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.

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

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.

[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).

CrossMark

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
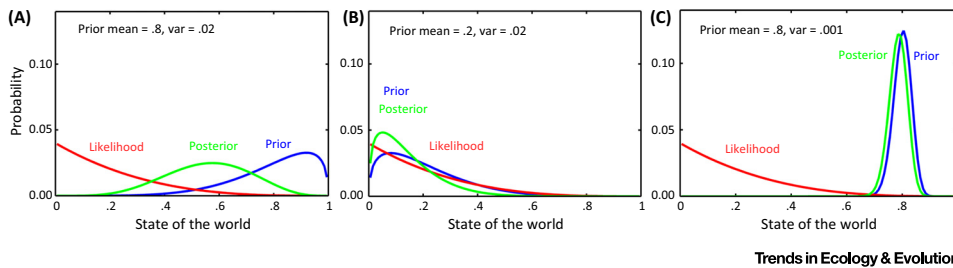
Figure 1. The Effects of The Means and Variances of Prior Distributions on Bayesian Updating in A Continuous Model, in which The State of The World (e.g., The Level of Danger) Can Take On Any Value between 0 and 1. Each distribution (the prior distribution, the likelihood function, and the posterior distribution) indicates the probability of each of the possible values of the state. Individuals with three different prior distributions (blue lines) are exposed to the same cue, with a likelihood function (red lines) that indicates that low values of the state are more likely than high values of the state. (A) and (B) illustrate the 'discrepancy rule'. In (A) and (B), the priors have different means but the same variance. In (A), the prior distribution and the likelihood function have very different means, resulting in a posterior (green line) displaced to the left of the prior. In (B), the prior and the likelihood have similar means, resulting in a posterior similar to the prior. (A) and (C) illustrate the effects of the confidence of the prior on updating. Here, the priors have the same mean, but the variance of the prior is higher in (A) than in (C). In (A), the confidence of the prior is relatively low, so exposure to the cue results in a posterior displaced to the left of the prior. In (C), the confidence of the prior is relatively high, so exposure to the same cue results in a posterior similar to the prior. As a result of these differences, ΔE (i.e., the difference between the mean of the prior and the mean of the posterior) is higher in (A) than in (B) or (C).

population begins life with the same naive prior, but some models do not [9,10]. The assumption that every individual in a population has the same naive prior facilitates study of the evolution of the optimal developmental program for organisms with that prior, whereas the assumption that individuals begin life with different naive priors allows for analyses of the developmental trajectories expected when organisms with different naive priors are exposed to the same cues. Finally, some models assume that different individuals can be exposed to the same cues for the same period of time [9,10], while others assume that exposure to cues varies stochastically across individuals at the same place and time (e.g., [7,8]). The former assumption is more applicable to controlled experiments in the laboratory, while the latter is more applicable to development in the field, where cue exposures can vary widely, even among individuals living near one another. For instance, in fluvial habitats, the chaotic effects of turbulence result in unpredictable spatial and temporal variation in the frequency, intensity, and duration of exposures to the same chemical cues [22].

Factors Affecting Bayesian Updating

Here, we use '**E**' to refer to an individual's current point estimate of the state of the world, and '**ΔE**' to indicate the difference between the prior and the posterior as a result of exposure to a given cue, stimulus, or experience (Box 1). ΔE provides the foundation for any Bayesian model of development [2,15], which share the assumption that changes in E encourage changes in phenotype. Of course, many factors besides ΔE can also affect trait development, including the benefits and costs of being phenotypically matched or mismatched to the state of the world, the extent to which trait development is reversible, lags between ΔE and change in phenotype, costs of sampling, or costs of switching between different phenotypes. However, none of these factors is unique to Bayesian models, and assumptions about them vary across both Bayesian and nonBayesian models of development. Hence, here we focus on the factors that affect ΔE.

