Identifying a developmental transition in honey bees using gene expression data

Bryan C. Daniels[1], Ying Wang[2], Robert E. Page Jr[3,4], and Gro V. Amdam[3,5]

[1]School of Complex Adaptive Systems, Arizona State University, Tempe, Arizona, USA

[2]Banner Health Corporation, Phoenix, Arizona, USA

[3]School of Life Sciences, Arizona State University, Tempe, Arizona, USA

[4]Department of Entomology and Nematology, University of California Davis, Davis, California, USA

[5]Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Aas, Norway

In many organisms, interactions among genes lead to multiple functional states, and changes to interactions can lead to transitions into new states. These transitions can be related to bifurcations (or critical points) in dynamical systems theory. Characterizing these collective transitions is a major challenge for systems biology. Here, we develop a statistical method for identifying bistability near a continuous transition directly from high-dimensional gene expression data. We apply the method to data from honey bees, where a known developmental transition occurs between bees performing tasks in the nest and leaving the nest to forage. Our method, which makes use of the expected shape of the distribution of gene expression levels near a transition, successfully identifies

the emergence of bistability and links it to genes that are known to be involved in the behavioral transition. This proof of concept demonstrates that going beyond correlative analysis to infer the shape of gene expression distributions might be used more generally to identify collective transitions from gene expression data.

1 Introduction

Social insects represent well-known examples of adaptive collective systems, combining the efforts of many individual actors to produce robust and adaptive aggregate behavior [1]. The allocation of tasks to individuals and, in some cases, specialization of individuals to particular tasks, often displays a sophisticated organization that promotes collective success [2]. This distributed coordination of effort is the result of a complicated process reaching from the level of gene regulation to social interactions. Mapping out the drivers of this process is a major challenge in understanding how collective behavior is successfully regulated in social insects, and connects to open questions more generally across living systems [3, 4].

In honey bees, specialization occurs typically among the worker daughters of a single queen. The largest distinction among workers is between those that mostly perform tasks within the nest (brood care, food processing, nest construction and maintenance, and colony defense) and those that leave the nest to forage. Bees typically begin their lives performing more in-nest tasks (nest bees), and in about their second to fourth week of life a major transition occurs in which they switch to foraging outside the nest (foragers) [5, 6]. Individual bees vary in the age at which they make this nest bee to forager transition, and while the transition can be reversed in individual bees if the colony's needs change, few normally do. This temporal bifurcation leads to an age-related division of labor [7].

The mechanisms behind this behavioral transition include interactions between genes and other factors expressed within each bee, interactions between bees, and interactions between bees and their local environment [5, 8–14]. At the individual level, the gene product Vitellogenin (VG, formerly known as an egg-yolk precursor) and the endocrine factor juvenile hormone (JH), and the interactions between the two, have been well studied [5, 12, 15]. The interaction takes the shape of a mutually

suppressive feedback control system that appears to be at the core of the bifurcation of behavior between nest activities and foraging. Increasing levels of VG suppress JH in early life, while lower levels of VG in later life leads to higher titers of JH that further suppresses VG by inhibiting vg gene expression [5]. The behavior of this positive feedback control is consistent with the phenomenology of the transition, with either increases to JH (e.g. resulting from fewer interactions with other foraging bees as well as exposure to a variety of stressors) or decreases to VG (e.g. resulting from reduced food intake or excessive consumption of physiological nutrient stores during nursing tasks) leading to reinforcing feedback and the initiation of foraging. This general picture has recently been refined by situating VG and JH within larger gene networks involved in target of rapamycin and insulin signaling, ecdysone response, ovarian and neural activation [13]. Still, little is known about the full spectrum of genes and network dynamics that drive the transition from nest bee to forager and at least temporarily "lock" worker honey bees into the different roles.

An analogous phenomenon occurs in cell differentiation in multicellular organisms. Here, genetically identical cells perform distinct roles in the larger organism. Distinct cell types are "locked in" by regulatory interactions typically understood at the level of gene transcription. The dynamics of these interactions are thought to produce separate attractor states with different gene expression patterns that correspond to distinct cell types [16, 17].

In the language of dynamical systems, transitions among separate attractors are understood as arising from bifurcations that change the number of attractor states or from noise-induced hopping among co-existing attractors. In the cellular differentiation literature, statistical tools have been developed for use with gene expression data to identify transitions that create new attractor states [17–20]. These studies have shown that multiple cases of cell state change are consistent with dynamics that pass through or near a continuous bifurcation, also known as a critical transition. This type of bifurcation is known to result from positive feedback loops in many biochemical networks [21,22]. More generally, such critical transitions are theorized to be ubiquitous across biology as mechanisms for controlling collective behavior [3, 4, 23–26].

In honey bees, studies linking gene expression to behavior have tended to focus on searching for individual genes whose expression correlates with behavior [14,27,28]. This contrasts with a collective-scale

3

75 view of multiple interacting genes across multiple individuals that we anticipate will be necessary to
76 find evidence of critical transitions.

77 In this study, we analyze gene expression data across an ensemble of individual honey bees and across
78 a series of developmental timepoints to examine genetic-scale indicators of the transition from in-nest
79 tasks to foraging. We develop a method for identifying critical transitions based on the Landau theory
80 of phase transitions in statistical physics and the theory of bifurcations in dynamical systems. At
81 each timepoint, using measurements of gene expression across multiple bees, we are able not only
82 to estimate mean expression values and their co-variance, but to infer the shape of the distribution
83 of expression values. This shape can be expected to begin as unimodal as bees initially emerge from
84 pupation, with gene expression values fluctuating around a common mean. As the adult bees develop,
85 increased variation in behavior could be connected to continuous variation in gene expression, or at
86 some point expression values could split into two well-defined groups (e.g. nest bees and foragers)
87 corresponding to discontinuous variation. Such a transition would be indicated by a change from
88 unimodality to bimodality in the distribution of expression values, which our method explicitly identifies.
89 Our method is designed to work with high-dimension-low-sample-size data as is common in transcriptomic
90 data; it does not require prior knowledge of important genes; and it is formulated using a Bayesian
91 interpretation to provide precise statistical meaning with respect to model selection and transition
92 identification. We use our method to answer the questions of whether and when the bees' transition
93 to foraging is visible in gene expression data, whether it is consistent with a continuous bifurcation
94 transition, and which genes are most associated with such a transition.

2 Results

96 We examined gene expression profiles of $N_{\mathrm{samples}} = 16$ bees at 5 developmental timepoints, from
97 1 to 15 days old, spanning the transition in behavioral development from all bees remaining in the
98 nest to some bees leaving the nest to forage. To critically test the concepts of our statistical model,
99 we focused on a set of genes $N_{\mathrm{genes}} = 91$ that we expected to be involved in distinguishing the nest
100 bee and forager behavioral phenotypes through their actions on and with VG, a known determinant

4

of foraging onset. Some of these genes have known correlations with VG production and with each other, while others were suspected of being involved through their action in common gene networks. Additional genes were selected because they are located within the genetic map confidence limits of known quantitative trait loci (QTL) shown to have effects on ovaries, VG production, and foraging behavior [29], or involved in stress or immune-associated pathways that have been correlated with the behavioral transition before [30]. We predicted that the expression of sets of genes affecting foraging should diverge from their expression in nest bees leaving a detectable bimodal distribution of gene expression representing the bifurcation.

We tested for bimodality consistent with a continuous transition using a Bayesian analysis. In particular, we compared the goodness-of-fit of a unimodal Gaussian distribution with the form of bimodal distribution expected near a continuous transition (see Methods). We were able to perform this test with low sample sizes ($N_{\mathrm{samples}} < N_{\mathrm{genes}}$) because bistability is expected to begin along the dimension with the largest variance (the first principal component), and we could thus focus on identifying bimodality along this single dimension.

We first demonstrate proof of the principle that this method can successfully infer a transition to bistability in simulated data from a simple model of noisy gene regulatory dynamics with $N_{\mathrm{genes}} = 91$. Fig. 1 displays the fraction of simulations in which evidence of a transition was successfully identified as a function of the strength of interactions that induce a transition. While a larger number of samples can identify the transition earlier ($N_{\mathrm{samples}} = 100$; red points), even with the smaller number of samples that we have here ($N_{\mathrm{samples}} = 16$; orange points), we expect to be able to identify such a transition once the bimodality is more pronounced.

