, Y., et al. (1999) *J. Biol. Chem.*, **274**, 8143-8152.
21. Ikram, M. S., Neill, G. W., Regl, G., et al. (2004) *J. Invest. Dermatol.*, **122**, 1503-1509.
22. Zugasti, O., Rajan, J., and Kuwabara, P. E. (2005) *Genome Res.*, **15**, 1402-1410.
23. Kuwabara, P. E., Lee, M. H., Schedl, T., and Jefferis, G. S. (2000) *Genes Dev.*, **14**, 1933-1944.
24. Aspock, G., Kagoshima, H., Niklaus, G., and Burglin, T. R. (1999) *Genome Res.*, **9**, 909-923.
25. Aleshin, V. V. (2004) *Zh. Obsch. Biol.*, **65**, 74-80.
26. Ozernyuk, N. D. (2004) *Ontogenez*, **35**, 441-450.
27. Gunbin, K. V., Omelyanchuk, L. V., Kogai, V. V., et al. (2007) *J. Bioinform. Comput. Biol.*, **5**, 491-506.
28. Ratushnyi, A. V. (2006) *Study of Dynamics of Gene Network Functioning by the Methods of Mathematic Modeling*: Candidate's dissertation [in Russian], Institute of Cytology and Genetics, SB RAS, Novosibirsk.
29. Tsumaki, N., Nakase, T., Miyaji, T., et al. (2002) *J. Bone Miner. Res.*, **17**, 898-906.
30. Li, X., and Cao, X. (2006) *Ann. N. Y. Acad. Sci.*, **1068**, 26-40.
31. Nohe, A., Keating, E., Knaus, P., and Petersen, N. O. (2004) *Cell Signal.*, **16**, 291-299.
32. Li, X., and Cao, X. (2003) *Front. Biosci.*, **8**, s805-s812.
33. Rentzsch, F., Zhang, J., Kramer, C., et al. (2006) *Development*, **133**, 801-811.
34. O'Connor, M. B., Umulis, D., Othmer, H. G., and Blair, S. S. (2006) *Development*, **133**, 183-193.
35. Chen, D., Zhao, M., Harris, S. E., et al. (2004) *Front. Biosci.*, **9**, 349-358.
36. Chen, D., Zhao, M., and Mundy, G. R. (2004) *Growth Factors*, **22**, 233-241.
37. Patterson, G. I., and Padgett, R. W. (2000) *Trends Genet.*, **16**, 27-33.
38. Savage-Dunn, C., Maduzia, L. L., Zimmerman, C. M., et al. (2003) *Genesis*, **35**, 239-247.
39. Umulis, D. M., Serpe, M., O'Connor, M. B., and Othmer, H. G. (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 11613-11618.
40. Shimmi, O., Umulis, D., Othmer, H., and O'Connor, M. B. (2005) *Cell*, **120**, 873-886.
41. Gunbin, K. V., Suslov, V. V., and Kolchanov, N. A. (2007) *Inform. Vestnik VOGiS*, **11**, 373-399.
42. Gunbin, K. V., Afonnikov, D. A., and Kolchanov, N. A. (2007) *In Silico Biol.*, **7**, 0047.
43. Moser, M., and Campbell, G. (2005) *Dev. Biol.*, **286**, 647-658.
44. Sutherland, D. J., Li, M., Liu, X. Q., et al. (2003) *Development*, **130**, 5705-5716.
45. Bennett, D., and Alphey, L. (2002) *Nat. Genet.*, **31**, 419-423.
46. Inoue, H., Imamura, T., Ishidou, Y., et al. (1998) *Mol. Biol. Cell.*, **9**, 2145-2156.
47. Ginsburg, E. Kh., and Nikoro, Z. S. (1982) *Decomposition of Dispersion and Selection Problems* [in Russian], Nauka, Novosibirsk.
48. Falconer, D., and Mackay, T. (1996) *Introduction to Quantitative Genetics*, 4th Edn., Longman, N. Y.
49. Afonnikov, D. A., Oschepkov, D. Yu., and Kolchanov, N. A. (2001) *Bioinformatics*, **17**, 1036-1046.
50. Wyckoff, G. J., Wang, W., and Wu, C. I. (2000) *Nature*, **403**, 304-309.
51. Swanson, W. J., and Vacquier, V. D. (2002) *Nat. Rev. Genet.*, **3**, 137-144.
52. Roony, A. P., Zhang, J., and Nei, M. (2000) *Mol. Biol. Evol.*, **17**, 278-283.
53. Yang, Z., Swanson, W. J., and Vacquier, V. D. (2000) *Mol. Biol. Evol.*, **17**, 1446-1455.
54. Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., et al. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 7375-7379.
