

Adaptation – not by sweeps alone

There has been recent progress in identifying selective sweeps underlying a range of adaptations. Jonathan Pritchard and Anna Di Rienzo argue that many adaptive events in natural populations may occur by polygenic adaptation, which would largely go undetected by conventional methods for detecting selection.

“ We believe that polygenic models have received too little attention in the population genetics community. ”

A major goal in evolutionary biology is to understand the genetic basis of how organisms adapt to their environments. Within the population genetics community, adaptation is typically viewed as involving selective sweeps that drive beneficial alleles from low to high frequency in a population.

In a highly influential paper, Maynard Smith and Haigh modelled selection acting on a new mutation that is rapidly driven to fixation¹. Recent work has also emphasized the possibility of sweeps on recessive variants or from standing variation, as well as partial sweeps^{2,3} (here, we use the term ‘selective sweep’ to refer to a large, rapid shift in frequency of a beneficial allele, regardless of whether there is an associated hitchhiking signal).

There are now many statistical tests that aim to detect sweep signals^{4,5}. These tests have been used to identify adaptive alleles at high frequency that underlie a range of beneficial phenotypes in humans and other species. Examples include human adaptations to infectious agents (for example, *DARC* and caspase 12), to reduced ultraviolet light (pigmentation loci such as *SLC24A5* and *KITLG*), to milk consumption (lactase) and to high altitude (*EPAS1*).

Despite these recent successes, we argue here that this view of adaptation that proceeds by selective sweeps at key loci is too limited. By contrast, classical selection models of artificial and natural selection in the quantitative genetics literature emphasize the importance of modest changes in allele frequencies at many loci^{6–8} — that is, ‘polygenic adaptation’. We believe that polygenic models have received too little attention in the population genetics community.

It seems likely to us that, as in traditional quantitative genetic models, many — possibly even most — adaptive events in natural populations occur by polygenic adaptation. Polygenic adaptation could allow rapid adaptive shifts, yet would often go undetected using conventional methods for detecting selection^{9–11}. Moreover, polygenic adaptation is qualitatively different from the models of adaptive substitutions that dominate the population genetics literature.

It has long been known that many traits in humans and other species are both phenotypically variable and

fairly heritable, and are controlled by large numbers of loci of small effect. Recent genome-wide association (GWA) studies have confirmed that many quantitative traits in humans are extremely polygenic, including height, body mass index, age at menarche, blood pressure, blood lipid levels, basal metabolic rate and glucose tolerance.

Suppose that a population is well-adapted to its environment until the environment changes (for example, the population moves to a new location with a different climate), thereby changing the optimal value of one or more phenotypes. If there is already considerable heritable variation underlying the phenotype in question, then the population can adapt rapidly to the new conditions at a speed that depends on the strength of selection and the heritability of the trait^{6,7}.

The key point is that we should expect such an adaptation to occur by small allele frequency shifts spread across many loci. As a hypothetical example, consider the adaptation of human height — a trait for which there are probably hundreds of SNPs that each affect height by a few millimeters¹². Strong selection for increased height could be very effective, as height is extremely heritable. But at the level of individual SNPs, the effect of selection would be rather weak, exerting just a small upward pressure in favour of each of hundreds of ‘tall’ alleles. Suppose that at 500 SNPs, the tall alleles each increase the expected height of a person by 2 mm. Then, an average shift of just 10% in the population allele frequency of each tall allele would increase average height in the population by 20 cm (assuming that SNPs contribute additively). Although these numbers are hypothetical, they illustrate that, for a highly polygenic trait, a dramatic adaptive response could result from modest allele frequency changes at many loci. This model is different from classical sweep models. Most importantly, adaptation could occur without dramatic allele frequency changes and without adaptive fixation events.

If polygenic adaptation is indeed an important mode of adaptation, then what types of trait would be most likely to evolve in this way? Traits that are strongly heritable prior to selection (that is, that have standing variation for selection to act on) with many loci of small

Jonathan K. Pritchard is at the Department of Human Genetics and Howard Hughes Medical Institute, University of Chicago, Chicago, Illinois 60615, USA.

