

BMB Reports – Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-17-226

Title: Prevalence of negative frequency-dependent selection, revealed by incomplete selective sweeps in African populations of *Drosophila melanogaster*

Article Type: Perspective (Invited Only)

Keywords: Positive selection; *Drosophila melanogaster*; African; DNA sequence polymorphism; sweeping haplotypes

Corresponding Author: Yuseob Kim

Authors: Yuseob Kim^{1,*}

Institution: ¹Department of Life Science and Division of EcoScience, Ewha Womans University, Seoul 03760, Korea,

교정 파일을 두 가지 version으로 송부합니다.

한 편은 교정표시가 되어 있는 파일이며, 또 다른 한편은 교정자의 수정을 모두 적용하여 특별한 표시 없이 드리는 파일입니다.

가급적, 교정표시가 되어 있는 파일을 사용하시기 바랍니다. (교정표시가 있는 edit version을 사용하지 않을 것을 권장합니다. clean version은 참고로만 사용하시면 감사하겠습니다.)

-

선생님께서 직접 교정된 사항들을 일일이 확인함으로써, 선생님 기준과 교정내용이 부합한지를 알 수 있기 때문입니다.

아울러 교정표시가 있는 파일에 저희가 보내드리는 교정표시형식과 달리 보여질 때가 있습니다.

첫 번째-간혹 논문 교정본에 교정자국이 없고 좌측에 실선으로 연결되어 풍선모양 안에 교정된 말들이 나열되어 있는 경우가 있습니다. 이런 현상은 MS 워드 Version 차이에서 생기는 것입니다. MS 워드 2007, 2010 Version인 경우, 워드 파일 상단의 메뉴에서 검토-변경내용추적-옵션-풍선도움말-사용안함 순서로 설정하시고, 2003 version인 경우 상단의 메뉴에서 도구-옵션-변경내용추적-옵션-풍선도움말을 선택하지 않기로 설정하시면 저희 교정 자국이 교정본에 바로 나타날 것 입니다.

두 번째-좌측이나 우측에 세로로 선이 보여지는 경우가 있습니다. 이 직선을 없애는 방법은 MS워드 2007, 2010 Version인 경우, 검토-변경내용추적-옵션-태두리-없음으로 설정하시고, 2003 Version인 경우 도구-옵션-변경내용추적-아래쪽 변경된 부분을 '태두리 없음'으로 설정하시면 세로선이 없어질 것입니다.

세 번째-버전의 차이로 인하여 edit version의 교정자국이 간혹 보이지 않는 경우가 있습니다. 오른쪽 혹은 왼쪽의 붉은 색 세로 실선을 클릭 혹은 더블클릭 하시면, 교정자국이 보이실 것입니다.

문의 사항 있으시면 언제든지 연락 주시기 바랍니다.

선생님의 의뢰와 관심에 좋은 논문교정으로 보답하겠습니다

하리스코 드림

Harrisco-

Tel: 02-557-1810~1 Fax: 02-555-0752

Web: www.harrisco.net

BMB Reports

Perspective

Multiple modes of positive selection shaping the patterns of prevalence of negative frequency-dependent selection, revealed by incomplete selective sweeps in *ever* African populations of *Drosophila melanogaster*

서식 있음: 강조

서식 있음: 강조

~~[Editor's Note: Please double check the highlighted new title and make sure it is technically correct. The original one was unclear.]~~

Yuseob Kim

Department of Life Science and Division of EcoScience, Ewha Womans University, Seoul 03760, Korea

email) yuseob@ewha.ac.kr

Perspective to: Multiple modes of positive selection shaping the patterns of incomplete selective sweeps over African populations of *Drosophila melanogaster*. Molecular

메모 포함[y1]: Please keep the original title

서식 있음: 글꼴: 기움임꼴

Biology and Evolution 2017, doi: 10.1093/molbev/msx207

Abbreviations:

