

Медицинская геномика

Василий Евгеньевич Раменский
Анастасия Александровна Жарикова и Мария Ильинична Зайченко

ramensky@gmail.com, azharikova89@gmail.com

НМИЦ Терапии и профилактической медицины
Факультет биоинженерии и биоинформатики МГУ
Институт искусственного интеллекта МГУ

2024

Мутации во времени:

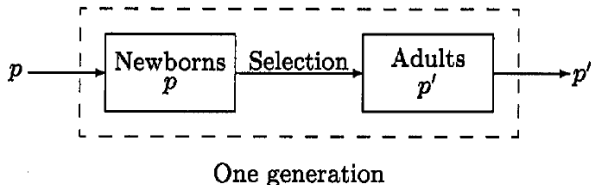
Основы популяционной генетики, ч.2

- 1 Естественный отбор
- 2 Мутационно-селекционное равновесие
- 3 Положительный отбор на фоне дрейфа
- 4 Вредные, слабовредные и нейтральные аллели

Определения

Естественный отбор – различное выживание и воспроизводство индивидуумов, происходящее благодаря различиям фенотипов.

Естественный отбор меняет частоту аллелей:



Приспособленность – способность индивидуума передать свои аллели \approx выживаемость [+ фертильность + время развития + способность найти пару + ...]

Вредные аллели уменьшают приспособленность (\neq патогенные, или болезнетворные)

Gillespie – *Population genetics. A concise guide*

Абсолютная и относительная приспособленность

	Генотип			Всего
Поколение $t - 1$	AA	Aa	aa	
Частота до отбора	p^2	$2pq$	q^2	$1 = p^2 + 2pq + q^2$
Приспособленность (выживаемость)	$w_{11} > 0$	$w_{12} \geq 0$	$w_{22} \geq 0$	Вероятность дожить до репр. возраста
После отбора	$p^2 w_{11}$	$2pq w_{12}$	$q^2 w_{22}$	$\bar{w} = p^2 w_{11} + 2pq w_{12} + q^2 w_{22}$
Нормализованные значения	$\frac{p^2 w_{11}}{\bar{w}}$	$\frac{2pq w_{12}}{\bar{w}}$	$\frac{q^2 w_{22}}{\bar{w}}$	$w_{11} > 0, w_{12} \geq 0, w_{22} \geq 0$
Поколение t	$p' = \frac{p^2 w_{11} + pq w_{12}}{\bar{w}}$ $q' = \frac{pq w_{12} + q^2 w_{22}}{\bar{w}}$			

$$\Delta p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\bar{w}}$$

Задача

Выведите формулу для Δp

Абсолютная и относительная приспособленность

Генотип	A_1A_1	A_1A_2	A_2A_2
Выживаемость (приспособленность)	w_{11}	w_{12}	w_{22}
Относительная приспособленность	1	w_{12}/w_{11}	w_{22}/w_{11}

Абсолютная и относительная приспособленность

Генотип	A_1A_1	A_1A_2	A_2A_2
Выживаемость (приспособленность)	w_{11}	w_{12}	w_{22}
Относительная приспособленность	1	w_{12}/w_{11}	w_{22}/w_{11}
Относительная приспособленность	1	$1 - hs$	$1 - s$

где $0 \leq s \leq 1$ – коэффициент отбора,
 h – эффект гетерозиготности, определяющий степень доминантности

$h = 0$	A_1 доминантный, A_2 рецессивный	$1, 1, 1 - s$
$h = 1$	A_1 рецессивный, A_2 доминантный	$1, 1 - s, 1 - s$
$0 < h < 1$	Неполное доминирование	
$h = 1/2$	Аддитивность	$1, 1 - s/2, 1 - s$
$h < 0$	Сверхдоминантность (overdominance)	
$h > 1$	Ослабленное доминирование	

Задача

Получите результаты для случаев $h < 0, h > 1$

Абсолютная и относительная приспособленность

$$\Delta p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\tilde{w}}$$

Переходя к относительной приспособленности: $w_{12}/w_{11} = 1 - hs$, $w_{22}/w_{11} = 1 - s$