Anna Di Rienzo is at the Department of Human Genetics, University of Chicago, Chicago, Illinois 60615, USA.

e-mails: pritch@uchicago.edu; dirienzo@bsd.uchicago.edu

“To understand polygenic adaptation, we need progress both in empirical approaches to detecting it and in theoretical modelling.”

effect, are likely to evolve by the polygenic model. These conditions probably hold for most quantitative traits.

By contrast, some traits are controlled by relatively few genes and this predisposes them to adaptation by selective sweeps. For example, the lactase persistence and Duffy-null mutations, both of which have undergone selective sweeps in humans, affect the expression of the key proteins in milk digestion and the entry of *Plasmodium vivax* into erythrocytes, respectively. Similarly, there may be traits that are heritable and affected by variants at many different loci, but for which many of the relevant variants would have strongly negative pleiotropic effects¹³. In these cases, sweeps might occur specifically at sites in the genome that can be modified without competing negative effects.

Finally, the sweep and polygenic models are not mutually exclusive. It may be that often the alleles with largest effect sizes (and without strongly negative pleiotropic effects) sweep to fixation, whereas much of the adaptive response is caused by smaller allele frequency shifts at many loci⁹.

To understand polygenic adaptation, we need progress both in empirical approaches to detecting it and in theoretical modelling.

To detect polygenic adaptation in real data will require different empirical approaches than those that are used to find selective sweeps. Unlike sweeps, the small allele frequency shifts underlying polygenic adaptations are unlikely to leave signals that stand out above neutral patterns of variation (at least at individual loci). Although selective sweeps can be detected using genetic data alone, studies of polygenic adaptations will be likely to require information about the adaptive phenotype or the relevant environmental pressures.

One type of approach will be to identify phenotypes that may have undergone adaptive changes in particular environments, such as adaptations to cold climate, high altitude or novel ecological conditions. To dissect the genetic basis of such adaptations, one might collect phenotyped samples from closely related populations that have and have not experienced the selective pressure of interest and use GWA mapping to identify relevant quantitative trait loci (QTLs). Additionally, one would want to measure the extent of phenotypic adaptation — estimated as the difference in average phenotype between the adapted and non-adapted populations when they are living under matched conditions (exact matching of conditions may be difficult in human studies). Then one could ask: what fraction of the phenotypic difference can be explained by alleles with large versus small frequency differences? Are the phenotypic effect sizes of QTLs with large allele frequency differences greater than those with subtle frequency shifts¹⁰? What fraction of the phenotypic difference cannot be explained by detected sweep signals or QTLs at all (and hence might result from the cumulative effect of many weak QTLs)?

In another type of scenario, one might hypothesize that a particular aspect of the environment is an important selective factor (for example, climate or diet) but

it is unclear what all the relevant phenotypes are. In this case, we might study adaptation by looking at sets of populations that have independently adapted to the same selective pressures^{14,15}. One type of signal would be alleles that show parallel frequency shifts in response to similar environmental pressures in distantly related populations (although this type of approach is unlikely to be powerful for alleles with very small effects)¹⁵.

On the theoretical side, there is a need to bridge the gap between the classical quantitative perspective and the population genetics approach (which has focused more explicitly on quantities that can be measured in molecular data).

First, we should understand better the conditions that can lead to a purely polygenic process. Does it result from sweeps or from a combination of sweeps along with smaller shifts in allele frequencies at many loci⁹? Key parameters in such models — some of which are not well known — include the numbers, allele frequencies and effect sizes for variants that affect a selected phenotype; the extent of pleiotropic effects; the effective population size; the mutation rates for new mutations that affect the phenotype; and the strength of selection¹⁶. More complex scenarios would result from selection regimes that fluctuate over time owing to environmental oscillations¹⁷.

We also need more modelling of the ways in which polygenic adaptation, if it is widespread, might affect genome-wide patterns of variation. For example, to what extent would widespread polygenic adaptation reduce levels of polymorphism or divergence near genes and other functional elements¹⁸? If this is likely to be a strong effect for realistic parameter values, then genome-wide patterns of polymorphism might allow us to measure the extent of polygenic adaptation. However, such signals might be difficult to distinguish from the predictions of selective sweep and background selection models, and may in fact contribute to patterns of variation that are commonly attributed to sweeps or background selection¹⁹.