Applying this analysis to the honey bee expression data, we find a transition into two distinct groups that is statistically visible at age 10 days and more apparent at age 15 days (Fig. 2). In Fig. 3, we plot for these two ages the gene expression data and inferred transition distribution along the bistable dimension (corresponding to the first principal component; see Methods). The data and the inferred distribution in this dimension are characterized by bees that fall roughly into two clusters, with a few bees that are transitional between the two clusters.

Different genes contribute differentially to defining the bistability. Sorting by the fraction s of each
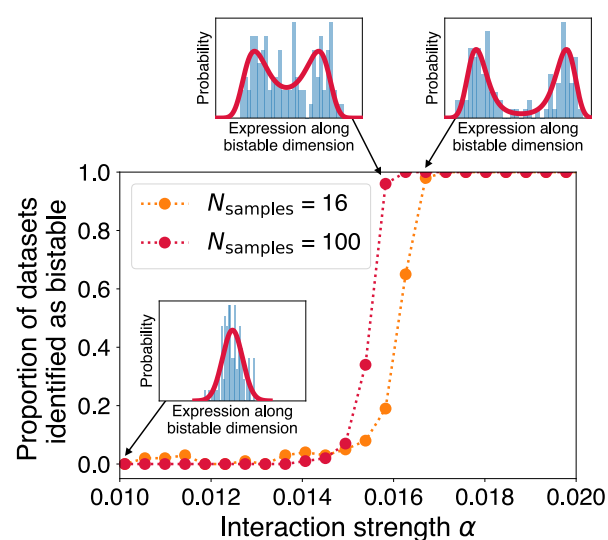
Figure 1: **The method successfully identifies the transition state in simulated data.** As interaction strength α increases in a simple dynamical model of gene regulation (see Eq. (19) in the Methods), gene expression levels projected along the first principal component transition from a unimodal to a bimodal distribution (three insets show simulated data as blue histograms and the best-fit distributions as red curves). A statistical comparison between a Gaussian distribution and the distribution shape expected near a continuous transition reliably identifies the transition state once the bimodality is sufficiently pronounced (red and orange points). Increasing the number of data samples allows identifying the transition state when bimodality is less pronounced (compare red and orange points).
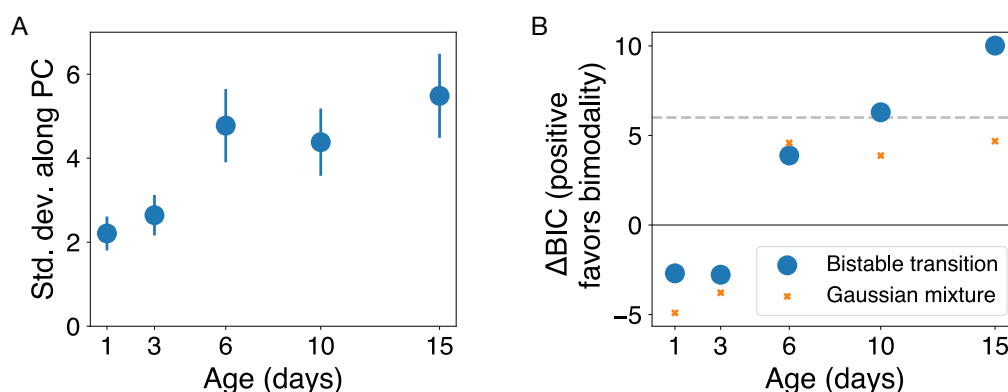
Figure 2: **In honey bees, both the variance of gene expression and the strength of evidence for bistability grow over developmental time.** A: The standard deviation of gene expression along the principal component increases during development. Error bars show standard errors ($N_{\mathrm{samples}} = 16$). B: The Bayesian Information Criterion measure ΔBIC quantifies the strength of evidence in favor of the bistable transition distribution as compared to a unimodal Gaussian, with positive values favoring bistability (blue circles; see Methods Eq. (16)). We interpret ΔBIC values larger than 6 (horizontal dashed line) as strong evidence for bistability. We also compute ΔBIC that compares a Gaussian mixture ($n = 2$) with the unimodal Gaussian, which identifies weaker evidence for bimodality (orange Xs; see section 3).
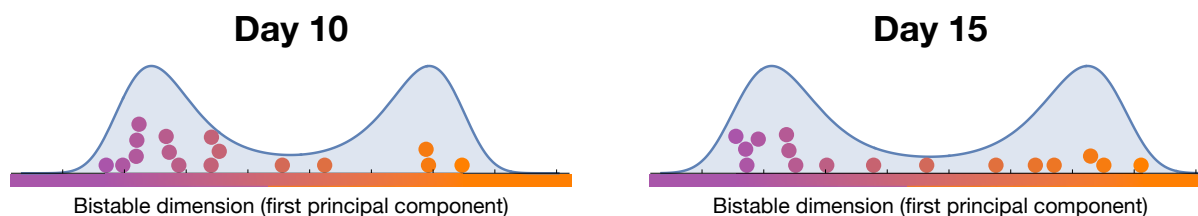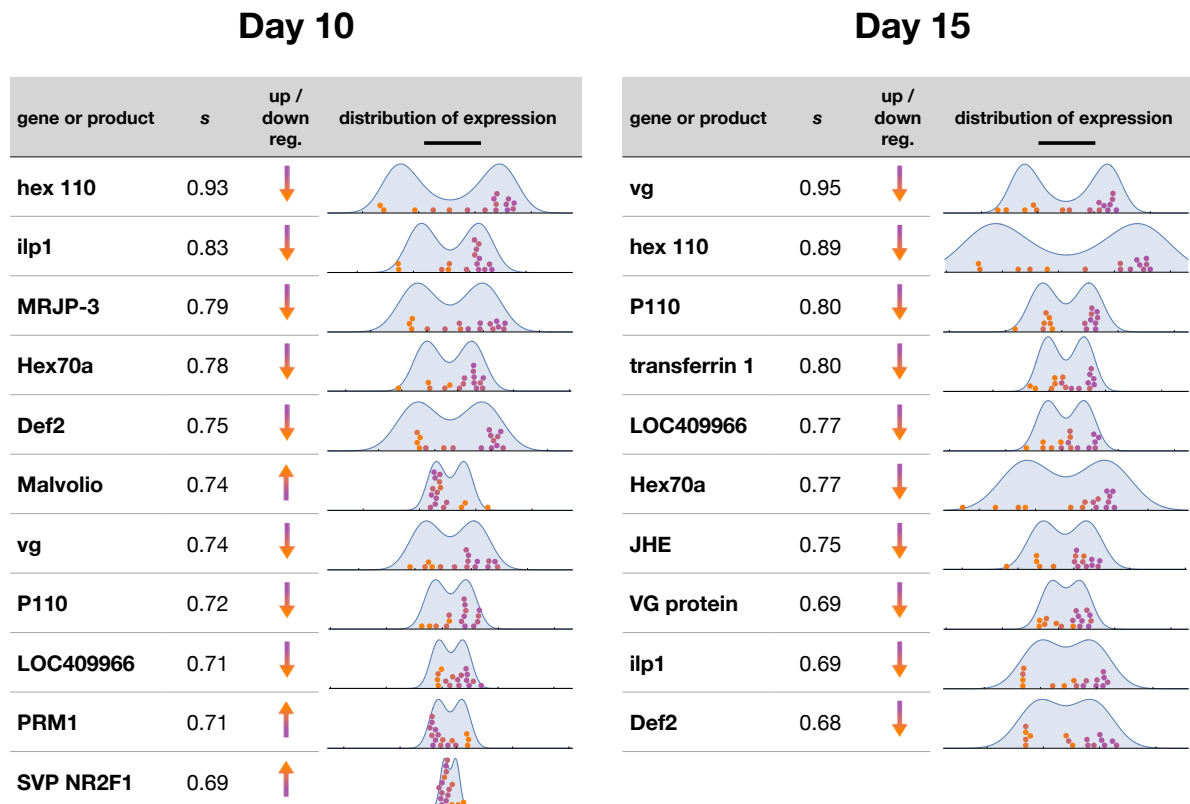

Figure 3: **Bistability along the first principal component.** Gene expression data from 16 bees at age 10 days and 15 days projected along the first principal component (colored circles; log-transformed data). Colors correspond to distance along this dimension, with orange chosen to represent the low *vg* state indicative of foragers. The fit Landau distribution is shown in blue.