Cue Reliability

Although many nonBayesian models have considered the effects of cue reliability on the evolution of developmental plasticity, with few exceptions (e.g., [23–25]) most have compared organisms with access to perfect cues to organisms with no access to cues. By contrast, recent

phenotypes in response to cues with a wide range of likelihoods.
Prior: a distribution specifying an individual's assessment of the probability of all possible states of the world before exposure to a given cue.
Realized developmental plasticity: the extent to which the phenotype of an individual or a genotype changes as a result of past exposure to a particular cue or set of cues. Empiricists typically study realized developmental plasticity using replicate-individual experimental designs (Box 2, main text).
Replicate-individual design: an experimental design for measuring developmental plasticity in which individuals with the same genotype are exposed to two or more different cues. For each genotype, differences between the mean trait values expressed by the different treatment groups at the end of an exposure period provide an index of the developmental plasticity of that genotype in response to those cues.
State of the world: the value of a variable in the external environment. In two-state models, there are only two possible states (e.g., two different habitats); in continuous models, there is a range of possible states (e.g., different levels of danger).
Within-individual design: an experimental design for measuring individual differences in developmental trajectories. Different individuals are exposed to the same cues over the same period of time, and changes in their trait values across ontogeny are used to describe their developmental patterns in response to those cues.

Bayesian models show that many important effects of cues on development occur when individuals are repeatedly exposed to cues that are moderately reliable, as opposed to either very reliable or very unreliable. If a cue is very reliable (e.g., close to 100% accurate), then every individual is expected to quickly adjust its estimate of the environmental state to the value indicated by the cue. Conversely, if a cue is very unreliable (close to 0% accurate) it will have little or no discernable effect on the estimates of any individual. By contrast, repeated exposure to moderately reliable cues encourages gradual (limited) changes in estimates of the environmental state across ontogeny, as well as variation across individuals in their estimates of the same environmental state at the same age and time [7–12].

A cue can be moderately reliable because it is more likely (but not much more likely) to occur for some environmental states than others (e.g., [26,27]), or as a result of errors in perceptual systems (e.g., [28,29]). For example, a particular color might be a moderately reliable indicator of fruit quality because that color can occur on fruits with a range of nutrient levels and/or because individuals have difficulty detecting that color against background vegetation [30]. To the extent that conditions in the external environment (e.g., noise or low light conditions) increase perceptual errors, the reliability of the same cue might be lower in field studies than under controlled laboratory conditions.

Recent Bayesian models show that repeated exposure to moderately reliable cues results in more gradual changes in phenotypic traits (i.e., developmental trajectories with shallower slopes) than is the case for repeated exposure to cues with high reliability [7–12], thus confirming earlier suggestions that cue reliability limits developmental plasticity [31]. However, these models suggest another reason why moderate cue reliabilities are important: differences among individuals or among individuals with different genotypes in developmental trajectories are more likely to occur when cues are moderately reliable than when cue reliability is either very low or very high. On the one hand, if cues are moderately reliable, stochastic variation in cue exposure leads to variation among individuals in their estimates of the environmental state at any given time. As a result, even if every individual began with the same naive prior, one would expect to observe variation among individuals in their developmental trajectories and trait values at the end of the developmental period [7,8,11]. On the other hand, when cues are moderately reliable, variation among individuals in their naive priors encourages predictable individual differences in developmental trajectories; this occurs even if every individual is exposed to the same cues [9,10].

Priors

In contrast to many Bayesian learning models (e.g., [32–34]), Bayesian models of development assume that priors can vary across individuals, genotypes, or populations. Priors have two important effects on ΔE. First, the discrepancy between the prior distribution and the **likelihood** function for a given cue is positively related to ΔE (Figures 1A,B and 2). If the prior and the likelihood for a cue are highly congruent (e.g., if both indicate that the level of danger is 0.2 on a scale of 0 to 1), then there would be little reason to expect individuals to respond to repeated exposure to that cue by changing their phenotype. By contrast, if the prior and the likelihood are very discrepant (e.g., if the prior indicates that the level of danger is 0.8, but the cue indicates that it is 0.2), then repeated exposure to that cue would change an individual's estimate of the environmental state and, hence, encourage change in its phenotype.