$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}}$$

$$\tilde{w} = 1 - 2pqhs - q^2s$$

Задача

Получите эту формулу

Gillespie – *Population genetics. A concise guide*

Естественный отбор

1. Направленный (движущий) {положительный; отрицательный, или очищающий} отбор

Рецессивный аллель: $w_{11} = 1, w_{12} = 1, w_{22} = 1 - s // w_{12} = 1$

Доминантный аллель: $w_{11} = 1, w_{12} = 1 - s, w_{22} = 1 - s // w_{12} = w_{22}$

Неполное доминирование:

$w_{11} = 1, w_{12} = 1 - hs, w_{22} = 1 - s, 0 < h < 1 // w_{11} > w_{12} > w_{22}$

Задача

Получите эти формулы

2. Балансирующий отбор

Сверхдоминантность: $w_{11} = 1, w_{12} = 1 - hs, w_{22} = 1 - s, h < 0 // w_{12} > w_{11}, w_{22}$

3. Разрывающий отбор

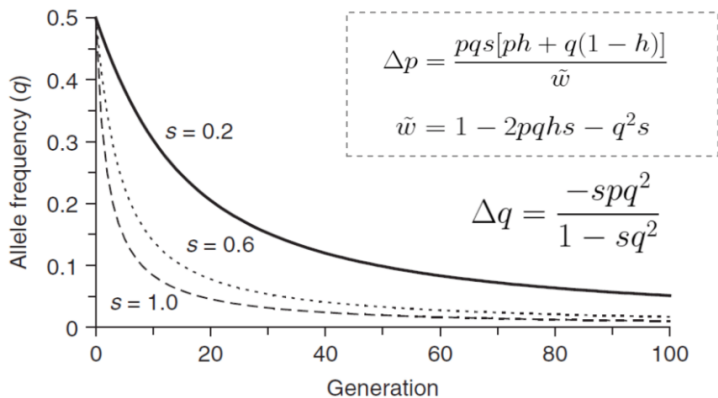
Ослабленная доминантность: $w_{11} = 1, w_{12} = 1 - hs, w_{22} = 1 - s, h > 1$

Задача

Каковы допустимые значения h ?

Направленный отбор

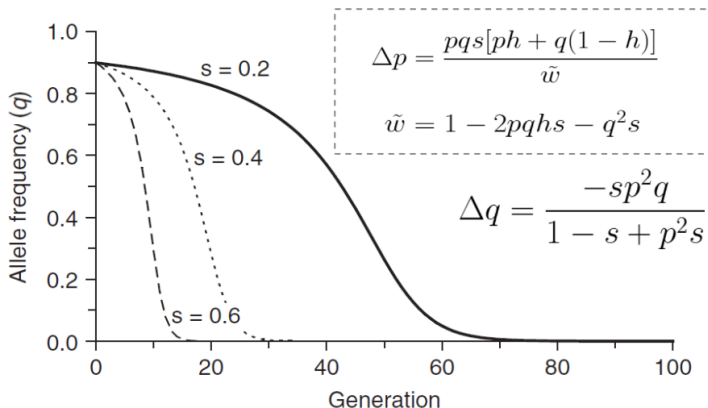
Направленный отбор против рецессивного аллеля: $w_{11} = w_{12} = 1, w_{22} = 1 - s$.



Relethford – *Human Population Genetics*

Направленный отбор

Направленный отбор против доминантного аллеля: $w_{11} = 1$, $w_{12} = w_{22} = 1 - s$.

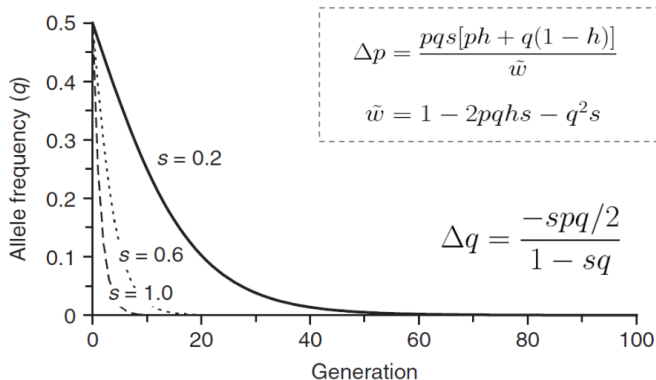


Relethford – *Human Population Genetics*

Направленный отбор

Направленный отбор против кодоминантного аллеля:

