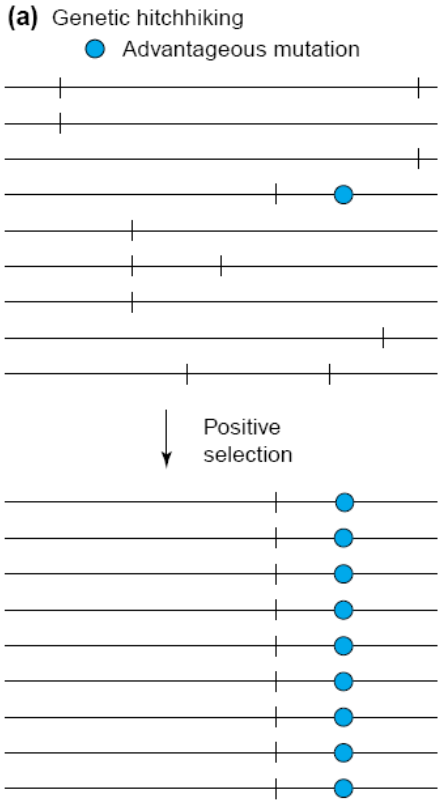
Detekce selekce na molekulární úrovni a genetická podstata adaptací

Typy selekce



Pozitivní selekce

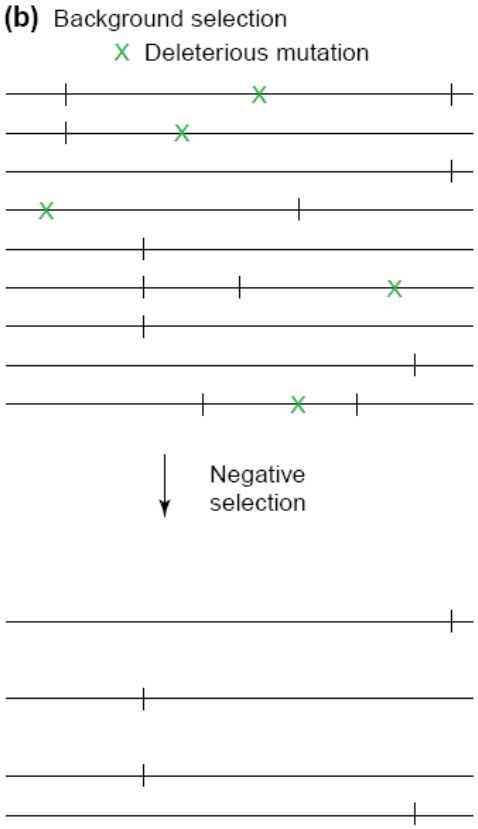
- snižuje genetickou variabilitu v populaci (“selective sweep”)
- zvyšuje míru divergence mezi druhy

Negativní selekce

- snižuje genetickou variabilitu v populaci (selekce na pozadí)
- snižuje míru divergence mezi druhy

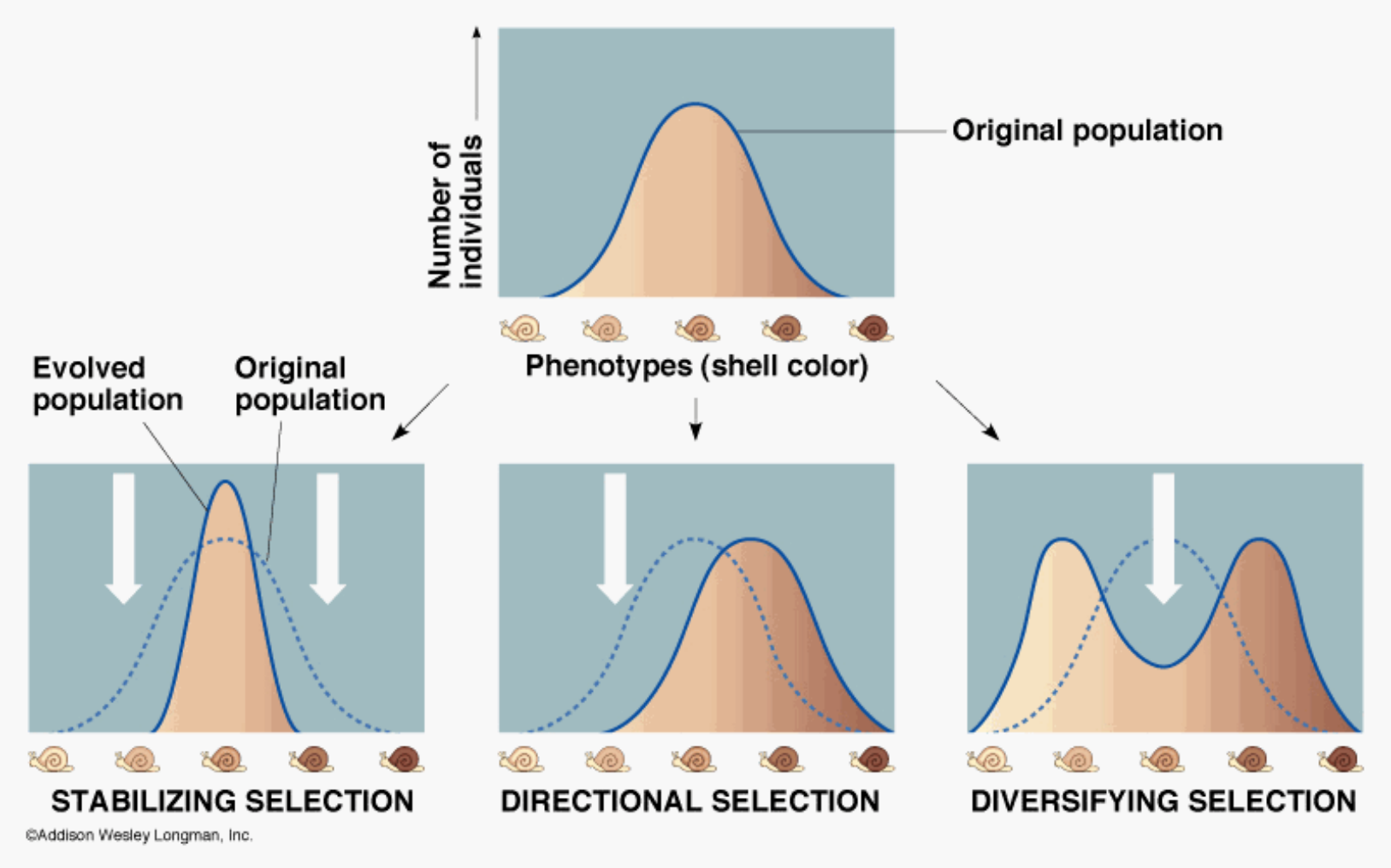
Balancing selekce

- selekce ve prospěch heterozygotů, frekvenčně závislá selekce
- udržuje v populaci trvale polymorfismus



Typy selekce

stabilizující **usměrňující** **disruptivní**



Metody detekce pozitivní selekce na molekulární úrovni

- Netřeba znát předem fenotyp.
- Lze detekovat i selekci, která působila v minulosti.

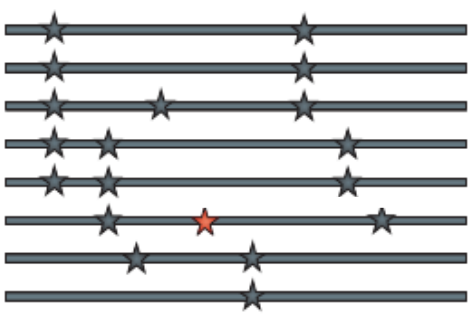
Metody detekce nedávné pozitivní selekce

- Selekce před max. ~ N_e generacemi (u člověka ~ 250 000 let).
- Doba fixace pozitivní mutace: **$2\ln(2N_e)/s$** generací (u člověka, N_e = 10 000, generační doba 25 let, s = 5%, bude trvat 10 000 let)

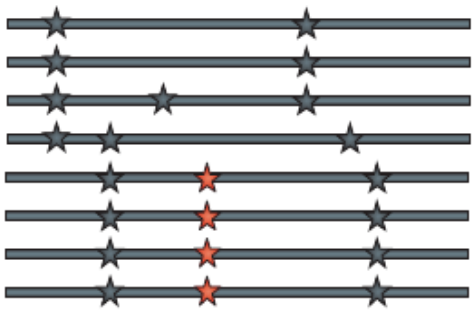
Založené na hledání průvodních jevů „selective sweep“

- snížení genetické variability
- zvýšení vazebné nerovnováhy
- změna v distribuci frekvence alel

vznik výhodné mutace



neúplný selective sweep



úplný selective sweep

Hudson-Kreitman-Aguadé (HKA) test

- Srovnává míru polymorfismu uvnitř druhu (θ) a divergence mezi druhy (D) pro dva či více lokusů.
- Pro neutrální sekvence je poměr θ a D pro konstantní
- Pozitivní selekce vede ke snížení θ, ale neovlivní D.
- Selekci nelze detekovat srovnáváme-li mezi sebou dva lokusy, které jsou oba pod působením pozitivní selekce.
- Pozitivní HKA test může být způsoben selekcí na lokus, který je ve vazbě s námi studovaným lokusem.