7

Day 10

gene or product	s	up / down reg.	distribution of expression
hex 110	0.93	↓	
ilp1	0.83	↓	
MRJP-3	0.79	↓	
Hex70a	0.78	↓	
Def2	0.75	↓	
Malvolio	0.74	↑	
vg	0.74	↓	
P110	0.72	↓	
LOC409966	0.71	↓	
PRM1	0.71	↑	
SVP NR2F1	0.69	↑	

Day 15

| gene or product	s	up / down reg.	distribution of expression
vg	0.95	↓	
hex 110	0.89	↓	
P110	0.80	↓	
transferrin 1	0.80	↓	
LOC409966	0.77	↓	
Hex70a	0.77	↓	
JHE	0.75	↓	
VG protein	0.69	↓	
ilp1	0.69	↓	
Def2	0.68	↓	

Table 1: **Individual genes associated with the bistability.** A handful of individual genes or gene products have a large proportion s of their variance along the bistable dimension. Here we highlight all genes with $s > 2/3$. Orange and purple colors correspond to bees on two sides of the bistability, with orange chosen to represent the low *vg* state that is indicative of foragers. In the column "up/down reg.", we indicate whether the gene is up or down regulated in the orange state as compared to the purple state. Expression data for these genes are shown with the same colors for individual bees as in Figure 3, along with the marginalized fit distribution in blue (log-transformed data, with scale bar corresponding to an expression ratio of 10).

129 gene's variance that lies along the bistable dimension highlights those genes whose expression values

130 give the most information about the side of the bistability on which each bee lies (Table 1 and SI

131 Table 2). In Table 1, we also display the measured expression levels of the most informative genes,

132 with datapoints from each bee colored according to their positions along the bistable dimension.

133 Comparison with known gene expression patterns suggests that purple points correspond to bees

134 predisposed to in-nest behavior, and orange points to those predisposed to foraging behavior [5, 31].

135 3 Comparison to other potential methods for identifying a transition
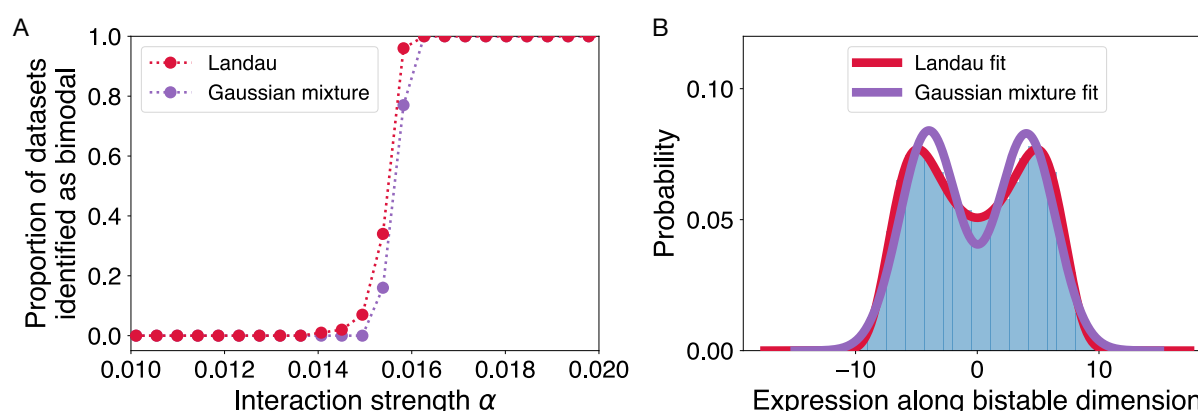
Figure 4: **Fitting to the Landau probability density function performs better than a mixture of two Gaussians.** In tests with simulated data, the Landau method (A) identifies the transition state sooner (at smaller μ, when the bistable states are closer to one another; shown here fitting 100 samples), and (B) fits the transition distribution much more closely (shown here fitting 50,000 samples at $\mu = 0.0158$; simulated data in blue histogram compared to best fits of Landau distribution and Gaussian mixture distribution shown as solid curves).

136 A family of related methods for detecting bimodality would correspond to fitting bimodal distributions

137 of different shapes along the first principal component. One simple choice would be a mixture model

138 that combines two Gaussian distributions. We compare our derived "Landau" distribution to this

139 Gaussian mixture distribution (fitting the location of each Gaussian, a common width, and their relative

140 weight, producing 2 extra degrees of freedom compared to the single Gaussian model). In Fig. 4, we

141 demonstrate using simulated data that this distribution performs noticeably worse in detecting the

142 transition to bimodality and in fitting the distribution of gene expression values. Furthermore, the

143 orange points in Fig. 2B show that the Gaussian mixture model is not as statistically favored as the
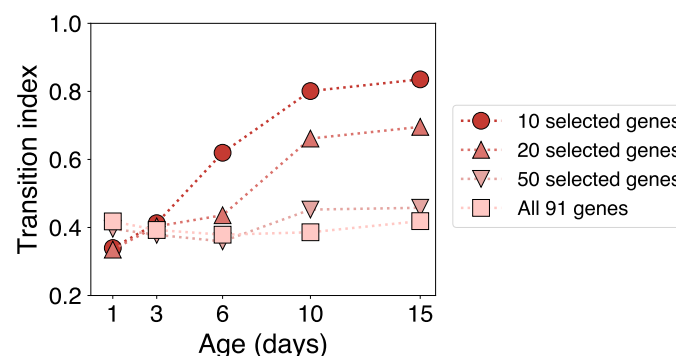
9

Figure 5: **Existing "transition index" measure also suggests a transition when focused on particular genes.** The transition index defined in Ref. [17] requires selecting a set of genes that are known *a priori* to be involved in a transition. Applying this measure to our data and restricting to genes identified by our method, the largest values occur at age 10 and 15 (red circles). This corroborates our method, which does not require selecting specific genes, yet locates a transition at a similar time. In contrast, when including data from all 91 genes (pink squares), the signal in the transition index is washed out.

144   Landau model in fitting our experimental data.

145   Another related technique, explored in the context of cell differentiation in Ref. [17], looks for signatures

146   of a continuous transition related to increased variance in the expression of a given subset of genes.

147   The method is simpler than ours in that it only requires measuring pairwise correlations among

148   genes and bees, and does not explicitly measure the shape of the distribution of expression values.

149   The method has the advantage of explicitly incorporating the increase in variance one expects to

150   find when an existing attractor is destabilized, which can help to distinguish the case of attractor

151   destabilization from the case of hopping between co-existing attractors. Applied to our data, a relative

152   increase of the resulting "transition index" at days 10 and 15 (see Fig. 5) is also suggestive of a transition

153   state, but only when limited to genes that we identify as most informative through our measure s.

154   Though our method is somewhat more computationally involved, it is able to identify individual

155   genes that are related to transition states, infer the shape of the transition distribution, and produce a

156   statistically interpretable measure (ΔBIC) of the likelihood of bistability given the data.

4 Discussion

We show that expression data from 91 genes, taken from 16 honey bee workers sampled at 5 different time intervals that span the first 15 days of life, demonstrate bimodality within time intervals that becomes more pronounced with age. We hypothesize that the observed bimodal distributions of gene expression correspond to known temporal dynamics of behavior. Days 10 to 15 correspond to a known developmental time point at which individual bees transition or prepare to transition from in-nest activities to foraging [6]. Days 10 and 15 demonstrate the strongest evidence for bimodality, suggesting that the two modes represent the transitioning of bees from those that engage in nest activities into those that forage.

