Towards a global view of phenotype space

Global constraints within the developmental program of the *Drosophila* wing
Authors: Vasyl Alba, James E Carthew, Richard W Carthew, and Madhav Mani
eLife 10, e66750 (2021)

Recommended with a Commentary by David K. Lubensky,
Department of Physics, University of Michigan

Robustness is one of the most widely invoked, but also one of the most slippery, concepts in the physics of living systems. Certainly, living things are able to carry out their basic functions in the face of thermal noise and intrinsic molecular stochasticity, of environmental variability, and even—up to a point—of genetic mutations, and it seems reasonable to describe this ability as a form of robustness. Things quickly become harder, however, if one tries to translate this general impression into something more precise and useful. To articulate a widely applicable biological design principle, one needs to know something about which traits, exactly, are robust, and against which perturbations [1]. (Both common sense and theorems from control theory [2] suggest that it is impossible to make everything perfectly robust to all conceivable insults.) In those cases where the answers to these questions are known, robustness analysis has proven to be a powerful tool to infer biological mechanisms [3, 4], but it is often difficult to generalize these examples to obtain broader insights into the nature of biological robustness. The situation becomes still more complicated when one realizes that robustness is not even necessarily always desirable. For example, in a system that was completely robust to the effects of mutations, natural selection would have no phenotypic variation to act on, and evolution as we currently understand it could not occur. Thus, two apparently fundamental features of living systems, their robustness and their evolvability, would seem to be in conflict, and it is not clear how biology balances their competing demands [5].

The questions of what traits are robust and what traits show natural variation upon which selection can act can be viewed as two aspects of the larger problem of understanding an organism's genotype to phenotype (GP) map. (In some contexts, it may make more sense to talk instead of a genotype and environment to phenotype map.) It is here that the thought-provoking new paper from Alba and coworkers enters the discussion. Although the theoretical idea of a GP map is well-established, comparatively little is known experimentally about what these maps actually look like. There are both technical and conceptual reasons for this gap. For example, biological studies tend to focus on only a few traits or variables rather than looking at phenotypes in all of their multi-dimensional richness, and collecting enough data to adequately sample the GP map remains difficult in many living systems. Moreover, classical
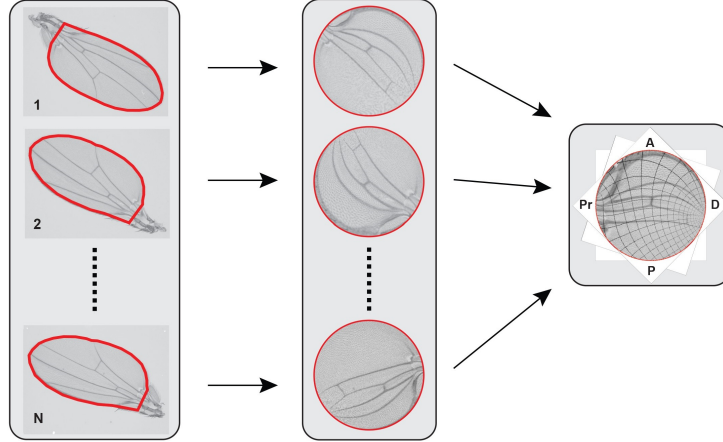
Figure 1: Schematic of wing pattern alignment using conformal maps. The outlines each individual wing blade are first detected (red curves, left), and the image of the blade boundary and interior are then conformally mapped to the unit circle (middle). Finally, a conformal transformation of the circle to itself is applied to each image to achieve maximal alignment with a reference image (right). The prominent linear objects in each image are the wing veins. A, Anterior; P, Posterior; Pr, Proximal; D, Distal. Adapted with permission from Alba *et al.*, 2021.

genetics can sometimes seem akin to taking a large hammer to an object and trying to infer its workings from the fragments that result [6]: It tells us something about the extreme, and often difficult to access, corners of the phenotype space but does not necessarily provide easy tools to explore the bulk of the space where weaker perturbations drive most natural variation. To address these difficulties, Alba *et al.* focus on the vein pattern in the wing of the fruit fly *Drosophila melanogaster*. This system has the advantages that the phenotype is simply captured by a two-dimensional image, facilitating data acquisition and processing, and that its developmental genetics is extremely well-studied. (Even so, the effort required to dissect and mount all of the wings needed for this study borders on the heroic.) In particular, previous work on the locations of a few landmarks where different veins cross has shown that their pattern is quite robust, with reproducibility on the order of a single cell diameter in highly inbred, essentially isogenic populations [7]. Rather than focusing only on the standard landmarks, however, Alba *et al.* consider the whole of the pattern and seek to work out, if not the entire GP map, then at least the image of the map, i.e. the manifold of accessible phenotypes in the high-dimensional pattern space. Strikingly, they find that this manifold is to a first approximation one-dimensional—thereby supporting a famous argument due to Waddington that the process of animal development "canalizes" genetic and environmental diversity into limited phenotypic variability [8].