$\theta = 4N_e\mu$

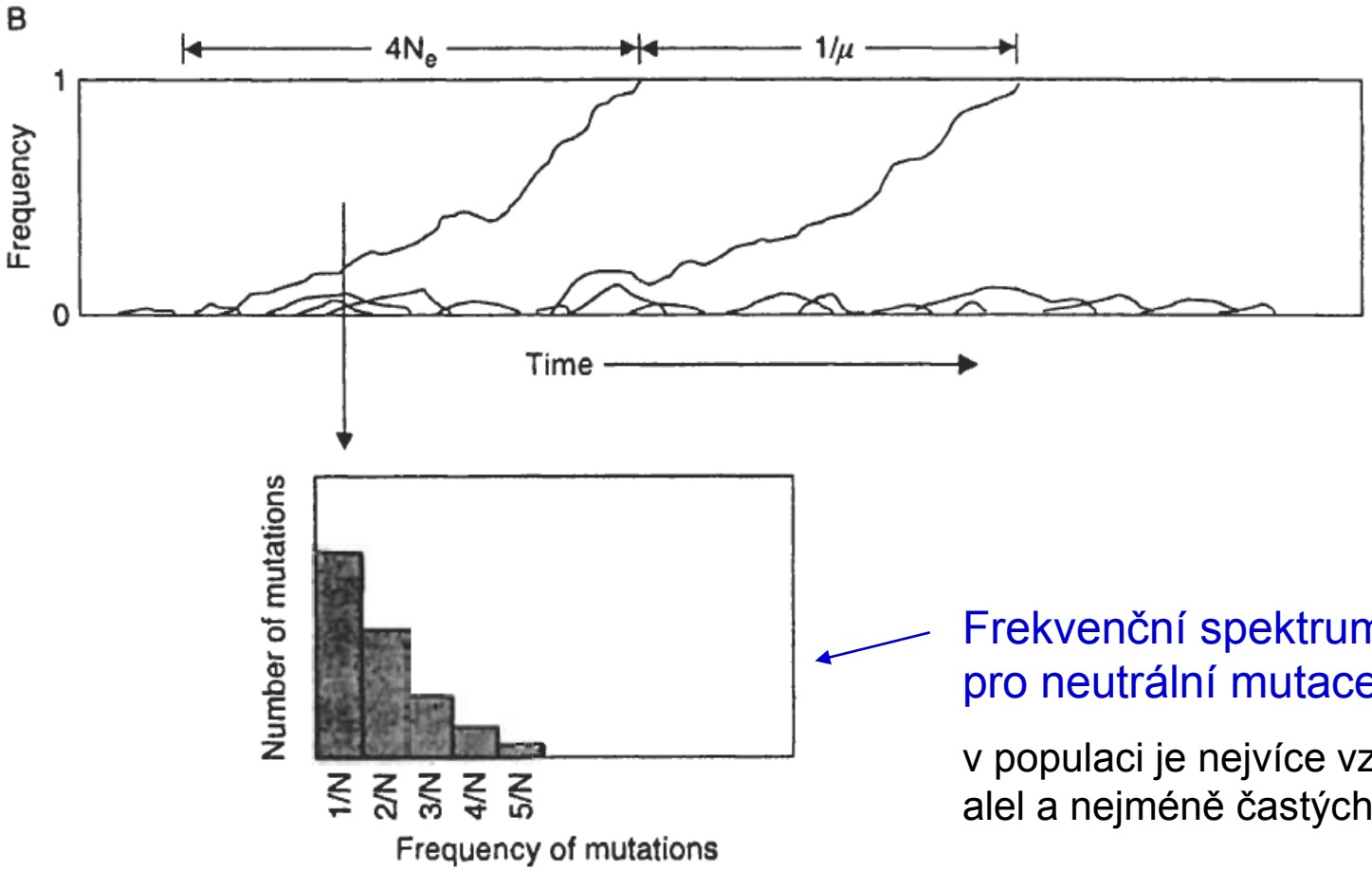
$D = 2\mu t$

	θ	D
Lokus 1	θ_1	D_1
Lokus 2	θ_2	D_2
Lokus 3	θ_3	D_3
Lokus 4	θ_4	D_4

Software: http://genfaculty.rutgers.edu/hey/software#HKA

Tajima's D test

- Založen na frekvenčních spektrech mutací

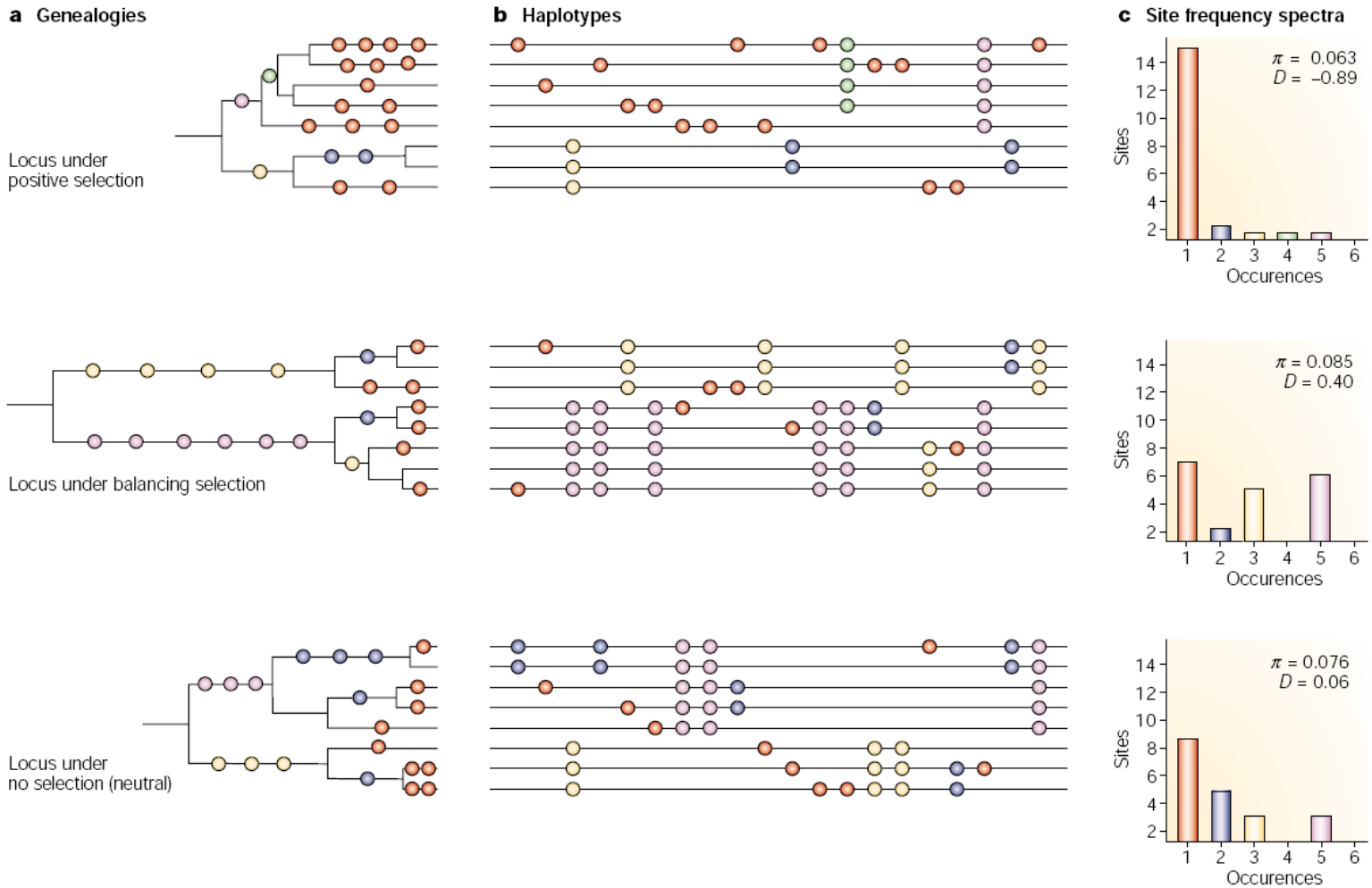


Frekvenční spektrum pro neutrální mutace

v populaci je nejvíce vzácných alel a nejméně častých alel.

