

Primer

Stochasticity,
individuality and
behaviorKyle Honegger and Benjamin de
Bivort

No two individuals are exactly alike. More than a simple platitude, this observation reflects the fundamentally stochastic nature of biological systems. The term ‘stochastic’ describes features that cannot be predicted *a priori* from readily measurable variables. In the dichotomous framework in which biological variation arises from genetic or environmental effects, stochastic effects are classified as environmental because they are not passed on to offspring — any non-heritable cause is, by definition, environmental. But non-heritable effects can be subdivided into those which can be predicted from measurable variables, and those that cannot. These latter effects are stochastic.

The existence of unpredictable stochastic effects on biological phenotypes is, oddly enough, guaranteed by deterministic physical laws. The vast numbers of nonlinear inter-molecular interactions involved in cellular function, coupled with thermodynamic instability, make it impossible that the chain of causal events driving organismal development and function will proceed in completely identical ways across organisms. In practice, the outcomes of stochastic events are exceedingly common and feasible to measure at the organismal level, even if their provenance at the molecular level is hidden.

The consequences of stochasticity for an organism are vast and span levels of biological organization from gene expression in single cells through complex patterns of behavior. There are striking stories of monozygotic (colloquially ‘identical’) twins, separated at birth, who share remarkably specific behavioral traits and quirks as adults, despite having been raised in different environments. These individuals share nothing but their genotype and *in utero* environment. To many, the co-occurrence of such unlikely

complex behavioral phenotypes presents a *prima facie* case for genetic determinism — the traits are fully predicted from genotype alone. But despite surprising idiosyncrasies rooted in genetics, there are invariably many other traits that differentiate identical twins, even if they are raised in the same environment. This is evident in discrepancies between twins in personality traits or disease outcomes. In such scenarios, both genotype and environment have been essentially matched, therefore these differences may be the result of stochastic influences. Simply put, even identical twins reared in the same environment will inevitably vary.

Before diving into the discussion of stochastic individuality, we must define these terms in the context of animal behavior. We use the term ‘individuality’ to mean behavioral characteristics that vary among conspecifics and persist on the timescale of a lifetime. This excludes characteristics that are shared by all individuals being compared, as well as characteristics that vary on short timescales but fill out the same distribution on long timescales (which are called ‘ergodic’ in the physical sciences). While some features of individuality have a genetic component, others are largely stochastic and, thus, non-heritable. Stochastic individuality, therefore, leads to non-heritable, inter-organism behavioral variation, observable as contrasts between pairs of individuals, or, at the level of populations, as variance in a measure of behavior. In practice, this means that even with a complete understanding of the basis of a trait and a thorough catalogue of all environmental factors an individual has encountered, its behaviors will remain beyond reliable prediction.

Intangible variance

The topic of inter-individual phenotypic variance has long occupied geneticists. Early efforts to place principles of population genetics within a quantitative framework focused on partitioning observed phenotypic variance into constituent parts. Because phenotypic variance was originally appreciated for its necessary role as a substrate for natural selection, understanding how heritable genetic

variation responds to selective pressures was an important piece for understanding evolution.

The simplest quantitative genetic formulation of phenotypic variance is as the sum of genetic and environmental variances: $V_{\text{phenotype}} = V_{\text{genotype}} + V_{\text{environment}}$. Of these terms, $V_{\text{phenotype}}$ and V_{genotype} can be measured directly, for example using a behavioral assay and pedigree analysis, respectively. In contrast, $V_{\text{environment}}$ is not measured, but defined as the difference in the measured quantities. Less rudimentary frameworks add terms for other factors that can be measured, for example: $V_{\text{phenotype}} = V_{\text{genetic}} + V_{\text{E}} + V_{\text{GxE}} + V_{\text{e}}$. Here, V_{E} is ‘general environmental variance’ and captures deterministic responses to environmental differences shared among individuals, and V_{GxE} accounts for measurable interaction effects between genotype and general environment. But V_{e} , ‘special environmental variance’, like the original $V_{\text{environment}}$, is defined as the difference between $V_{\text{phenotype}}$ and all the measurable terms.

Whichever variance term accounts for unmeasurable effects can be considered the stochastic component, V_{s} . Several studies have found that a substantial amount of phenotypic variance remains when holding genotype and environment as constant as possible. These studies have used highly-inbred, isogenic, monozygotic or clonal organisms to minimize genetic variability, and employed highly standardized animal husbandry practices to minimize variation in the environment. In a now classic result (Gärtner, 1990), 30 years of inbreeding experiments on laboratory mice and rats in shared environments eliminated only 20–30% of observed variance in a number of phenotypes. The remaining 70–80% was referred to as the ‘intangible variance’. As a number of studies in other organisms have also shown, there seems to be a lower limit to phenotypic variance that often exceeds what may be accounted for by measurable genetic and environmental factors.