Specifically, Alba *et al.* begin by crossing and backcrossing flies from different strains (with slightly different genotypes) to construct a highly outbred population containing considerable natural genetic variation. This strategy of using an outbred population, rather than the more standard inbred laboratory strains, allows the researchers to efficiently sample a substantial region of the phenotype space without having to introduce a large number
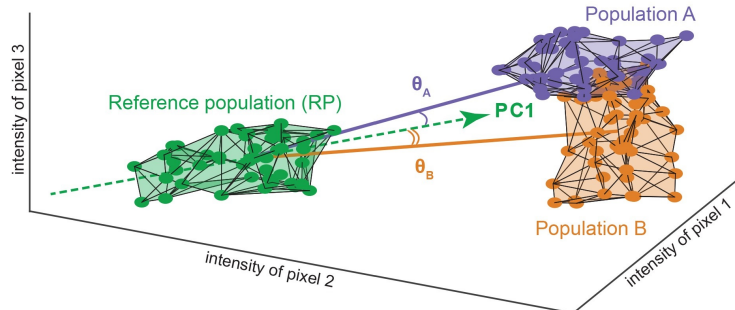
Figure 2: Schematic of a hypothetical phenotype space, projected down to three dimensions (corresponding to the intensities of 3 individual pixels in an aligned wing image). Each point represents the phenotype (i.e. the wing pattern) of one individual. A principal component analysis of the reference (wildtype, outbred) population (RP) shows that the cloud of observed phenotypes is highly elongated along the direction of the first principal component (PC1). Moreover, the vectors joining the centroid of the RP cloud to those of other populations (A and B, corresponding either to mild mutations or to different environmental conditions from the RP) are nearly parallel with PC1—that is, the angles θ_A and θ_B are close to zero. Thus, most phenotypic variation occurs along a quasi-one-dimensional manifold. Note that the real phenotype space in this study has many more than three dimensions, and both the dominance of PC1 over the other principal components and the alignment of the vectors between populations with PC1 are correspondingly more pronounced than can be represented visually here. Adapted with permission from Alba *et al.*, 2021.

of different perturbations by hand; it is essential to their ability to obtain interesting results. They then use conformal mapping to align different wing images so that their patterns may be compared pixel by pixel (Fig. 1) and look at the cloud of points that an ensemble of wings from their fly population generates in the roughly 30,000-dimensional space spanned by all of the individual pixel intensities. A principal component analysis reveals that variation along a single direction in this space (PC1, Fig. 2) accounts for more than 60% of the overall phenotypic variability. (The next largest principal component, accounting for a few percent, corresponds to sexual dimorphism between male and female flies.) More remarkable still, the average differences between the wildtype population and various perturbed populations lie along essentially the same direction as PC1. These perturbed populations were created either by raising flies in more stressful than normal conditions (high or low temperatures or nutritional restriction) or by studying animals heterozygous for mutations known to affect wing development. Importantly, in this latter case all of the alleles studied are classically defined as *recessive* to the wildtype. That is, the differences between the wing patterns of the heterozygotes and of homozygous wildtype flies are too subtle to be readily apparent to the human eye, but the mutations nonetheless introduce weak pattern deformations that can be detected by the quantitative methods introduced in the paper. Thus variation in response to a variety of genetic and environmental perturbations all collapses approximately onto a single line in phenotype space.