We propose that the sets of genes defining these bimodal structures are working together in networks with bistability associated with the distinct behavioral states [5, 15, 32]. The gene networks have two stable states that are driven by the JH/VG positive feedback control system. The JH/VG feedback system drives gene expression that sculpts the physiological and behavioral changes in worker behavior. The resulting switch is steep as a consequence of the feedback mechanisms affecting blood titers of VG and JH. During the nest stage, blood levels of VG initially increase but eventually decline due to dynamics including depletion (its use in brood food production) and/or suppression of synthesis due to reduced food intake/availability, disease or stress. As VG levels decline, JH titer increases, driving the production of VG even lower. This positive feedback is behind the steep switch and a route to the transition of the bee into the forager state. Note that VG may directly affect expression levels of other genes via DNA-binding capability [33]. There is also a JH and VG independent pathway to the forager state, but its initiator(s) is unknown [5, 34].

The transition is affected by the external environment through feeding activities and stressors: Feeding more larvae can result in more rapid VG depletion and an earlier onset of foraging. Workers in colonies that have stored insufficient pollen, the source of protein, initiate foraging earlier in life, presumably due to a deficiency in VG levels. Workers in colonies with fewer foragers accelerate the onset of foraging, presumably by having a reduction of some unknown inhibitor passed from foragers to younger bees [35] or by the lack of nutrients to support colony growth. Similarly, pathogens that

11

184 reduce VG production by, for instance, reducing nutrient uptake over the gut or consumption of

185 resources in immune activation lead to earlier foraging onset [36].

186 Given this context, our list of genes that align most strongly with with the gene network bistability

187 (Table 1 and SI Table 2) is one of "usual suspects." Most of the genes listed have known involvement

188 in lipid and lipoprotein synthesis and metabolism, are lipoproteins found in the hemolymph or hypopharyngeal

189 glands of nest bees involved directly in larval and adult nutrition, regulate the JH/VG switch between

190 nest bees and foragers, or are known "immunity genes" involved in responses to pathogenic organisms.

191 This result is also an outcome of our initial selection of the 91 genes, which in essence focused on

192 "usual suspects" to provide a strong basis for testing our statistical methodology.

193 More specifically, VG, major royal jelly protein 3 (MRJP-3), and the hexamerins (hex 100 and Hex70a)

194 are lipoproteins that are metabolized prior to the onset of the foraging state. Elements of the insulin

195 signaling pathway include insulin like peptide 1 (ilp1), its receptor (InR1), and P110. Insulin-like

196 signaling (IIIS) acting in the fat body has been shown to be involved in VG production and the transition

197 from nest bee to forager. The target of rapamysin (TOR) pathway cross-talks with IIS signaling in the

198 regulation of JH, while at least TOR is influencing VG levels more directly presumably by communicating

199 nutritional context information to the VG production system [37], which is highly nutrient sensitive

200 [38]. Transferrin 1, hymenoptaecin, argonaute 2 (AGO2), and defensin-2 (Def2) are genes associated

201 with immunity. Their role may be one of adjusting to the change in pathogenic challenges that take

202 place when a bee transitions from the protected nest environment to foraging. Malvolio is a manganese

203 transporter that affects sucrose sensitivity, and through its effect on mn2+ in the brain presumably

204 influences neuromodulators involved in the onset of foraging. Octopamine (OA) is involved in the

205 onset of foraging and increases sensory motor responses involved in foraging; OAR is its receptor.

206 The pattern we reveal for juvenile hormone esterase (JHE), strongly correlated with bistability in

207 bees of age 15 days, might be particularly novel and intriguing. JHE provides opportunity to metabolize

208 JH inside cells, and this action may block cellular responses to bursts in circulating (blood) JH titers.

209 We show that JHE transcript levels decrease in the low vg state indicative of foragers, thereby making

210 these individuals potentially more sensitive to JH. Such sensitivity would add to the positive feedback

211 control system of VG and JH, which secures that transitioning individuals stay committed to their

12

212   new behavioral role.

213   What causes this split into two groups? The correlation of these genes and the bistability can be

214   explained in multiple ways. The transition could be controlled by an internal temporal program, or

215   "physiological clock", running independently within individual bees [39]. Alternatively, the bifurcation

216   could arise solely via external positive feedback, with amplification of individual differences in foraging

217   thresholds through interactions with the nest, the environment, and other bees, similar to that seen

218   in other examples of division of labor [9]. Perhaps most likely is a combination of these two mechanisms:

219   internal dynamics that create a steep switch between nest bees and foragers, with the timing of this

220   switch greatly influenced by the environment [5, 14]. The toy model of gene interactions that we

221   use to test the statistical inference is one way to represent dynamical mechanisms that can amplify

222   differences in gene expression to create distinct types, which could be further amplified at the behavioral

223   scale.

224   Our statistical method of Bayesian model selection for identifying bistability is designed to impose

225   the least structure possible to explain the data: we start from the simplest possible probability distribution

226   (Gaussian) and compare this to the case in which we add the lowest order correction term consistent

227   with a transition. This simple representation is useful when we have limited data (in this case, 16

228   bees per timepoint). In future cases with more data, more precise versions may be fruitful, including

229   those that test whether expression data are consistent with a continuous or discontinuous transition.

230   Existing methods for identifying critical transitions from gene expression data have largely focused

231   on the increase of correlated variability near a transition, and before the transition induces a split

232   into two clusters [17, 18, 40]. In contrast, and similarly to Ref. [19], the method we develop here goes

233   beyond pairwise correlation and looks for a characteristic shape of bimodality during and after this

234   split. Our method remains statistically robust, working in the regime of "high dimension low sample

235   size" [41] that is common in gene expression studies. The method has no free parameters, unlike

236   many clustering methods, and it is interpretable in a Bayesian analysis to quantify the evidence favoring

237   bistability.

238   We anticipate that our approach could be useful in interpreting existing gene expression data, both in

239   the context of social insects and more generally across biology. First, in the development of multicellular

13

240 organisms, the divergence of progenitor cells into differentiated cells can also be well-represented

241 using bifurcation theory [17, 19, 42, 43]. Current approaches for identifying cell types from gene

242 expression data typically separate the data into distinct types using clustering methods (e.g. [19, 41,

243 44]). While a clustering approach is suitable for cases in which cell types are highly distinguishable,

244 clustering will break down as cell types become less distinguishable near a continuous transition.

245 Our method, in contrast, is designed to work near the transition point. This type of analysis could be

246 used, for instance, to use gene expression data to more precisely define the time in development at

247 which one cell type becomes two.

248 Other collective decisions, such as in neural systems, are thought to arise from continuous bifurcations

249 as well [45], and our method could be used directly on activity data to identify the individual components

250 (e.g. neurons) that may drive these transitions.

251 To summarize, in the context of animal behavior, our results add a new layer to understanding how

252 the strategy of task specialization plays out in biological collectives. Even when individuals share

253 the same genetic code and environment, their traits can diverge into separate tracks. Similarly to

254 differentiation in multicellular organisms, this divergence is visible both at the scale of functional

255 characteristics and at the scale of gene expression. Furthermore, the simplest model of critical bifurcation

256 provides a good description of how this divergence occurs. The generality of this phenomenology

257 suggests that such critical transitions may be a common mechanism within biology, making use

258 of the emergent properties of strongly interacting dynamical networks to generate reproducible

259 diversity. Further work is needed to detail how those networks are successfully constructed and

260 regulated in an evolutionary context.