As a method of reducing genetic variation, inbreeding sexual organisms has caveats, in particular the fixation of deleterious recessive alleles. A striking display of stochastic individuality,



Figure 1. Genetically identical individuals reared in the same environment have different behaviors.

Here, a parthenogenetic pea aphid (*Acyrtosiphon pisum*) produces clonal daughters who show individual variation in startle responses. When startled by identical stimuli, some daughters cling to their perch on vegetation, while other genetically identical sisters leap off.

without such concerns, may be found in the behavior of the parthenogenetic marbled crayfish. In this species, all individuals are female and all siblings are clonal (this mode of reproduction is termed apomictic thelytoky). But despite being genetically identical, marbled crayfish siblings raised in the same environment display dramatic variation in social, reproductive and locomotor behaviors. Genetically identical siblings develop different preferred resting positions, display varying degrees of gregariousness, and establish pronounced, persistent social hierarchies with dominant and subordinate individuals. This underscores stochasticity's impact in producing ecologically relevant behavioral differences within a population.

Other species reveal similar stochastic individuality. Genetically identical pea aphids, which also reproduce by apomictic thelytoky exhibit conspicuous variability in their escape responses (Figure 1). When a large object looms towards aphids feeding on vegetation, some individuals jump away quickly, while others remain feeding. Even though this variation is not heritable, it contributes to the fitness of each aphid. Beyond the susceptibility to predation, individual predisposition to escape behaviors correlates with life-history strategies. Fearful aphids that jump early and often are more likely to have a long period of fertility, birthing new daughters over

many days. For these slow-and-steady breeders, jumping at the slightest threat of a predator lets them live and reproduce another day. Their fast-living siblings that reproduce predominantly early in life have less total reproductive productivity to lose if they are eaten by a predator after having achieved that early productivity.

Stochastic individuality is not specific to invertebrates, though they offer practical advantages for measuring it, including low cost and high cost fecundity. Individual mice from an inbred lab strain, reared in identical cages, vary in their propensity to explore their environment. Those mice that explore the most have more prolific neurogenesis in their hippocampi, where neural activity correlates with navigational cues. The causal relationship between exploration and neurogenesis is not known, but these findings established a link between neurobiological mechanisms and stochastic individuality. It is likely that stochastic individuality is a general phenomenon affecting essentially all behaviors in all species.

The quantitative genetic framework facilitates the quantification of the effects of stochasticity, but it does not illuminate its mechanistic underpinnings. What factors may contribute to this intangible variance (V_g)? Though Gärtner ascribed this third component to unobserved ooplasmic factors — varying molecular constituents of the cytoplasm of the egg which cause predictable phenotypes in the organisms that develop from these eggs — much, if not most, of that variance likely arises from stochasticity at the molecular and cellular level.

Stochasticity from molecules to the brain

How does one study the mechanisms of a phenomenon that, by definition, arises in unobservable variables? Despite this seeming intangibility, modern molecular and systems biology approaches have begun to reveal the mechanisms by which stochastic processes may influence the function and developmental trajectories of individual organisms. Experimentally, the ability to measure dynamics of single cells and single molecules has made it apparent that genetically

identical cells in the same environment still display a high degree of variability in gene expression and other cellular phenotypes.

Mathematical modeling highlights a number of circumstances in which physical processes will produce stochasticity. First, when the number of discrete components involved in a molecular process is small, fluctuations will be relatively large (Figure 2). For example, there may only be a few copies of a transcription factor present in the nucleus. At a given gene promoter site, these proteins will bind and initiate transcription only occasionally and with unpredictable timing. The result is variability in the amount of messenger-RNA present in the cell over time. The immune system exploits the inherent stochasticity of small numbers in the 'VDJ recombination process' which gives rise to T-cell receptor diversity. Here, dedicated enzymes induce genomic rearrangements to concatenate three gene segments, one from each of three sets of segments. Each of these recombination events happens exactly once, maximizing the small number effect, and precluding any averaging that would smooth out the fluctuations in these molecular events.

Second, positive feedback in gene networks (and dynamical systems in general) can amplify fluctuations, leading cells to jump between discrete states, a phenomenon called bistability. Third, because biological processes are multidimensional, nonlinear, and rife with feedback, their dynamics are often chaotic. Math tells us that, in nonlinear dynamical systems, small differences in initial conditions, such as the small number sampling effects described above or distribution of molecular configurations caused by thermal fluctuations, will be amplified into large, even qualitative, differences at high levels of biological organization. At present, observing this multi-level causal cascade remains an experimental aspiration.