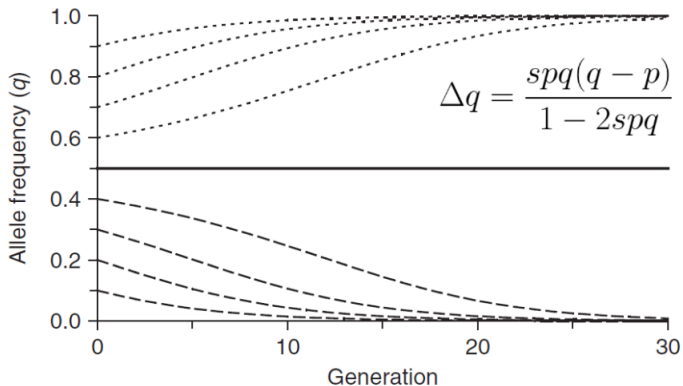
$w_{11} = 1, w_{12} = 1 - s/2, w_{22} = 1 - s$ // неполное доминирование.



Relethford – *Human Population Genetics*

Направленный отбор

Направленный отбор против гетерозиготы: $w_{11} = 1$, $w_{12} = 1 - s$, $w_{22} = 1$ // ослабленная доминантность.



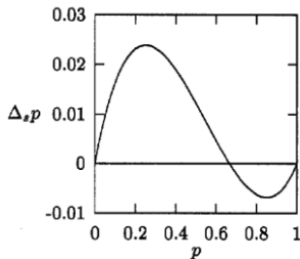
Relethford – *Human Population Genetics*

Балансирующий отбор

Балансирующий отбор гетерозиготы: $w_{11} = 1, w_{12} = 1 - hs, w_{22} = 1, h < 0 //$
сверхдоминантность.

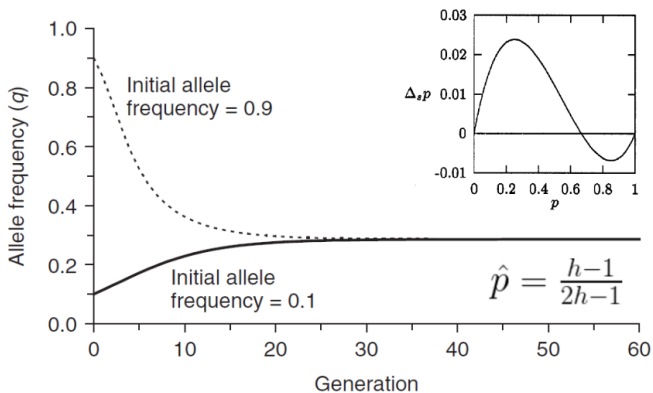
$$\Delta p = \frac{pqs[ph + q(1 - h)]}{\tilde{w}}$$

$$\tilde{w} = 1 - 2pqhs - q^2s$$



Балансирующий отбор

Балансирующий отбор гетерозиготы: $w_{11} = 1, w_{12} = 1 - hs, = w_{22} = 1, h < 0 //$
 сверхдоминантность.



Задача

Получите \hat{p}

Балансирующий отбор: случай муковисцидоза 🏠

BOX 3.7 SELECTION IN FAVOR OF HETEROZYGOTES FOR CYSTIC FIBROSIS

For CF, the disease frequency in Denmark is about one in 2000 births.

Phenotypes:	Unaffected		Affected
Genotypes:	AA	Aa	aa
Frequencies:	p^2	$2pq$	$q^2 = 1/2000$

q^2 is 5×10^{-4} ; therefore $q = 0.022$ and $p = 1 - q = 0.978$.

$p/q = 0.978/0.022 = 43.72 = s_2/s_1$.

If $s_2 = 1$ (affected homozygotes never reproduce), $s_1 = 0.023$.

The present CF gene frequency will be maintained, even without fresh mutations, if Aa heterozygotes have on average 2.3% more surviving children than AA homozygotes.

Strachan, Read – *Human Molecular Genetics*

Задача

Выразите эффект гетерозиготности h как функцию от \hat{p} , подтвердите оценку выше.

