A simple model reveals why complex evolutionary innovations follow predictable paths

Daohan Jiang[1], Matt Pennell[2], Lauren Sallan[1*]

[1]Macroevolution Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa 904-0497 Japan.

[2]Department of Computational Biology, Cornell University, Ithaca, NY 14853, USA.

*Corresponding author. Email: lauren.sallan@oist.jp

Determining the principles underlying the origin of novel characters has been a fundamental goal of evolutionary biology. Yet, key mechanisms remain poorly understood, hindered by the lack of a general, mechanistic model that unites genotypic and phenotypic change and predicts outcomes. Here, we present a model founded in developmental biology, where phenotype is controlled by a hierarchical gene regulatory network (GRN) consisting of regulators specifying character identity and effectors producing specific states. While our model is simplified, evolutionary simulations for divergence between repeated characters and switching between alternative identities easily recreated empirical patterns. While observations of the emergence of novel identities are lacking due to their rarity or multi-step nature, our simulations reveal that the most complex characters exhibit the strongest convergence in regulatory pathways (deep homology). Our model provides insights into the mechanisms underlying evolutionary novelties and offers a framework for the developmental evolution of a variety of traits.

Introduction

To understand mechanisms underlying the origin of morphological characters with novel forms and functions has been a fundamental goal of evolutionary biology. Central to the problem of evolutionary novelties is how developmental processes can constrain and bias the trajectory of of evolution. It has been proposed that developmental constraints, in general, can have a dispositional effect on phenotypic evolution by channeling the effect of mutations on the phenotype and thereby biasing the availability of genetic variations (*1–9*). Advancements in developmental biology over the past few decades have shed light on developmental mechanisms underlying phenotypic divergence between species and between body parts of the same organism. For example, phenotypic similarity between homologous body parts of different species and that between repeated body parts of the same organism (serial homologs) results from similar gene expression programs during their development—i.e., which genes are expressed, their expression levels, and timing of their expression; by the same principle, difference in the gene expression programs is responsible for phenotypic variation among the otherwise identical body parts (*10–15*).

There is abundant evidence that phenotypic changes result from the interplay between genes at different levels of an organizational hierarchy during development. This suggests that mutations in some genes have larger effects on character form than others depending on their position within this hierarchy. Observed morphological differences between species or between serial homologs have, in many cases, been hypothesized or confirmed as attributable to differential local expression of upstream ”master regulator” genes that activate or suppress many other genes and/or pathways. Expression of such regulatory genes can specify the identity of a body part, but not necessarily specific character states (the observable realized phenotype, such as size, shape, or color). That is, the expression profile of these genes is shared by of homologous body parts of different species, regardless of how different their forms are (*11–13*). A well documented case of such regulator is the *Ultrabithorax* (*Ubx*) gene: *Ubx* expression specifies the identity of hindwings in insects, and the loss of *Ubx* expression causes forewings to grow at the position of hindwings, regardless of specific size or morphology of the wings (*16–24*). For instance, in fruit flies, the perturbation causes the halteres to be replaced by wing blades, whereas in beetles, it is the hindwings being replaced by elytra. Another example of well-studied character identity determination mechanisms is that

of tetrapod limbs, in which case two *T-box* genes, *Tbx4* and *Tbx5*, specify forelimb and hindlimb identities respectively but not specific limb morphologies (*25–31*). In addition to morphological characters, identities of cell types are also found to be regulated in similar ways (e.g., (*32–39*)). This suggests that modeling the interactions between master regulators and their downstream effectors can be used to infer the resultant phenotypes and phenotypic divergence for many traits.

Despite the progress in understanding the underlying developmental mechanisms linking genotypes and phenotypes, there has been little systematic study of their implications for morphological evolution and particularly the origin of novel characters. One major obstacle is that there is no existing quantitative, mechanistic model linking genetic changes and developmental processes. In addition, developmental models that make quantitative predictions about developmental evolution are rare to our knowledge. The few examples of such models include those for the growth mollusk shells (*40, 41*), which lacked a genetic basis, and for mammalian molars (*42*), which was based on a well characterized pathway. The latter allowed developmental constraints to be incorporated in phylogenetic comparative analyses to study the evolutionary dynamics (*43*), thus suggesting the utility of such models for evolutionary study of phenotypic change.

In the absence of quantitative models and sufficient information on the developmental basis for most traits, empirical observations have routinely been used to infer processes underlying the emergence of new forms. These inferences then form the basis of qualitative models and hypotheses, which are often used to make predictions about the plausibility of specific and potential character states. Qualitative models for novelty are then assumed to be supported if predicted intermediate forms exist in the fossil record or matches normal ontogenetic stages (recapitulation) or mutant phenotypes arising from mutations. They are typically rejected if forms predicted to be implausible are observed in any such data. For instance, the qualitative model for feather development and evolution proposed by Prum (*44*) makes predictions about possible forms of early feathers, which are supported by findings in various early-diverging coelurosaur groups (*45*). Another case of a qualitative model (un)supported by fossil evidence is the teleost caudal fin, where a growth series of the non-teleost ray-finned fish *Aetheretmon valentiacum* overturned a classic model of recapitulation of evolutionary stages during development while supporting an alternative model of derived states for non-homologous structures (*46*). The notochord-bearing "tail" in early embryos and adult fossils was proposed to be serially homologous with tetrapod limbs in terms of shared dependence on

homeobox (*Hox*) gene-regulated pathways, based on existing qualitative models of limb and other vertebrate outgrowth (*46*). This new hypothesis is in turn supported by subsequent developmental observations; genetic manipulation of *Hox13* paralogs in zebra fish changes the growth of caudal structures in expected ways (*47*), and outgrowth of structures such as the zebrafish cloaca, fin rays and tetrapod digits originated from co-option of *Hox13* and are serially homologous (*48*). There are many other qualitative models, hypotheses, and proposed mechanisms which remain supported only by empirical observations. For most, there has been no way objectively test the validity and logic of their underlying assumptions, simulate outcomes, identify the minimum requirements, or make predictions of quantitative data should be structured like if the model were accurate.

To address these issues, we need a generalizable quantitative model under which one can investigate and analyze how features of developmental processes shape phenotypic evolution. Here, we present a mathematical model that incorporates key general features of gene regulatory networks (GRNs) underlying the development of morphological characters and cell types and explore its implications for phenotypic evolution and the emergence of novelties. Under this model, the character state of a body part is controlled via a hierarchical gene regulatory network consisting of two classes of genes: the regulatory genes (”regulators” for short) and the effector genes (”effectors” for short). The regulators encode *trans*-regulatory factors (e.g., transcription factors) that interact with *cis*-elements (e.g., promoters and enhancers) of the effect genes and regulate the effectors' expression. A regulator expression program (combination of expression levels of regulators) that specifies a character identity is referred to as a character identity network (ChIN) (*12, 13, 49, 50*), which is also known as a core regulatory complex in the context of cell type identity determination (*51*).

Using our quantitative modeling framework, we investigated three types of evolutionary novelties. The first type of novelty is the divergence in character states between serially homologous body parts with conserved identities. These characters have distinct ChINs that are conserved during evolution, and divergence in character states is due to substitutions in *cis*-elements of the effectors. Prime examples of such characters include the aforementioned cases of tetrapod limbs and insect wings, among others. The second type of evolutionary change we examined is that a body part switches to an alternative identity by expressing a pre-existing alternative ChIN. As each regulator can regulate the expression of many effectors, alternation of their expression can lead to substantial transcriptomic change, potentially leading to drastic phenotypic changes. The last type

of novelty is the emergence of a novel ChIN via evolutionary changes in local expression levels of multiple regulators. To illustrate principles underlying these three types of evolutionary novelties, we performed evolutionary simulations and examined evolutionary dynamics under a variety of selection regimes. We explored conditions under which each type of novelty is likely to evolve, demonstrating how patterns of evolution are connected to the underlying mechanisms.