Second, the **confidence** of the prior is positively related to ΔE (Figures 1A,C and 2). The confidence of a naive prior indicates an individual's level of confidence in its estimate of the environmental state before it is exposed to any cue. In two-state models, the confidence of the prior is lowest when $p = 0.5$, and highest when $p = 0$ or 1. In continuous models, the mean of the prior can provide an estimate of the environmental state, E, and the variance of the prior can provide an estimate of an individual's level of confidence in that estimate [9] (Figure 1A,C).

(A) Two-state model

(B) Continuous model

Likelihood
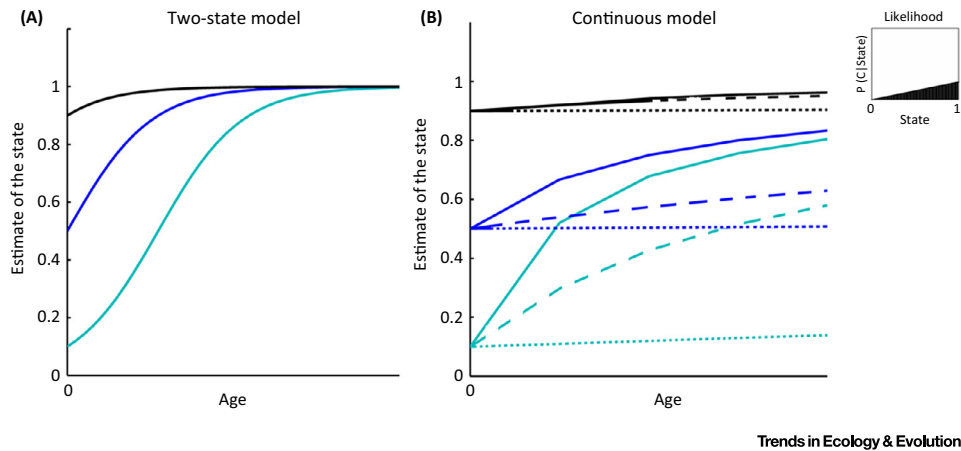
Trends in Ecology & Evolution

Figure 2. How Estimates of The State Change over Time as A Function of Different Priors in Two-State and Continuous Models. (A) In a two-state model, the state of the world (e.g., the type of habitat) is either A or B. Each individual's current estimate of the state, E, is indicated by its current value of p, where p is its current estimate of the probability that the state is A. Three individuals with naive priors of 0.1, 0.5 and 0.9 (indicated by their p values at age 0) are repeatedly exposed to the same cue, C, with a likelihood function that indicates that the state is more likely to be A than B [p (C|A) = 0.55, p(C|B) = 0.45]. E changes the most across ontogeny when p at age 0 = 0.1 (a result of the discrepancy rule). However, the rate of change in E early in life (i.e., the slope of each line immediately after age 0) is highest for individuals with a naive prior of 0.5. This is because a naive prior of 0.5 is less confident than a naive prior of either 0.1 or 0.9. (B) In a continuous model, the state of the world (e.g., the level of danger) can take on any value from 0 to 1. In this model, each individual's current estimate of the state, E, is indicated by the mean of its prior or posterior. Nine individuals with naive priors with different means (0.1, 0.5 or 0.9) and different variances (high, unbroken lines; medium, broken lines; and low, dotted lines) are repeatedly exposed to the same cue, C. Cue C has a right-biased likelihood function (small box), indicating that the state is more likely to be higher than lower. Across ontogeny, the change in E is highest for individuals whose naive priors had a mean of 0.1 (a result of the discrepancy rule). For naive priors with the same mean, the change in E across ontogeny is positively related to the variance of the prior, because priors with high variances are less confident than priors with low variances. Modified from [35] (A) and [9] (B).

The confidence of the prior is important because, all else being equal, ΔE is higher when the prior is less confident than when it is more confident (Figure 1A,C). The effects of prior confidence on ΔE make intuitive sense: if an individual was very sure about its initial estimate of the environmental state, it should be less likely to revise that estimate based on a moderately reliable cue than if it was unsure about its initial estimate.