NFDS, negative frequency-dependent selection

SH, sweeping haplotypes

ABSTRACT

Positive selection on a new beneficial mutation generates a characteristic pattern of DNA sequence polymorphism when it reaches an intermediate allele frequency. On ~~the~~ genome sequences of African *Drosophila melanogaster*, we detected such signatures of selection at 37 candidate loci and identified “sweeping haplotypes (SHs)” that are increasing or have increased rapidly in frequency due to hitchhiking. ~~From the~~Based on geographic distribution of SH frequencies, we could infer whether selective sweeps occurred starting from *de novo* beneficial mutants under simple constant selective pressure. Single SHs were identified at more than half of loci. However, at many other loci, we observed multiple independent SHs, ~~which implying ies~~ soft selective sweeps ~~either~~ due to a high beneficial mutation rate or parallel evolution across space. Interestingly, SH frequencies ~~we~~are intermediate across multiple populations at about a quarter of the loci despite relatively low migration rates inferred between African populations. ~~This, which~~ invokes a certain form of frequency-dependent selection such as heterozygote advantage. At one locus, we observed a complex pattern of multiple independent that ~~was~~is compatible with recurrent frequency-dependent positive selection on new variants. In conclusion, ~~the~~genomic patterns of positive selection are very diverse, with equal contributions of hard and soft sweeps and a surprisingly large proportion of frequency-dependent selection, in *D. melanogaster* populations.

Explosive growth of next-generation sequencing data allows us to observe an enormous amount of genetic variation within and between species. Evolutionary biologists are now moving briskly to discover genomic patterns that can reveal species' phylogenetic and demographic history, ~~the~~ evidences of both positive and negative Darwinian selection, and ~~the~~ origins of important biological processes. It is also revising our basic understanding of evolutionary processes: mainly ~~which has long been~~ guided by ~~the~~ classical theories of population genetics in the past.

One of the most fundamental questions in evolutionary biology is how beneficial mutations ~~which are defined to increase the carriers' fitness (i.e. expected number of offspring or how well the individual under consideration can survive and reproduce)~~ arise and spread in natural populations. Beneficial mutations are defined as those that can increase carriers' fitness (i.e., expected number of offspring or how well the individual under consideration can survive and reproduce). After a beneficial mutation arises and avoids an initial stochastic loss, its frequency in the population is expected to increase exponentially. Under a simple assumption that the selective advantage of this beneficial allele is constant over time, it will eventually reach fixation (reaching 100% frequency) as described by the standard mathematical theory of positive selection (Figure 1A). However, such a frequency trajectory has ~~was~~ not been demonstrated in natural populations. Furthermore, there are multiple biological scenarios under which the strength of selection may gradually diminish toward zero as the beneficial allele increases to an intermediate frequency, a process termed negative frequency-dependent selection (NFDS). The fitness of a heterozygote with one copy of mutant allele might ~~ay~~ be higher than that of a mutant homozygote. In addition, a new phenotype might ~~ay~~ be

advantageous only when it is rare under resource competition and other ecological conditions.

We tested whether beneficial alleles in the genome of African *Drosophila melanogaster* would propagate according to the standard theory by observing the distribution of “sweeping haplotypes (SHs)” over different populations. An SH is produced by recent positive selection in a process called genetic hitchhiking or selective sweep (Figure 1B). When a beneficial allele increases rapidly up to an intermediate frequency, neutral variants at nearby loci on the same chromosome will “hitchhike” along, and thus forming a distinct block of homogeneous sequences among randomly sampled haploid genomes. The frequency of SH in a sample provides a good estimate for of the frequency of the putative beneficial allele.