Acknowledgments

Methods

Gene expression data

The bee collection was conducted during March 18th to April 7th, 2016. Five strong wild type colonies were chosen as sources of newly emerging bees. Two small nuclear colonies with 9 frames in each were used as recipient colonies that had similar supplies of pollen and honey, and similar numbers of brood and adult bees. The emerging frames from source colonies were put into incubator (37°C) and about 200 newly emerged bees were collected, marked in a color on the bee thoraxes and distributed evenly to the recipient colonies. This procedure was repeated for three consecutive days. About 600 newly emerged bees were introduced in the recipient colonies from March 19th to March 21st, 2016. Ten each of 1 day old, 3 day old, 5 day old, 10 day old and 15 day old bees from each recipient colony were collected in the following days. Therefore, there were total 20 sample bees for further experiment from two recipient colonies for each age. RNA was extracted from 16 randomly selected bees of each age using TRIZOL method. The quality of RNA was measured by Nanodrop and 260/280 was between 1.85-2.00, indicating good quality of RNA. RNA samples were sent to the University of Arizona Core lab for Nanostring analysis, a method to quantify the copy of RNA for the target genes. 96 genes (see SI Table) were selected based on prior knowledge and the potential association between these genes, VG production, and social behavior. Four housekeeping genes, actin (GB17681), rp49 (GB10903), GAPDH (LOC410122) and RPS18 (LOC552726) were used for sample normalization. Geometric mean of four housekeeping genes were used to calculate normalization factors to minimize the noise from individual genes using nSolver software provided by the company. We use log-transformed values of the raw expression data for all analyses.

Conceptual overview of statistical method

The expression of gene products across individual bees is correlated, with correlations arising from a complex set of dynamics. These dynamics include both direct interactions among measured genes and common factors of influence external to the measured genes. Changes to these dynamics (e.g. caused by stronger interactions among genes during certain phases of development, or shared interactions

15

291 with a common environment) can lead to increased variance in gene expression. This additional

292 variance is also correlated, forming patterns consisting of sets of genes with expression that moves

293 in tandem. In the case of a transition into two distinct types, we expect that this correlated variance

294 will additionally display bimodality. That is, two separate patterns of gene expression will coexist,

295 with few individuals having gene expression levels that are a mixture between the two types.

296 In the following description of methods, we first build a simple model of gene expression dynamics

297 that displays this characteristic transition to bimodality. We demonstrate that the separate patterns

298 of gene expression corresponding to this bimodality are given by the dimension of largest variance

299 (first principal component), as can be extracted from gene expression data by principal components

300 analysis. We then show that the distribution of gene expression values along the principal component

301 has a specific shape near the transition, which we derive. By fitting the observed data to this distribution,

302 we determine whether gene expression across a population of individuals of the same age displays

303 evidence of a bistable transition.

Dynamical model of gene interactions

305 Near a transition from a system with continuous variation into one displaying distinct types, one

306 expects the emergence of a bimodal distribution with a particular shape. In this section, we derive to

307 lowest order the form of this distribution for the case of the dynamics of a densely connected gene

308 regulatory network.

309 We start with a general form of dynamics for the (log-transformed) concentrations of the measured

310 gene products, arranged in the N_{genes}-dimensional vector $\vec{x}$:

$$\frac{d\vec{x}}{dt} = F(\vec{x}, \vec{y}, t), \tag{1}$$

311 where $\vec{y}$ consists of other unmeasured variables and t is time. We assume a separation of timescales

312 such that, at each developmental timepoint, we are measuring an equilibrium distribution, which

313 slowly changes as a function of developmental time. This means F depends on the long developmental

314 time T (on the timescale of days), but the distribution of $\vec{x}$ at each t will not depend on t defining

16

gene transcription dynamics (on the timescale of seconds):

$$\frac{d\vec{x}}{dt} = F_T(\vec{x}, \vec{y}). \tag{2}$$

The $\vec{x}$ we observe will then be restricted to be near attractors of F_T, determined by taking $t \to \infty$. Without the ability to measure the dynamics of more complicated attractors on this fast timescale (limit cycles, strange attractors, etc.), we will focus our attention on stable fixed point attractors $\vec{x}^*$, where

$$F_T(\vec{x}^*, \vec{y}^*) = 0. \tag{3}$$

If we were to observe only these fixed point attractors, and for fixed hidden variables $\vec{y}^*$, we would erase all information about hidden variables and about dynamics at the fast timescale. This corresponds to a description of the system we could get if we measured one bee per developmental timepoint.

We now want to incorporate information about variance in the observed $\vec{x}$ coming both from the effects of hidden variables $\vec{y}$ and from other fast-timescale dynamics. When these effects are small and temporally uncorrelated, we can approximate them as adding uncorrelated noise to F_T:

$$\frac{d\vec{x}}{dt} = F_T(\vec{x}) + \vec{\xi}, \tag{4}$$

where $\vec{\xi}$ is chosen from a multidimensional Gaussian with some fixed covariance. Assuming this variance is sufficiently small (compared to nonlinearities in F_T), the system stays near the fixed point $\vec{x}^*$. In this case, dynamics can be approximated using a Taylor expansion of F_T around $\vec{x}^*$ (related to the Hartman–Grobman theorem [46])

$$F_T(\vec{x}) = \Lambda\, \delta\vec{x} + O(|\delta\vec{x}|^2), \tag{5}$$

where $\delta\vec{x} = \vec{x} - \vec{x}^*$. The assumption that we are at a stable fixed point is equivalent to the statement that the eigenvalues of the matrix Λ are all negative. The solution of these dynamics is a multidimensional Gaussian. This corresponds to the analysis performed in Ref. [17], where the existence of a critical point corresponds to an eigenvalue of Λ becoming close to 0, such that there is a measurable large

17

334   variance in $\delta\vec{x}$ along the corresponding eigenvector.

335   We take the next step in the expansion by looking at what happens when all of the above assumptions

336   hold but we additionally want to treat the super-critical case, when an eigenvalue of Λ becomes

337   larger than 0. Our goal is to find the form of the simplest possible symmetry breaking transition that

338   leads to two distinct modes of gene expression. Still assuming the noise is small means $\vec{x}$ stays near

339   the fixed point except along the eigenvector $\hat{\nu}$ with eigenvalue greater than 0. Along that direction,

340   the variance is largest, so this dimension will correspond to the first principal component. Further,

341   because the dynamics are locally unstable to lowest order, the behavior along $\hat{\nu}$ is no longer determined

342   only by the lowest-order expansion given in Eq. (5). It will instead be determined by the lowest-order

343   term that cures the divergence of the variance in that direction.[1] Next will generally be a term of

344   order $|\delta\vec{x}|^2$, but this still allows a divergence of $\delta\vec{x}$ along either positive or negative $\hat{\nu}$. In addition,

345   we restrict our analysis to retain the symmetry of $F_T(\delta\vec{x}) = F_T(-\delta\vec{x})$, so that the system will be

346   equally likely to move in the positive or negative $\hat{\nu}$ direction at the transition, corresponding to a

347   symmetry breaking that is unbiased. It is also possible to include a bias term that leads to one of the

348   final attractors being favored over the other [19], but we omit that possibility here to produce the

349   simplest possible symmetry breaking transition. For these reasons, we assume the second-order term

350   is zero. The lowest-order term that does cure the divergence is of order $|\delta\vec{x}|^3$; we write

$$F_T(\vec{x}) = \Lambda\,\delta\vec{x} + \Gamma(\delta\vec{x}\cdot\hat{\nu})^3\hat{\nu} + O(|\delta\vec{x}\cdot\hat{\nu}|^4), \tag{6}$$

351   with $\Gamma < 0$. This corresponds to a so-called "supercritical pitchfork" bifurcation in dynamical systems

352   theory, where one stable fixed point changes continuously into three fixed points, two stable separated

353   by one unstable [47, 48]. The resulting two stable fixed points represent two distinct types of gene

354   expression — phenotypes or "cell fates" in the cell differentiation literature — into which the single

355   fixed point in gene expression splits. Finally, we note that even higher-order terms would be necessary

356   if $\Gamma > 0$, which corresponds to a subcritical pitchfork [42], but we again omit that possibility to

357   produce the simplest possible symmetry breaking transition.

---

[1]There are other ways the transition could happen such that this divergence is not cured (or not cured in a way that keeps $\delta\vec{x}$ relatively small). This corresponds, for instance, to a saddle-node bifurcation, in which the transition would cause the system to suddenly change to some other distant attractor. That is, there are transitions that would not be identified by our method. Our method is designed to detect only the simplest symmetric, continuous transitions.

