Cryptic variation in butterfly eyespot development: the importance of sample size in gene expression studies

Robert D. Reed,* Po-Hao Chen, and H. Frederik Nijhout

Department of Biology, Duke University, Durham, NC 27708, USA

*Author for correspondence (email: reed@uci.edu)

SUMMARY  Previous studies have shown that development can be robust to variation in parameters such as the timing or level of gene expression. This leads to the prediction that natural populations should be able to host developmental variation that has little phenotypic effect. Cryptic variation is of particular interest because it can result in selectable phenotypes when "released" by environmental or genetic factors. Currently, however, we have little idea of how variation is distributed between genes or over time in pattern formation processes. Here we survey expression of Notch (N), Spalt (Sal), and Engrailed (En) during butterfly eyespot determination to better understand how pattern formation may vary within a population. We observed substantial heterochronic variance in the progress of spatial expression patterns for all three proteins, suggesting some degree of developmental buffering in eyespot development. Peak variance for different proteins was found at both early and late stages of development, contrasting with previous models suggesting that the distribution of variance should be more temporally focused during pattern formation. We speculate that our observations are representative of a standing reservoir of cryptic variation that may contribute to phenotypic evolution under certain circumstances. Our results also provide a strong cautionary message that gene expression studies with limited sample sizes can be positively misleading in terms of inferring expression pattern time series, as well as for making cross-species phylogenetic comparisons.

INTRODUCTION

Recent theoretical and experimental insights into the robustness of development to variation in gene expression and/or protein distribution (von Dassow et al. 2000; Houchmandzadeh et al. 2002; Meir et al. 2002; Lucchetta et al. 2005; Horikawa et al. 2006; Veitia and Nijhout 2006) lend support to the idea that some cryptic variation could be expressed developmentally. To date, however, this hypothesis remains poorly tested because few direct measurements of intraspecific developmental variation have been made. To address this issue, and to better understand how a developmental process can vary in a population, we conducted a survey of spatiotemporal gene expression patterns associated with eyespot color pattern determination in a population of Junonia (Precis) coenia butterflies.

Butterfly eyespot determination occurs in final-instar imaginal wing disks when several signaling molecules and transcription factors are expressed in a group of cells that will become the eyespot center, or "focus" (Carroll et al. 1994; Brakefield et al. 1996; Keys et al. 1999; Reed and Serfas 2004). Shortly after pupation, the focal cells send out a diffusible signal that induces surrounding cells to express various transcription factors and take on the identity of color rings in the adult wing pattern (Nijhout 1980; Brunetti et al. 2001). The eyespot-associated proteins so far identified belong to develop-
developmental regulatory networks that are well known from *Drosophila* (Beldade and Brakefield 2002; McMillan et al. 2002; Evans and Marcus 2006). Their specific functions in butterfly eyespot development have not yet been elucidated experimentally, but modeling studies (Evans and Marcus 2006), mutants (Brunetti et al. 2001; Reed and Serfas 2004), and DNA sequence/phenotype associations (Beldade et al. 2002) strongly imply that they function in the development of eyespot patterns. These proteins thus provide convenient markers for the processes underlying eyespot pattern formation and have been used extensively in comparative studies (Brakefield et al. 1996; Brunetti et al. 2001; Reed and Gilbert 2004; Reed and Serfas 2004). Here we use Notch (N), Spalt (Sal), and Engrailed (En) as markers to study intraspecific variation in the early focal determination process.

**MATERIALS AND METHODS**

**Antibody stains**

*J. coenia* were collected from field sites around Durham, North Carolina and maintained in a lab colony at 28°C, on a 16L:8D photoperiod as previously described (Nijhout 1980). Larval wing disks were removed, fixed, and stained as previously described (Brunetti et al. 2001). Anti-En mouse monoclonal 4F11 (Patel et al. 1989) (1:5 dilution), anti-N mouse monoclonal C17.9C6 (Fehon et al. 1990; 1:200 dilution), and anti-Sal rabbit polyclonal (Barrio et al. 1999; 1:200 dilution) primary antibodies were detected with Cy2-conjugated goat anti-rabbit or Cy3-conjugated goat anti-mouse secondary antibodies (Jackson ImmunoResearch Laboratories Inc., West Grove, PA, USA). All dorsal epithelial expression patterns were visualized on a fluorescent light microscope and digitally photographed.

