Footprints of polygenic adaptation of a quantitative trait under stabilizing selection

Benjamin Wölfl, Ilse Höllinger and Joachim Hermisson

Vienna Grad. School of Pop. Gen. & Faculty of Mathematics, University of Vienna

Abstract

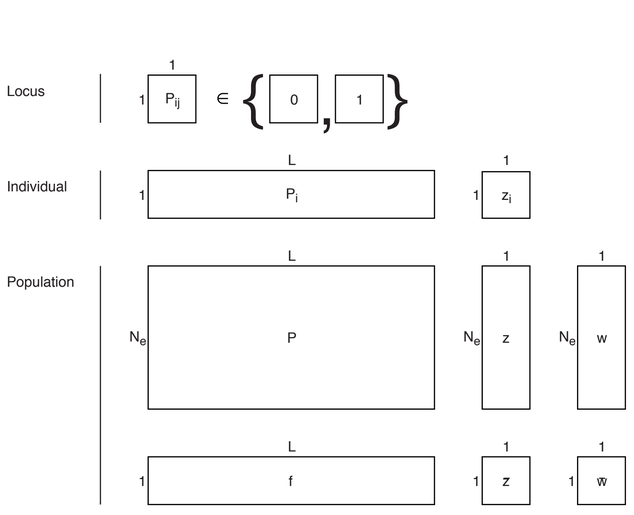
We study the **architecture of a quantitative trait to a new optimum after a sudden shift in the environment**. Using analytical theory and individual-based computer simulations we characterize the conditions under which adaptation occurs either due to sweeps at few loci or subtle frequency shifts at many loci. In particular, we analyze the impact of mutation rates, starting allele frequencies, number of loci, linkage and strength of selection on the adaptive scenario. In a final step, we compare theoretical predictions with data from replicated evolution experiments of Drosophila subject to thermal stress.

Introduction

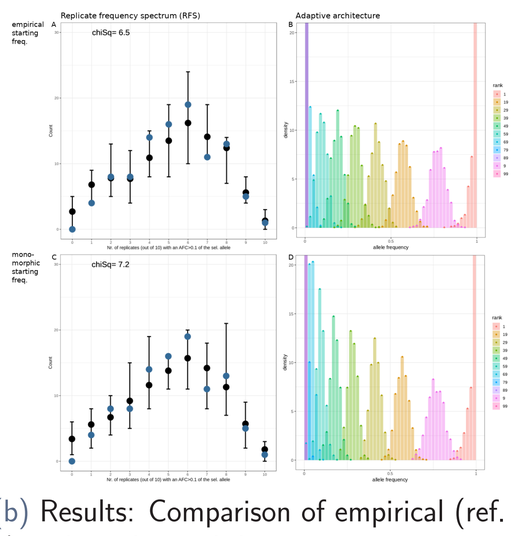
Which patterns of derived allele frequencies (footprints) are expected in nucleic acid polymorphism data, given:

- Quantitative trait: quasi-continuous variation in the trait value z
- Polygenic adaptation: > 1 contributing loci ($L > 1$)
- Stabilizing selection: Intermediate trait value optimum z_{opt}

Framework



(a) Representation of a population.



(b) Results: Comparison of empirical (ref. 3) and synthesized data.

- Genotype-phenotype map: equal ($\gamma_i = \gamma = \frac{1}{L} \rightsquigarrow z_i \in [0, 1]$) and additive allele effects $z_\iota = \sum_i \left(\delta_{\mathcal{L}_i A_i} \gamma_i\right)$
- Phenotype-fitness map: Gaussian fitness function (diminishing-returns epistasis) $w(z_\iota) = exp\left(-\frac{\sigma}{2}(z_\iota - z_{opt})^2\right)$

How can one express the population state (derived allele frequencies) f as a function of $\bar{z} = \frac{1}{N_e}\sum_\iota (z_\iota)$?

Models

Full selection model

$$\dot{p}_{A_i} = p_{A_i}(1 - p_{A_i})\sigma\gamma_i\Big(\underbrace{(z_{opt} - \bar{z})}_{\mathcal{A}} - \underbrace{\frac{\gamma_i}{2}(1 - 2p_{A_i})}_{\mathcal{B}}\Big) \quad (1)$$

Directional selection model

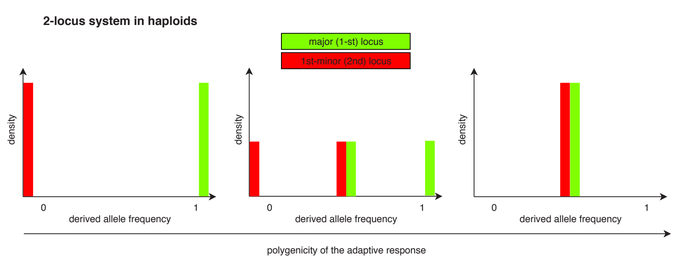
$$\dot{p}_{A_i} = p_{A_i}(1 - p_{A_i})\sigma\gamma_i \underbrace{(z_{opt} - \bar{z})}_{\mathcal{A}}. \quad (2)$$

For simulations both models are made discrete.

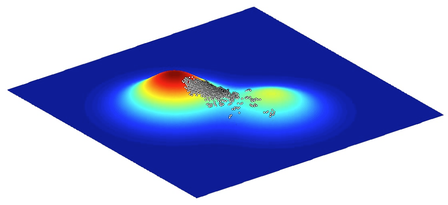
Adaptive architecture of a trait

- Joint distribution of derived allele frequencies f
- Express the adaptive architecture as a function of z
- But: $(L1)$-dimensional event space - hence: marginalization

Order f decreasingly and call the $(1^{st}, 2^{nd}, \ldots, L^{th})$ rank: major, 1^{st} minor, $\ldots$ and $(n-1)^{th}$ minor locus. Compare the derived allele frequency distributions of these ranks.