Tajima's D test

- Pozitivní a negativní selekce vede ke zvýšení počtu vzácných alel. Balancing selekce vede ke zvýšení počtu středně četných alel.



Tajima's D test

- Srovnání hodnot dvou odhadů genetické diverzity, θ a π
- **Tajima's D = (π - θ)/SD (π - θ)**
- Pro neutrální sekvence: θ = π. D = 0.
- **Při nadbytku vzácných alel θ > π. D < 0.**
  (pozitivní a negativní selekce)

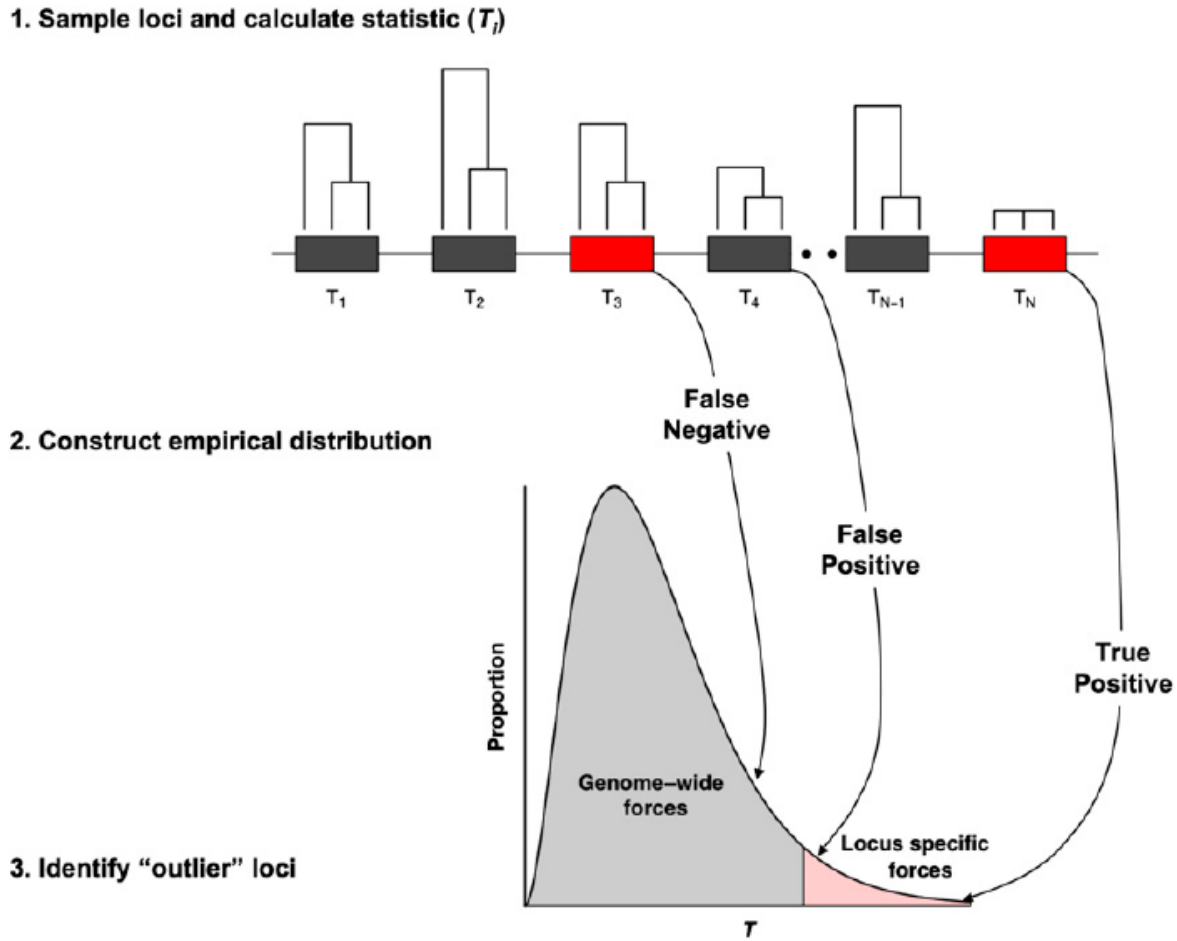
  Při nadbytku středně četných alel π > θ. D > 0.
  (balancing selekce)
- Selekci je možné detekovat jen v omezeném časovém okně (ne ihned po působení pozitivní selekce – není žádná variabilita, ani moc pozdě – dojde k obnovení variability).

POZOR! Populační expanze má podobný vliv na frekvenční spektrum jako pozitivní či negativní selekce. Naopak bottle-neck má podobný vliv jako balancing selekce.

Jak odlišit působení selekce od vlivu demografických faktorů?

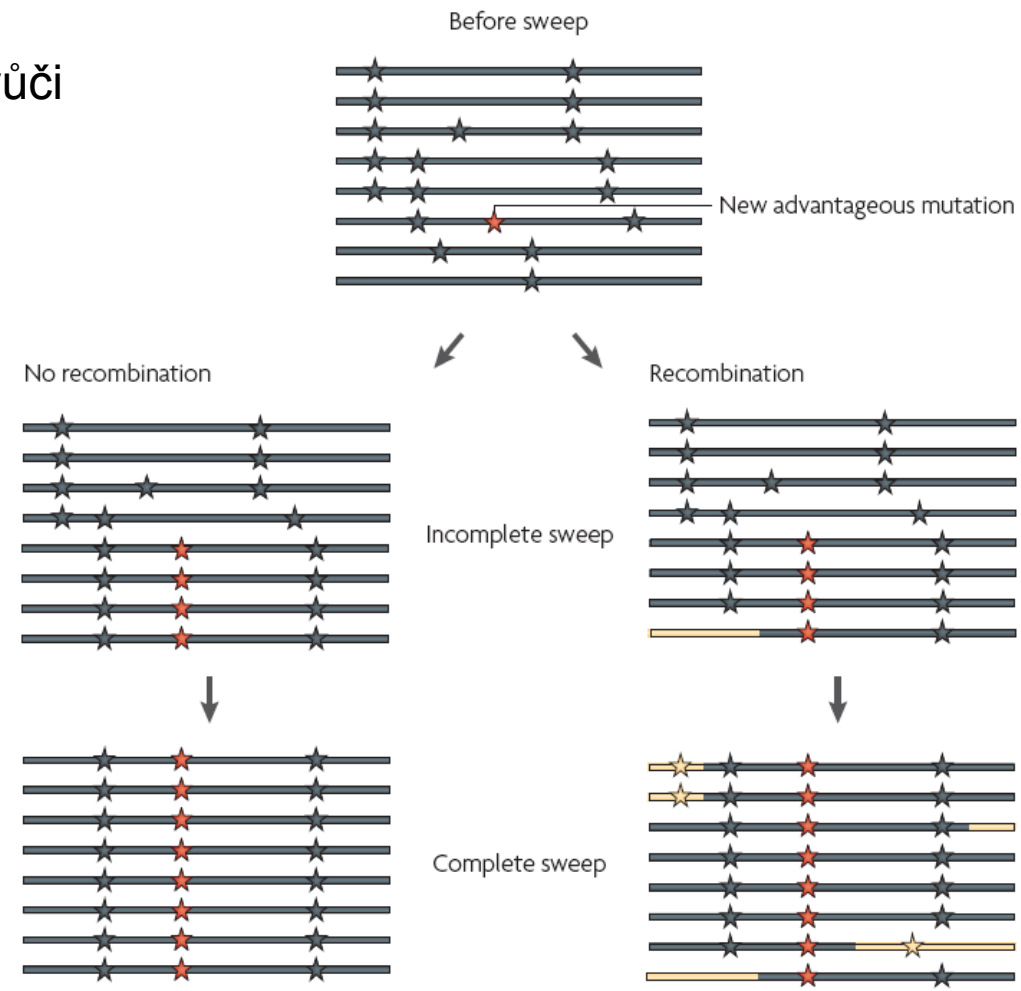
Tajima's D test pro mnoho lokusů.