Specific Predictions of Bayesian Models of Development
Current Bayesian models make several interesting predictions about patterns of **potential** and **realized developmental plasticity** and developmental trajectories (Box 2).

Predictions about Age-Dependent Changes in Developmental Plasticity
Several Bayesian models predict that developmental plasticity will typically be higher earlier in life than later in life [7–9,11,12]. This general prediction is consistent with empirical studies indicating that sensitive periods (in which cues shape phenotypic development to a larger extent than other periods [35]) often occur early during ontogeny ([36–40], but see [8,41] for exceptions). Another interesting, as yet untested, prediction is that sensitive periods will last longer (i.e., end at older ages) in populations or species in which cues have low reliability than in taxa in which comparable cues have moderate to high reliability, if the confidence of the naive prior is moderate to low [11].

Predictions about Individual or Genotypic Differences in Developmental Trajectories
A continuous model predicts that, if neonates with different initial scores for boldness are subsequently reared in 'safe' environments in the laboratory, the intercepts and slopes of their

Box 2. Experimental Designs for Studying Developmental Plasticity and Developmental Trajectories

In discussing developmental plasticity, it is useful to distinguish between potential plasticity and realized plasticity [10]. Potential plasticity refers to the ability of an individual to generate different phenotypes in response to different cues with a wide range of likelihoods; it is a hypothetical attribute of an individual. Realized plasticity indicates the extent to which an individual's phenotype changes or varies in response to a specific cue or set of cues; it is what empiricists measure in their studies.

Empiricists typically use two experimental designs to study developmental trajectories and developmental plasticity [59]. In 'within-individual' designs (Figure IA), different individuals are repeatedly exposed to the same cue or series of cues, and their trait values are measured at different ages over ontogeny. This design is routinely used to describe individual differences in the developmental trajectories of subjects as a function of exposure to particular sets of cues.

In 'replicate-individual' designs, the goal is to generate individuals with comparable internal states before they are exposed to different cues. This is accomplished by using individuals with the same or similar genotypes (e.g., isolines, clones, or full sibs), raised under the same conditions. Individuals with each genotype are then divided into treatment groups, each of which is exposed to a different cue for the same period of time. At the end of one or more exposure periods, for each genotype, the mean trait value of each treatment group is measured, and the difference between mean trait values is used as an index of developmental plasticity of that genotype (Figure IB). Scaled-up versions of replicate-individual designs have been used for many years to describe differences among populations or species in realized developmental plasticity.

Bayesian models underscore the distinction between potential and realized plasticity by demonstrating that the realized developmental plasticity of a given individual or genotype is expected to vary as a function of interactions between its naive prior and the likelihoods of the cues to which it is exposed over ontogeny [10] (see also [7,9,11] and Figure 1, main text). Furthermore, by specifying how different naive priors and different cues might interact to encourage changes in E across a lifetime (e.g., Figure 2, main text), Bayesian models generate predictions about developmental trajectories and developmental plasticity amenable to empirical tests using within-individual or replicate-individual designs (e.g., see the section 'Specific Predictions of Bayesian Models of Development' in the main text).
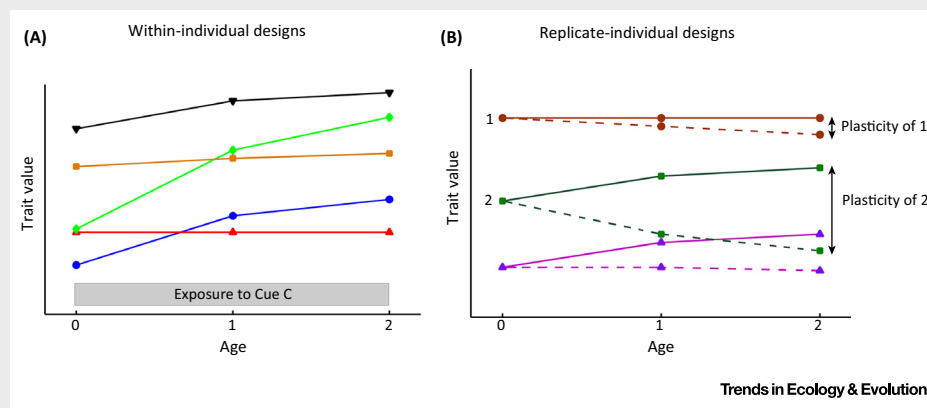