Cellular differences in gene expression will influence numerous other cellular-level phenomena, including post-transcriptional regulation. The axon guidance gene *DSCAM* has over 38 thousand splice-variants in *Drosophila*. Each neuron expresses a stochastic subset of

these variants that is distinct from its neighbors' subsets, and this provides a basis by which their connections to other neurons will differ. Stochasticity in gene expression and post-transcriptional regulation will alter patterns of chromatin modification — this mechanism is now commonly referred to as 'epigenetics'. (Confusingly for the study of stochastic individuality, the original meaning of this term, as coined by Conrad Waddington, referred to the developmental dynamics that produce traits and determine their variability.) Chromatin modification will in turn affect transcriptional and post-transcriptional dynamics, in an example of a feedback loop that can amplify and lock in stochastic outcomes at the cellular level, which are rooted at the molecular level. Thus, it may not be surprising that up to a quarter of all genes are differentially expressed among genetically identical individual fruit flies reared in nominally identical laboratory conditions.

Another potential source of stochastic individuality is somatic mosaicism. While all cells in an animal's body are ultimately derived from the single genotype present in the zygote, as cellular proliferation and differentiation proceeds, many opportunities arise for genetic alteration to individual cells and their descendants. One of these opportunities is the activation of mobile DNA elements (transposons), that excise and re-insert themselves in new positions in a cell's genome. Insertion sites are not all equally probable, but the destination of any single hopping transposon is a stochastic outcome. While the notion that somatic mosaicism is specifically rampant in neural tissue has come into question, it is likely that the brain, like most if not all tissues, is subject to significant genomic rearrangements across development, and that these rearrangements can contribute to diversity in neural physiology and even disease.

Neural tissue, which generates behavior through its interconnected wiring, may be particularly susceptible to stochastic effects. The cellular-level developmental fate of a neuron will determine where it sends its projections, and thus its role in the

circuits that coordinate behavior. This wiring occurs through the guidance, via molecular cues, of axonal growth cones, small pouches of cytoplasm where small-number effects in the counts of signalling receptors, cell adhesion molecules, ion channels, and local transcriptional events can have a large impact. Similarly, the variations of lower-level stochastic mechanisms, like the composition of *DSCAM* splice variants, will be magnified in the small-number regime of the growth cone.

Once a brain is wired, the effects of stochasticity on individuality are not finished. The brain itself appears to be a regulator of individuality, as neuromodulatory tone and the physiological state of specific neural circuit elements alters the degree of variability exhibited by a collection of individual animals. For example, silencing a small subset of neurons in the fruit fly central complex, a structure involved in sensory integration, navigation, and pre-motor coordination, alters the inter-individual variability of locomotor behavior without altering the mean. The brain may have an active role as a regulator and generator of stochasticity. This raises the possibility that stochasticity can be adaptively tuned on the timescale of behavioral decision-making — not just on longer developmental or evolutionary timescales — when the brain detects changes in the environment, or the organism's internal states, via its sensory inputs.

Stochastic individuality: developmental bug or adaptive feature?

Darwin and Wallace established the importance of heritable phenotypic variation as a major component of evolutionary change. The role, if any, of non-heritable variation in evolution has remained less clear. One possibility is that its only consequence is to muddy the correspondence between genotype and phenotype, thereby masking an individual's 'true' phenotype and reducing the efficiency of evolution by natural selection. In this view, stochastic individuality is a bug stemming from the physical impossibility of flawless self-construction.

The other possibility is that stochastic individuality is actually



Figure 2. Small number effects promote stochasticity at the molecular level.

Molecules that exist in low copy numbers, like gene transcriptional start sites, will exhibit relatively large, random fluctuations in activity over time. This is illustrated in the binding (or non-binding) of RNA polymerase (red) to DNA (blue). Alternative splicing of transcripts (olive) by splicing complexes (orange) is another molecular event subject to random fluctuations. As gene products are nodes in complex, nonlinear regulatory networks rife with feedback, these small fluctuations can be amplified to lock in stochastic effects at the cellular level.

an adaptive feature which provides an evolutionary benefit compared to lower variability. This possibility is supported, albeit circumstantially, by the observation that the endogenous role of some genes (or neural circuit elements) seems to be to promote stochasticity. When these genes (or neural circuit elements) are mutated (or silenced transgenically) behavioral variability goes down. Evolution appears not to have purged these mechanisms that increase stochastic individuality.