(a) Hard sweeps vs. small allele frequency shifts.



(b) The process of interest: Rapid phenotypic adaptation on a fitness landscape.

Methods

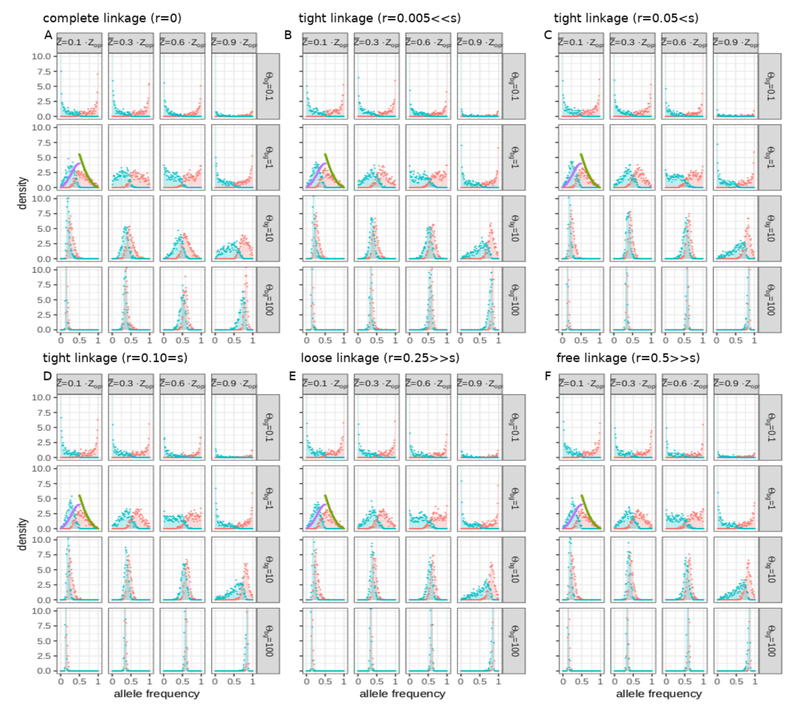
- **(A)** Stochastic analytical model (Yule process): Expression for marginal derived allele frequency distribution is an inverted Dirichlet distribution and Θ_{bg} is decisive (ref. 1 and 2)
  $\Theta_{bg} = \Theta_\ell(L-1)$ is the population-scaled background mutation rate and measures the "genetic redundancy of a trait"
- **(B)** Numerical simulations to validate (A) (focus):
  - Haploid Wright-Fisher simulations: independent alleles in `C/C++` (approximation)
  - Individual-based simulations in `C/C++` and `SLiM v3` (ref. 4)
- **(C)** Application to evolve & resequencing experiment of *Drosophila. sim.* (ref. 3) to estimate Θ_{bg} of the "heat adaptation" trait

Parameters

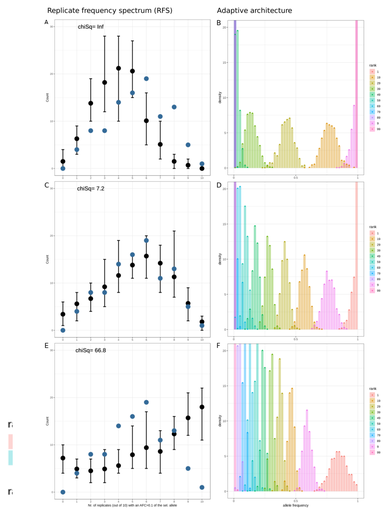
- Number of loci $L = 10$
- Allele effects: $\gamma^{hap} = \frac{1}{10}$, $\gamma^{dip} = \frac{1}{20}$
- Selection coefficient $s = \frac{1}{10}$
- Eff. population size: $N_e^{hap} = 10^4$, $N_e^{dip} = 5 \cdot 10^3$
- Optimal trait value $z_{opt} = 1$

So, the the individual fitness $z_\iota \in [0, 1]$.
We observe at different $\bar{z}$, as well as variable Θ_{bg} and r.

Results



(a) Adaptive architectures: Haploids are robust w.r.t. r. The same is true for diploids (not shown).



(b) Θ_{bg} determines the phenomenology of the replicate frequency spectrum and the adaptive architecture.

Conclusion & future work

- Θ_{bg} is decisive for the phenomenology
- Invariance to initial conditions: De-novo mutations vs.mutation-selection-drift balance
- Invariance to the recombination rate r
- Invariance to ploidy: haploidy vs. diploidy
- Experimental evolution experiment could be fitted
- Efficient individual-based simulator in $C++$ was written

Future work:
- Unequal locus effects
- Population structure
- Truncation selection
- Dominance in diploids
- Time-variant selection $s(t)$

References

1. Höllinger I, Pennings P and Hermisson J (2018). Polygenic Adaptation: From sweeps to subtle frequency shifts. PLoS Genetics 118:4251.
2. Höllinger I (2018). Mathematical Models of Speciation and Polygenic Adaptation. PhD thesis. University of Vienna, Austria.
3. Barghi N, Tobler R, Nolte V, Jakšić AM, Mallard F, Otte KA, et al. (2019) Genetic redundancy fuels polygenic adaptation in Drosophila. PLoS Biol 17(2): e3000128.
4. Haller BC and Messer PW. SLiM 3 (2019): Forward genetic simulations beyond the Wright-Fisher model. Molecular Biology and Evolution 36(3):632-637.

E-mail: b.wolfl@student.vu.nl, ORCID: 0000-0001-6447-6583, Twitter: @BenjaminWolfl, Date: 2019-08-20