Figure I. Experimental Designs for Studying Individual Differences in Developmental Trajectories and Developmental Plasticity. (A) Different individuals (indicated by different symbols and colors) are repeatedly exposed to the same cue, C, from age 0 to age 2. The developmental trajectory of each individual describes how the trait values of that individual change over ontogeny when it is reared in the presence of a particular cue. (B) Three genotypes (indicated by different symbols and colors) are each exposed to two different cues (indicated by either unbroken or broken lines) from age 0 to age 2. For each genotype, realized plasticity at a given age is indicated by the difference in the mean trait values expressed by the treatment groups after exposure to their respective cues. In this example, genotype 1 is less plastic than genotype 2.

developmental trajectories will be negatively related to one another [9] (Figure 2B). This prediction is supported by data on the development of boldness in individual pigtailed macaques (*Macaca nemestrina*) [42]) and lines of mangrove killifish (*Kryptolebias marmoratus*) [43]. A related model predicts that the developmental trajectories of different individuals will 'fan in' (i.e., tend to converge) across ontogeny if individuals that initially express a range of trait values are reared in the same environment [10]. This prediction is supported by field data on the developmental trajectories of activity and aggressiveness in red squirrels

(*Tamiasciurus hudsonicus*) [44] and the developmental trajectories of stress responsiveness in house sparrows (*Passer domesticus*) [45].

Predictions about Individual Differences in Potential Developmental Plasticity

The models predict that individuals with more-confident naive priors will be less developmentally plastic across a range of ages and in response to cues with a range of likelihoods and reliabilities, than individuals with less-confident priors. These results are consistent with studies of humans, which suggest that individuals carrying specific alleles are less developmentally plastic than others in response to different experiences, assays, and situations, and at different ages ([46,47], but see [48]).

Predictions about Relations between Intragenotypic Variability and Realized Developmental Plasticity

Intragenotypic variability (IGV) is the variability in trait values of individuals with the same genotype, reared in the same environment [49,50]. A recent model suggests that the IGV of naive individuals might be negatively related, across genotypes, to the confidence of their naive priors [9]. In that case, we would expect a positive relation across genotypes between their IGV before exposure to cues and their plasticity in response to those cues. This contrasts with previous predictions of positive relations between developmental plasticity and IGV, which measure IGV after (not before) genotypes have been exposed to cues (e.g., [51,52]).

Limited Plasticity and Individual Differences in Plasticity

Bayesian models help shed light on one of the most contentious questions in modern biology: why organisms do not exhibit 'perfect' plasticity [31,53]. Most theoreticians have explained limited plasticity and individual differences in developmental plasticity by invoking costs of plasticity [31,54], but, despite considerable effort, empiricists have had difficulty documenting such costs [53–55]. Less attention has been paid to the ways that information might limit the evolution and expression of developmental plasticity. Recent Bayesian models build upon earlier findings that cue reliability can limit plasticity (e.g., [18,23,56]), by demonstrating that exposure to cues with moderate reliability results in slower rates of trait development than exposure to perfect cues, and that even cue-users that follow optimal developmental programs might not attain the optimal trait values for a given age in a given environment [7–9,11]. They also expand upon earlier findings by showing that the confidence of naive priors can limit plasticity: limited developmental plasticity is not only expected when cue reliability is low to moderate, but also when the confidence of the naive prior is moderate to high [9–11]. Finally, the models predict that individual differences in developmental trajectories and developmental plasticity will be widespread, because every individual in a population is unlikely to begin life with the same naive prior, and because every individual need not be exposed to an identical series of cues over ontogeny. Thus, even in the absence of any costs of plasticity, Bayesian models that assume that individuals do the best they can, based on the information they have, predict the limited plasticity and individual differences in plasticity observed in many organisms.