Form of Landau distribution

We now derive the expected equilibrium distribution under the dynamics given by Eq. (4) with F_T as in Eq. (6). Writing the noise term more explicitly, we have

$$d(\delta\vec{x}) = \left[\Lambda\,\delta\vec{x} + \Gamma(\delta\vec{x}\cdot\hat{\nu})^3\hat{\nu}\right] dt + \sigma\,d\vec{W}_t, \tag{7}$$

where σ is a tensor that produces the noise when applied to the standard N_{genes}-dimensional Wiener process $\vec{W}_t$. The time derivative of $p(\vec{x})$, the distribution of $\vec{x}$ over multiple realizations of this dynamics, is given by the Fokker–Planck equation:

$$\frac{\partial p(\vec{x})}{\partial t} = -\vec{\nabla}\cdot\left[\left(\Lambda\,\delta\vec{x} + \Gamma(\delta\vec{x}\cdot\hat{\nu})^3\hat{\nu}\right) p\right] + \frac{1}{2}\nabla^2\left(\sigma\sigma^T p\right). \tag{8}$$

We then obtain equilibrium solutions by setting $\frac{\partial p}{\partial t} = 0$. Assuming a solution of the form

$$p_L(\vec{x}) = Z^{-1}\exp\left[-\frac{1}{2}\delta\vec{x}\cdot A\cdot\delta\vec{x}^T - \frac{B}{4}\left(\delta\vec{x}\cdot\hat{\nu}\right)^4\right], \tag{9}$$

we solve for A and B in terms of the dynamical parameters Λ, Γ, and σ:

$$A = -2\left(\sigma\sigma^T\right)^{-1}\Lambda \tag{10}$$

$$B = -2\left[\hat{\nu}^T\cdot\left(\sigma\sigma^T\right)^{-1}\cdot\hat{\nu}\right]\Gamma. \tag{11}$$

We call p_L from Eq. (9) the "Landau distribution" as we expect this shape near any continuous phase transition that can be described by simple Landau theory. In particular, Landau's theory of continuous phase transitions describes a Gibbs free energy of the form [49, 50]

$$\mathcal{F}(y) - \mathcal{F}_0 = \frac{a}{2}y^2 + \frac{b}{4}y^4. \tag{12}$$

Along the bistable dimension $\hat{\nu}$, the corresponding equilibrium distribution maps exactly onto the distribution $p_L(\vec{x})$ in Eq. (9) that we derived from system dynamics near a transition, with $y = \delta\vec{x}\cdot\hat{\nu}$, $a = \hat{\nu}^T\cdot A\cdot\hat{\nu}$, $b = B$, and $p_L(y) \propto \exp-\mathcal{F}(y)$.

19

372 Note that the above derivation assumes that the bistable dimension $\hat{\nu}$ is also an eigenvector of the

373 inverse noise tensor $(\sigma\sigma^T)^{-1}$. This assumption holds in our simulation tests, but would not hold for

374 aribitrary dynamics, in which case the noise can affect the direction of the lowest-order bistability

375 displayed by the equilibrium distribution. In this case, we expect that the equilibrium distribution

376 near the transition will still have the form of Eq. (9) (so that our fitting procedure remains unchanged),

377 but with a modified $\hat{\nu}$ that is no longer simply an eigenvector of Λ, and with an effective Γ that contains

378 information about the shape of F_T along that dimension.

Fitting and model selection

380 To fit the Landau distribution to data, we rewrite it in terms of more easily managed parameters. We

381 first find the Gaussian approximation corresponding to PCA:

$$p_G(\vec{x}) \propto \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T \cdot J \cdot (\vec{x}-\vec{\mu})\right],\qquad(13)$$

382 with J equal to the inverse of the data's covariance matrix Σ. We then express $p_L(\vec{x})$ in terms of

383 parameters c and d that set the shape along the bistable dimension $\hat{\nu}$, which we assume to be the first

384 principal component of the data (the dominant eigenvector of Σ):

$$p_L(\vec{x}) = Z^{-1}\exp\left[\begin{array}{l} -\frac{1}{2}(\vec{x}-\vec{\mu})^T \cdot J \cdot (\vec{x}-\vec{\mu}) \\[2mm] -\frac{c-1}{2}J_\nu((\vec{x}-\vec{\mu})\cdot\hat{\nu})^2 \\[2mm] -\frac{d}{4}J_\nu^2((\vec{x}-\vec{\mu})\cdot\hat{\nu})^4 \end{array}\right],\qquad(14)$$

385 where $J_\nu = \hat{\nu}^T \cdot J \cdot \hat{\nu}$ and the normalization factor Z is given by

$$Z = \sqrt{\frac{(2\pi)^{N-1}}{\det J}\frac{|c|}{2d}}\exp\left(\frac{c^2}{8d}\right)\mathcal{B},\qquad(15)$$

386 with $\mathcal{B} = K_{1/4}(\frac{c^2}{8d})$ when $c > 0$, $\mathcal{B} = \frac{\pi}{\sqrt{2}}[I_{-1/4}(\frac{c^2}{8d}) + I_{1/4}(\frac{c^2}{8d})]$ when $c < 0$, and I_α and K_α are the

387 modified Bessel functions of the first and second kind, respectively. Note that c and d are related to A

388 and B in Eq. (9) (specifically $A = J + (c-1)J_\nu\hat{\nu}\hat{\nu}^T$ and $B = dJ_\nu^2$), the original Gaussian model p_G

389 corresponds to $c = 1$ and $d = 0$, and bistability corresponds to $c < 0$ and $d > 0$.

390 In the case in which $N_{\text{samples}} < N_{\text{genes}}$, the covariance matrix Σ does not have full rank, and the

391 above equations can be interpreted by remaining in the subspace with variance. This does not affect

392 model selection because we only look for bistability along the single dimension with largest variance

393 (the principal component).

394 We use the Bayesian Information Criterion (BIC) to determine whether to select the null distribution

395 p_G (Gaussian along all components) or the Landau distribution p_L (bimodal along the principal component).

396 Note that if the number of degrees of freedom for the simple Gaussian is n_G, the Landau distribution

397 has $n_L = n_G + 1$ (d is the extra parameter; c and μ_ν do not add any additional freedom because these

398 correspond to degrees of freedom that are also in the Gaussian model). The difference in BIC values

399 between the two models,

$$\Delta\text{BIC} = -(n_L - n_G)\log N_{\text{samples}} + 2 \sum_{i=1}^{N_{\text{samples}}} \left(\log p_L(\vec{x}_i) - \log p_G(\vec{x}_i) \right), \qquad (16)$$

400 produces our criterion for model selection (see Figure 2). We choose a threshold $\Delta\text{BIC} > 6$ to

401 indicate strong evidence in favor of a bistable distribution [51].

402 The maximum likelihood mean μ will be the same as the sample mean of the data in all dimensions

403 except for the bistable dimension, so we vary three parameters in numerically fitting the Landau

404 distribution: c, d, and $\mu_\nu = \mu \cdot \hat{\nu}$. We constrain d to have a minimum value of 10^{-3} to avoid numerical

405 issues near $d = 0$.

406 For the two cases in which we find strong evidence for bistability, we find the following best-fit parameter

407 values: On day 10, $c = -4.75$ and $d = 3.19$, and on day 15, $c = -6.09$ and $d = 4.95$. (Note that these

408 values were computed with respect to data in which we took the natural logarithm of the original

409 expression values.)

21

Interpreting results with respect to individual genes

In Fig. 3, we interpret the bistable dimension $\hat{\nu}$ (principal component dimension) in terms of individual genes by highlighting those genes whose variance lies mostly along $\hat{\nu}$. We accomplish this using the measure

$$s_i = \frac{\hat{\nu}_i^2}{J_\nu \Sigma_{ii}}, \tag{17}$$

a number between 0 and 1 that indicates the proportion of gene i's variance that lies along $\hat{\nu}$. Genes with largest s are highlighted in Table 1 and SI Table 2.