Haploid genome sequences, made publically available by The Drosophila Population Genomics Projects, were sampled from eleven populations ~~that are~~ scattered over Sub-Saharan Africa (Figure 1C). Genome-wide genetic differentiation among populations indicate ~~that small rates with which~~ individuals will migrate between them ~~per unit time at small rates~~. ~~This leads to a prediction that U~~nder the standard model of constant selection with a small migration rate, if a beneficial allele arises in one population, it will increase to a large frequency by the time it arrives at another population and start to propagating there. ~~However, s~~Then, similar intermediate frequencies of SHs (thus beneficial alleles), as observed in a locus shown in Figure 1C, cannot be explained by the standard model, ~~although whereas~~. ~~It is however~~ compatible with the model of NFDS in which a beneficial allele stays very long at an intermediate frequency, therefore allowing migrations to equilibrate the pattern of variation among populations. Of 37 loci ~~that at which where~~ we detected clear SHs, at least eight loci exhibited such

evidences of NFDS. Actual proportion of such loci might be even higher because NFDS ~~could an~~ not be detected if selective pressures ~~aw~~ere spatially localized.

We also ~~fou~~ind that ~~at about half loci~~ positive selection produced multiple SHs ~~,~~ arising from independent mutation events at about half loci. Such results, termed “soft” selective sweeps, are explained by adaptive evolution proceeding in parallel over space or a high rate of mutation to beneficial alleles per locus. In one locus located in chromosome 2R, three to five independent SHs per population, the sum of which however ~~didees~~ not reach fixation, were observed. This complex haplotype pattern suggests an interplay between soft sweeps and NFDS.

Overall, this study suggests that diverse modes of positive selection, not anticipated in the classical theory of population genetics, are operating in the African population of *D. melanogaster*. ~~It should be confirmed if Whether this is~~ a similar set of selection modes operate in other species including humans needs to be determined in the future. Further investigations are also needed to elucidate the cause of NFDS and the fate of those beneficial alleles, namely ~~and determine~~ whether and how fixations at such loci would eventually occur.

Acknowledgements

This work was supported by grant (NRF-2015R1A4A1041997) ~~,~~ from the National Research Foundation of Korea.

Figure 1

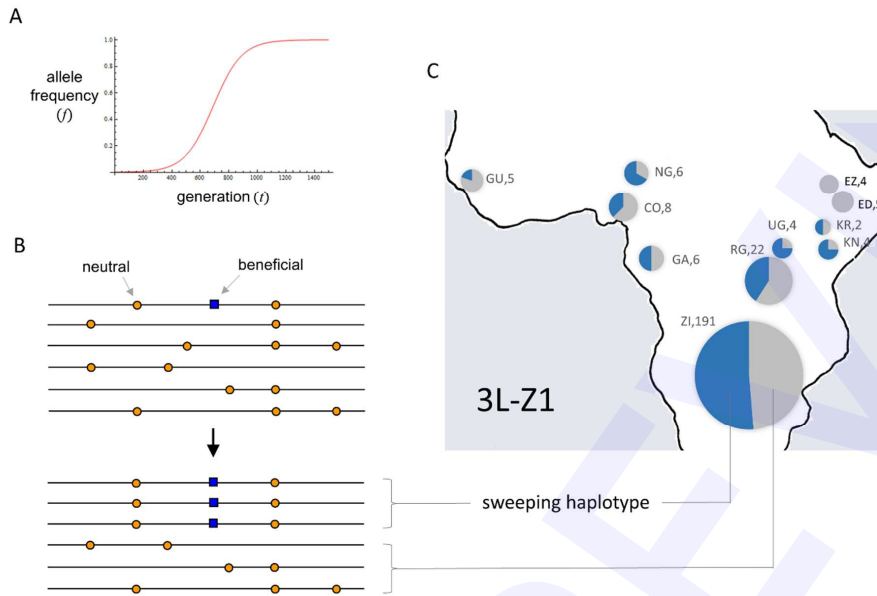


Figure legend:

Figure 1. **A.** Allele frequency (f) of a beneficial allele with selective advantage (s) over time (t) is given by the classical equation, $df/dt = sf(1 - f)$. **B.** On an interval of chromosome small enough to ignore recombination (cross-over) during the exponential growth phase of a beneficial allele, DNA sequences carrying the beneficial allele (blue square) have identical neutral variants (orange circles), and thus define a “sweeping haplotype”. **C.** The geographical distribution of SH at a candidate locus (named 3L-Z1) under selection, named 3L-Z1. The frequencies of SH at 11 African populations are shown by the blue portions of pie charts. The sizes of pie charts reflect sample sizes.