- Signifikance Tajima's D statistiky testovaná na základě simulací, pomocí kterých se vytvoří distribuce D statistiky za platnosti nulové hypotézy.
  Lokusy na které působí selekce by měli vykazovat extrémní hodnoty Tajima's D.



Testy selekce založené na míře vazebné nerovnováhy, linkage disequilibrium (LD)

- Detekce v populaci četného haplotypu s vysokou mírou LD vůči ostatním haplotypům.
- Detekce velmi recentní selekce, kdy ještě nedošlo v úplné fixaci mutace. Nebo kdy k selective sweep došlo jen v jedné subpopulaci.



SweepFinder

Program na detekci selective sweeps

RASMUS NIELSEN

SweepFinder

SweepFinder is a program implementing the method described in Nielsen et al. 2005. Genomic scans for selective sweeps using SNP data. Genome Research 1566-1575. It can be used to detect the location of a selective sweep based on SNP data. It will also estimate the frequency spectrum of observed SNP data in the presence of missing data.

MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 142–156 doi: 10.1111/mec.13351

DETECTING SELECTION IN NATURAL POPULATIONS: MAKING SENSE OF GENOME SCANS AND TOWARDS ALTERNATIVE SOLUTIONS

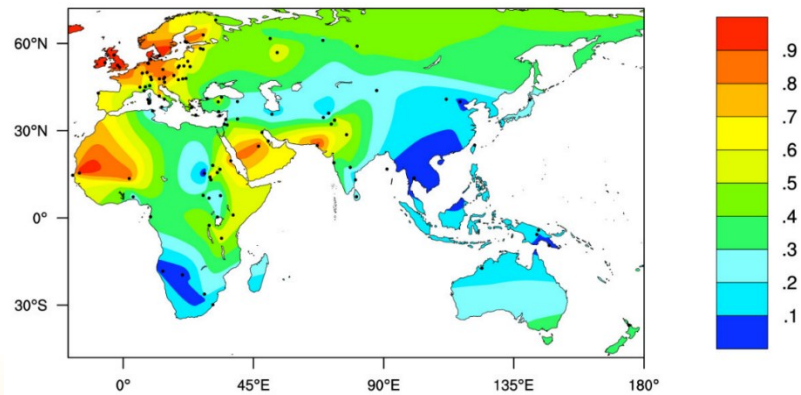
Detecting recent selective sweeps while controlling for mutation rate and background selection

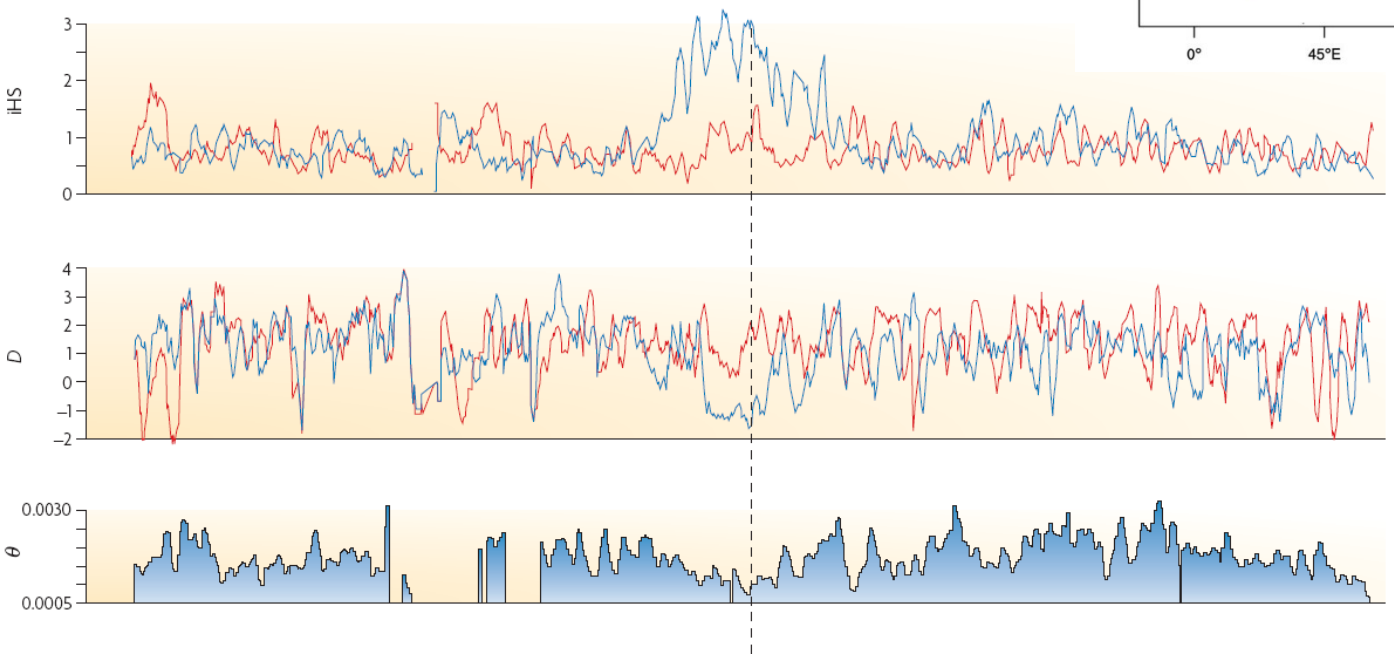
CHRISTIAN D. HUBER,*†‡ MICHAEL DEGIORGIO,§¶ INES HELLMANN** and RASMUS NIELSEN††

Laktázová persistence

- V lidské populaci vzniklo nezávisle několik mutací v LCT genu (kóduje laktázu), které umožňují trávit mléko i v dospělostí.
- Souvisí s rozšířením pastevectví.
- s = ~1,4-15% (jeden z nejsilněji selektovaných lokusů v lidském genomu).

Výskyt laktázové perzistence





Evropská populace

Asijská populace

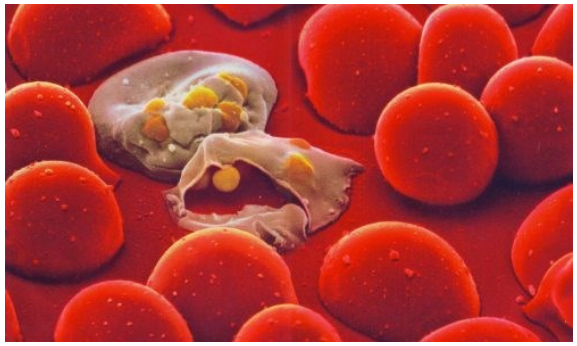
Recentní lidské adaptace

Rozšíření člověka (H. sapiens) z Afriky
(před cca 60 tis. lety).

- Adaptace na nové biotopy, nové podnebí zejména v severských zeměpisných šířkách, nové patogeny.

Přechod od lovecko-sběračské společnosti na usedlý zemědělský způsob života
(před cca 12 tis. lety).

- Změny ve stravování
- Zvýšení populační denzity – zvýšené vystavení parazitům a patogenům.



Referát: Historie lidských adaptací

APPLICATIONS OF NEXT-GENERATION SEQUENCING

Harnessing ancient genomes to study the history of human adaptation

Stephanie Marciniak[1] and George H. Perry[1,2,3]

Abstract | The past several years have witnessed an explosion of successful ancient human genome-sequencing projects, with genomic-scale ancient DNA data sets now available for more than 1,100 ancient human and archaic hominin (for example, Neandertal) individuals. Recent 'evolution in action' analyses have started using these data sets to identify and track the spatiotemporal trajectories of genetic variants associated with human adaptations to novel and changing environments, agricultural lifestyles, and introduced or co-evolving pathogens. Together with evidence of adaptive introgression of genetic variants from archaic hominins to humans and emerging ancient genome data sets for domesticated animals and plants, these studies provide novel insights into human evolution and the evolutionary consequences of human behaviour that go well beyond those that can be obtained from modern genomic data or the fossil and archaeological records alone.

Mezi pozitivně selektovanými geny jsou i geny podmiňující různé lidské choroby

- *CFTR* gen (cystická fibróza)
- ALMS1 gen (Alstromův syndrom)
- *GBA* gen *(*Gaucherova choroba*)*
- PCDH15 (Uscherův syndrom)

atd.