Results

An generalizable model for the developmental control of character identity and state

In our model, we assume that there are n regulatory genes that specify the character identity and m effector genes that encode the character state. Expression levels of the effector genes are then given by $\mathbf{Y} = \exp(\mathbf{AX})$, where $\mathbf{Y}$ is a vector of length m and $\mathbf{A}$ is an $m \times n$ matrix summarizing per-unit-expression regulatory effects of the regulators on the expression of the effectors. Element $A_{i,j}$ of $\mathbf{A}$ represents the effect of the j-th regulator on the expression of the i-th effector, and its values captures both strength and direction of interactions between the regulator and *cis*-regulatory elements of the effector: a positive value represents an activation effect, whereas a negative value represents repression. The $\mathbf{A}$-matrix can further be decomposed as a product of two matrices $\mathbf{A} = \alpha\beta$, where α captures effects of *cis*-elements on the effector genes and β captures binding affinity between gene products of the regulators and *cis*-elements of the effectors. At last, the realized character state, which is d-dimensional and represented by a vector of length d, is given by $\mathbf{z} = \mathbf{B} \ln \mathbf{Y}$, where $\mathbf{B}$ is a $d \times m$ matrix characterizing the per-unit-expression-level effect of the effectors on the phenotype (see also Materials and Methods). Homologous body parts of different organisms are expected to have the same $\mathbf{X}$, but may have different states due to differences in their $\mathbf{A}$-matrices. By contrast, serially homologous body parts with distinct identities within the same organism will have difference(s) in $\mathbf{X}$, and potentially different character states depending on the structure of $\mathbf{A}$.

Character state divergence between body parts with conserved identities

We considered two serially homologous body parts with distinct identities, each specified by a respective regulator. Character states of the two body parts (denoted by z_1 and z_2, respectively) are produced by the same set of effectors. Each effector is regulated either by a single *cis*-element that interacts with both regulators, or two *cis*-elements that respectively interact with the two regulators. We considered scenarios where different proportions of effectors fall into these two categories. Expression levels of the regulators in the two body parts are conserved over time, and evolutionary changes in character states are mediated by mutations in *cis*-elements that affect the strength of *cis*-*trans* interactions (β). Mutations in a *cis*-element that is a shared target of two regulators will tend to have concordant effects on its binding affinity with both regulators, thereby causing concordant effects on effector gene expression in two body parts. By contrast, mutations in *cis*-elements that are not shared targets of the two regulators can result in differential expression of the effectors between two body parts, thereby character state divergence between them (Fig. 1A).

We considered three scenarios of directional selection. In the first scenario, z_1 is under directional selection for a greater value and z_2 is under stabilizing selection. In the second scenario, which is referred to as a scenario of diversifying selection, the optimum of z_1 is greater than the ancestral state and whereas that of z_2 is smaller than the ancestral state. In the last scenario, which is referred to as a scenario of concordant selection, greater values of both z_1 and z_2 are selected for. In the absence of any genetic constraint, phenotypic divergence between serial homologs is expected to increase in the first two scenarios but not the third one. In addition, we considered scenarios of no correlational selection, positive correlational selection (optimal scaling line has a positive slope), and negative correlational selection (optimal scaling line has a negative slope). Together, nine combinations of directional and correlational selection were examined (see also Materials and Methods). For each regime of selection, we simulated 10 replicate populations for 10000 generations in SLiM (*52*), and examined the mean level character state divergence between two body parts across populations and the mean evolutionary change in z_1 relative to the ancestral state.

We found that, when two body parts had different optimal states and correlational selection is either absent or positive, the degree character state divergence between body parts decreased with the fraction of effector genes regulated through shared targets of two regulators (Fig. 1B-C, E-F).

In other scenarios, the level of character state divergence is low and rather insensitive to sharing of targets by two regulators (Fig. 1D, G, H-J). Evolutionary change of z_1, which has the same optimum in all scenarios, shows consistent patterns (Fig. S1); z_1 also shows less adaptive evolution when the the gradient of directional selection and optimal scaling line are aligned, as the strength of directional selection is effectively weaker in such scenarios (Fig. S1F, H). We also examined how these quantities changed over time, and found that difference in their rates of change between different parameter combinations generally agree with those in end-point values (Fig. S2-S3). Together, our results demonstrate the impact of developmental constraint on adaptive divergence between character states of serial homologs.

Switching between pre-existing identities

We considered a single body part that can assume one of two potential identities (denoted I_1 and I_2, of which I_1 is ancestral) by expressing two different regulators that respectively activate two distinct sets of effector genes. Effectors regulated by two regulators affect the character state differently, so the two identities correspond to two distinct character states. Which regulator to express in a given body part is determined by the local concentration of a morphogen, denoted c. When c is equal to or above a cutoff, the regulator specifying I_1 is expressed; otherwise, the one specifying I_2 is expressed. Mutations that alter c can turn off one regulator while turning on the other, resulting in change in expression states of their targets and drastic change in character state (Fig. 2A). The ancestral value of c is just near the cutoff to produce I_1, and a mutation that lowers c will switch the identity; that is, there is minimal genetic constraint on identity switching on an ancestral background.

We performed simulations in SLiM under five regimes of selection characterized by different optimal states (z_{opt}), including the character state produced by the ancestral identity I_1 (given ancestral regulatory parameters; the same for the rest), the phenotype produced by the non-ancestral identity I_2, an intermediate state right in between those produced by I_1 and I_2, an intermediate phenotype that is closer to that produced by I_1, and an intermediate phenotype that is closer to that produced by I_2. For each regime of selection, we simulated 20 replicate populations and examined the proportion of populations where I_2 is fixed. As predicted, I_2 was fixed in most, albeit not all

populations when z_{opt} is the character state it produces (Fig. 2B, leftmost column). Similar results were observed when the optimum is not the same but relatively close to the character state produced by I_2 (Fig. 2B, second column), indicating an identity shift can happen even if a shift does not lead to the exact optimal phenotype. Identity shift happened relatively infrequently under other regimes of selection as well as under neutrality, as expected (Fig. 2B). Mean expression of the regulator underlying I_2 showed consistent results (Fig. 2C). In the two scenarios where identity switching was frequent, the mean character states across populations ended up close to the respective optima (S4), indicating complementation by evolutionary changes in regulator-effector interactions. Together, our results show that identity shift can facilitate adaptation when the new optimum is in the vicinity of a character state that can potentially be produced by an alternative character identity.

Origin of novel character identities

We considered the evolution of a body part toward an optimal character state ($\mathbf{z}_{opt}$). We assuming that each regulator, together with effectors it activates, is expressed many pre-existing body parts and cell types other than the focal one, such that mutations affecting interactions between the regulator and its targets will be highly pleiotropic. As a result, adaption can only be mediated by changes in local expression levels of the regulators (the ChIN) and achieved upon the emergence of a ChIN that produce an adaptive character state (Fig. 3A).