Балансирующий отбор: случай β -гемоглобина 🏠

The most thoroughly studied example of overdominance is the sickle-cell hemoglobin polymorphism found in many human populations in Africa. Hemoglobin, the oxygen-carrying red protein found in red blood cells, is a tetramer composed of two alpha chains and two beta chains. In native West and Central African populations, the S allele of beta hemoglobin reaches a frequency as high as 0.3 in some areas. The more common A allele is found at very high frequency in most other areas of the world. The two alleles differ only in that the S allele has a glutamic acid at its sixth amino position while the A allele has a valine. The glutamic acid causes the hemoglobin to form crystal aggregates under low partial pressures of oxygen, as occur, for example, in the capillaries. As a result, SS homozygotes suffer from sickle-cell anemia, a disease that is often fatal.

The S allele could not have reached a frequency of 0.3 unless AS heterozygotes are more fit than AA homozygotes. This is precisely the case in regions where malaria is endemic, for there the heterozygotes are somewhat resistant to severe forms of malaria. The resistance is due to the sickling phenomena, which makes red blood cells less suitable for *Plasmodium falciparum*. In an old study from 1961, it was shown that the viability of AS relative to AA is 1.176 in regions with malaria. Assuming that the fitness of SS is zero ($s = 1$), $h = -0.176$. Plugging this into Equation 3.4 gives $\hat{p} = 0.87$ or $\hat{q} = 0.13$ for the S allele, which is nestled right in the middle of allele frequencies in regions with endemic malaria.

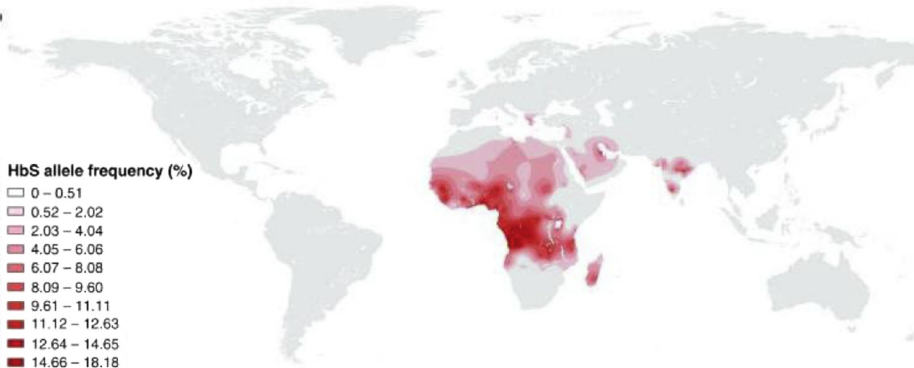
Балансирующий отбор: случай β -гемоглобина 🏠

FIGURE 14.2

The global distribution of the sickle-cell allele.

From Piel, F.B., Patil, A.P., Howes, R.E., Nyangiri, O.A., Gething, P.W., Williams, T.N., et al., 2010. Global distribution of sickle cell gene and geographical confirmation of the malaria hypothesis. Nature Communications 1, 104–111.

Изменение приспособленности со временем

$$\Delta p = \frac{qs[ph+q(1-h)]}{\tilde{w}} \quad \tilde{w} = 1 - 2pqsh - q^2$$

$$\text{Сьюэл Райт: } \Delta p = \frac{pq}{2\tilde{w}(p)} \frac{d\tilde{w}(p)}{dp}$$

Что происходит с $\tilde{w}(p)$?

«Естественный отбор всегда стремится увеличить среднюю приспособленность популяции и скорость этого пропорциональна величине генетической вариабельности»

Gillespie – *Population genetics. A concise guide*

Баланс мутаций и отбора

- Множество новых аллелей вредны и не полностью доминантны.
- Они появляются в популяции за счет мутации и исчезают за счет отрицательного отбора. $A_1(p \approx 1) \xrightarrow{\mu} A_2(q \approx 0)$
- Баланс: скорость появления мутаций равняется скорости их потери из-за отбора.

$$\Delta_{mut} p = -\mu p \approx -\mu$$

$$\Delta_{sel} p = \frac{pqs[ph+q(1-h)]}{1-2pqhs-q^2s} \approx qhs$$

$$\Delta_{mut} p + \Delta_{sel} p = 0$$

$$\hat{q} \approx \frac{\mu}{hs}$$

Баланс мутаций и отбора

- Множество новых аллелей вредны и не полностью доминантны.
- Они появляются в популяции за счет мутации и исчезают за счет отрицательного отбора. $A_1(p \approx 1) \xrightarrow{\mu} A_2(q \approx 0)$
- Баланс: скорость появления мутаций равняется скорости их потери из-за отбора.

$$\Delta_{mut} p = -\mu p \approx -\mu$$

$$\Delta_{sel} p = \frac{pqs[ph+q(1-h)]}{1-2pqhs-q^2s} \approx qhs$$

$$\Delta_{mut} p + \Delta_{sel} p = 0$$

$$\hat{q} \approx \frac{\mu}{hs}$$

Сильный эффект → маленькая частота

Задача

Получите \hat{q} для рецессивного аллеля

Случайный дрейф и благоприятный аллель

Отбор в конечных популяциях очень слаб для аллелей *de novo*.