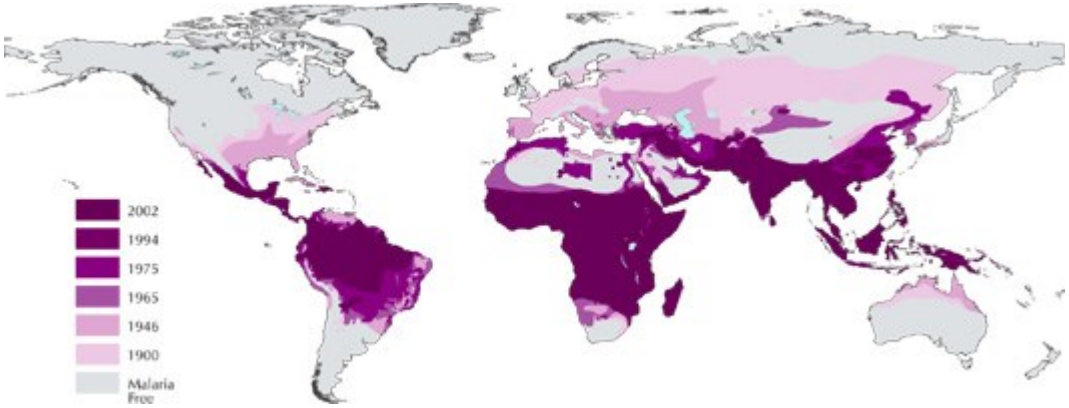
Jak může frekvenci mutací způsobujících lidské choroby ovlivňovat pozitivní selekce?

1. Mutace způsobující chorobu jsou recesivní a v heterozygotním stavu přinášejí nějakou výhodu (balancing selekce).
2. V historii došlo ke změně ve směru působení selekce

Rezistence vůči malárii

- Mutace v hemoglobinu mohou způsobit resistenci vůči malárii.
- Srpkovitá anémie. Způsobena recesivní mutací. Recesivní homozygoti nepřežijí. Heterozytoti jsou v pořádku, navíc jsou imunní proti malárii. Proto častý výskyt této choroby v tropické Africe.

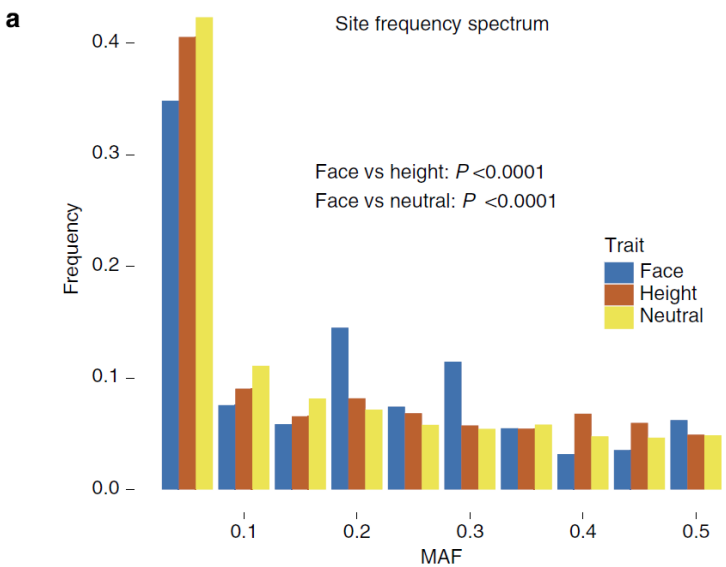
Srpkovitá anémie

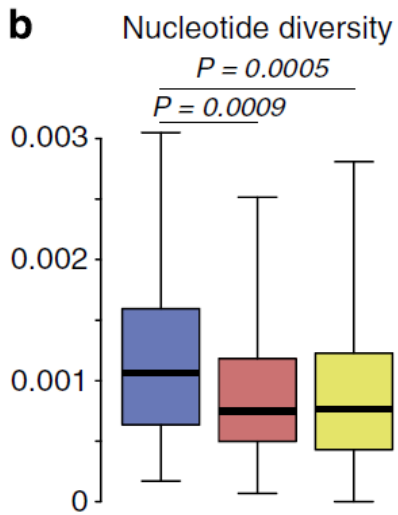


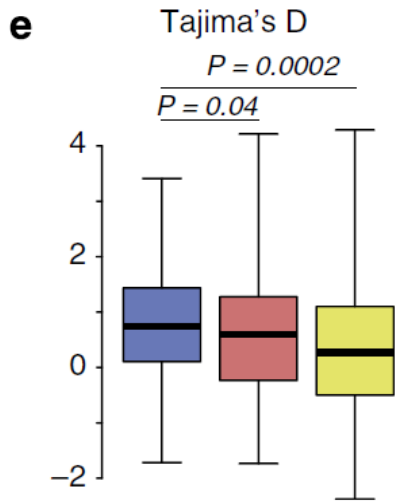
Distribuce malárie 1900 - 2002

Balancing selekce

Udržuje velkou rozmanitost lidských obličejů







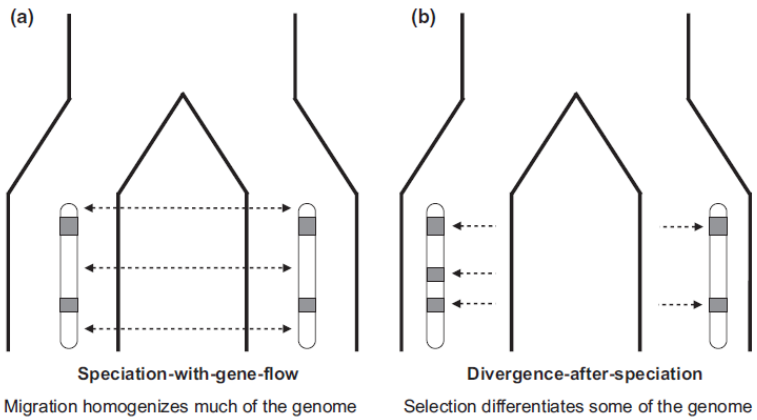
Sheeman and Nachman. 2014. Nature.

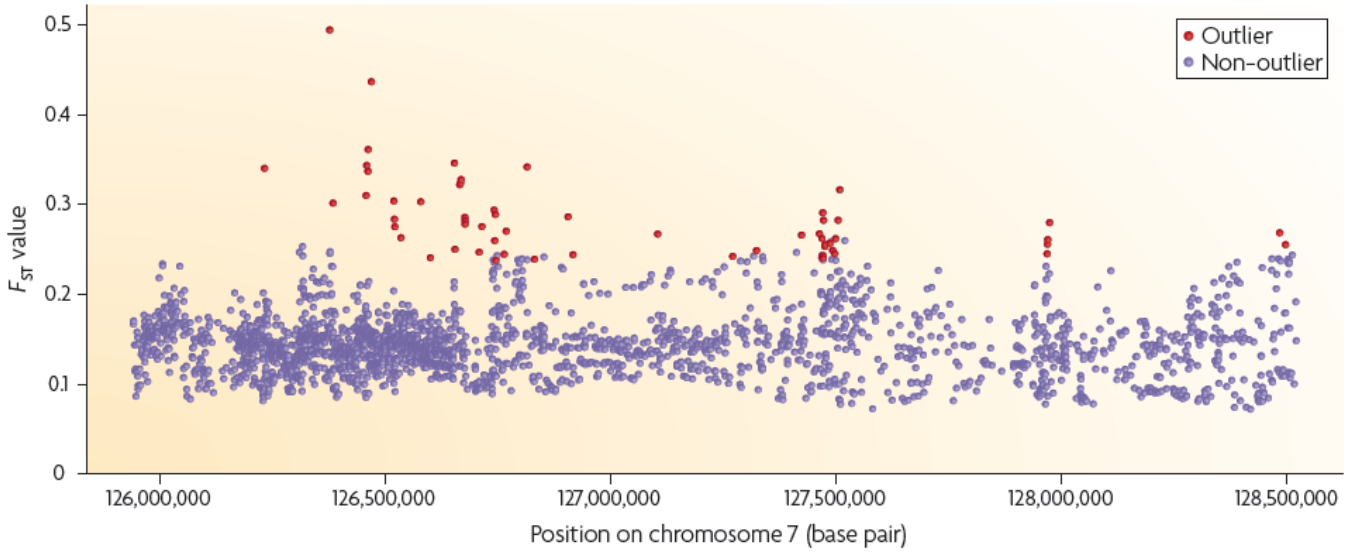
Detekce genů odpovědných za lokální adaptace

- Geny pro lokální adaptace vykazují zvýšenou míru genetické diferenciace (F_{ST}) mezi populacemi.