55. Swanson, W. J., Wong, A., Wolfner, M. F., and Aquadro, C. F. (2004) *Genetics*, **168**, 1457-1465.
56. Swanson, W. J., Yang, Z., Wolfner, M. F., and Aquadro, C. F. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 2509-2514.
57. Turner, L. M., and Hoekstra, H. E. (2006) *Mol. Biol. Evol.*, **23**, 1656-1669.
58. Wildman, D. E., Uddin, M., Liu G., et al. (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 7181-7188.
59. *Chimpanzee Sequencing and Analysis Consortium* (2005) *Nature*, **437**, 69-87.
60. Arbiza, L., Dopazo, J., and Dopazo, H. (2006) *PLoS Comput. Biol.*, **2**, 0288-0300.
61. Gilbert, S. F. (2003) *Developmental Biology*, 7th Edn., Sinauer Associates, Inc., Sunderland, MA.
62. Wallis, M. (2000) *J. Mol. Evol.*, **50**, 465-473.
63. Opazo, J. C., Palma, R. E., Melo, F., and Lessa, E. P. (2005) *Mol. Biol. Evol.*, **22**, 1290-1298.
64. Ryynanen, H. J., and Primmer, C. R. (2006) *Genome*, **49**, 42-53.
65. Wang, Y. Q., and Su, B. (2004) *Hum. Mol. Genet.*, **13**, 1131-1137.
66. Zhang, J. (2003) *Genetics*, **165**, 2063-2070.
67. Etingen, L. E. (1997) *Chelovek*, **4**, 19-28.
68. Brumm, A., Aziz, F., van den Bergh, G. D., et al. (2006) *Nature*, **441**, 624-628.
69. Oppenheimer, S. (2004) *Out of Eden* [Russian translation], Eksmo, Moscow.
70. Shishkin, M. A. (1988) in *Modern Paleontology* (Menner, V. V., and Makridin, V. P., eds.) [in Russian], Nedra, Moscow, pp. 142-169.
71. Golichenkov, V. A., Veselago, I. A., Levina, M. N., and Popov, L. V. (1986) in *Morphology and Evolution of Animals*

(Vorob'eva, E. I., and Lebedkina, N. S., eds.) [in Russian], Nauka, Moscow, pp. 75-86.

72. Kolchanov, N. A., and Shindyalov, I. N. (1991) in *Problems of Genetics and Evolution Theory* (Shumnyi, V. K., and Ruvinskii, A. O., eds.) [in Russian], Nauka, Novosibirsk, pp. 268-279.
73. Shmalhausen, I. I. (1968) *Factors of Evolution: The Theory of Stabilizing Selection* [in Russian], Nauka, Moscow.
74. Waddington, C. H. (1975) *The Evolution of an Evolutionist*, Cornell University Press, New York.
75. Jordansky, N. N. (1986) in *Morphology and Evolution of Animals* (Vorob'eva, E. I., and Lebedkina, N. S., eds.) [in Russian], Nauka, Moscow, pp. 38-49.
76. Jordansky, N. N. (2004) *Zh. Obsch. Biol.*, **65**, 451-463.
77. Gorlov, I. P., and Gorlova, O. Yu. (2007) *Inform. Vestnik VOGiS*, **11**, 363-372.
78. Malko, D. B., and Ermakova, E. O. (2007) in *Proc. of the Moscow Conf. on Computational Molecular Biology 2007* (Gelfand, M. S., ed.) Moscow State University Press, Moscow, pp. 193-194.
79. Stedman, H. H., Kozyak, B. W., Nelson, A., et al. (2004) *Nature*, **428**, 415-418.
80. Simpson, G. G. (1948) *Tempo and Mode in Evolution* [Russian translation], Izd-vo Inostrannoi Literatury, Moscow.
81. Eldrege, N., and Gould, S. J. (1972) in *Models in Paleobiology* (Schopf, T. J. M., ed.) Freeman Cooper & Co, San Francisco, pp. 82-115.
82. Markov, A. V. (2001) *Zh. Obsch. Biol.*, **62**, 460-471.
83. Rozhnov, S. V. (2005) in *Evolutionary Factors of Formation of the Animal World Diversity* (Vorob'eva, E. I., and Striganova, B. R., eds.) KMK Press, Moscow, pp. 156-170.