Новый аллель: $\Delta p \approx (1 + s)p - p = sp = s/2N \ll 1/2N$ (дрейф), кроме случаев $s \approx 1$

$$P_F(p) = \frac{1 - e^{-2Nsp}}{1 - e^{-2Ns}}, \text{ если } h = 1/2$$

$$P_F(p) = \frac{1 - e^{-s}}{1 - e^{-2Ns}}, P_F \approx s \text{ если } s \approx 1 \text{ и } 2Ns \gg 1$$

Случайный дрейф и благоприятный аллель

Отбор в конечных популяциях очень слаб для аллелей *de novo*.

Новый аллель: $\Delta p \approx (1 + s)p - p = sp = s/2N \ll 1/2N$ (дрейф), кроме случаев $s \approx 1$

$$P_F(p) = \frac{1 - e^{-2Nsp}}{1 - e^{-2Ns}}, \text{ если } h = 1/2$$

$$P_F(p) = \frac{1 - e^{-s}}{1 - e^{-2Ns}}, P_F \approx s \text{ если } s \approx 1 \text{ и } 2Ns \gg 1$$

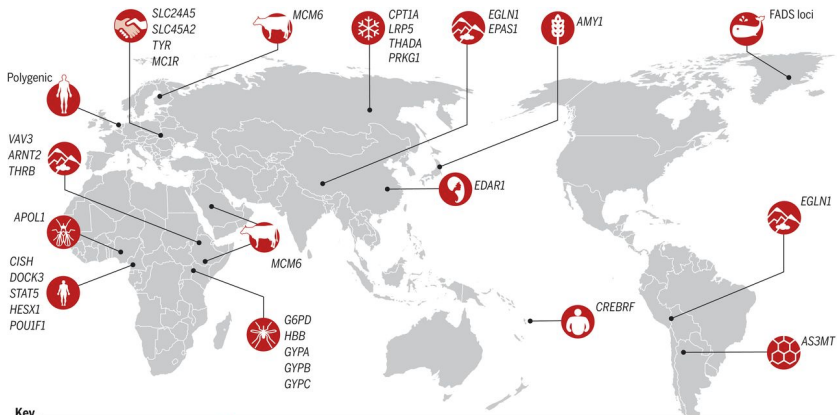
- Самые благоприятные аллели теряются.
- Адаптивная эволюция случайна.

Задача

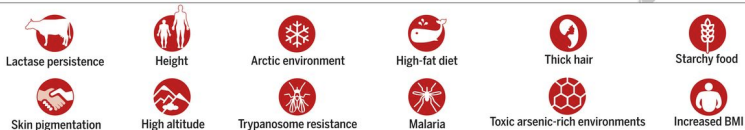
Получите P_f для случая $s, 2Ns \approx 0$

Gillespie – Population genetics. A concise guide

Примеры локальной адаптации человека 🏠



Key



Fan (2016) *Science*

Примеры локальной адаптации человека 🏠

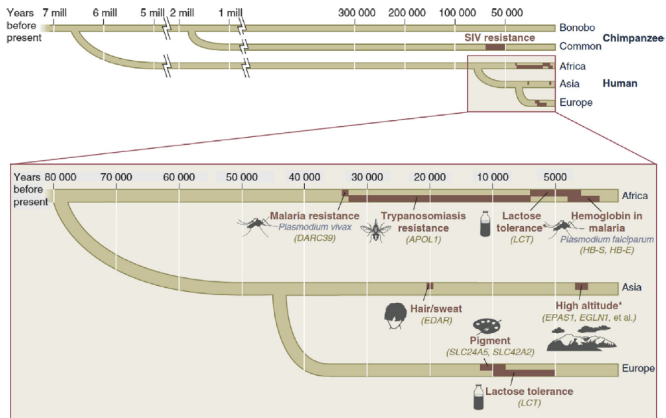


FIGURE 14.4

A portrayal of human evolution in a combined species and intraspecific population phylogenetic tree. Various points on the human “branches” indicate the estimated times at which various positively selected human adaptations arose.