Ostrovy se zvýšenou genetickou diferenciací (F_{ST})

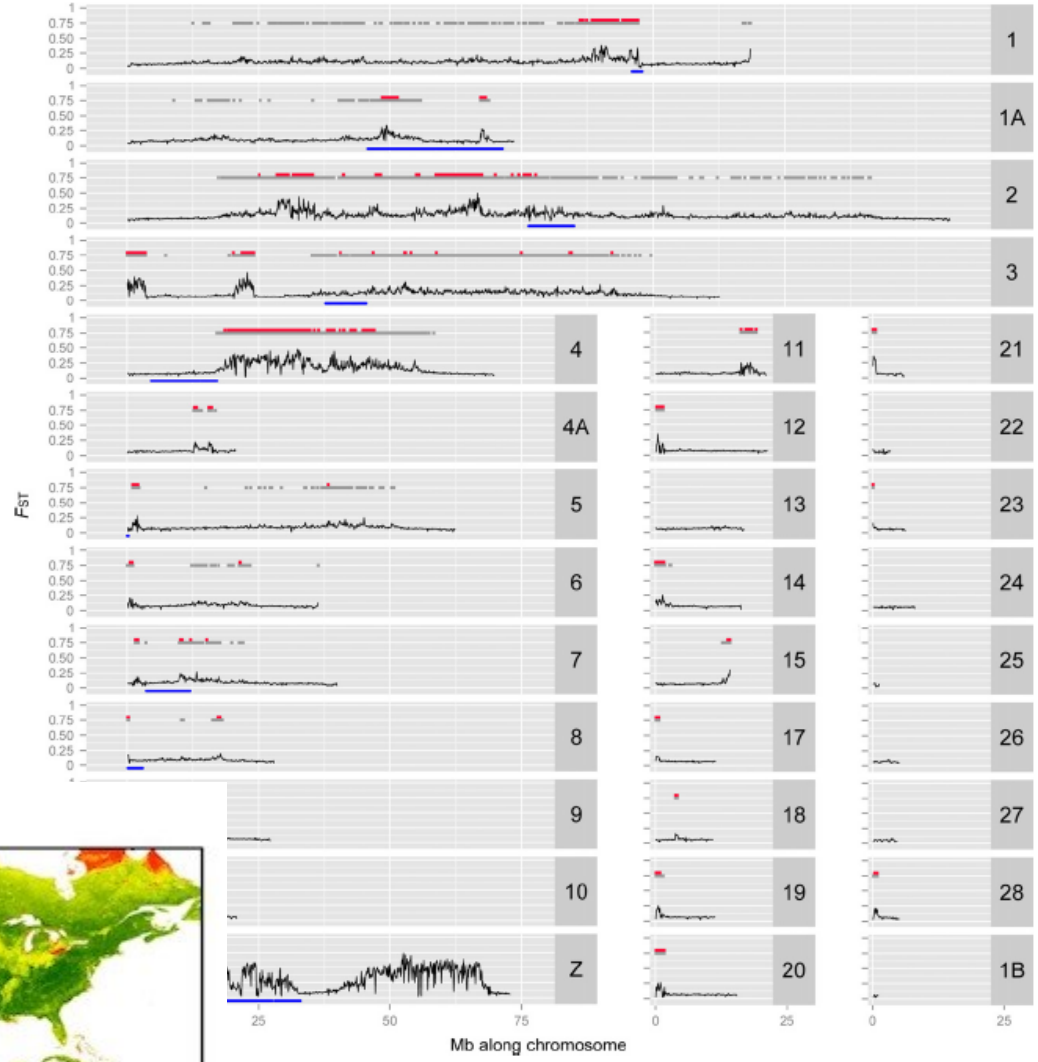
- pozitivní selekce v subpopulacích
  (snížení genetické diverzity v rámci subpopulace)
- snížený genový tok mezi subpopulacemi

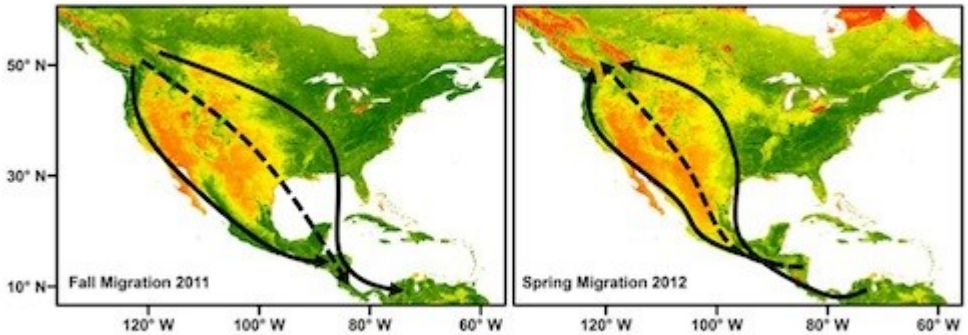




Ostrovy diferenciace u drozda malého (*Catharus ustulatus*)

- Odlišné tahové cesty u dvou poddruhů





Delmore et al. 2015

Referát: Genomické ostrovy diferenciace

Making sense of genomic islands of differentiation in light of speciation

Jochen B. W. Wolf[1,2] and Hans Ellegren[1]

Abstract | As populations diverge, genetic differences accumulate across the genome. Spurred by rapid developments in sequencing technology, genome-wide population surveys of natural populations promise insights into the evolutionary processes and the genetic basis underlying speciation. Although genomic regions of elevated differentiation are the focus of searches for 'speciation genes', there is an increasing realization that such genomic signatures can also arise by alternative processes that are not related to population divergence, such as linked selection. In this Review, we explore methodological trends in speciation genomic studies, highlight the difficulty in separating processes related to speciation from those emerging from genome-wide properties that are not related to reproductive isolation, and provide a set of suggestions for future work in this area.

Received: 5 October 2016 | Revised: 15 December 2017 | Accepted: 16 December 2017

DOI: 10.1111/mec.14479

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Genomic islands of differentiation in two songbird species reveal candidate genes for hybrid female sterility

Libor Mořkovský[1] | Václav Janoušek[1] | Jiří Reif[2] | Jakub Rídl[3] | Jan Pačes[3] | Lukáš Choleva[4,5] | Karel Janko[4] | Michael W. Nachman[6] | Radka Reifová[1]

MOLECULAR ECOLOGY

Molecular Ecology (2017)

Correlated patterns of genetic diversity and differentiation across an avian family

BENJAMIN M. VAN DOREN,*† LEONARDO CAMPAGNA,*† BARBARA HELM,‡ JUAN CARLOS ILLERA,§ IRBY J. LOVETTE*† and MIRIAM LIEDVOGEL¶

Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA, †Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA, ‡Animal Health and Comparative Medicine, Institute of Biodiversity, University of Glasgow, Glasgow G12 8QQ, UK, §Research Unit of Biodiversity (UO-CSIC-PA), Oviedo University, Campus of Mieres, Research Building, 5th Floor, c/ Gonzalo Gutiérrez Quirós s/n, 33600 Mieres, Asturias, Spain, ¶Max Planck Institute for Evolutionary Biology, AG Behavioural Genomics, August-Thienemann-Str. 2, 24306 Plön, Germany

Metody detekce dlouhodobě působící selekce

Založené na srovnání počtu nesynonymních (K_A) a synonymních (K_S) substitucí

"The Red Queen has to run faster and faster in order to keep still where she is. That is exactly what you all are doing!"

K_A/K_S test

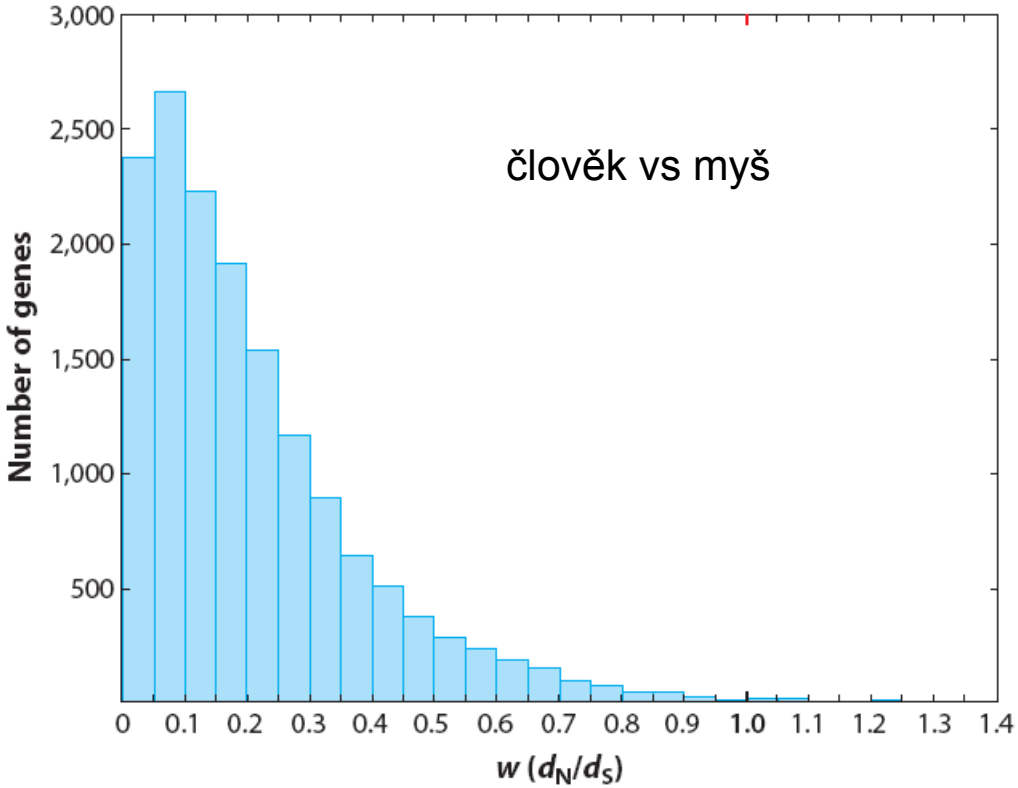
- Narozdíl od předchozích testů nevyžaduje data o genetické variabilitě v rámci druhu, ale jen data o divergenci mezi druhy.
- Detekce dlouhodobě působící pozitivní či negativní selekce
- K_A = počet nesynonymních mutací (na celkový počet nesynonymních míst)
  K_S = počet synonymních mutací (na celkový počet synonymních míst)

$K_A/K_S = 1$	neutrální evoluce
$K_A/K_S > 1$	pozitivní selekce
$K_A/K_S < 1$	negativní selekce

K_A/K_S test

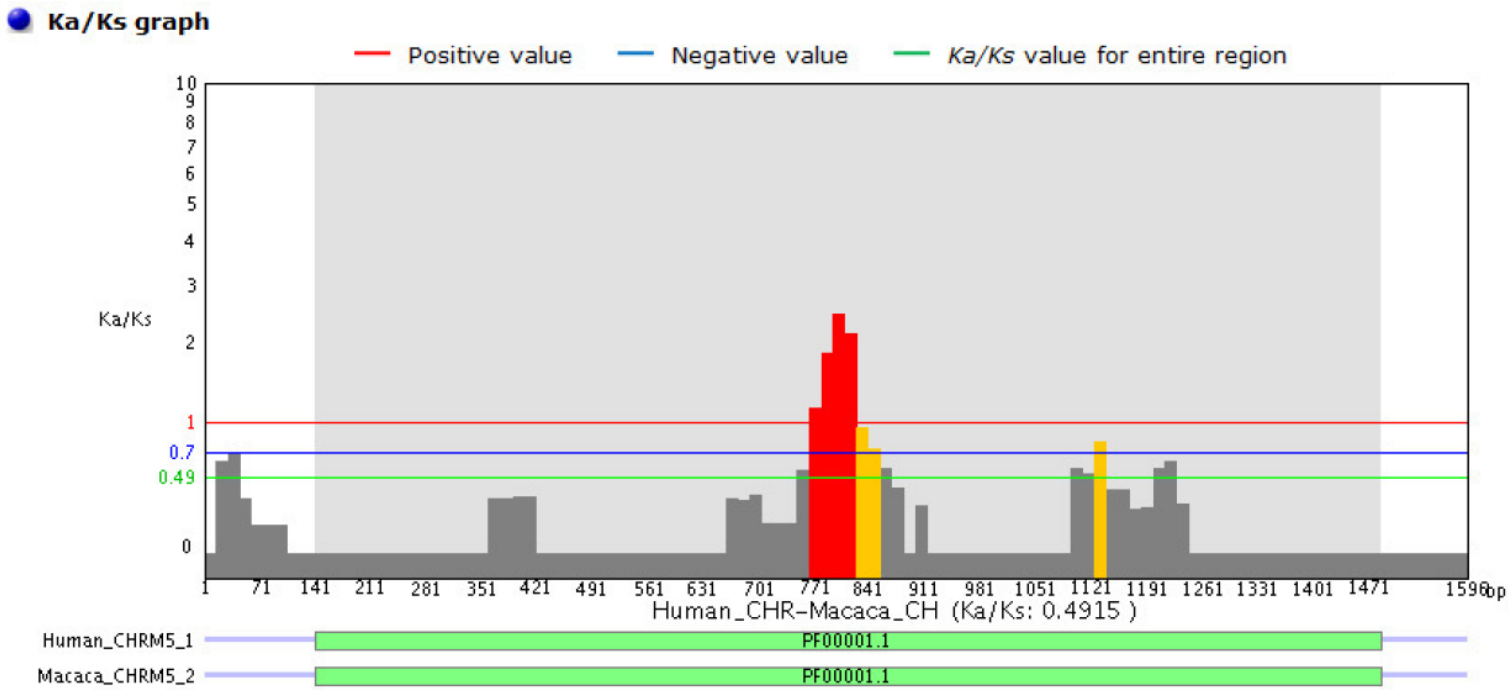
- Většina genů má K_A/K_S zhruba 0,1 – 0,2.
  Tzn. na většinu genů působí negativní selekce.

- Geny s vysokým poměrem K_A/K_S často důležitou roli v **reprodukci** (zvláště ve spermatogenezi), **imunitní odpovědi** (např. MHC geny), u savců také v **čichu** (např. OBP geny).



K_A/K_S test

- V rámci jednoho genu může na některé kodóny působit pozitivní selekce, na jiné negativní a jiné se mohou vyvíjet neutrálně.
- Lze počítat K_A/K_S pro jednotlivé kodóny.
- Užitečná je analýza „sliding window“ umožnující počítat K_A/K_S pro různé oblasti genu.



McDonald-Kreitman (MK) test

- Porovnává počet nesynonymních a synonymních mutací uvnitř druhu (P) a mezi druhy (D) pro daný gen.
- Pro neutrální lokusy by měl být poměr Ds/Ps a Dn/Pn stejný.
- Pozitivní selekce způsobí zvýšení počtu nesynonymních mutací fixovaných mezi druhy (Dn/Pn >> Ds/Ps).
- Test je velmi robustní vůči porušení demografických předpokladů
- Pozitivní MK test znamená, že selekce působí přímo na zkoumaný gen.

	fixované	polymorfní
synonymní	Ds	Ps
nesynonymní	Dn ↑	Pn ↑

Pozitivní selekce

Genetický drift

MK test umožňuje odhadnout proporci aminokyselinových substitucí fixovaných pozitivní selekcí (α).

Organismy s vyšší Ne mají vyšší odhad α.
Selekce je ve velkých populacích účinnější.

- Drosophila (Ne ~ 10^6) α > 50%
- Mus musculus castaneus (Ne ~ 5 x 10^5) α ~ 40-60%
- Topol osika (Ne ~ 10^5) α ~ 30-40%
- Mus musculus domesticus (Ne ~ 10^5) α ~ 13%
- Human (Ne ~ 10^4) α ~ 10 – 20%

Genetická podstata adaptací

Způsobují adaptace spíše mutace v kódujících oblastech genů nebo regulačních oblastech genů?

Vznikají adaptace pomocí de novo mutací nebo ze stávající genetické variability? Jakou roli hraje při vzniku adaptací mezidruhová genová introgrese?

Jakou roli hrají při adaptacích genové duplikace?

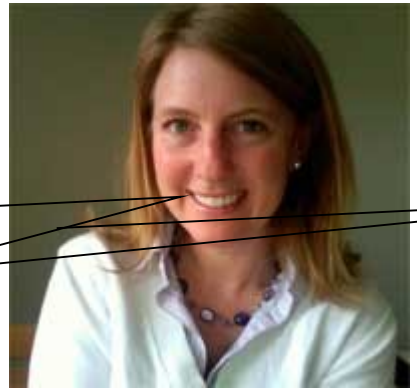
Jsou adaptace podmíněny epigenetickými změnami?

Význam kódujících vs. regulačních změn při vzniku adaptací

Sean B. Carroll

Důležitější při vzniku adaptací jsou mutace v regulačních oblastech genů, způsobující rozdíl v expresi genu. Nemají tak velké pleiotropické účinky.

Ne, ne, většina adaptací je podmíněna mutacemi v kódující části proteinu.