Under this setting, effects of the regulatory genes on the character state can be summarized by a matrix $\mathbf{C} = \mathbf{BA}$, where $C_{i,j}$ denotes the effect of the j-th regulator on the i-th phenotypic dimension per unit expression. The ChIN(s) that can give rise to a given optimal state $\mathbf{z}_{opt}$ can be obtained by solving the linear system $\mathbf{CX} = \mathbf{z}_{opt}$. This means that a character state of low complexity (dimensionality) can be produced a variety of ChINs, and evolution towards such a state can take diverse courses (Fig. 3B). As $\mathbf{z}_{opt}$ becomes more complex, the solution space for $\mathbf{X}$ becomes smaller, such that evolutionary trajectories from different ancestral states are likely to lead to similar $\mathbf{X}$ (Fig. 3C), resulting in the phenomenon of deep homology (the same genes repeatedly recruited to specify identities of independently evolved characters).

To confirm the above general predictions, we simulated the evolution of a novel character whose identity is specified via the expression of 50 regulators. We considered scenarios where the

number of phenotypic dimensions (traits) subject to selection ($d_{selection}$) varied. For each regime of selection, we simulated 20 replicate lineages with the same ancestral expression state $\mathbf{X}_a$. At the end of the simulation, we examined variances of regulator expression levels among lineages as well as their mean expression divergence from the ancestor. We found a non-monotonic relationship between expression variance and $d_{selection}$, where expression variance first increased, peaking when 15 to 20 traits are under selection, and then decreased (Fig. 3D). The degree of expression divergence from the ancestor mostly increased in the range of parameters examined, only dipping when more than 40 traits were selected (Fig. 3E). Together, when $d_{selection}$ is within an intermediate range, the adapting lineages evolved farther away from the ancestral regulatory expression program but closer to each other, showing a trend of convergent evolution. The mean fitness across replicate lineages ($\bar{\omega}$) at the end generally decreases with $d_{selection}$, especially when $d_{selection}$ exceeds 20 (Fig. S5), demonstrating a cost of complexity (*53*) and stronger constraint on adaptation at higher $d_{selection}$.

Discussion

In this study, we present a generalizable model of developmental evolution and used it to explore principles of evolutionary innovations. We focused on a simplified, two-level model where two classes of genes, namely regulators that specify character identities and effector genes that produce character states, and regulatory effect of the regulators on the effectors, were considered. This abstraction, while allowing the mathematical models and simulations to be simpler and more manageable, also captures essential properties of hierarchical GRNs underlying the development of morphological characters formation of cell types. Within-group regulation and feed-back loops, on the other hand, can potentially have their own evolutionary consequences, as structural properties of GRNs do in general (*54–56*), and are of interest to future studies. Importantly, interactions between regulators, such as mutual suppression that mutually exclusive expression and positive feedback loops that drive co-expression, are involved in specification identities of morphological characters and cell types (*31, 36*). Our simulations, while not explicitly modeling interactions between regulators, can be interpreted as having some incorporated implicitly. For instance, interactions driving co-expression of regulators can be readily modeled if each entry of $\mathbf{X}$ is interpreted as a set of co-expressed genes instead of a single gene. One can also model expression levels of two or

more regulators as functions of a latent variable (morphogen concentration) to recapitulate specific patterns of co-expression. In addition, it should be noted that, the that the same gene can be involved in the development of different body parts that are not serially homologous and have different roles in the corresponding GRNs (e.g., various *Hox* genes (*57*) and *Shh* (*58*)). Such pleiotropy is beyond the scope of this study, but will be essential to modeling the coevolution of distinct types of organs.

In our simulations, structure of GRNs underlying development underwent minimal change. Specifically, *de novo* birth of new *cis*-elements (new column of α and new row of β) was not modeled. Modeling such events requires information on their frequency the distribution of effects of new *cis*-elements. Recent years have also seen progress made in predicting regulatory potential of putative *cis*-elements and detecting selection on regulatory activities (*59–66*); findings in this area may shed light on how gene regulatory networks underlying development could potentially be rewired during evolution and help better parameterize models for developmental evolution. Modeling the birth of new *cis*-elements and the resulting change in GRN structure could also help understand the role of second-order selection in evolutionary innovations. It has been hypothesized that second-order selection for mutational robustness and/or evolvability has been responsible for modularity of GRNs controlling development (*67–71*), and understanding the mutational input for this evolutionary process is key to understanding its dynamics.

We examined models of selection where the relationship between phenotype and fitness is specified while mechanisms underlying the relationship are not. However, such simple models can also represent more specific scenarios of potential interest. For example, negative correlational selection could be mediated by selection for division of labor in response to trade-off between functions performed by the same body part (*72*). Our modeling framework can readily be developed to explore implications of specific mechanisms of selection; one can introduce variables that mediate the effect of the phenotype on fitness—e.g., locomotion and foraging performance and model them as custom-define functions of the character state. Models of adaptation examined in this study are simple models concerning a single episode of adaptive evolution with the optimal character state pre-specified. Another class of models that would be of interest is models with moving adaptive landscapes, which may better explain variation in the distribution of phenotypes across divergent clades. In addition, movement of the adaptive landscape can potentially interact with developmental constraints to shape patterns macroevolutionary divergence—e.g., when there is

strong developmental (and thereby genetic) constraint, populations track the moving optimum less well and may end up stranded near an ancestral state and display less evolutionary divergence (*73*). It remains an open question, however, whether adaptive landscapes' movement is generally predictable, and if so, how they move (*74, 75*). For body parts with positional identities, there could also be correlation between their identities and functional constraints they are subject to; that is, how well a body part can perform a given task depends on where it is located (e.g., position along an axis). Such constraints could bias the movement of the adaptive peak and might explain some observed patterns of evolution: for example, bipedal tetrapods typically support their weight using their hindlimbs rather than forelimbs (*76*), and arthropod appendages specialized for assisting feeding need to posited near the anterior end (*77–79*).

Under our modeling framework, we examined three types of evolutionary novelties, the first being character state divergence given conserved identities, mediated by evolutionary changes in interactions between regulators and *cis*-elements. Under certain regimes of selection, the rate of adaptive divergence in character states of two body parts is faster when regulators specifying different identities share a smaller portion of targets (Fig. 1B-J, Fig. S2). This is consistent with the notion that decoupling of developmental regulation of different serial homologs enhances evolvability and allows more diverse combinations to evolve (*80*). Our model for such characters can potentially help us understand evolutionary mechanisms underlying a broad range of novel character states. A potential case of this type of evolutionary change the evolution of long arms in bats: substitutions in an enhancer increased the expression level of *Prx1*, a gene that does not specify limb identities, in developing forelimb bones, resulting in forelimb elongation (*81*). Similarly, limb size reduction, which has taken place repeatedly in various tetrapod lineages (*82, 83*), could have been mediated by similar mechanisms as well (we note that the complete loss of limbs is more appropriately a loss of character identity (*84*) rather than an extreme character state, albeit potentially driven by similar selection regimes).

The second type of novelty we examined is the switching of character identity via turning on and off regulators. We found such evolutionary changes are likely to occur when the optimum is closer to the phenotype produced by the derived identity than that produced by the ancestral identity and can be complemented. Scenarios examined in our simulations can potentially explain drastic evolutionary changes between apparently discrete states arising from the same developmental cell

lineages in similar positions. A putative example is the transition between feathers and scales states in birds as well as non-avian dinosaurs: mechanistically, such a conversion can be achieved by simply activating or suppressing *Sonic Hedgehog* (*Shh*) signaling (in given body regions) (*85–87*). On the other hand, there is indeed substantial diversity in the spatial distribution of feathers among theropods (*45, 88*), which could have been mediated by genetic changes that altered the expression of master regulators like *Shh*.