From Vitti, J.J., Cho, M.K., Tishkoff, S.A., Sabeti, P.C., 2012. Human evolutionary genomics: ethical and interpretive issues. *Trends in Genetics* 28, 137–145.

Случайный дрейф и вредный аллель

Может ли вредный аллель зафиксироваться в популяции конечного размера?

$$P_F(q) = 1 - P_F(1 - q) = \frac{e^{2Ns q} - 1}{e^{2Ns} - 1}$$

$$P_F(1/2N) \approx \frac{s}{e^{2Ns} - 1} \quad P_F \approx \text{если } 2Ns \gg 1$$

Случайный дрейф и вредный аллель

Может ли вредный аллель зафиксироваться в популяции конечного размера?

$$P_F(q) = 1 - P_F(1 - q) = \frac{e^{2Ns q} - 1}{e^{2Ns} - 1}$$

$$P_F(1/2N) \approx \frac{s}{e^{2Ns} - 1} \quad P_F \approx \text{если } 2Ns \gg 1$$

Частота фиксации вредного аллеля:

$$k = 2N\mu P_F(1/2N) = \frac{2N\mu s}{e^{2Ns} - 1}$$

Задача

- Получите P_F для $s \rightarrow 0$.
- Получите k для $s \rightarrow 0$.

Gillespie – *Population genetics. A concise guide*

Слабовредные vs нейтральные мутации

Мутации можно отнести к трем основным категориям:

- Подверженные отбору (положительному или отрицательному);
- Нейтральные (не имеющие эффекта на приспособленность);
- мутации с маленькими коэффициентами отбора, таким образом ведущее себя как нейтральные в малых популяциях (где доминируют эффекты дрейфа) или подверженные отбору в больших популяциях, где возобладают эффекты отбора.

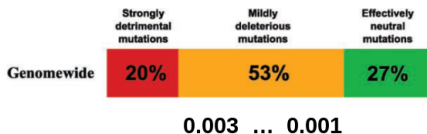
Meyer, Diogo, and Eugene E. Harris. "Selection Operating on Protein-coding Genes in the Human Genome." eLS

Слабовредные vs нейтральные мутации

Most Rare Missense Alleles Are Deleterious in Humans:
Implications for Complex Disease and Association Studies

Gregory V. Kryukov, Len A. Pennacchio, and Shamil R. Sunyaev

The American Journal of Human Genetics Volume 80 April 2007

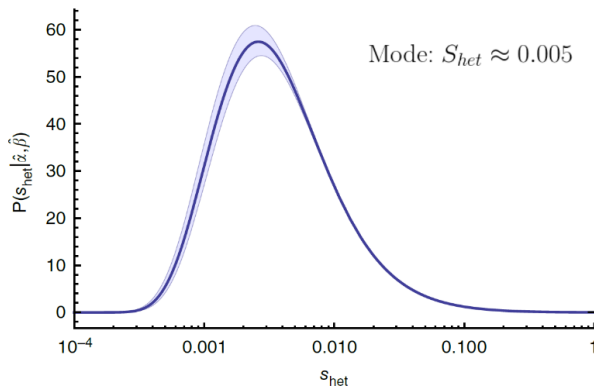


We combined analysis of mutations causing human Mendelian diseases, of human-chimpanzee divergence, and of systematic data on human genetic variation and ... estimated that $>50\%$ of de novo missense mutations in an average human gene and 70% of missense SNPs detected only once among 1,500 chromosomes are mildly deleterious. **Such mildly deleterious mutations are associated with selection coefficients within a surprisingly narrow range of 0.001–0.003.**

