Current Biology, Vol. 13, R958–R960, December 16, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.cub.2003.11.042

Evolutionary Genetics: The Nature of Hidden Genetic Variation Unveiled

Dispatch

Günter P. Wagner

It has long been known that wild-type phenotypes harbor considerable amounts of 'hidden' genetic variation. A new study has mapped this variation at the nucleotide level and revealed some unexpected properties.

All life on earth owes its existence to genetic accidents — mutations. Without genetic variation, no evolution would be possible and hence life as we know it would not exist. Also existing life forms would be doomed in the long run if there were not a steady stream of genetic novelties that allow species to adapt to changing environmental conditions, such as ice ages or global warming. Yet genetic variation is perhaps most intensely studied because it usually has negative effects. These effects attract our attention in the form of congenital diseases and disease susceptibilities. The study of the negative effects of genetic variation gave rise to the fields of genetic epidemiology and human medical genetics, which may now be the most intensely studied area of genetics.

So, in the face of a constant stream of deleterious genetic accidents, how does life continue? Decrease the mutation rate and run the risk of being unable to adapt to unforeseen environmental changes? Produce armies of offspring, in the hope that a few will have the right genes to survive and have children of their own? Neither of these seems to be the answer chosen by nature, at least not exclusively. Mutation rates are somewhat different between bacteria and eukaryotes, but among eukaryotes, the rates are not so different between 'lower' and 'higher' forms. More complex organisms, such as humans, tend to have fewer offspring than simpler forms, such as fungi. Hence a large number of offspring also does not seem to be the solution of choice.

A third solution to the problem of deleterious genetic variation was suggested by early genetic experiments. It was observed that a mutation with a major impact on a phenotypic character generally has two kinds of effects: there is its primary effect, changing the average appearance of the character; but there is also a secondary effect, as the mutant phenotype is also more variable than the wild type. A well studied example is the mutation *Scute*, which influences, among other things, the number of bristles on the back of the fruit fly [1]. *Drosophila melanogaster* usually has four bristles, with very rare deviations from this number. After the mutation *Scute* is introduced, this number is increased, but also the amount of variation is many orders of magnitude higher than in the wild type (Figure 1).

Surprisingly, the variation of the *Scute* mutant phenotype turned out to be partly genetic. This is shown, for instance, by the way artificial selection leads to a strong response in the mutant populations, while having little, if any, effect on the average number of bristles in a wild-type population. Even though these experiments did not reveal the precise nature of the genetic variation, they undeniably demonstrated genetic variation for the character that is not expressed in the wild type, but becomes visible in the mutant background. A paper published very recently in *Current Biology* by Greg Gibson and colleagues [2] has now provided the first detailed molecular portrait of this elusive form of genetic variation.

The discovery of hidden genetic variation caught the attention of experimental geneticists early on and led to the idea that the wild-type genotype may be more robust against the effects of mutations than the mutant phenotype. This idea was called 'canalization' by Waddington [3], one of the early enthusiasts for this field of research. Furthermore, the possibility was raised that this 'robustness' might be an evolved trait of the organism that protects the organism against the harmful effects of mutations. These ideas led to a quite intense research effort in the 1950s and 1960s (reviewed in [4]).

Needless to say, hidden variation and its corollary, robustness or canalization, is very hard to study, either experimentally or theoretically. The techniques at the disposal of geneticists in the middle of the 20th century turned out to be wholly inadequate for this task. The

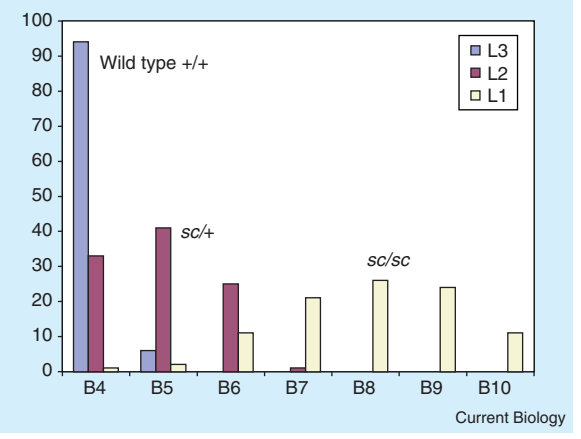

Figure. 1. The effect of the *Scute* mutation on the number of bristles on the *Drosophila* scutellum.