The last type of novelty we examined is the origin of novel character identities by assembling a novel ChIN. We found a non-monotonic relationship between complexity of the selected character state and parallel evolution of the underlying regulator expression program. When the selected state has low complexity, evolving lineages with the same ancestral expression program underwent relatively minimal evolutionary change, thereby showing rather low expression variance, echoing observed cases of parallel adaptation where common genetic mechanisms were repeatedly involved (e.g., (*89*)). Expression variance among lineages peaked at an intermediate level of complexity, which interpret as ecological or biomechanical adaptation achievable by diverse anatomical or physiological mechanisms (many-to-one mapping between form and function) (*90*). As complexity of the selected state further increases, adaptation becomes more constrained, resulting in pervasive parallel evolution. This finding has major implications for deep homology, the phenomenon that the same genes or pathways are independently recruited for the development of non-homologous characters in different lineages (*50, 91*); prime examples of non-homologous characters sharing underlying developmental regulators include body appendages of different bilaterian groups (*92*) and fin rays and tetrapod digits (*93*). Our finding indicates that such evolutionary convergence can be attributed to limitation imposed by the regulator repertoire: while the number of effectors that can potentially contribute to novel character states can be large, the number of available regulators is much more limited. We also note that the degree of constraint on expression evolution was reduced by the lack of interaction between regulators in our simulations; such interaction can potentially lead to greater limit on the solution space and result in stronger tendency of parallelism.

When interpreting results of our evolutionary simulations, we can disentangle evolutionary changes of character identity and state, as we know the truth of our simulations. However, in empirical studies, it is not always clear if the observed states resulted from character identity changes, character state changes, or acquired through loss and replacement rather than fully homologous,

which can potentially confound further evolutionary analyses. Thus, more developmental studies will be necessary to determine how broad a range of novelties are explained by either of the two types of scenarios examined in this study. A general approach to distinguishing evolutionary changes of character identity and state is to compare the effect of the same genetic or gene expression perturbation across species. Specifically, one can test if the same perturbation produce similar character states in different species, or affect homologous characters of different species while producing dissimilar character states. A prime example of such investigations is the aforementioned case of *Ubx*. In both fruit flies and beetles, hindwings are replaced by forewings upon loss of *Ubx* expression while the resulting character state changes differ in the two groups, indicating the gene is responsible for specifying hindwing identity but not specific character states of halteres or 'normal' wings (*12, 16–23*). Hence, evolutionary changes mediated by *Ubx* expression evolution shall be interpreted as identity instead of state changes. More such investigations will be necessary to clarify the nature of other instances of novelties and understand the underlying evolutionary mechanisms using our model.

Our modeling framework can be extended to study the dynamics of evolutionary innovations mediated by a broader range of developmental genetic mechanisms. One type of evolutionary change of interest is change in the number of serially homologous body parts within an organism; specifically, how the duplication of body parts could potentiate the evolution of novel character identities under selection for novel functions or division of labor. Mechanisms mediating developmental and selective constraints on the number of serial homologs, however, are poorly understood and likely vary among different types of characters, making it difficult to model the evolution of the number of serial homologs. In particular, it is challenging to distinguish constraints resulting from selection on the number *per se* and indirect selective constraint mediated by pleiotropic effects of mutations that alter the number of them (e.g., mutations responsible for Pallister–Hall syndrome cause polydactyly along with other symptoms (*94*)). Pleiotropy is known to constrain the evolution of traits that are not themselves under selection, resulting in reduced level of within-population phenotypic variation and evolutionary rate (*95–98*). For discrete traits showing little variation among species, however, it can be much more challenging to tease apart from direct and indirect effects of selection. A type of character that may potentially offer implications for this problem are serial homologs whose number is conserved within some clades but variable in other clades and between

clades. For example, the number of cervical vertebrae is highly conserved among mammals (typically seven, with few exceptions) but much more variable in sauropsids, likely due to different levels constraint resulting from pleiotropy (*99–103*). It would be of interest for future studies to compare developmental processes as well as selective constraints in different clades to provide insights into mechanisms underlying constraints on the number of serial homologs. Another type of genetic change that could lead to novelties is gene or genome duplication. Specifically, duplication of regulators that are expressed in different body parts may allow different duplicates to undergo subfunctionalization and become expressed locally in specific body parts, resulting in distinct character identities and potentiating further character state divergence (*104*). It is hypothesized that genome duplication facilitated the evolutionary innovations and ultimately diversification in certain clades, such as in teleost fishes (*105, 106*). Indeed, genome duplication in the common ancestor of teleosts made it possible for a novel axial patterning pathway to be established (*107–109*), but whether it is responsible for a broader range of novelties as well as species diversity has not been rigorously tested. Gene and genome duplications can be readily modeled in our framework as duplications of entries of **X**, **A**, and/or **B**, and one can examine and compare dynamics of phenotypic evolution in genotypes containing different numbers of gene copies.

Together, in this study, we present a generalizable model for the evolution of novel morphological characters and cell types, incorporating their underlying developmental mechanisms. Using this model, we investigated three types of evolutionary novelties and demonstrate the interplay of selective and developmental constraints in shaping the evolutionary dynamics. Readily extendable to a broader range of novel characters, our model offers a framework for understanding principles of evolutionary innovations.