**Wing disk staging**

Timing of larval development in butterflies is only generally associated with wing disk development and is a poor predictor of the exact developmental state of the disks (H. F. Nijhout unpublished). We therefore followed previous studies (Miner et al. 2000; Reed and Serfas 2004) in using objective landmarks in the progress of vein development as a reference for staging disk development (Fig. 1):

*Stage 0*: Wing disk as found after the fourth-instar molt.
*Stage 0.25*: Vein lacunae discernable.
*Stage 0.5*: Subtle thread-like tracheoles extended from basal mass are discernable.
*Stage 0.75*: Early extension of tracheae.
*Stage 1.0*: Extension of two prominent anterior tracheae.
*Stage 1.25*: Beginning extension of posterior trachea.
*Stage 1.5*: Extension of all major tracheal branches.
*Stage 1.75*: A few tracheae reach the border lacuna.
*Stage 2.0*: Majority of tracheae reach border lacuna.
*Stage 2.25*: A few tracheae begin extending into border lacuna.
*Stage 2.5*: Extension of most tracheae into border lacuna.
*Stage 2.75*: Early extension of tracheoles into peripheral tissue.

*Stage 3.0*: Tracheae form continuous line in border lacuna, with moderate extension of tracheoles into peripheral tissue.
*Stage 3.25*: Early extension of tracheoles into intervenous tissue.
*Stage 3.5*: Moderate levels of tracheoles growth into intervenous and peripheral tissue.
*Stage 3.75*: Extensive tracheole growth in intervenous tissue. Disk has a brownish tint when dissected (not visible in figure).
*Stage 4.0*: Large, brownish wing disk with extensive tracheation. As found in larvae just before the prepupal phase.

**Gene expression patterns**

The N expression patterns we observed were consistent with previous work (Reed and Serfas 2004), and could be divided into five categories: (1) broad expression in intervenous regions without a noticeable upregulation along intervein midlines (Fig. 2A), (2) upregulation along intervein midlines with no obvious expansion of focal expression (Fig. 2B), (3) upregulation along intervenous midlines with an obvious expansion of focal expression (Fig. 2C), (4) upregulation in five well-defined foci, with little or no midline expression (Fig. 2D), and (5) strong upregulation in posterior-most focus, with four anterior foci being greatly reduced or undetectable (Fig. 2E).

We observed previously undescribed expression patterns implicating Sal in early eyespot determination. Expression of Sal in final-instar wing disks could be defined in five categories: (1) broad intervenous expression with no upregulation in foci (Fig. 2F), (2) broad intervenous expression with expansion and upregulation in a column of foci running antero-posteriorly (Fig. 2G), (3) the appearance of two parallel lines of expression running from foci to the border lacuna (Fig. 2H), (4) expression in seven foci without distal parallel lines (Fig. 2I), and (5) expression in the five posterior foci corresponding to the foci of N expression (Fig. 2J).

**Scoring and analysis of expression patterns**

To score complex gene expression patterns, we defined five gross expression states for each protein based on easily diagnosable characteristics scored from whole-mount stains (see “Results” for specific descriptions). To assess how expression patterns for each protein varied over time, we treated the expression pattern designations as continuous values from 1 to 5, and calculated variance for each protein along the developmental stage scale using a sliding 0.75-stage window. We found that using sliding windows of different interval sizes yielded similar results, but a 0.75-stage window worked best for producing continuous variance plots given our sampling density and scoring scheme. Temporal variance profiles were compared with a null model threshold (\(\sigma^2 = 0.5\)), representing the maximum variance for situations where expression patterns are allowed to overlap at one time point.

**RESULTS**

**Gene expression patterns**

The N expression patterns we observed were consistent with previous work (Reed and Serfas 2004), and could be divided into five categories: (1) broad expression in intervenous regions without a noticeable upregulation along intervein midlines (Fig. 2A), (2) upregulation along intervein midlines with no obvious expansion of focal expression (Fig. 2B), (3) upregulation along intervenous midlines with an obvious expansion of focal expression (Fig. 2C), (4) upregulation in five well-defined foci, with little or no midline expression (Fig. 2D), and (5) strong upregulation in posterior-most focus, with four anterior foci being greatly reduced or undetectable (Fig. 2E).

We observed previously undescribed expression patterns implicating Sal in early eyespot determination. Expression of Sal in final-instar wing disks could be defined in five categories: (1) broad intervenous expression with no upregulation in foci (Fig. 2F), (2) broad intervenous expression with expansion and upregulation in a column of foci running antero-posteriorly (Fig. 2G), (3) the appearance of two parallel lines of expression running from foci to the border lacuna (Fig. 2H), (4) expression in seven foci without distal parallel lines (Fig. 2I), and (5) expression in the five posterior foci corresponding to the foci of N expression (Fig. 2J).
Fig. 1. Developmental staging for final-instar *Junonia coenia* wing disks. Stages are given in the bottom right corners of panels. Images shown are not to scale. The varying opacities of disks in different panels are due to different preparation methods, and are not related to developmental stage. See “Materials and Methods” for specific stage descriptions.
Here we present a time series of En expression patterns divided into five states: (1) broad intervenous expression in posterior of wing (Fig. 2K), (2) broad intervenous expression in posterior of wing, with diffuse upregulation in region of major eyespot focus (Fig. 2L), (3) diffuse upregulation along anterior-to-posterior eyespot column (Fig. 2M), (4) discrete, well-defined upregulation in a row of crescent-like foci (Fig. 2N), and (5) strong upregulation in a discrete, round focus associated with the position of the major eyespot (Fig. 2O).

Variation in the progress of pattern formation
To determine the relationship between developmental stage and expression patterns we assessed both characteristics...
for all wing disks surveyed (Fig. 3, A–D). We calculated the mean developmental stage for each expression pattern (Fig. 3A), and found that our scoring system produced co-linear associations ($r^2 = 0.84$ for N, $r^2 = 0.86$ for Sal, and $r^2 = 0.85$ for En). This orderly relationship among average patterns of expression, however, masked a great amount of underlying individual variation in the absolute and relative timing of the expression patterns of these three proteins (Fig. 3, B–D).

Sal variance peaked earliest of the three proteins, at Stage 1.0 ($\sigma^2 = 0.8$, Fig. 4), around the time when the anteroposterior column of nascent eyespot foci was observed in most disks. N variance peaked at Stage 1.25 ($\sigma^2 = 0.9$, Fig. 4), when the intervein midline became apparent in most wing disks. En variance peaked between Stages 3.25 and 3.5 ($\sigma^2 = 0.6$, Fig. 4), when the row of crescent-like focal expression patterns was typically seen. Variance peaks for all three proteins crossed the $\sigma^2 = 0.5$ null model threshold.

Fig. 3. Relationships between developmental stage and expression patterns of Notch (N), Spalt (Sal), and Engrailed (En). (A) Mean developmental stages observed for each expression pattern. Wing disks were assigned Stages 0–4 based on progress of vein development as in Fig. 1. Expression patterns were scored as ordered states 1–5, as in Fig. 2. Bar and whisker plots illustrate the medians, ranges, and 25th and 75th percentiles of expression patterns for N (B, $n = 53$), Sal (C, $n = 34$), and En (D, $n = 27$). Mean stages are shown by black points. The colored fields are smoothed contour landscapes generated around the raw data.

**DISCUSSION**

**Early versus late variation in pattern formation**

It is controversial whether or not developmental variation is expected to be greater in early or late stages of pattern formation, and informative data are scarce. One body of theory predicts that individual variation in early development, attributable to stochastic starting conditions or individual differences in genetic or environmental conditions, is gradually reduced or canalized by various noise filtering functions (Wagner and Misof 1993; von Dassow et al. 2000; Meir et al. 2002). This results in a final phenotype that is substantially less variable than it was at earlier ontogenetic stages (Veitia and Nijhout 2006). In contrast, scenarios favoring late variation are based on the idea that early developmental variation is less likely because it will have a greater effect on later development (Wagner and Misof 1993), and that many late-expressing genes will have passed their critical periods of
function and are freer to exhibit neutral variation in expression patterns (True and Carroll 2002).

In support of the early-variation hypothesis, our data show that N and Sal expression patterns are most variable in the early stages of prepattern elaboration, with an overall decrease in variance as the eyespot focal patterns mature. Furthermore, temporal associations of N and Sal expression patterns (Fig. 3A), as well as variance peaks (Fig. 4) imply that the regulation of these genes may be tightly coordinated by a shared process. En expression patterns, in contrast to N and Sal, showed highest variance late in the pattern formation process (Fig. 4). We speculate that this observation could be explained if En expression was a downstream effect of N and/or Sal activity, in which case one could imagine in a two-stage delayed “pulse” of variance moving through the pattern formation cascade. Indeed, previous eyespot development simulations (Evans and Marcus 2006) and our mean stage data (Fig. 3A) both suggest En is a later-responding component of the eyespot determination pathway. Whatever the functional relationships between N, Sal, and En may be, it is of general interest to find that molecules involved in the same pattern formation process can show fairly different patterns of early versus late variance.

**Implications for wing pattern evolution**

The variation we report here can be described as heterochronic, in that it consists of differences in the timing of developmental events relative to each other—in our case gene expression versus wing disk stage. Variation in relative timing is of particular interest because developmental heterochrony has been proposed as a mechanism for butterfly wing pattern polymorphisms (Koch et al. 2000a, b), as well as wing pattern evolution at a phylogenetic level (Reed and Serfas 2004).