In a wild-type fruitfly, the average number of bristles on the scutellum is four, and there is very little variation from this. With one or two mutant *Scute* alleles, however, the average number of bristles increases and so does the amount of variation. Much of this variation is genetic variation, that is 'hidden' in the wild-type condition. This is a classical example of this phenomenon. (Data from [1].)

Department of Ecology/Evolutionary Biology, Osborn Memorial Labs, Yale University, 165 Prospect Street, New Haven, Connecticutt 06520-8106, USA.
E-mail: gunter.wagner@yale.edu

field consequently collapsed, as no progress seemed possible at the time. But since about the mid 1990s, newly developed molecular techniques have led to a renaissance of studies into the nature of phenotypic robustness and hidden variation [5,6]. At the experimental end of things, most attention has been directed into the issues of whether robustness actually exists and what factors might lead to phenotypic robustness. For instance, it was found that the molecular chaperone Hsp90 plays a role in hiding genetic variation in organisms as diverse as yeast, *Arabidopsis* and *Drosophila* [7,8]. But Hsp90 is certainly not alone in having such a role [9].

Until now, however, very little was known about the nature of hidden genetic variation itself. Molecular marker studies showed that some of the hidden variation for the *bithorax* phenotype in *Drosophila* is associated with the same region in the genome as a known gene with this mutant phenotype [6], but the precise molecular basis of this variation was not known. Now Gibson and colleagues [2] have reported the results of a massive study which has pushed the association between phenotype to the level of single nuceotide polymorphisms.

Gibson and colleagues [2] studied eye development in *Drosophila*. It has been known that a dominant mutation of the epidermal growth factor receptor gene, *EgfrE1*, perturbs eye development, leading to a characteristic 'roughening' of the eye's surface. The Gibson lab crossed this mutant allele into 210 isogenic wild-type lines and scored for the association of single nucleotide polymorphisms (SNPs) with the phenotype. Of 267 SNPs from a 10.9 kilobase sequence covering the *Egfr* locus, ten showed significant association with the phenotypes and are likely responsible for part of the hidden variation revealed by the presence of one copy of the *EgfrE1* allele.

A cynic may say that we already knew that there was genetic variation and that it has in one way or the other to be based on DNA sequence variation, thus the large amount of work by Gibson and colleagues [2] only confirms the obvious. But actually the results are anything but expected, at least not by me. It turns out that the polymorphisms are found both in the coding region as well as in the flanking non-coding regions of the gene, but only one polymorphism in the coding region leads to an amino-acid substitution. All the others are changes in the synonymous sites, changing the codon but not the encoded amino acid. By implication, the gene product is not changed by these mutations. Variation at the synonymous sites of a coding region has been and often is the paradigm of a 'neutral' genetic change, one that is irrelevant to the life of the organisms [10]. But the synonymous SNPs at the *EgfrE1* locus have a measurable effect on eye development. What does this tell us about the nature of phenotypic stability?

Gibson and colleagues [2] are careful not to express a definite opinion on why synonymous substitutions contribute to the hidden variation at the *EgfrE1* locus. But I will take the commentator's license to speculate a little beyond what their data directly say. It has been known for a while that, under certain circumstances, synonymous codons are not equally likely to occur, as one would expect if they would be really neutral with respect to fitness. This phenomenon is called codon bias, as some codons are more likely to be used in the genome than others. One explanation for codon bias is that the translational efficiency of a gene depends on the availability of tRNAs: if the protein coding gene is using a lot of codons which require rare tRNAs, the rate of translation can be slowed down considerably. This is particularly a problem for genes that have a high level of expression. Accordingly, codon bias is found most strongly in genes with high rates of protein expression, as for instance the alcohol dehydrogenase gene in *Drosophila*.

Carlini and Stephan [11] have shown recently that it is possible to detect selection acting on synonymous polymorphisms of the alcohol dehydrogenase locus in *Drosophila*. The degree of codon bias can even be used as a measure of the intensity of gene expression [12]. All that is well established, but the genes that control development are usually not highly expressed, and codon bias tends to be low among genes encoding transcription factor, such as the *Hox* genes (my unpublished data). In the eye, for instance, the Egfr protein is present for a relatively short period of time and in small amounts. But this does not imply that the rate of translation does not matter for this protein. Assume, for instance, that Egfr is required during a specific short period of time, but that it has to be absent before and after that period, to avoid unwanted side effects. It stands to reason, then, that the rate at which this protein becomes available can be of critical importance for its function. If this is the case, translational efficiency, and hence codon usage, can be of importance even for a protein that is not expressed in large amounts. And that is what Gibson and colleagues [2] found for *Egfr*.