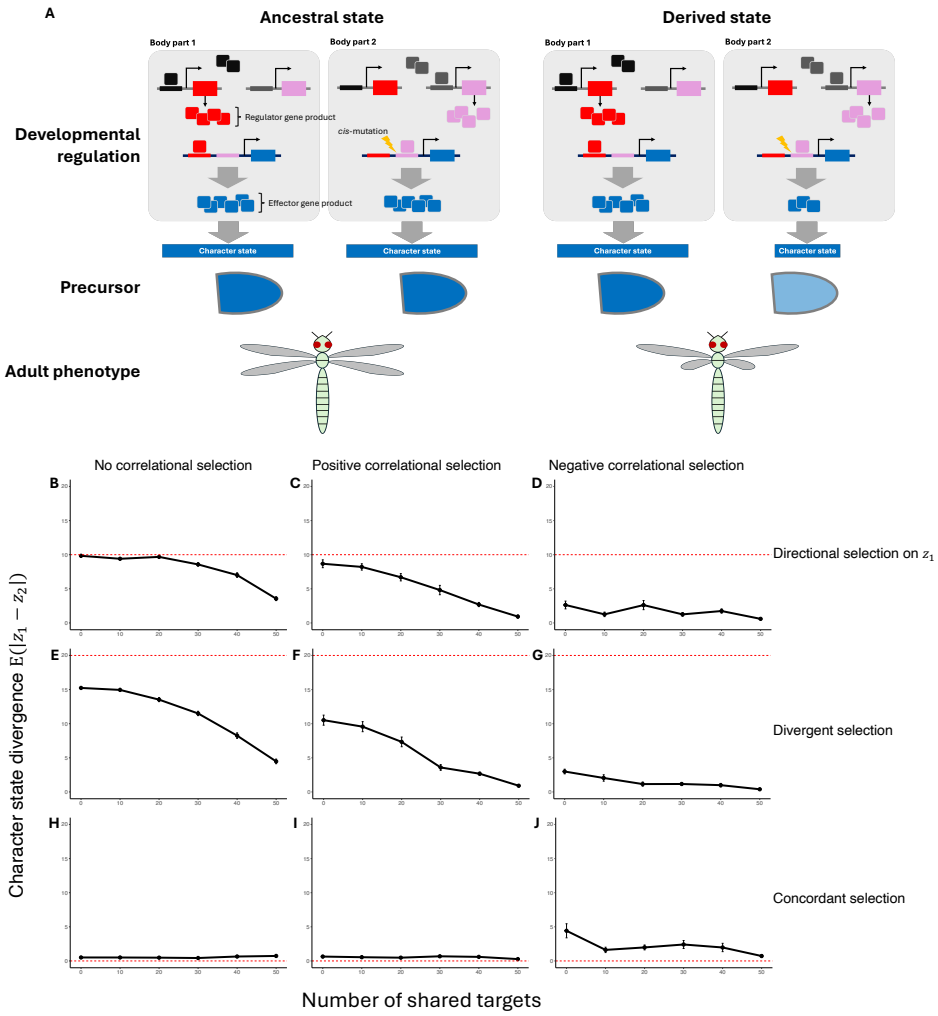

Figure 1: Character state divergence between body parts with conserved identities. (A) Schematic illustration of the scenario, where two regulators specifying distinct identities activate the same effector gene via different *cis*-elements, and the effector gene's expression level determines the character state (i.e., length of wings). The *cis*-mutation shown affects the interaction between a regulator and a *cis*-element, thereby effector expression level in body part 2. Upper row: developmental regulation in two body parts. Middle row: the developing precursor part, with color corresponding to concentration of the effector gene product. Lower row: character state in the adult organism. (B-J) Relationship between state divergence between two body parts at the end of the simulation (calculated as $|z_1 - z_2|$) and the number of effector genes regulated via shared targets of two regulators. Red dashed lines correspond to the level of divergence expected upon full optimization by selection. Height of each error bars is twice the standard error (sample size $n = 10$).

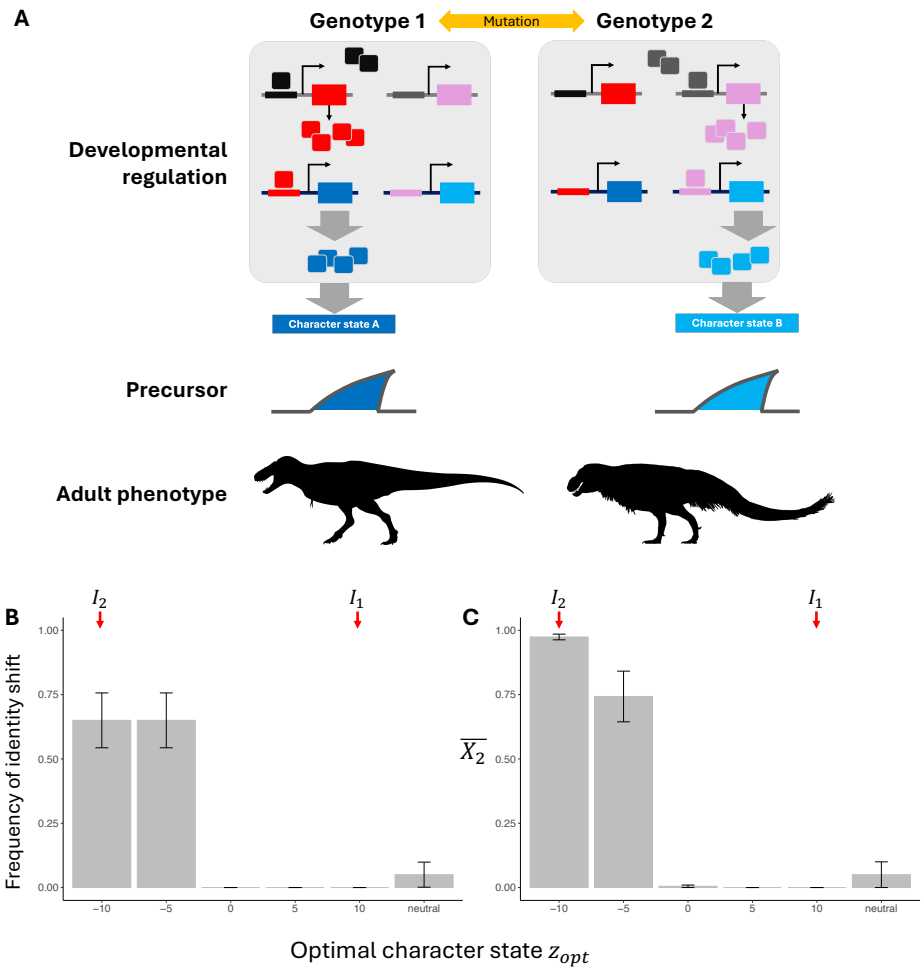

Figure 2: Switching between pre-existing identities. (A) Schematic illustration of the scenario. Left and right columns represent two genotypes where different regulators are expressed, which results in different character identities and states. Upper row: developmental regulation in two body parts. Middle row: the developing precursor part, with color indicating which effector genes are expressed. Lower row: character state in the adult organism. (B) Fraction of simulated populations where the non-ancestral identity I_2 is fixed. (C) Mean expression level of the regulator underlying I_2 across populations. Height of each error bar in (B) and (C) is twice the standard error (sample size $n = 20$). Character states produced by I_1 and I_2 give ancestral regulatory parameters are marked with arrows in (B) and (C). *T. rex* silhouettes in (B) are from PhyloPic (https://www.phylopic.org/), and were created by Matt Dempsey (left, non-feathered) and Matt Martyniuk (right, feathered), respectively.

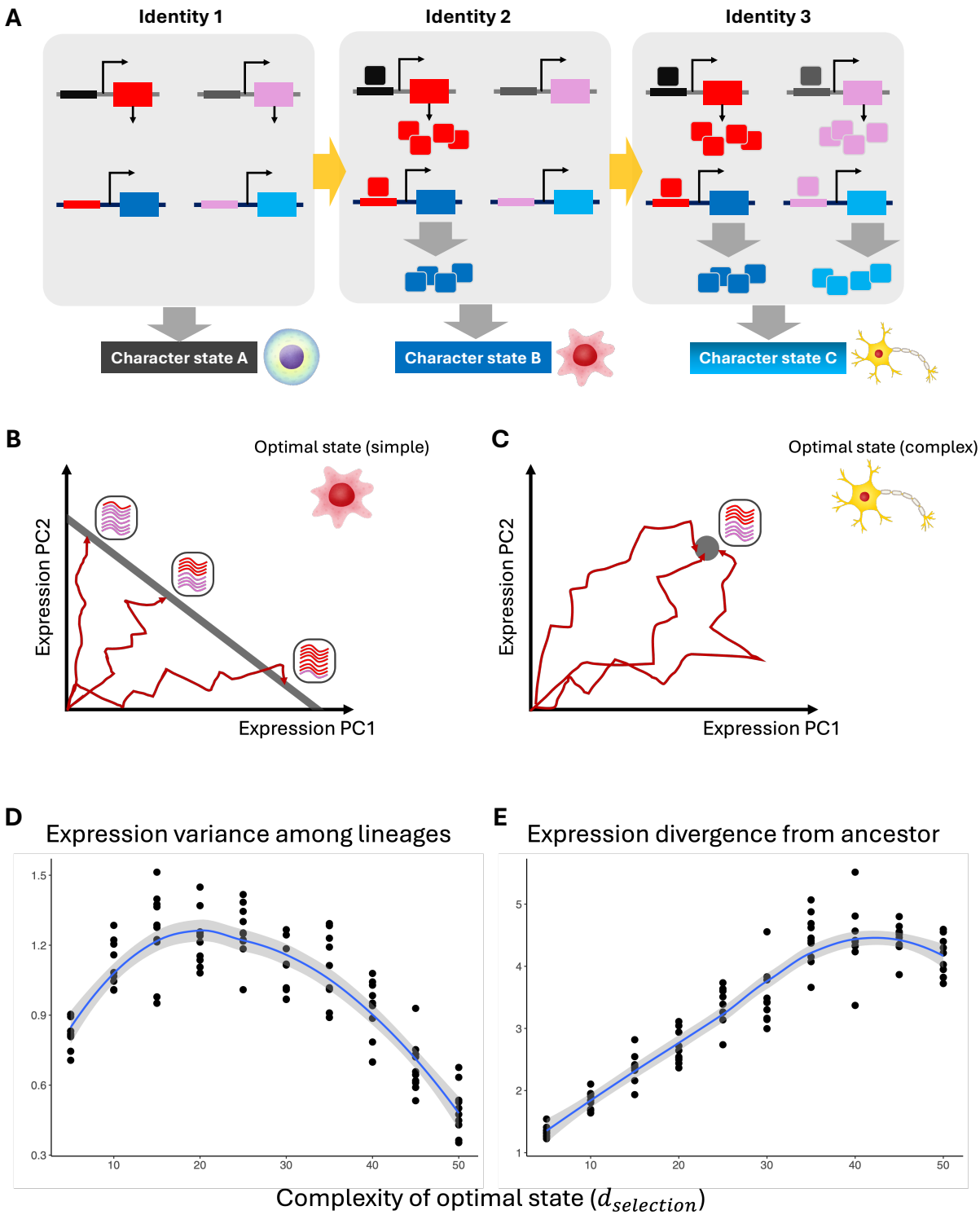
A
Identity 1
Identity 2
Identity 3
Character state A
Character state B
Character state C
B
Optimal state (simple)
Expression PC2
Expression PC1
C
Optimal state (complex)
Expression PC2
Expression PC1
D
Expression variance among lineages
E
Expression divergence from ancestor
Complexity of optimal state ($d_{selection}$)

Figure 3: (A) Schematic illustration of the origin of a novel character identity network (ChIN) via multiple evolutionary changes in regulator expression. (B-C) Schematic illustrations of different scenarios of selection, where the axes are two principal components of regulator expression. (B) The optimal character state can be produced by various ChINs (gray line) such that lineages evolving from a common ancestor can end up acquiring diverse ChINs (red arrows). (C) The optimal state can only be produced by a specific ChIN, in which cases different lineages evolve towards the ChIN (gray point). (D) Relationship between regulator expression variance across replicate lineages at the end of the simulation and complexity of the optimal character state. (E) Relationship between mean regulator expression divergence from the ancestral expression program at the end of the simulation and complexity of the optimal character state. Each data point in (D) and (E) correspond to a unique set of traits subject to selection, which are randomly drawn before the simulation. Each curve is a locally estimated scatterplot smoothing (LOESS) curve, with the shadowed area being the 95% confidence interval. Cell images in (A)-(C) are from Irasutoya (https://www.irasutoya.com/

References and Notes

1. S. J. Gould, R. C. Lewontin, The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Conceptual Issues in Evolutionary Biology* **205**, 79 (1979).

2. J. M. Smith, *et al.*, Developmental constraints and evolution: a perspective from the Mountain Lake conference on development and evolution. *The Quarterly Review of Biology* **60** (3), 265–287 (1985).

3. L. Y. Yampolsky, A. Stoltzfus, Bias in the introduction of variation as an orienting factor in evolution. *Evolution & development* **3** (2), 73–83 (2001).

4. J. L. Hendrikse, T. E. Parsons, B. Hallgrímsson, Evolvability as the proper focus of evolutionary developmental biology. *Evolution & development* **9** (4), 393–401 (2007).

5. K. D. Kavanagh, *et al.*, Developmental bias in the evolution of phalanges. *Proceedings of the National Academy of Sciences* **110** (45), 18190–18195 (2013).

6. J. A. Fritz, *et al.*, Shared developmental programme strongly constrains beak shape diversity in songbirds. *Nature Communications* **5** (1), 3700 (2014).

7. A. Stoltzfus, *Mutation, randomness, and evolution* (Oxford University Press) (2021).

8. P. T. Rohner, D. Berger, Developmental bias predicts 60 million years of wing shape evolution. *Proceedings of the National Academy of Sciences* **120** (19), e2211210120 (2023).

9. P. T. Rohner, D. Berger, Macroevolution along developmental lines of least resistance in fly wings. *Nature Ecology & Evolution* pp. 1–13 (2025).

10. S. B. Carroll, Homeotic genes and the evolution of arthropods and chordates. *Nature* **376** (6540), 479–485 (1995).

11. G. P. Wagner, P. F. Stadler, Quasi-independence, homology and the unity of type: A topological theory of characters. *Journal of Theoretical Biology* **220** (4), 505–527 (2003).

12. G. P. Wagner, The developmental genetics of homology. *Nature Reviews Genetics* **8** (6), 473–479 (2007).

13. G. P. Wagner, *Homology, genes, and evolutionary innovation* (princeton university press) (2014).

14. J. M. Musser, G. P. Wagner, Character trees from transcriptome data: origin and individuation of morphological characters and the so-called "species signal". *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **324** (7), 588–604 (2015).

15. J. Parker, M. Pennell, The cellular substrate of evolutionary novelty. *Current Biology* **35** (12), R626–R637 (2025).

16. R. W. Warren, L. Nagy, J. Selegue, J. Gates, S. Carroll, Evolution of homeotic gene regulation and function in flies and butterflies. *Nature* **372** (6505), 458–461 (1994).

17. S. D. Weatherbee, G. Halder, J. Kim, A. Hudson, S. Carroll, Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the Drosophila haltere. *Genes & development* **12** (10), 1474–1482 (1998).

18. S. D. Weatherbee, *et al.*, Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. *Current Biology* **9** (3), 109–115 (1999).

19. J. Deutsch, Hox and wings. *Bioessays* **27** (7), 673–675 (2005).

20. Y. Tomoyasu, S. R. Wheeler, R. E. Denell, Ultrabithorax is required for membranous wing identity in the beetle Tribolium castaneum. *Nature* **433** (7026), 643–647 (2005).

21. B. M. Hersh, *et al.*, The UBX-regulated network in the haltere imaginal disc of D. melanogaster. *Developmental biology* **302** (2), 717–727 (2007).

22. Y. Tomoyasu, Y. Arakane, K. J. Kramer, R. E. Denell, Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. *Current Biology* **19** (24), 2057–2065 (2009).

23. X. Tong, S. Hrycaj, O. Podlaha, A. Popadic, A. Monteiro, Over-expression of Ultrabithorax alters embryonic body plan and wing patterns in the butterfly Bicyclus anynana. *Developmental biology* **394** (2), 357–366 (2014).

24. Y. Matsuoka, A. Monteiro, Ultrabithorax modifies a regulatory network of genes essential for butterfly eyespot development in a wing sector-specific manner. *Development* **149** (23), dev200781 (2022).

25. H. Ohuchi, *et al.*, Correlation of wing-leg identity in ectopic FGF-induced chimeric limbs with the differential expression of chick Tbx5 and Tbx4. *Development* **125** (1), 51–60 (1998).

26. P. Agarwal, *et al.*, Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development* **130** (3), 623–633 (2003).

27. L. Naiche, V. E. Papaioannou, Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois. *Development* **130** (12), 2681–2693 (2003).

28. C. Minguillon, J. Del Buono, M. P. Logan, Tbx5 and Tbx4 are not sufficient to determine limb-specific morphologies but have common roles in initiating limb outgrowth. *Developmental cell* **8** (1), 75–84 (2005).

29. J.-F. Ouimette, M. L. Jolin, A. L'honoré, A. Gifuni, J. Drouin, Divergent transcriptional activities determine limb identity. *Nature communications* **1** (1), 35 (2010).