서식 있음: 글꼴: 기움임꼴 없음

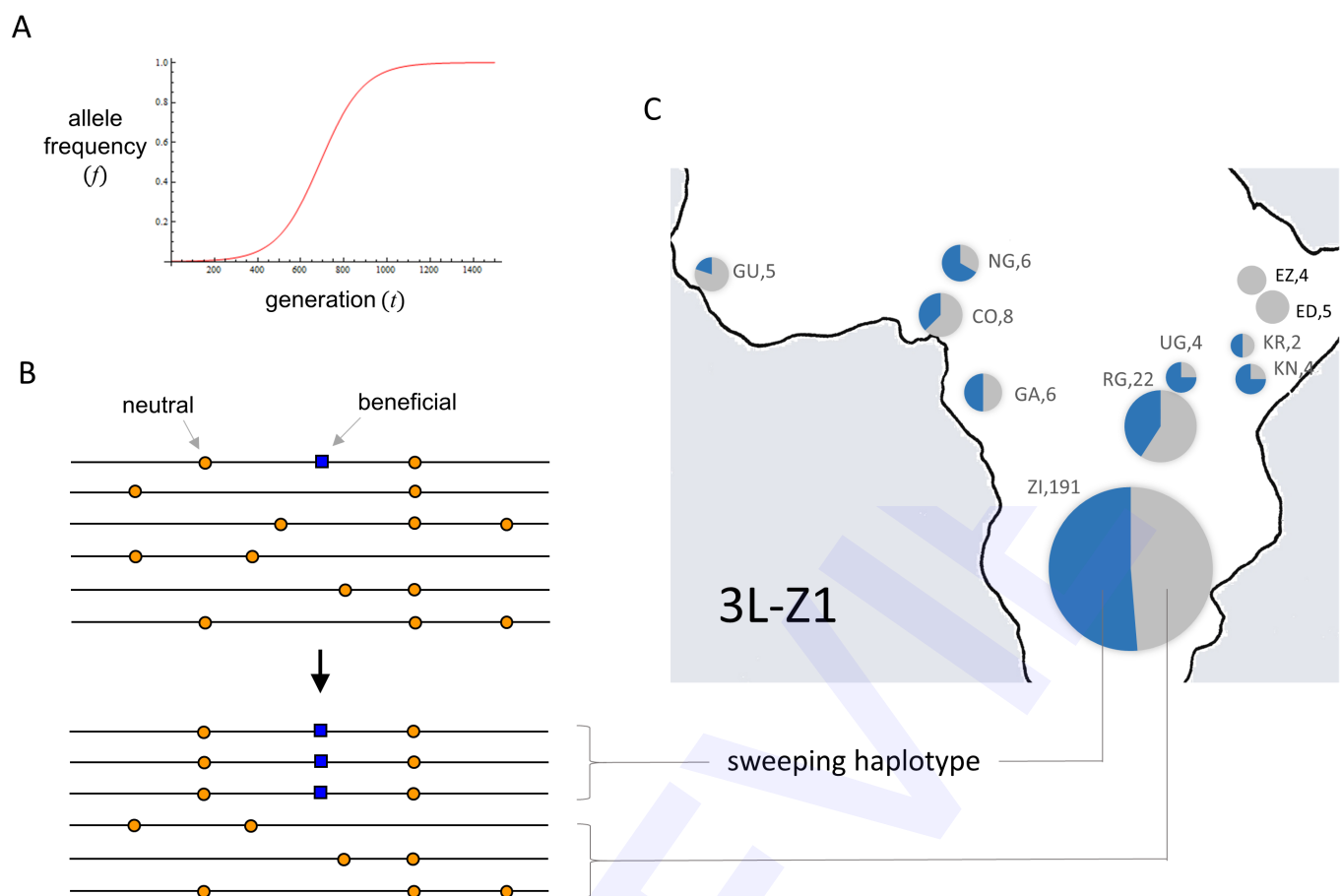


Fig. 1.