We speculate that the heterochronic variation we observed in this study may be representative of a standing resource of variation that has the potential to contribute to phenotypic evolution. In this respect, it is notable that cold shock experiments on carefully synchronized cohorts of *J. coenia* and other butterflies typically result in Poisson distributed, quantitatively continuous series of aberrant wing pattern phenotypes (Nijhout 1984). The biological basis of this consistent pattern of variation remains a mystery, but could possibly be explained by the kind of pattern formation variance we report here. For instance, if the stage of color pattern formation is not perfectly correlated with the external features used for timing temperature shocks, a range of aberrant color pattern phenotypes would be expected.

One finding from this study is that there are transient eyespot expression patterns over the course of development in *J. coenia*. In particular, Sal expression at Stage 4 occurs in seven spots, which decreases to five spots at Stage 5. Additionally, N at Stage 4 is expressed in five spots, which then decreases to one major spot at Stage 5. We speculate that these transient prepatterns of focal expression may represent developmental potential for eyespot formation. If variation in the timing of the transition from “more foci” to “fewer foci” permits some individuals to pupate before focal downregulation is completed, variation in eyespot number might be predicted. Indeed, supernumerary eyespots coincident with the observed transient expression patterns are found at extremely low frequencies in natural populations of *J. coenia*, as are individuals lacking the small anterior eyespot (H. F. Nijhout unpublished).

Future studies to address the heritability and genetic underpinnings of developmental variation would be welcomed. There are some examples of developmental processes whose logic relies to some degree on stochasticity to produce stereotyped outcomes, with one of the best-known examples being the N/Delta interaction in establishing neural precursors in *Drosophila* (Kaern et al. 2005). While it is unlikely that many developmental events are truly random, untangling the relationship between hard-wired genetic regulatory variation (i.e., gene-intrinsic variation) and process stochasticity (i.e., gene-extrinsic variation) would be useful for better understanding the evolutionary roles of developmental variation. Indeed, our evolutionary speculations require that at least a portion of developmental variation be heritable.

Another avenue for future work would be to improve methods for quantitative analysis of spatiotemporal gene expression data. While qualitative scoring systems like the one we used here are sufficient for analyzing gross expression pattern characteristics, quantitative treatment of antibody stain images would facilitate a better understanding of variation in subtle pattern elements like gradients. Importantly, careful quantitative analysis can also help compensate for the effects of experimental error caused by variability in detection techniques (Crauk and Dostatni 2005). While the hetero-
chronic variation in pattern formation we describe in this study has no immediately obvious correlation with phenotypic variation (beyond possibly the rare variation in eyespot number discussed above), there does exist heritable variation in *J. coenia* eyespot size (Paulsen 1994) that could potentially be associated with quantitative differences in the intensity and/or diameter of focal expression patterns.

**Importance of temporal sampling**

The variation we describe here has practical implications for comparative studies of gene expression, because it illustrates the importance of sample size when characterizing expression patterns. The majority of published studies consider only one or a few expression patterns to inform specific conclusions, particularly where in situ hybridization or antibody stain data are concerned. This approach may be problematic because transient prepatterns, unsampled patterns, and variation in timing can all impose a significant bias on studies with limited temporal sampling. These biases can lead to incorrect conclusions regarding expression pattern time series within species, as well as the presence or absence of specific expression patterns in multi-species comparative studies.

For example, if we sampled three wing disks at Stage 2.5 and assayed one each for expression of N, Sal, and En, there is an 11% chance that all three genes would have the same score, a 33% chance that Sal expression would appear to be advanced relative to the other two genes, a 22% chance that N expression would appear to be advanced relative to the other two, and a 33% chance that En expression would appear to lag behind the other two. In each of these cases, chance alone would lead to a different interpretation about the hierarchy of gene expression during pattern development.

As well, significant problems would also arise if only one or more time points were chosen to represent each gene in a phylogenetic analysis. For example, N could be scored as occurring in zero, one, or five focal expression patterns depending on the individual and/or developmental stage sampled. Development is a variable and time-dependent process, and conscientious studies would do well to take both of these characteristics into account.

**CONCLUSION**

All three proteins we assessed in this study showed variance in the progression of pattern formation. This suggests that developmental variation may be common across genes, and that butterfly wing pattern development in particular is developmentally buffered to some extent. Furthermore, our data offer a preliminary view that variation may not be consistently associated with either early or late stages of pattern formation, or even protein function (N is a receptor, while Sal and En are transcription factors). Our findings also clearly demonstrate the importance of temporal sampling in studies of gene expression patterns.

**Acknowledgments**

This work was supported by the National Science Foundation and Duke University. We thank Nipam Patel for sharing the En antibody, Rosa Barrio for sharing the Sal antibody, Developmental Studies Hybridoma Bank for providing the N antibody, Laura Grunert for rearing butterflies, and the anonymous reviewers for their helpful comments.

**REFERENCES**