Another important type of data in studying molecular evolution comes from between-species differences. To what extent would widespread polygenic adaptation affect genome-wide substitution rates at functional sites²⁰? How much would it contribute to signals of positive selection in between-species tests? For a polygenic trait at the selective optimum, we might imagine a kind of stabilizing selection on the constellation of variants segregating in the genome. That is, for a population near the selective optimum, a random change in the allele frequency at one locus would be approximately neutral if it is balanced by opposing change(s) in allele frequency elsewhere in the genome, such that the average phenotype remains unchanged. Now consider an adaptive shift. Suppose that a species adapts to be taller by small increases in the allele frequencies of many tall alleles. This would result in a small increase in the long-term fixation probabilities of the tall alleles. However, the effect of this mode of adaptation on substitution rates would likely be modest (except for extreme selective shifts), as the tall alleles would include a mixture of ancestral and derived alleles.

In summary, polygenic adaptation is likely to be an important mechanism of adaptive change. Polygenic adaptation would allow effective adaptation to environmental changes but would do so without selective

sweeps and with little effect on substitution rates. A key challenge for the coming years will be to understand the extent and role of polygenic adaptation in evolution.

1. Maynard Smith, J. & Haigh, J. The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**, 23–35 (1974).
2. Hermisson, J., Pennings, P. S. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).
3. Teshima, K. M. & Przeworski, M. Directional positive selection on an allele of arbitrary dominance. *Genetics* **172**, 713–718 (2006).
4. Sabeti, P. C. *et al.* Positive natural selection in the human lineage. *Science* **312**, 1614–1620 (2006).
5. Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C. & Clark, A. G. Recent and ongoing selection in the human genome. *Nature Rev. Genet.* **8**, 857–868 (2007).
6. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics*. (Longman, New York, 1996).
7. Gillespie, J. H. *Population Genetics: a Concise Guide*. (Johns Hopkins Univ. Press, Baltimore, USA, 2004).
8. Barton, N. H. & Keightley, P. D. Understanding quantitative genetic variation. *Nature Rev. Genet.* **3**, 11–21 (2002).
9. Chevin, L. M. & Hospital, F. Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics* **180**, 1645–1660 (2008).
10. Hancock, A. M., Alkorta-Aranburu, G., Witonsky, D. B. & Di Rienzo, A. Adaptations to new environments in humans: the role of subtle allele frequency shifts. *Philos. Trans. R. Soc. Lond. B* **365**, 2459–2468 (2010).
11. Pritchard, J. K., Pickrell, J. K. & Coop, G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**, R208–R215 (2010).
12. Weedon, M. N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genet.* **40**, 575–583 (2008).
13. Otto, S. P. Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proc. Biol. Sci.* **271**, 705–714 (2004).
14. Perry, G. H. & Dominy, N. J. Evolution of the human pygmy phenotype. *Trends Ecol. Evol.* **24**, 218–225 (2009).
15. Hancock, A. M. *et al.* Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proc. Natl Acad. Sci. USA* **107** (Suppl. 2), 8924–8930 (2010).
16. Martin, G. & Lenormand, T. The distribution of beneficial and fixed mutation fitness effects close to an optimum. *Genetics* **179**, 907–916 (2008).
17. Gillespie, J. H. *The Causes of Molecular Evolution*. (Oxford Univ. Press, Oxford, 1994).
18. Barton, N. H. Genetic hitchhiking. *Philos. Trans. R. Soc. Lond. B* **355**, 1553–1562 (2000).
19. McVicker, G., Gordon, D., Davis, C. & Green, P. Widespread genomic signatures of natural selection in hominid evolution *PLoS Genet.* **5**, e1000471 (2009).
20. Kopp, M. & Hermisson, J. The genetic basis of phenotypic adaptation II: the distribution of adaptive substitutions in the moving optimum model. *Genetics* **183**, 1453–1476 (2009).

Acknowledgements

We thank G. Coop, J. Pickrell, M. Przeworski and G. Sella for comments, and the National Institutes of Health grants DK56670, GM79558 and MH84703 and the Howard Hughes Medical Institute for funding support. We apologize to authors whose work could not be cited owing to space constraints.

Competing interests statement

The authors declare no competing financial interests.

FURTHER INFORMATION

Jonathan Pritchard's homepage: <http://pritch.bsd.uchicago.edu>
 Anna Di Rienzo's homepage: <http://genapps.uchicago.edu/labweb/index.html>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF