CellPress

Opinion

Bayesian Models of Development

Judy A. Stamps[1,*] and Willem E. Frankenhuis[2]

Until recently, biology lacked a framework for studying how information from genes, parental effects, and different personal experiences is combined across the lifetime to affect phenotypic development. Over the past few years, researchers have begun to build such a framework, using models that incorporate Bayesian updating to study the evolution of developmental plasticity and developmental trajectories. Here, we describe the merits of a Bayesian approach to development, review the main findings and implications of the current set of models, and describe predictions that can be tested using protocols already used by empiricists. We suggest that a Bayesian perspective affords a simple and tractable way to conceptualize, explain, and predict how information combines across the lifetime to affect development.

Trends

Bayesian models of development offer a simple and tractable way to model how information from ancestors (e.g., via genes or parental effects) combines with information from a series of personal experiences over the lifetime to affect the development of phenotypic traits.

Bayesian models show how individuals' naive prior distributions and subsequent cue exposures limit developmental plasticity and generate individual differences in plasticity.

Current Bayesian models make novel predictions about developmental plasticity and developmental trajectories, some of which are already supported by empiricists.

Even in the absence of any costs of plasticity, Bayesian models predict that limited developmental plasticity and individual differences in plasticity will be widespread if individuals make optimal developmental decisions based on the information that is available to them.

Why A Bayesian Framework for Development?

A basic premise in biology is that the phenotype of an organism is, at least to some extent, based on its estimates of variables in the external environment [1,2]. Theory suggests that information about the external environment can come from the genes of an organism [3,4], parental effects [5,6], and the many types of personal experience that can occur over the course of a lifetime. The question, then, is how information from all of these sources combines across ontogeny to affect the development of phenotypic traits.

Over the past few years, investigators have begun to address this question, using models that incorporate Bayesian updating to study the evolution of **developmental plasticity** and **developmental trajectories** (see Glossary) [7–12]. These models are based on the assumptions that Bayes' theorem provides the most logically consistent way to combine probabilistic information from different sources at different times [13–15], and that one can model an individual's current assessment of conditions in the external environment ('the **state of the world**') using a probability distribution (Box 1). The models assume that even before individuals have been personally exposed to any **cues** from the environment, they already have 'naive' **prior** distributions, based on information from their distant ancestors (e.g., via genes) and from their immediate ancestors (e.g., via parental effects or inherited epigenetic factors). These naive prior distributions are then updated as individuals are exposed to a series of potentially informative cues over the course of their lives, yielding a series of **posterior** distributions. The models readily accommodate situations in which individuals are repeatedly exposed to the same cues or are exposed to different cues across ontogeny. Finally, the models assume that the phenotypic traits expressed by individuals are affected by their assessments of the state of the world, as reflected by their posterior distributions. Thus, Bayesian models offer a way to make predictions about the developmental trajectories of different individuals and the developmental plasticity of individuals with different genotypes as a function of their naive priors and the series of cues to which they were exposed across their lives.

Although Bayesian models provide a 'benchmark' for information updating against which observations can be compared, this does not imply that organisms necessarily compute full

[1]Section of Evolution and Ecology, Division of Biological Sciences, University of California at Davis, Davis, CA 95616, USA
[2]Behavioural Science Institute, Radboud University, Nijmegen, Montessorilaan 3, PO Box 9104, 6500 HE Nijmegen, The Netherlands

*Correspondence: jastamps@ucdavis.edu (J.A. Stamps).

 http://dx.doi.org/10.1016/j.tree.2016.01.012

Box 1. Bayesian Basics for Models of Development

A prior distribution ('prior') specifies an individual's assessment of the probability of all possible states of the world before it is exposed to a given cue. Bayesian models of development assume that individuals begin life with a naive prior, based on information from their genes and parental effects. For instance, if there are only two possible habitats, an individual might initially assess, based on information from its ancestors, that it is more likely to be in habitat A (p = 0.7) than in habitat B (1–p = 0.3).

A cue is a stimulus, experience, or event that can provide information about the state of the world. A likelihood function ('likelihood') specifies the conditional probability that a given cue will occur, given each of the possible states of the world. The likelihood determines the reliability of a cue, where reliability indicates the extent to which a given cue is differentially associated with different states of the world. For instance, a cue, C, would provide a moderately reliable indication that the habitat was A if p(C|A) = 0.7 and p(C|B) = 0.2. Bayesian models of development typically assume that organisms 'know' the likelihoods of naturally occurring cues rather than learning them, because relations between those cues and states of the world have been a recurrent feature of their evolutionary environments.

A prior is updated based on exposure to a given likelihood, yielding a posterior distribution (a 'posterior'), where the posterior provides a new assessment of the state of the world, conditional on exposure to the cue. Formally, this is accomplished using Bayes' theorem, followed by normalization to ensure that the probabilities of all possible states add up to 1. 'E' refers to an individual's point estimate of the state at a given time, and 'ΔE' refers to the difference in an individual's estimate of the state before and after exposure to a cue. If there are only two states, E is indicated by p, and ΔE by the difference between the p values of the prior and the posterior. If states are continuously distributed, the means of the prior and the posterior can provide useful estimates of E, in which case ΔE is indicated by the difference between those two means.

The posterior for one cue becomes the prior for the next cue, which allows Bayesian models to predict how E would change across ontogeny, in response to exposure to cues from different sources and at different times.

Bayesian solutions. Instead, organisms might use heuristics or rules of thumb that approximate 'optimal' Bayesian solutions under natural conditions, but which are computationally simpler or less expensive (e.g., [14–17]).

Here, we characterize the diverse array of recent models of Bayesian development, and describe what these models tell us about the ways that **cue reliability** and prior distributions affect an individual's estimates of the state of the world over ontogeny. We outline specific testable predictions generated by these models, and highlight their general prediction that limited developmental plasticity and individual differences in plasticity will be widespread, even in the absence of any costs of plasticity. Finally, we describe outstanding problems in development that might profit from a Bayesian perspective.

Variation among Bayesian Models of Development

Although Bayesian models of development are based on shared assumptions (Box 1), they also differ in important ways. Two-state models assume that all possible states of the world fall into two discrete categories (e.g., high food versus low food) [7,8,11,12], whereas continuous models assume that many possible states vary continuously between minimum and maximum possible values (e.g., the level of danger) [9,10]. Two-state models are analytically simpler and more tractable, and provide a useful first approximation of the patterns expected under Bayesian updating. Continuous models allow for greater biological realism, and provide a way to examine how the means and the variances of priors separately contribute to Bayesian updating (Figure 1). All of the models assume that offspring can develop in different environments than their parents; if this were not the case, one would not expect plasticity or information updating to evolve [18,19]. However, most of the current crop of models assume that the environment is stable within an individual's lifetime [7,9,11,12]. The sole exception [8] assumes that the state of the world can change within generations and that, in response, organisms have evolved mechanisms that devalue information obtained earlier in ontogeny. This latter model is comparable to many learning models, which routinely assume that environmental conditions change within generations, and that animals have evolved mechanisms that allow them to detect and respond to such changes (e.g., [20,21]). Most of the current models assume that every individual in a

Glossary

ΔE: the difference between E for the prior and E for the posterior as a result of exposure to a given cue.
Confidence: an individual's level of confidence (degree of belief) in its current estimate of the state, E. In two-state models, confidence is determined by the value of p: lowest for p = 0.5, and highest for p = 0 or p = 1. In continuous models, confidence can be represented by the variance of the prior or the posterior.
Cue: a stimulus, experience, or event that can provide information about the 'state of the world'.
Cue reliability: the extent to which a given cue is differentially associated with all of the possible states of the world. A perfectly reliable cue would only occur for one state of the world; a very unreliable cue would be nearly equally likely to occur for every possible state of the world.
Developmental plasticity: the effects of stimuli, cues or experiences in the past on the current phenotype. Learning is often viewed as a special case of developmental plasticity [59,60].
Developmental trajectory: a description of how the values of a given phenotypic trait change within a given individual as a function of age or time. Developmental trajectories are investigated using within-individual experimental designs [61].
E: a point estimate of an individual's prior or posterior distribution. In two-state Bayesian models of development, E is the probability of one of the two states, p. In continuous Bayesian models of development, E can be represented by the mean of the prior or the posterior distribution.
Intragenotypic variability (IGV): interindividual variation in the phenotypes expressed at a given age by individuals with the same genotype, reared under the same conditions prior to measurement.
Likelihood: a distribution specifying the conditional probability that a given cue will occur, given each of the possible states of the world.
Posterior: a distribution specifying an individual's assessment of the probability of all possible states of the world after exposure to a given cue.
Potential developmental plasticity: the ability of an individual or genotype to generate a wide range of