30. V. Duboc, M. P. Logan, Regulation of limb bud initiation and limb-type morphology. *Developmental Dynamics* **240** (5), 1017–1027 (2011).

31. F. Petit, K. E. Sears, N. Ahituv, Limb development: a paradigm of gene regulation. *Nature Reviews Genetics* **18** (4), 245–258 (2017).

32. R. Shirasaki, S. L. Pfaff, Transcriptional codes and the control of neuronal identity. *Annual review of neuroscience* **25** (1), 251–281 (2002).

33. L. A. Boyer, *et al.*, Core transcriptional regulatory circuitry in human embryonic stem cells. *cell* **122** (6), 947–956 (2005).

34. S. Lee, *et al.*, A regulatory network to segregate the identity of neuronal subtypes. *Developmental cell* **14** (6), 877–889 (2008).

35. D. Arendt, The evolution of cell types in animals: emerging principles from molecular studies. *Nature Reviews Genetics* **9** (11), 868–882 (2008).

36. T. Graf, T. Enver, Forcing cells to change lineages. *Nature* **462** (7273), 587–594 (2009).

37. T. Hashimshony, M. Feder, M. Levin, B. K. Hall, I. Yanai, Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature* **519** (7542), 219–222 (2015).

38. K. Kin, M. C. Nnamani, V. J. Lynch, E. Michaelides, G. P. Wagner, Cell-type phylogenetics and the origin of endometrial stromal cells. *Cell reports* **10** (8), 1398–1409 (2015).

39. Y. Dai, *et al.*, Evolutionary origin of the chordate nervous system revealed by amphioxus developmental trajectories. *Nature Ecology & Evolution* pp. 1–18 (2024).

40. D. M. Raup, Geometric analysis of shell coiling: general problems. *Journal of paleontology* pp. 1178–1190 (1966).

41. D. M. Raup, Theoretical morphology of echinoid growth. *Journal of Paleontology* **42** (S2), 50–63 (1968).

42. I. Salazar-Ciudad, J. Jernvall, A computational model of teeth and the developmental origins of morphological variation. *Nature* **464** (7288), 583–586 (2010).

43. F. A. Machado, *et al.*, Rules of teeth development align microevolution with macroevolution in extant and extinct primates. *Nature Ecology & Evolution* pp. 1–11 (2023).

44. R. O. Prum, Development and evolutionary origin of feathers. *Journal of Experimental Zoology* **285** (4), 291–306 (1999).

45. M. A. Norell, X. Xu, Feathered dinosaurs. *Annu. Rev. Earth Planet. Sci.* **33** (1), 277–299 (2005).

46. L. Sallan, Fish ‘tails’ result from outgrowth and reduction of two separate ancestral tails. *Current Biology* **26** (23), R1224–R1225 (2016).

47. N. Cumplido, *et al.*, Hox genes control homocercal caudal fin development and evolution. *Science Advances* **10** (3), eadj5991 (2024).

48. A. Hintermann, *et al.*, Co-option of an ancestral cloacal regulatory landscape during digit evolution. *Nature* (2025).

49. J. DiFrisco, A. C. Love, G. P. Wagner, Character identity mechanisms: a conceptual model for comparative-mechanistic biology. *Biology & Philosophy* **35** (4), 44 (2020).

50. J. DiFrisco, G. P. Wagner, A. C. Love, Reframing research on evolutionary novelty and co-option: character identity mechanisms versus deep homology, in *Seminars in Cell & Developmental Biology* (Elsevier), vol. 145 (2023), pp. 3–12.

51. D. Arendt, *et al.*, The origin and evolution of cell types. *Nature Reviews Genetics* **17** (12), 744–757 (2016).

52. B. C. Haller, P. W. Messer, SLiM 4: multispecies eco-evolutionary modeling. *The American Naturalist* **201** (5), E127–E139 (2023).

53. H. A. Orr, Adaptation and the cost of complexity. *Evolution* **54** (1), 13–20 (2000).

54. T. D. Hether, P. A. Hohenlohe, Genetic regulatory network motifs constrain adaptation through curvature in the landscape of mutational (co) variance. *Evolution* **68** (4), 950–964 (2014).

55. D. Jiang, A. L. Cope, J. Zhang, M. Pennell, On the decoupling of evolutionary changes in mRNA and protein levels. *Molecular Biology and Evolution* **40** (8), msad169 (2023).

56. S. Pouzet, A. Le Rouzic, Gene network topology drives the mutational landscape of gene expression. *Evolution* p. qpaf068 (2025).

57. M. Mallo, Reassessing the role of Hox genes during vertebrate development and evolution. *Trends in Genetics* **34** (3), 209–217 (2018).

58. P. W. Ingham, M. Placzek, Orchestrating ontogenesis: variations on a theme by sonic hedgehog. *Nature Reviews Genetics* **7** (11), 841–850 (2006).

59. L. Tang, *et al.*, Predicting unrecognized enhancer-mediated genome topology by an ensemble machine learning model. *Genome research* **30** (12), 1835–1845 (2020).

60. J. Liu, M. Robinson-Rechavi, Robust inference of positive selection on regulatory sequences in the human brain. *Science advances* **6** (48), eabc9863 (2020).

61. E. D. Vaishnav, *et al.*, The evolution, evolvability and engineering of gene regulatory DNA. *Nature* **603** (7901), 455–463 (2022).

62. G. Schweizer, A. Wagner, Both binding strength and evolutionary accessibility affect the population frequency of transcription factor binding sequences in Arabidopsis thaliana. *Genome Biology and Evolution* **13** (12), evab273 (2021).

63. X. Zhang, B. Fang, Y.-F. Huang, Transcription factor binding sites are frequently under accelerated evolution in primates. *Nature communications* **14** (1), 783 (2023).

64. H. Xu, C. Li, C. Xu, J. Zhang, Chance promoter activities illuminate the origins of eukaryotic intergenic transcriptions. *Nature Communications* **14** (1), 1826 (2023).

65. G. D. Smith, W. H. Ching, P. Cornejo-Páramo, E. S. Wong, Decoding enhancer complexity with machine learning and high-throughput discovery. *Genome biology* **24** (1), 116 (2023).

66. D. Jiang, J. Zhang, Ascertainment bias in the genomic test of positive selection on regulatory sequences. *Molecular Biology and Evolution* **41** (1), msad284 (2024).

67. G. P. Wagner, L. Altenberg, Perspective: complex adaptations and the evolution of evolvability. *Evolution* **50** (3), 967–976 (1996).

68. G. P. Wagner, M. Pavlicev, J. M. Cheverud, The road to modularity. *Nature Reviews Genetics* **8** (12), 921–931 (2007).

69. J. Clune, J.-B. Mouret, H. Lipson, The evolutionary origins of modularity. *Proceedings of the Royal Society b: Biological sciences* **280** (1755), 20122863 (2013).

70. A. Goswami, J. B. Smaers, C. Soligo, P. D. Polly, The macroevolutionary consequences of phenotypic integration: from development to deep time. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369** (1649), 20130254 (2014).

71. S. Ciliberti, O. C. Martin, A. Wagner, Innovation and robustness in complex regulatory gene networks. *Proceedings of the National Academy of Sciences* **104** (34), 13591–13596 (2007).

72. C. Rueffler, J. Hermisson, G. P. Wagner, Evolution of functional specialization and division of labor. *Proceedings of the National Academy of Sciences* **109** (6), E326–E335 (2012).

73. A. Holstad, *et al.*, Evolvability predicts macroevolution under fluctuating selection. *Science* **384** (6696), 688–693 (2024).