Future Directions: Somatic State in Bayesian Models

Most current Bayesian models of development focus exclusively on the effects of information on development. However, experiences early in life can affect development not only because they provide information about the external world, but also because they have direct effects on an individual's somatic state (i.e., its body size, strength, motor skills, or other enduring aspects of the phenotype with broad implications for fitness) [57]. Thus far, only one Bayesian model [12] has predicted the developmental trajectories expected if the same experience (food intake) simultaneously affects an individual's estimate of the state of the world (food level in the local habitat) and its somatic state (growth rate). Another question is how to model situations in which an individual's own somatic state provides it with information about conditions in the external

world. For instance, if the probability of being attacked by predators declines as a function of body size in the natural habitat [58], then an individual's current body size might provide it with a moderately reliable cue to its current risk of predation [9]. Finally, although it is typically assumed that individuals have highly reliable cues to their own somatic state, this assumption need not always be valid. In such cases, Bayesian approaches could analyze situations in which personal experiences with moderately reliable likelihoods provide individuals with cues to their current somatic state [7].

Concluding Remarks

An emerging trend in ecology and evolution is to model developmental plasticity and developmental trajectories using Bayesian updating. Here, we have reviewed the current set of Bayesian models and shown that these make several predictions, some of which are already supported by empirical data. As more Bayesian models of development are generated and tested, the value of the approach will become increasingly clear. Our prior estimate is that Bayesian approaches might become as central to the study of development as they already are to other fields, such as learning theory and statistics.

References

1. McNamara, J.M. and Dall, S.R.X. (2010) Information is a fitness enhancing resource. *Oikos* 119, 231–236
2. Dall, S.R.X. *et al.* (2005) Information and its use by animals in evolutionary ecology. *Trends Ecol. Evol.* 20, 187–193
3. Leimar, O. *et al.* (2006) A new perspective on developmental plasticity and the principles of adaptive morph determination. *Am. Nat.* 167, 367–376
4. Dall, S.R.X. *et al.* (2015) Genes as cues: phenotypic integration of genetic and epigenetic information from a Darwinian perspective. *Trends Ecol. Evol.* 30, 327–333
5. English, S. *et al.* (2015) The information value of non-genetic inheritance in plants and animals. *PLoS ONE* 10, e0116996
6. Leimar, O. and McNamara, J.M. (2015) The evolution of transgenerational integration of information in heterogeneous environments. *Am. Nat.* 185, E55–E69
7. Frankenhuis, W.E. and Panchanathan, K. (2011) Balancing sampling and specialization: an adaptationist model of incremental development. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 278, 3558–3565
8. Fischer, B. *et al.* (2014) The evolution of age-dependent plasticity. *Am. Nat.* 183, 108–125
9. Stamps, J.A. and Krishnan, V.V. (2014) Combining information from ancestors and personal experiences to predict individual differences in developmental trajectories. *Am. Nat.* 184, 647–657
10. Stamps, J. and Krishnan, V.V. (2014) Individual differences in the potential and realized developmental plasticity of personality traits. *Front. Ecol. Evol.* Published online April 10, 2015. http://dx.doi.org/10.3389/fevo.2014.00069
11. Panchanathan, K. and Frankenhuis, W.E. (2016) The evolution of sensitive windows in a model of incremental development. *Proc. R. Soc. Lond. B Biol. Sci.* 283, 20152439
12. English, S. *et al.* (2016) Adaptive use of information during growth can explain long-term effects of early-life experiences. *Am. Nat.* Published online 2016. http://dx.doi.org/10.1086/685644
13. Trimmer, P.C. *et al.* (2011) Decision-making under uncertainty: biases and Bayesians. *Anim. Cogn.* 14, 465–476
14. Lange, A. and Dukas, R. (2009) Bayesian approximations and extensions: optimal decisions for small brains and possibly big ones too. *J. Theor. Biol.* 259, 503–516
15. McNamara, J.M. *et al.* (2006) Bayes' theorem and its applications in animal behaviour. *Oikos* 112, 243–251
16. Sanborn, A.N. *et al.* (2010) Rational approximations to rational models: Alternative algorithms for category learning. *Psychol. Rev.* 117, 1144–1167
17. Sanborn, A.N. (2015) Types of approximation for probabilistic cognition: sampling and variational. *Brain Cognition.* Published online July 27, 2015. http://dx.doi.org/10.1016/j.bandc.2015.06.008
18. Botero, C.A. *et al.* (2015) Evolutionary tipping points in the capacity to adapt to environmental change. *Proc. Natl. Acad. Sci. U.S.A.* 112, 184–189
19. Scheiner, S.M. (2014) The genetics of phenotypic plasticity. XIII. Interactions with developmental instability. *Ecol. Evol.* 4, 1347–1360
20. Lloyd, K. and Leslie, D.S. (2013) Context-dependent decision-making: a simple Bayesian model. *J. R. Soc. Interface* 10, 20130069
21. Mangel, M. (1990) Dynamic information in uncertain and changing worlds. *J. Theor. Biol.* 146, 317–332
22. Edwards, D.D. and Moore, P.A. (2014) Real exposure: field measurement of chemical plumes in headwater streams. *Arch. Environ. Contam. Toxicol.* 67, 413–425
23. Tufto, J. (2000) The evolution of plasticity and nonplastic spatial and temporal adaptations in the presence of imperfect environmental cues. *Am. Nat.* 156, 121–130
24. Wong, T.G. and Ackerly, D.D. (2005) Optimal reproductive allocation in annuals and an informational constraint on plasticity. *New Phytol.* 166, 159–171
25. Donaldson-Matasci, M.C. *et al.* (2013) When unreliable cues are good enough. *Am. Nat.* 182, 313–327
26. Furness, A.I. *et al.* (2015) Adaptation in a variable environment: phenotypic plasticity and bet-hedging during egg diapause and hatching in an annual killifish. *Evolution* 69, 1461–1475
27. Lyons, I.M. *et al.* (2014) The influence of cue reliability and cue representation on spatial reorientation in young children. *J. Cogn. Dev.* 15, 402–413
28. Geisler, W.S. (2008) Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.* 59, 167–192
29. Ord, T.J. and Stamps, J.A. (2008) Alert signals enhance animal communication in "noisy" environments. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18830–18835

Outstanding Questions

How does information interact with somatic state across ontogeny to affect developmental plasticity and developmental trajectories?

Can Bayesian models be used to analyze the development of traits whose fitness depends on the trait values of other organisms (i.e., frequency dependence)? For instance, there is empirical evidence that the development of various types of social behavior (aggressiveness or mating behavior) depends on cues from both the physical and the social environment, but currently we lack models to describe and predict the development of such traits.

How can empiricists estimate the naive prior distributions of individuals and genotypes? Some authors have suggested ways to estimate the means and variances of naive priors for traits that are expressed before organisms have been exposed to cues from the external environment (e.g., see discussion of IGV in the main text). However, it is currently unclear how we might estimate naive priors for traits whose development depends on exposure to cues early in life, but are first expressed later in life.

30. Schaefer, H.M. and Braun, J. (2009) Reliable cues and signals of fruit quality are contingent on the habitat in black elder (Sambucus nigra). *Ecology* 90, 1564–1573

31. DeWitt, T.J. *et al.* (1998) Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* 13, 77–81

32. Holyoak, K.J. and Cheng, P.W. (2011) Causal learning and inference as a rational process: the new synthesis. *Annu. Rev. Psychol.* 62, 135–163

33. Dayan, P. and Daw, N.D. (2008) Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* 8, 429–453

34. Kikuchi, D.W. and Sherratt, T.N. (2015) Costs of learning and the evolution of mimetic signals. *Am. Nat.* 186, 321–332

35. Fawcett, T.W. and Frankenhuis, W.E. (2015) Adaptive explanations for sensitive windows in development. *Front. Zool.* 12 (Suppl. 1), S3

36. Barton, K.E. and Koricheva, J. (2010) The ontogeny of plant defense and herbivory: characterizing general patterns using meta-analysis. *Am. Nat.* 175, 481–493