The distributions for individual gene expression shown in Fig. 3 are marginals over the distribution in Eq. 14:

$$
\begin{aligned}
p(x_i) &= \int p_L(\vec{x}) \prod_{j \neq i} dx_j \\
&\propto \int N(x_i, \mu_i + \eta(\hat{\nu} \cdot \hat{i}), \Sigma'_{ii}) \exp\left[-\frac{c}{2} J_\nu \eta^2 - \frac{d}{4} J_\nu^2 \eta^4\right] d\eta,
\end{aligned}
\tag{18}
$$

where Σ' is the covariance matrix Σ modified to remove variance along the bistable dimension $\hat{\nu}$, $N(x, \mu, \sigma^2)$ indicates a normal distribution over random variable x with mean μ and variance σ^2, $J_\nu = \hat{\nu}^T \cdot J \cdot \hat{\nu}$, and we evaluate the final integral over η numerically using Mathematica.

Simulation data and tests

We create test data using a discrete time simulation of a simple model of gene regulation:

$$\frac{dx_i}{dt} = -x_i + \alpha \sum_j w_{ij} \tanh(x_j) + \xi, \tag{19}$$

here assuming the simplest all-to-all network structure for simplicity: $w_{ij} = 1$ for all i and j. We simulate the dynamics until equilibrium using a simple Euler timestep of $\Delta t = 10^{-3}$. We ensure equilibrium by simulating to time $t_f = 100$: As in Ref. [52], we expect the maximal relaxation timescale (here with τ and σ effectively 1 and the participation ratio $p = N_{\text{genes}}$) $t_{\max} \approx \sqrt{N_{\text{genes}}}$, so with $N_{\text{genes}} = 91$ our chosen $t_f > 10\, t_{\max}$. In the limit of small noise, the model has a transition to

22

428 bistability at $\alpha_{\text{crit}} = 1/N_{\text{genes}}$.

429 We then test our method's ability to detect the resulting bistability (Fig. 1). At each value of α, we

430 run 100 independent trial simulations with $N_{\text{genes}} = 91$ and run the fitting analysis on the final states

431 from each case. When the difference in BIC exceeds a threshold (here set at $\Delta\text{BIC} > 6$ [51]), we

432 count this as a positive identification of a bistable distribution. In the case analogous to ours, with

433 $N_{\text{samples}} = 16$ (orange points), we see that misidentifications are rare far below the transition,

434 and identifications become easy sufficiently above the transition. As expected, with more samples

435 ($N_{\text{samples}} = 100$, red points), it is possible to identify the incipient transition sooner, when the two

436 modes of the distribution remain closer to one another.

<table>
<tr><td colspan="2" align="center">Day 10</td><td colspan="2" align="center">Day 15</td></tr>
<tr><td>gene</td><td>s</td><td>gene</td><td>s</td></tr>
<tr><td>hex 110</td><td>0.93</td><td>vg</td><td>0.95</td></tr>
<tr><td>ilp1</td><td>0.83</td><td>hex 110</td><td>0.89</td></tr>
<tr><td>MRJP-3</td><td>0.79</td><td>P110</td><td>0.80</td></tr>
<tr><td>Hex70a</td><td>0.78</td><td>transferrin 1</td><td>0.80</td></tr>
<tr><td>Def2</td><td>0.75</td><td>LOC409966</td><td>0.77</td></tr>
<tr><td>Malvolio</td><td>0.74</td><td>Hex70a</td><td>0.77</td></tr>
<tr><td>vg</td><td>0.74</td><td>JHE</td><td>0.75</td></tr>
<tr><td>P110</td><td>0.72</td><td>VG protein</td><td>0.69</td></tr>
<tr><td>LOC409966</td><td>0.71</td><td>ilp1</td><td>0.69</td></tr>
<tr><td>PRM1</td><td>0.71</td><td>Def2</td><td>0.68</td></tr>
<tr><td>SVP NR2F1</td><td>0.69</td><td>PRM1</td><td>0.65</td></tr>
<tr><td>AGO2</td><td>0.61</td><td>TOR</td><td>0.64</td></tr>
<tr><td>Hymenoptaecin</td><td>0.59</td><td>Hymenoptaecin</td><td>0.62</td></tr>
<tr><td>SmG</td><td>0.58</td><td>Malvolio</td><td>0.59</td></tr>
<tr><td>TOR</td><td>0.54</td><td>TYR1</td><td>0.58</td></tr>
<tr><td>OA1 or OAR</td><td>0.54</td><td>E74</td><td>0.45</td></tr>
<tr><td>VG protein</td><td>0.53</td><td>Kr-h1</td><td>0.44</td></tr>
<tr><td>USP (RXR)</td><td>0.53</td><td>cad</td><td>0.42</td></tr>
<tr><td>AKHR</td><td>0.53</td><td>InR1</td><td>0.39</td></tr>
<tr><td>Kr-h1</td><td>0.52</td><td>AGO2</td><td>0.35</td></tr>
</table>

Table 2: **Genes whose variance in expression is most aligned with the transition dimension.** Here we list the top 20 genes, at age 10 and 15 days, ordered according to s, the fraction of the gene's variance that lies along the bistable dimension (see Methods Eq. (17)).

References

[1] Iain D Couzin. Collective cognition in animal groups. *Trends in cognitive sciences*, 13(1):36–43, jan 2009.

[2] Deborah M. Gordon. From division of labor to the collective behavior of social insects. *Behavioral Ecology and Sociobiology*, 70(7):1101–1108, 2016.

[3] Bryan C Daniels, Christopher J Ellison, David C Krakauer, and Jessica C Flack. Quantifying collectivity. *Current Opinion in Neurobiology*, 37:106–113, 2016.

[4] Bryan C. Daniels. Inferring the Logic of Collective Information Processors. In Min Chen, J. Michael Dunn, Amos Golan, and Aman Ullah, editors, *Advances in Info-Metrics: Information and Information Processing across Disciplines*. Oxford University Press, 2020.

[5] Gro Vang Amdam and Stig W. Omholt. The hive bee to forager transition in honeybee colonies: The double repressor hypothesis. *Journal of Theoretical Biology*, 223(4):451–464, 2003.

[6] Benjamin Wild, David M. Dormagen, Adrian Zachariae, Michael L. Smith, Kirsten S. Traynor, Dirk Brockmann, Iain D. Couzin, and Tim Landgraf. Social networks predict the life and death of honey bees. *Nature Communications*, 12(1):1–12, 2021.

[7] Gene E. Robinson. Regulation of division of labor in insect societies. *Annual Review of Entomology*, 37(1):637–665, 1992.

[8] Tanya Pankiw, Robert E. Page, and M. Kim Fondrk. Brood pheromone stimulates pollen foraging in honey bees (Apis mellifera). *Behavioral Ecology and Sociobiology*, 44(3):193–198, 1998.

[9] Robert E. Page and Sandra D. Mitchell. Self-organization and the evolution of division of labor. *Apidologie*, 29(1-2):171–190, 1998.

[10] Jennifer M. Tsuruda and Robert E. Page. The effects of young brood on the foraging behavior of two strains of honey bees (Apis mellifera). *Behavioral Ecology and Sociobiology*, 64(2):161–167, 2009.

[11] Kirsten S. Traynor, Yves Le Conte, and Robert E. Page. Age matters: Pheromone profiles of

462 larvae differentially influence foraging behaviour in the honeybee, Apis mellifera. *Animal*
463 *Behaviour*, 99:1–8, 2015.

464 [12] Kirsten S. Traynor, Ying Wang, Colin S. Brent, Gro V. Amdam, and Robert E. Page. Young and
465 old honeybee (Apis mellifera) larvae differentially prime the developmental maturation of their
466 caregivers. *Animal Behaviour*, 124:193–202, 2017.

467 [13] Ying Wang, Gro V Amdam, Bryan C Daniels, and Robert E Page. Tyramine and its receptor
468 TYR1 linked behavior QTL to reproductive physiology in honey bee workers (Apis mellifera).
469 *Journal of Insect Physiology*, 126(August):104093, 2020.

470 [14] Robert E. Page. Societies to genes: Can we get there from here? *Genetics*, 219(1), 2021.

471 [15] C. Mindy Nelson, Kate E. Ihle, M. Kim Fondrk, Robert E. Page, and Gro V. Amdam. The
472 gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biology*,
473 5(3):0673–0677, 2007.

474 [16] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford
475 University Press, New York, 1993.