An interesting corollary of this finding is that the kind of variation revealed by a sensitized phenotype might be qualitatively different from the kind of genetic variation that is the grist for the mill of natural selection. Adaptive evolution at the molecular level is usually detectable by an increased rate of non-synonymous substitutions, but rarely associated with an increased rate of synonymous substitutions [10]. So it is important to know the molecular nature of hidden variation, particularly in view of the idea that robustness, or canalization, allows the accumulation of potentially useful genetic variation out of view of the judging eyes of natural selection. This idea requires that the hidden variation is, in large part, of the same kind as the known adaptive variation, namely either regulatory differences or amino acid substitutions. If, however, the hidden variation is of a kind that is only expressed under very special circumstances, the adaptive value of this variation might not be very high.

What does this tell us about the nature of canalized developmental pathways? Most likely the stable wild-type phenotype is resistant to slight variations in the translation efficiency of this gene, as it is likely to be produced in the presence of the kind of synonymous substitutions detected by Gibson and colleagues [2]. One way the organism can achieve such stability is by providing an amount of mRNA with a safety margin which

ensures a translation rate above the required threshold. Only if the wild-type protein level is diminished by mutation, below the threshold necessary for full function, will the rate of translation make a difference. So it seems likely that threshold behavior is an important way that phenotypes can become robust against small genetic changes. The same conclusion was reached by researchers who modeled molecular genetic regulatory networks and discovered that they exhibit an 'exuberance' of robustness against variations in the rate parameters. Intrinsic threshold effects are what shelters much of the network against deleterious effects of rate changes and presumably against mutational effects [13]. This is a nice convergence of theoretical and experimental work rare in this field of research.

Far from only confirming an existing model for the genetic basis of complex phenotypic traits, the new work of Gibson and colleagues [2] has the potential to fundamentally challenge how we think about the role of genes and genetic variation. Organisms sustain their functional organization against the onslaught of genetic mutations by providing the phenotype with a physiological margin of safety that makes many small effects irrelevant. Hence the effects that become visible when the organisms loses its safety margins in a developmental process are not necessarily the same kind of mutations that are the stuff natural selection sees in producing adaptations.

References

1. Rendel, J. M. (1967). Canalization and Gene Control. Logos Press, Academic Press., New York.
2. Dworkin, I., Palsson, A., Birdsall, K., and Gibson, G. (2003). Evidence that Egfr contributes to cryptic genetic variation for photoreceptor determination in natural populations of *Drosophila*. Curr. Biol. October 28 issue.
3. Waddington, C. H. (1957). The Strategy of the Genes. MacMillan Co., New York.
4. Scharloo, W. (1991). Canalization: Genetic and developmental aspects. Annu. Rev. Ecol. Syst. 22, 65-93.
5. Stearns, S. C., and Kawecki, T.J. (1994). Fitness sensitivity and the canalization of life history traits. Evolution 48, 1438-1450.
6. Gibson, G., and Hogness, D.S. (1996). Effect of polymorphism in the *Drosophila* regulatory gene *Ultrabithorax* on homeotic stability. Science 271, 200-203.
7. Rutherford, S.L., and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. Nature *396*, 336-342.
8. Queitsch, C., Sangster, T. A. & Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. Nature *417*, 618-624.
9. Bergman, A., and Siegal, M.L. (2003). Evolutionary capacitance as a general feature of complex gene networks. Nature *424*, 549-552.
10. Graur, D., and Li, W.-H. (2000). Fundamentals of Molecular Evolution. Sinauer Ass. Inc., Sunderland, MA.
11. Carlini, D.B., and Stephan, W. (2003). *In vivo* introduction of unpreferred synonymous codons into the *Drosophila* Adh gene results in reduced levels of ADH protein. Genetics *163*, 239-243.
12. Ikemura, T. (1985). Codon usage bias and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 2, 13-34.
13. Dassow, G.V., and Odell, G.M. (2002). Design and constraints of the Drosophila segment polarity module: robust spatial patterning emerges from intertwined cell state switches. J. Exp. Zool. *294*, 179-215.