74. S. J. Arnold, *Evolutionary Quantitative Genetics* (Oxford University Press, Oxford, UK) (2023).

75. M. Pennell, D. Jiang, The macroevolutionary adaptive landscape: more than a metaphor? (2024).

76. A. E. Maher, *et al.*, Body size, shape and ecology in tetrapods. *Nature Communications* **13** (1), 4340 (2022).

77. G. Scholtz, G. D. Edgecombe, The evolution of arthropod heads: reconciling morphological, developmental and palaeontological evidence. *Development genes and evolution* **216**, 395–415 (2006).

78. E. L. Jockusch, Developmental and evolutionary perspectives on the origin and diversification of arthropod appendages. *Integrative and comparative biology* **57** (3), 533–545 (2017).

79. J. Ortega-Hernández, R. Janssen, G. E. Budd, Origin and evolution of the panarthropod head–a palaeobiological and developmental perspective. *Arthropod structure & development* **46** (3), 354–379 (2017).

80. N. M. Young, G. P. Wagner, B. Hallgrímsson, Development and the evolvability of human limbs. *Proceedings of the National Academy of Sciences* **107** (8), 3400–3405 (2010).

81. C. J. Cretekos, *et al.*, Regulatory divergence modifies limb length between mammals. *Genes & development* **22** (2), 141–151 (2008).

82. A. E. Greer, Limb reduction in squamates: identification of the lineages and discussion of the trends. *Journal of Herpetology* pp. 166–173 (1991).

83. L. Bejder, B. K. Hall, Limbs in whales and limblessness in other vertebrates: mechanisms of evolutionary and developmental transformation and loss. *Evolution & development* **4** (6), 445–458 (2002).

84. E. Z. Kvon, *et al.*, Progressive loss of function in a limb enhancer during snake evolution. *Cell* **167** (3), 633–642 (2016).

85. M. P. Harris, J. F. Fallon, R. O. Prum, Shh-Bmp2 signaling module and the evolutionary origin and diversification of feathers. *Journal of Experimental Zoology* **294** (2), 160–176 (2002).

86. F. Prin, D. Dhouailly, How and when the regional competence of chick epidermis is established: feathers vs. scutate and reticulate scales, a problem en route to a solution. *The International journal of developmental biology* **48** (2-3), 137–148 (2004).

87. R. L. Cooper, M. C. Milinkovitch, Transient agonism of the sonic hedgehog pathway triggers a permanent transition of skin appendage fate in the chicken embryo. *Science Advances* **9** (20), eadg9619 (2023).

88. M. J. Benton, *Vertebrate palaeontology* (John Wiley & Sons) (2014).

89. W. A. Cresko, *et al.*, Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences* **101** (16), 6050–6055 (2004).

90. P. C. Wainwright, M. E. Alfaro, D. I. Bolnick, C. D. Hulsey, Many-to-one mapping of form to function: a general principle in organismal design? *Integrative and comparative biology* **45** (2), 256–262 (2005).

91. N. Shubin, C. Tabin, S. Carroll, Deep homology and the origins of evolutionary novelty. *Nature* **457** (7231), 818–823 (2009).

92. G. Panganiban, *et al.*, The origin and evolution of animal appendages. *Proceedings of the National Academy of Sciences* **94** (10), 5162–5166 (1997).

93. T. Nakamura, A. R. Gehrke, J. Lemberg, J. Szymaszek, N. H. Shubin, Digits and fin rays share common developmental histories. *Nature* **537** (7619), 225–228 (2016).

94. S. Kang, J. M. Graham Jr, A. H. Olney, L. G. Biesecker, GLI3 frameshift mutations cause autosomal dominant Pallister-Hall syndrome. *Nature genetics* **15** (3), 266–268 (1997).

95. M. Turelli, Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* **111** (1), 165–195 (1985).

96. X.-S. Zhang, W. G. Hill, Genetic variability under mutation selection balance. *Trends in Ecology & Evolution* **20** (9), 468–470 (2005).

97. K. McGuigan, L. Rowe, M. W. Blows, Pleiotropy, apparent stabilizing selection and uncovering fitness optima. *Trends in Ecology & Evolution* **26** (1), 22–29 (2011).

98. D. Jiang, J. Zhang, Fly wing evolution explained by a neutral model with mutational pleiotropy. *Evolution* **74** (9), 2158–2167 (2020).

99. F. Galis, Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes, and cancer. *Journal of Experimental Zoology* **285** (1), 19–26 (1999).

100. F. Galis, *et al.*, Extreme selection in humans against homeotic transformations of cervical vertebrae. *Evolution* **60** (12), 2643–2654 (2006).

101. I. Varela-Lasheras, *et al.*, Breaking evolutionary and pleiotropic constraints in mammals: on sloths, manatees and homeotic mutations. *EvoDevo* **2**, 1–27 (2011).

102. C. Böhmer, O. W. Rauhut, G. Wörheide, Correlation between Hox code and vertebral morphology in archosaurs. *Proceedings of the Royal Society B: Biological Sciences* **282** (1810), 20150077 (2015).

103. R. T. Cerbus, I. Hiratani, K. Kawaguchi, Homeotic and nonhomeotic patterns in the tetrapod vertebral formula. *Proceedings of the National Academy of Sciences* **121** (47), e2411421121 (2024).

104. S. B. Carroll, Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134** (1), 25–36 (2008).

105. S. M. Glasauer, S. C. Neuhauss, Whole-genome duplication in teleost fishes and its evolutionary consequences. *Molecular genetics and genomics* **289**, 1045–1060 (2014).

106. L. C. Sallan, Major issues in the origins of ray-finned fish (A ctinopterygii) biodiversity. *Biological Reviews* **89** (4), 950–971 (2014).

107. B. Peskin, *et al.*, Notochordal signals establish phylogenetic identity of the teleost spine. *Current Biology* **30** (14), 2805–2814 (2020).

108. L. Sallan, Evolution: Spinal innovation enabled by genome duplication. *Current Biology* **30** (17), R1006–R1008 (2020).

109. K. E. Criswell, J. A. Gillis, Resegmentation is an ancestral feature of the gnathostome vertebral skeleton. *Elife* **9**, e51696 (2020).

110. L. Wolpert, Positional information and the spatial pattern of cellular differentiation. *Journal of theoretical biology* **25** (1), 1–47 (1969).

111. P. W. Holland, H. A. F. Booth, E. A. Bruford, Classification and nomenclature of all human homeobox genes. *BMC biology* **5** (1), 47 (2007).

112. E. A. Boyle, Y. I. Li, J. K. Pritchard, An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169** (7), 1177–1186 (2017).

113. X. Liu, Y. I. Li, J. K. Pritchard, Trans effects on gene expression can drive omnigenic inheritance. *Cell* **177** (4), 1022–1034 (2019).

114. D. M. McCandlish, A. Stoltzfus, Modeling evolution using the probability of fixation: history and implications. *The Quarterly review of biology* **89** (3), 225–252 (2014).

115. M. Kimura, On the probability of fixation of mutant genes in a population. *Genetics* **47** (6), 713 (1962).

Acknowledgments

We thank Gunter P. Wagner, Joanna Wolfe, and Mark Kim for helpful comments.

Funding: DJ and LS are supported by Okinawa Institute of Science and Technology Graduate University. MP was supported by National Institute of Health grant R35GM151348.

Author contributions: DJ, MP, and LS conceived the project. DJ designed and performed the study. DJ and LS wrote the paper with input from MP.

Competing interests: There are no competing interests to declare.

Data and materials availability: Code and data files are available at `https://github.com/RexJiangEvoBio/novelty_evo_model`.

Supplementary materials

Materials and Methods

Figs. S1 to S4

Table S1

References *(7-115)*