476 [17] Mitra Mojtahedi, Alexander Skupin, Joseph Zhou, Ivan G Castaño, Rebecca Y Y Leong-Quong,
477 Hannah Chang, Kalliopi Trachana, Alessandro Giuliani, and Sui Huang. Cell Fate Decision as
478 High-Dimensional Critical State Transition. *PLoS Biology*, 14(12):e2000640, 2016.

479 [18] Rui Liu, Meiyi Li, Zhi Ping Liu, Jiarui Wu, Luonan Chen, and Kazuyuki Aihara. Identifying
480 critical transitions and their leading biomolecular networks in complex diseases. *Scientific*
481 *Reports*, 2:1–9, 2012.

482 [19] Eugenio Marco, Robert L. Karp, Guoji Guo, Paul Robson, Adam H. Hart, Lorenzo Trippa, and
483 Guo-Cheng Yuan. Bifurcation analysis of single-cell gene expression data reveals epigenetic
484 landscape. *Proceedings of the National Academy of Sciences*, 111(52), December 2014.

485 [20] Sonya J. Ridden, Hannah H. Chang, Konstantinos C. Zygalakis, and Ben D. Macarthur. Entropy,
486 Ergodicity, and Stem Cell Multipotency. *Physical Review Letters*, 115(20):1–5, 2015.

[21] Wen Xiong and James E. Ferrell. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature*, 426:460–465, 2003.

[22] Amir Erez, Tommy A. Byrd, Robert M. Vogel, Grégoire Altan-Bonnet, and Andrew Mugler. Universality of biochemical feedback and its application to immune cells. *Physical Review E*, 99(2), 2019.

[23] Thierry Mora and William Bialek. Are Biological Systems Poised at Criticality? *Journal of Statistical Physics*, 144:268–302, jun 2011.

[24] Miguel A. Muñoz. Colloquium: Criticality and dynamical scaling in living systems. *Reviews of Modern Physics*, 90:031001, 2018.

[25] Bryan C. Daniels, David C. Krakauer, and Jessica C. Flack. Control of finite critical behaviour in a small-scale social system. *Nature Communications*, 8:14301, 2017.

[26] Bryan C Daniels, Hyunju Kim, Douglas Moore, Siyu Zhou, Harrison Smith, Bradley Karas, Stuart A Kauffman, and Sara I Walker. Criticality distinguishes the ensemble of biological regulatory networks. *Physical Review Letters*, 121(13):138102, 2018.

[27] Robert Kucharski and Ryszard Maleszka. Evaluation of differential gene expression during behavioral development in the honeybee using microarrays and northern blots. *Genome biology*, 3(2):1–9, 2002.

[28] Charles W. Whitfield, Anne Marie Cziko, and Gene E. Robinson. Gene expression profiles in the brain predict behavior in individual honey bees. *Science*, 302(5643):296–299, 2003.

[29] Greg J. Hunt, Gro V. Amdam, David Schlipalius, Christine Emore, Nagesh Sardesai, Christie E. Williams, Olav Rueppell, Ernesto Guzmán-Novoa, Miguel Arechavaleta-Velasco, Sathees Chandra, M. Kim Fondrk, Martin Beye, and Robert E. Page. Behavioral genomics of honeybee foraging and nest defense. *Naturwissenschaften*, 94(4):247–267, April 2007.

[30] Eva Marit Hystad, Heli Salmela, Gro Vang Amdam, and Daniel Münch. Hemocyte-mediated phagocytosis differs between honey bee (Apis mellifera) worker castes. *PLOS ONE*, 12(9):e0184108, September 2017.

[31] H. Michael G. Lattorff and Robin F.A. Moritz. Genetic underpinnings of division of labor in the honeybee (Apis mellifera). *Trends in Genetics*, 29(11):641–648, 2013.

[32] Gro V. Amdam, Karl Norberg, M. Kim Fondrk, and Robert E. Page. Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31):11350–11355, 2004.

[33] Heli Salmela, Gyan P. Harwood, Daniel Münch, Christine G. Elsik, Elías Herrero-Galán, Maria K. Vartiainen, and Gro V. Amdam. Nuclear translocation of vitellogenin in the honey bee (Apis mellifera). *Apidologie*, 53(1):13, March 2022.

[34] Joseph P. Sullivan, Omar Jassim, Susan E. Fahrbach, and Gene E. Robinson. Juvenile hormone paces behavioral development in the adult worker honey bee. *Hormones and Behavior*, 37(1):1–14, 2000.

[35] Zhi Yong Huang and Gene E. Robinson. Honeybee colony integration: Worker-worker interactions mediate hormonally regulated plasticity in division of labor. *Proceedings of the National Academy of Sciences of the United States of America*, 89(24):11726–11729, 1992.

[36] Clint J. Perry, Eirik Søvik, Mary R. Myerscough, and Andrew B. Barron. Rapid behavioral maturation accelerates failure of stressed honey bee colonies. *Proceedings of the National Academy of Sciences*, 112(11):3427–3432, March 2015.

[37] Navdeep S. Mutti, Adam G. Dolezal, Florian Wolschin, Jasdeep S. Mutti, Kulvinder S. Gill, and Gro V. Amdam. Irs and tor nutrient-signaling pathways act via juvenile hormone to influence honey bee caste fate. *Journal of Experimental Biology*, 214:3977–3984, 12 2011.

[38] Avani Patel, M. Kim Fondrk, Osman Kaftanoglu, Christine Emore, Greg Hunt, Katy Frederick, and Gro V. Amdam. The Making of a Queen: TOR Pathway Is a Key Player in Diphenic Caste Development. *PLoS ONE*, 2(6):e509, June 2007.

[39] Nicholas W. Calderone and Robert E. Page Jr. Temporal polyethism and behavioural canalization in the honey bee, Apis mellifera. *Animal Behaviour*, 51(3):631–643, March 1996.

[40] Rhishikesh Bargaje, Kalliopi Trachana, Martin N. Shelton, Christopher S. McGinnis, Joseph X. Zhou, Cora Chadick, Savannah Cook, Christopher Cavanaugh, Sui Huang, and Leroy Hood. Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. *Proceedings of the National Academy of Sciences*, 114(9):2271–2276, February 2017.

[41] Yufeng Liu, David Neil Hayes, Andrew Nobel, and J. S. Marron. Statistical significance of clustering for high-dimension, low-sample size data. *Journal of the American Statistical Association*, 103(483):1281–1293, 2008.

[42] Sui Huang, Yan-Ping Guo, Gillian May, and Tariq Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental Biology*, 305(2):695–713, May 2007.

[43] Jin Wang, Li Xu, Erkang Wang, and Sui Huang. The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophysical Journal*, 99(1):29–39, 2010.

[44] Gioele La Manno, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, Maria E. Kastriti, Peter Lönnerberg, Alessandro Furlan, Jean Fan, Lars E. Borm, Zehua Liu, David van Bruggen, Jimin Guo, Xiaoling He, Roger Barker, Erik Sundström, Gonçalo Castelo-Branco, Patrick Cramer, Igor Adameyko, Sten Linnarsson, and Peter V. Kharchenko. RNA velocity of single cells. *Nature*, 560(7719):494–498, August 2018.

[45] Bryan C. Daniels, Jessica C. Flack, and David C. Krakauer. Dual coding theory explains biphasic collective computation in neural decision-making. *Frontiers in Neuroscience*, 11:1–16, 2017.

[46] L. Perko. *Differential Equations and Dynamical Systems*. Texts in Applied Mathematics. Springer New York, 2013.

[47] Yuri A. Kuznetsov. *Elements of Applied Bifurcation Theory*. Springer-Verlag, New York, 1995.

[48] Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press, 2015.

[49] L. Landau. On the theory of phase transitions. *Ukr. J. Phys*, 53:25, 2008.

[50] Michael Plischke and Birger Bergersen. *Equilibrium Statistical Physics*. World Scientific, 2006.

[51] Adriatn E Raftery. Bayesian Model Selection in Social Research. *Sociological Methodology*, 25:111–163, 1995.

[52] Bryan C. Daniels and Pawel Romanczuk. Quantifying the impact of network structure on speed and accuracy in collective decision-making. *Theory in Biosciences*, 140:379–390, 2021.