Beyond this empirical evidence, there are ample theoretical arguments for the adaptive value of stochastic individuality. In the 'gene-saving' hypothesis, derived from studies of algorithmic genetic strategies, non-genetic variation may facilitate natural selection by supplementing genetic variation, thereby reducing the number of causal genes needed to create sufficient phenotypic variance. This will consequently reduce the number of generations before the first individual reaches an evolutionary optimum. To the extent that an organism's ecological and evolutionary interactions can be modeled as moves in a game theoretic framework, many optimal and evolutionarily stable strategies will be 'mixed'. This means that moves — behavioral interactions with

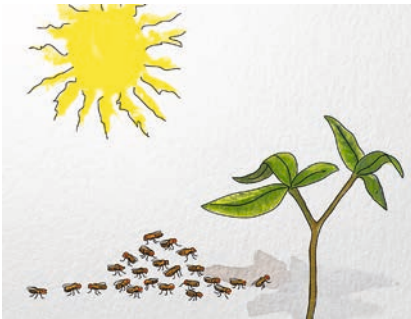


Figure 3. Thermal preference variability in fruit flies may reflect a bet-hedging strategy.

Individual fruit flies exhibit idiosyncratic thermal and light preference behavior (illustrated by the Gaussian pile of flies), with some flies preferring warmth or coolness, light or shade. Shade-seeking flies will have a fitness advantage in the summer or in heat-waves. A broad distribution of behaviors can increase the chance that some individuals will be well matched to unpredictable environmental fluctuations.

conspecifics and heterospecifics — are optimally chosen at random from the list of possible moves. For example, if an organism's interactions reduce to a game of rock–paper–scissors, the only strategy it can employ, which cannot possibly be exploited by an adversary, is to play rock, paper or scissors at random every turn. This will guarantee a win half the time. Mixed strategies can play out at the turn-to-turn level (behavior-to-behavior) and/or the individual-to-individual level (stochastic individuality).

'Diversified bet-hedging' is an evolutionary strategy in which a single genotype produces a distribution of phenotypes across offspring in order to increase the likelihood that at least some individuals are well-adapted to the selection pressures of unpredictable environments. While there is ample theoretical evidence indicating this strategy can be beneficial, experimental evidence in animals is scant. We combined experimental measurements of thermal preference variability with a mathematical model of how integrated thermal experience affects life-history, and concluded that the observed behavioral individuality might reflect bet-hedging against seasonal temperature fluctuations (Figure 3). In the early spring and late fall, warm-seeking flies have an

advantage; in the high summer, cool-seeking flies have the advantage. As an evolutionary strategy, offspring with thermal preferences drawn randomly from a broad cool-preferring to warm-preferring distribution (bet-hedging) outperformed strategies in which individual phenotypes were heritable. But this is just a test of the plausibility of a bet-hedging hypothesis, not an attempt to experimentally falsify it. The overall lack of evidence for bet-hedging in animals likely reflects the practical challenges of conducting such experiments, as there is substantial experimental evidence in plants and microorganisms consistent with the bet-hedging hypothesis.

Bet-hedging can be contrasted with another phenotypic strategy that increases variability: phenotypic plasticity. In this strategy, an organism can deterministically adjust its phenotype in response to the environment. Phenotypic plasticity is a flexible enough framework that it can, in principle, encompass the optimal environment-to-phenotype solutions; for whatever environment comes along, a fully plastic organism could, in theory, morph into the perfect phenotype. But implementing such morphing rules may be too costly or complex to evolve reliably. By analogy, the perfect financial investment strategy would be to read market variables and pivot all funds into whatever financial instrument will give the greatest returns at the moment. But in reality, predicting future trends is unreliable, and transaction fees penalize rapidly switching investments. Instead, a diversified portfolio of steady composition can succeed in most circumstances. Thus, bet-hedging may represent a solution to environmental fluctuations that is readily attainable by evolution.

If the formulations of gene-saving, mixed strategies, or bet-hedging are correct, then evolution will favor some level of stochastic individuality. But stochasticity can only be tuned by evolution if, in addition to affording a selective advantage, it varies across genotypes and is heritable. Both of these additional conditions appear to be true. Across different isogenic *Drosophila* lines, the magnitude of behavioral variability (stochastic individuality) in locomotor handedness, the tendency of individuals to turn

left or turn right during spontaneous exploration, itself varies. Some lines have low variability, with individuals exhibiting small (but significant) differences in locomotor bias. Other lines have high variability, with individuals exhibiting large differences in locomotor bias. These are heritable traits of their respective lines: crossing two low variability lines produces low variability hybrids, and crossing two high variability lines produces high variability hybrids.