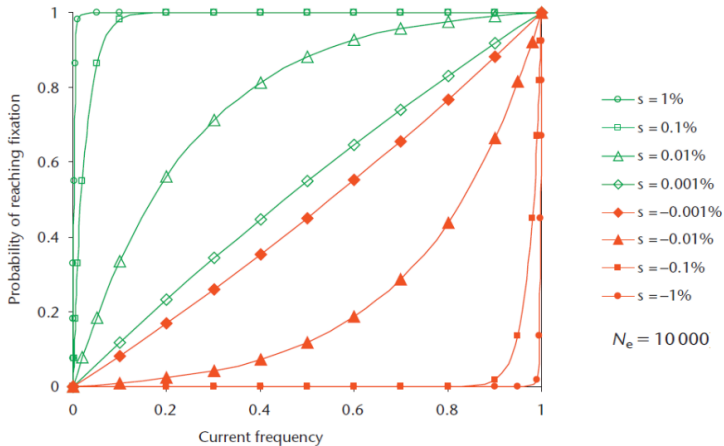
Estimating the selective effects of heterozygous protein-truncating variants from human exome data

Christopher A Cassa^{1,2,9}, Donate Weghorn^{1,9}, Daniel J Balick^{1,9}, Daniel M Jordan^{3,9}, David Nusinow¹, Kaitlin E Samocha^{4,5}, Anne O'Donnell-Luria^{4,6}, Daniel G MacArthur^{2,4}, Mark J Daly^{2,4}, David R Beier^{7,8} & Shamil R Sunyaev^{1,2}

VOLUME 49 | NUMBER 5 | MAY 2017 NATURE GENETICS

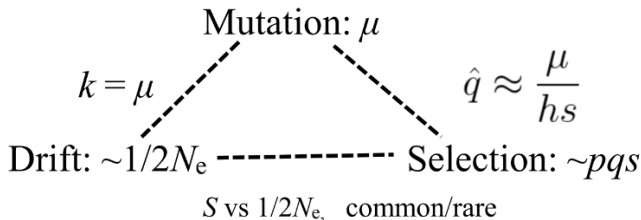


Вероятности фиксации для всех аллелей 🏠



Thomas, Paul D (July 2008) Single Nucleotide Polymorphisms in Human Disease and Evolution: Phylogenies and Genealogies. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0020763

Мутации, отбор и дрейф



The interaction of drift and selection is more complex than that of mutation and drift because the strength of selection changes with the frequency of the allele. (Recall the factor pq in $\Delta_s p$ or examine [Figure 3.3](#).) Natural selection is always a very weak force for rare alleles and, if $s \gg 1/(2N)$, is a much stronger force than drift for common alleles. Thus, the dynamics of rare alleles should be influenced by both drift and selection, while the dynamics of common alleles should be determined mainly by natural selection (again, if $s \gg 1/(2N)$).

Выводы

Что меняет частоты аллелей/генотипов?

- **Мутации:** появление новых аллелей в популяции
- **Случайный дрейф:** выборочная вариация передаваемых аллелей
- **Отбор:** различные вероятности выживания/размножения в зависимости от генотипов

Выводы

- Равновесие Харди-Вайнберга описывает как зиготы появляются из гамет.
- Случайный дрейф способствует потере или фиксации аллелей и уменьшает гетерозиготность.
- Нейтральная теория утверждает, что большинство внутри- и междувидовых изменений происходят из-за отрицательного отбора и случайного дрейфа.
- Естественный отбор меняет частоты аллелей. Он всегда увеличивает среднюю приспособленность со скоростью, пропорциональной генетической вариации.
- Большинство новых аллелей вредны и не полностью доминантны. Они появляются за счет мутаций и подвержены отрицательному отбору (баланс мутаций и отбора).
- В конечной популяции, новые благоприятные мутации обычно исчезают за счет случайного дрейфа. С другой стороны, вредные аллели могут зафиксироваться.

Список литературы

- Meyer, D., Harris, E. (2008) Selection Operating on Protein-coding Genes in the Human Genome. In: Encyclopedia of Life Sciences (ELS).
doi:10.1002/9780470015902.a0020791
- Nei, M., Suzuki, Y., and Nozawa, M. (2010). The neutral theory of molecular evolution in the genomic era. *Annu Rev Genomics Hum Genet* 11, 265–289
- Hurst, L.D. (2009). Genetics and the understanding of selection. *Nature Reviews Genetics* 10, 83–93.
- Fan, S., Hansen, M.E.B., Lo, Y., and Tishkoff, S.A. (2016). Going global by adapting local: A review of recent human adaptation. *Science* 354, 54–59.
- John H. Gillespie – Population Genetics. A concise guide
- John H. Relethford – Human population genetics