Like any good scientific manuscript, the work by Alba *et al.* raises as many questions as

it answers. On the highest level, it gives new articulation and precision to a time-honored hypothesis about the tendency of GP maps to canalize variation into only a few directions in phenotype space, but it of course strictly applies only to the single case of wing patterns; it will be necessary to find ways to generalize the paper's methods to other systems to test whether the findings hold more broadly. It might also be interesting to consider whether the phenotype manifold remains quasi-one-dimensional if the animal is subjected to insults—say, sustained mechanical confinement [9] or sharp temperature gradients in space [10]—that it would be less likely to encounter in the natural world. The suspicion is that, because the response to these sorts of perturbations is unlikely to have been subject to strong selection, they might tend to move the phenotype along a direction orthogonal to PC1. There is also the obvious problem of developing a mechanistic understanding of the particular version of robustness revealed by this study: Experimentally, what genes and pathways does it depend on, and theoretically, what models can reproduce a similar GP map? Fully answering these questions will likely require data not just on the final distribution of wing patterns, but also on its evolution with time during development [11]. It is interesting to note in this regard that PC1 defies easy description but appears to correspond in part to a shift of the longitudinal veins along the anterior-posterior axis. More generally, one might ask what broad classes of models have roughly one-dimensional phenotype manifolds. It is tempting here to draw a parallel with the well-documented phenomenon of "sloppiness" in models, whereby available data (e.g., the final wing pattern) poorly constrains all but a few combinations of model parameters [12]; turned around, this observation would also suggest that excursions in many directions in parameter (i.e. genotype) space have little effect on the final phenotype, so that the phenotype manifold has a much lower effective dimension than one might guess from the dimension of the parameter space. Methods that interpret model sloppiness in terms of the shape of an information geometric manifold might be especially valuable here [13]. Finally, on a more technical level, it is important to acknowledge that using conformal maps to align different wing images, while appealing in many ways–for example, by preserving angles, it also preserves many local pattern features, and moreover tissue growth is expected to be conformal under certain (fairly restrictive) hypotheses [14]–necessarily throws out some information and is not the only alignment procedure that can be imagined. It is entirely possible that other approaches might reveal that the phenotype manifold possesses additional dimensions beyond the single one found here. In short, the paper by Alba and coworkers is an important step towards the eventual goal of systematically understanding the nature of GP maps in living systems, but it is far from the last one.

References

[1] MA Felix and M Barkoulas. Pervasive robustness in biological systems. *Nat Rev Genet*, 16(8):483–496, 2015.

[2] K Zhou, JC Doyle, and K Glover. *Robust and Optimal Control.* Prentice Hall, 1996.

[3] N Barkai and S Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, 1997.

[4] A Eldar, R Dorfman, D Weiss, H Ashe, B-Z Shilo, and N Barkai. Robustness of the BMP morphogen gradient in *Drosophila* embryonic patterning. *Nature*, 419(6904):304–308, 2002.

[5] A Wagner. *Robustness and Evolvability in Living Systems*. Princeton UP, 2013.

[6] Y Lazebnik. Can a biologist fix a radio?—or, what I learned while studying apoptosis. *Cancer cell*, 2(3):179–182, 2002.

[7] L Abouchar, MD Petkova, CR Steinhardt, and T Gregor. Fly wing vein patterns have spatial reproducibility of a single cell. *J Royal Soc Interface*, 11(97):20140443, 2014.

[8] CH Waddington. Canalization of development and the inheritance of acquired characters. *Nature*, 150(3811):563–565, 1942.

[9] JM López-Gay, H Nunley, M Spencer, F di Pietro, B Guirao, F Bosveld, O Markova, I Gaugue, S Pelletier, DK Lubensky, and Y Bellaïche. Apical stress fibers enable a scaling between cell mechanical response and area in epithelial tissue. *Science*, 370(6514):eabb2169, 2020.

[10] EM Lucchetta, JH Lee, LA Fu, NH Patel, and RF Ismagilov. Dynamics of *Drosophila* embryonic patterning network perturbed in space and time using microfluidics. *Nature*, 434(7037):1134–1138, 2005.

[11] PA Haas, S Hohn, AR Honerkamp-Smith, JB Kirkegaard, and RE Goldstein. The noisy basis of morphogenesis: Mechanisms and mechanics of cell sheet folding inferred from developmental variability. *PLoS Biol*, 16(7):e2005536, 2018.

[12] RN Gutenkunst, JJ Waterfall, FP Casey, KS Brown, CR Myers, and JP Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comp Biol*, 3(10):e189, 2007.

[13] MK Transtrum, BB Machta, and JP Sethna. Geometry of nonlinear least squares with applications to sloppy models and optimization. *Phys Rev E*, 83(3):036701, 2011.

[14] K Alim, S Armon, BI Shraiman, and A Boudaoud. Leaf growth is conformal. *Phys Biol*, 13(5):05LT01, 2016.