The genetic basis of behavioral variability permits the mapping of genetic variants controlling variability as a trait. One implicated gene is *teneurin-a*, which encodes a cell-surface protein involved in axon guidance and synapse formation, developmental processes invoked in the wiring of the neural circuits mediating locomotor behavior. Strikingly, lines which are high variability for one behavior were not high variability for other behaviors. This implies there is modular control of the level of stochasticity exhibited in each separate behavior. Thus, stochastic individuality is a flexibly evolvable trait, which can vary across behaviors, vary across genotypes, be selected for by mechanisms such as bet-hedging, and be passed on to offspring.

These are still early days in the study of the mechanistic basis of stochastic individuality. But, principles have been identified by which stochasticity can arise even under Newtonian physical rules. There are plausible paths by which these fluctuations can be amplified and made manifest at the cellular, neural circuit, and behavioral levels. And there are theoretical frameworks, with some experimental evidence, in which the observed behavioral variability imparts evolvable selective advantages on those genotypes that produce it. Exciting future directions include the enumeration of the causal relationships from the molecular to the evolutionary levels in the case of a single case study of stochastic individuality. Such integrative understanding will likely exploit new cutting-edge tools that extract rich data sets from individual animals: single-cell sequencing, whole-brain neural recordings, and connectomic reconstructions of

whole-brain circuitry. Just as genomics entered a new era when sequencing added breadth across individuals to its original depth, these technologies of the individual will enable unprecedented insights into stochastic individuality and the biological basis of behavior, when applied comparatively across individuals.

FURTHER READING

- Ayroles, J., Buchanan, S., O'Leary, C., Kakaria, K., Grenier, J., Clark, A., Hartl, D., and de Bivort, B. (2015). Behavioral individuality reveals genetic control of phenotypic variability. *Proc. Natl. Acad. Sci. USA* *112*, 6706–6711.
- Buchanan, S., Kain, J., and de Bivort, B. (2015). Neuronal control of locomotor handedness in *Drosophila*. *Proc. Natl. Acad. Sci. USA* *112*, 6700–6705.
- Dubnau, D., and Losick R. (2006). Bistability in bacteria. *Mol. Microbiol.* *61*, 564–572.
- Elowitz, M., Levine, A., Siggia, E., and Swain, P. (2002). Stochastic gene expression in a single cell. *Science* *297*, 1183–1186.
- Freund, J., Brandmaier, A., Lewejohann, L., Kirste, I., Kritzler, M., Krüger, A., Sachser, N., Lindenberg, U., and Kempermann, G. (2013). Emergence of individuality in genetically identical mice. *Science* *340*, 756–759.
- Gärtner, K. (1990). A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals? *Lab. Anim.* *24*, 71–77.
- Hopper, K. (1999). Risk-spreading and bet-hedging in insect population biology. *Annu. Rev. Entomol.* *44*, 535–560.
- Kain, J., Zhang, S., Akhund-Zade, J., Samuel, A., Klein, M., and de Bivort, B. (2015). Variability in thermal and phototactic preferences in *Drosophila* may reflect an adaptive bet-hedging strategy. *Evolution* *69*, 3171–3815.
- Lin, Y., Chen, Z., Oliver, B., and Harbison, S. (2016). Microenvironmental gene expression plasticity among individual *Drosophila melanogaster*. *G3 (Bethesda)* *6*, 4197–4210.
- Lynch, M., and Walsh, B. (1997). *Genetics and Analysis of Quantitative Traits*. (Sinauer Associates, Sunderland: Sinauer Associates).
- Moris, N., Pina, C., and Arias, A. (2016) Transition states and cell fate decisions in epigenetic landscapes. *Nat. Rev. Genet.* *17*, 693–703.
- Muotri, A., Chu, V., Marchetto, M., Deng, W., Moran, J., and Gage, F. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* *435*, 903–910.
- Neves, G., Zucker, J., Daly, M., and Chess, A. (2004). Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nat. Genet.* *36*, 240–246.
- Schuetz, W., Dall, S., Kloesener, M., Baeumer, J., Beinlich, F., and Eggers, T. (2014). Life-history trade-offs mediate 'personality' variation in two colour morphs of the pea aphid, *Acyrtosiphon pisum*. *J. Anim. Ecol.* *84*, 90–101.
- Vogt, G., Huber, M., Thiemann, M., van den Boogaart, G., Schmitz, O., and Schubart C. (2008). Production of different phenotypes from the same genotype in the same environment by developmental variation. *J. Exp. Biol.* *211*, 510–523.

Department of Organismic and Evolutionary Biology and Center for Brain Science, Harvard University, Cambridge, MA USA 02138.
E-mail: debivort@oeb.harvard.edu