37. Shiojiri, K. and Karban, R. (2006) Plant age, communication, and resistance to herbivores: young sagebrush plants are better emitters and receivers. *Oecologia* 149, 214–220

38. Harvell, C.D. (1991) Coloniality and inducible polymorphism. *Am. Nat.* 138, 1–14

39. Andersen, S.L. (2015) Exposure to early adversity: points of cross-species translation that can lead to improved understanding of depression. *Dev. Psychopathol.* 27, 477–491

40. Keuroghlian, A.S. and Knudsen, E.I. (2007) Adaptive auditory plasticity in developing and adult animals. *Prog. Neurobiol.* 82, 109–121

41. Fuhrmann, D. *et al.* (2015) Adolescence as a sensitive period of brain development. *Trends Cogn. Sci.* 19, 558–566

42. Sussman, A. and Ha, J. (2011) Developmental and cross-situational stability in infant Pigtailed Macaque temperament. *Dev. Psychol.* 47, 781–791

43. Edenbrow, M. and Croft, D.P. (2011) Behavioural types and life history strategies during ontogeny in the mangrove killifish, Kryptolebias marmoratus. *Anim. Behav.* 82, 731–741

44. Kelley, A.D. *et al.* (2015) Changes in wild red squirrel personality across ontogeny: activity and aggression regress towards the mean. *Behaviour* 152, 1291–1306

45. Lendvai, A.Z. *et al.* (2015) Within-individual plasticity explains age-related decrease in stress response in a short-lived bird. *Biol. Lett.* 11, 20150272

46. Karg, K. *et al.* (2011) The serotonin transporter promoter variant (5-HTTLPR), stress, and depression: meta-analysis revisited evidence of genetic moderation. *Arch. Gen. Psychiatry* 68, 444–454

47. van Ijzendoorn, M.H. and Bakermans-Kranenburg, M.J. (2015) Genetic differential susceptibility on trial: meta-analytic support from randomized controlled experiments. *Dev. Psychopathol.* 27, 151–162

48. Duncan, L.E. and Keller, M.C. (2011) A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am. J. Psychiatry* 168, 1041–1049

49. Stamps, J. *et al.* (2013) Genotypic differences in behavioural entropy: unpredictable genotypes are composed of unpredictable individuals. *Anim. Behav.* 86, 641–649

50. Ayroles, J.F. *et al.* (2015) Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6706–6711

51. Tonsor, S.J. *et al.* (2013) Developmental instability is genetically correlated with phenotypic plasticity, constraining heritability and fitness. *Evolution* 67, 2923–2935

52. Scheiner, S.M. (2014) Bet-hedging as a complex interaction among developmental instability, environmental heterogeneity, dispersal, and life-history strategy. *Ecol. Evol.* 4, 505–515

53. Murren, C.J. *et al.* (2015) Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity* 215, 293–301

54. Auld, J.R. *et al.* (2010) Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 277, 503–511

55. Van Buskirk, J. and Steiner, U.K. (2009) The fitness costs of developmental canalization and plasticity. *J. Evol. Biol.* 22, 852–860

56. Utz, M. *et al.* (2014) Phenotypic plasticity with instantaneous but delayed switches. *J. Theor. Biol.* 340, 60–72

57. Nettle, D. and Bateson, M. (2015) Adaptive developmental plasticity: what is it, how can we recognize it and when can it evolve? *Proc. R. Soc. Lond. B Biol. Sci.* 282, 20151005

58. Sogard, S.M. (1997) Size-selective mortality in the juvenile stage of teleost fishes: a review. *Bull. Mar. Sci.* 60, 1129–1157

59. Stamps, J. (2015) Individual differences in behavioural plasticities. *Biol. Rev.* Published online April 10, 2015. http://dx.doi.org/10.1111/brv.12186

60. Snell-Rood, E.C. (2013) An overview of the evolutionary causes and consequences of behavioural plasticity. *Anim. Behav.* 85, 1004–1011

61. Frankenhuis, W.E. and Panchanathan, K. (2011) Individual differences in developmental plasticity may result from stochastic sampling. *Perspect. Psychol. Sci.* 6, 336–347