Hopi E. Hoekstra

Jerry A. Coyne

Hoekstra HE and Coyne JA (2007).
The locus of evolution:evo devo and the genetics of adaptation. *Evolution.*

Adaptace způsobené změnou v protein kódující oblasti genu

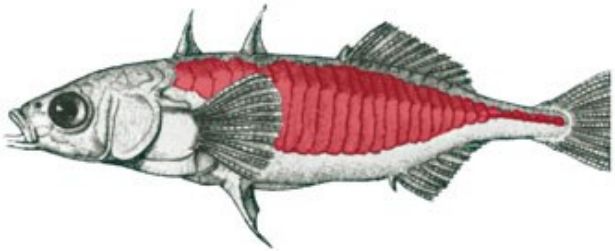
Zbarvení u pytloušů

Rezistence vůči insekticidu dieldrin u D. melanogaster

Adaptace způsobné změnou genové exprese

Skvrny na křídlech u samců Drosophily

Both coding and regulatory changes occur in the set of loci underlying marine-freshwator evolution, but regulatory changes appear to predominate.



marine stickleback

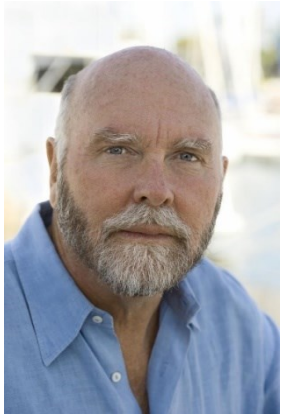
freshwater stickleback

Studium evoluce genové exprese

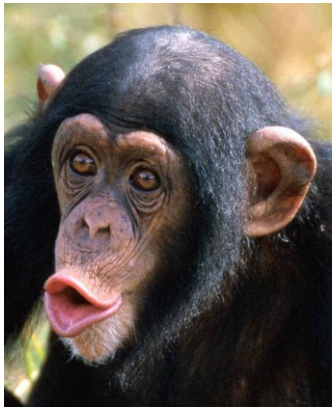
Komplikovanější než studium evoluce nukleotidové sekvence.

1. Variabilita v míře exprese ovlivněna nejen geneticky, ale i epigeneticky, velký vliv prostředí. Navíc častá chyba měření.
2. Pro každý gen máme jen jeden údaj o míře exprese.
3. Exprese jednotlivých genů na sobě závisí. Změna v expresi jednoho genu může vyvolat změnu v expresi mnoha dalších genů.
4. Změna v expresi může být způsobena změnou v cis či trans-regulační oblasti.

Variabilita v expresi častá v rámci druhů i mezi druhy



člověk



šimpanz

- Rozdíl v nukleotidové sekvenci ve ~ 4% genomu.
- Rozdíl v expresi vykazuje cca 8% genů exprimovaných v mozku, srdci, játrech a ledvinách a 32% genů exprimovaných v testes.

Khaitovich et al. 2004, Ranz et al. 2003, Nuzhdin et al. 2004

Jsou mezidruhové rozdíly v expresi neutrální či adaptivní?

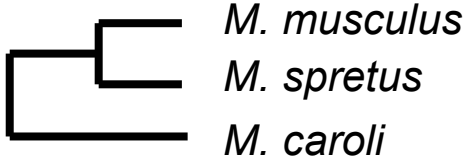
Za předpokladu neutrality

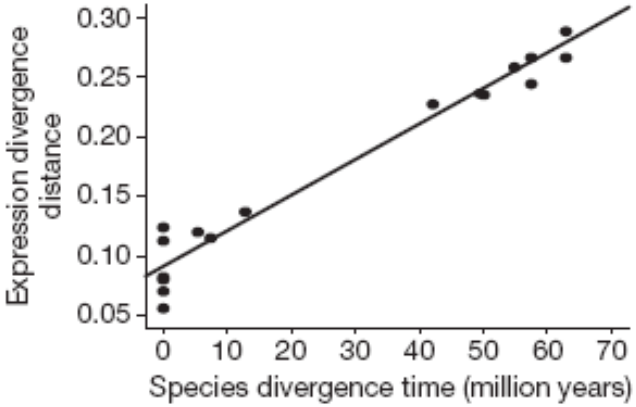
- Míra expresních rozdílů by měla vzrůstat lineárně s dobou divergence.
- Míra divergence mezi druhy by měla být přímoúměrná míře vnitrodruhové variability.
- Míra expresních rozdílů by neměla korelovat s fenotypickou divergencí (alespoň ne tou adaptivní).

Za předpokladu pozitivní selekce

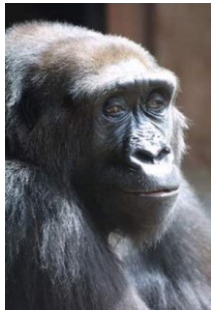
- Poměr mezidruhové divergence a vnitrodruhového polymorfismu by měl být vyšší než za předpokladu neutrality.
- Míra divergence v expresi by měla korelovat s mírou fenotypické divergence.

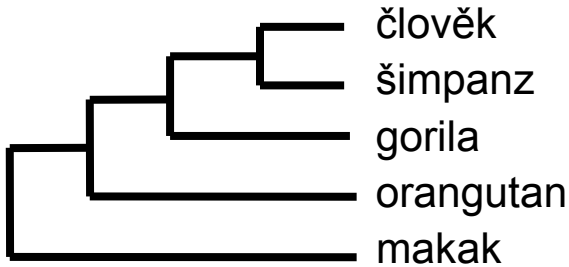
Míra expresních rozdílů vzrůstá přibližně lineárně s dobou divergence

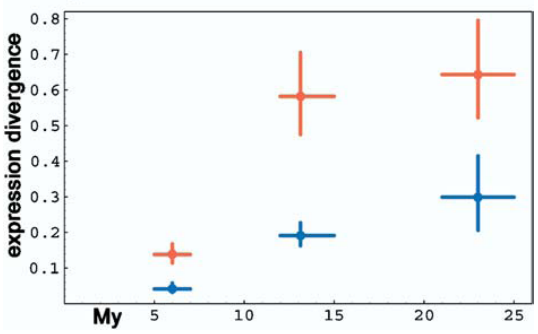




(Rifkin et al. 2003 Nat Genet
Zhang et al. 2007 Science)



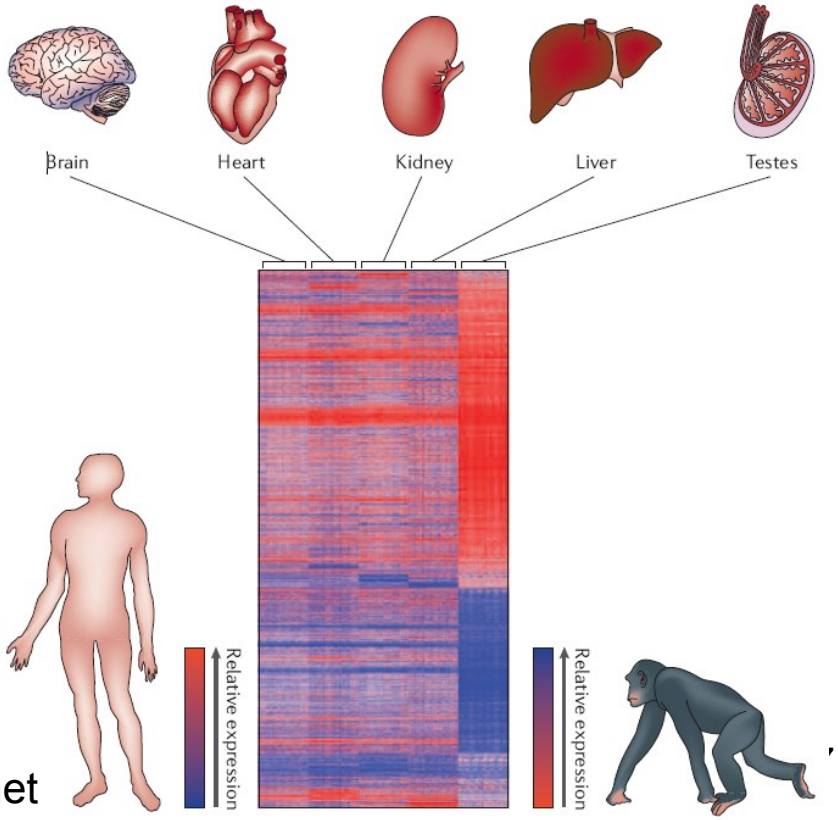




(Khaitovich et al. 2004 Plos Biol)

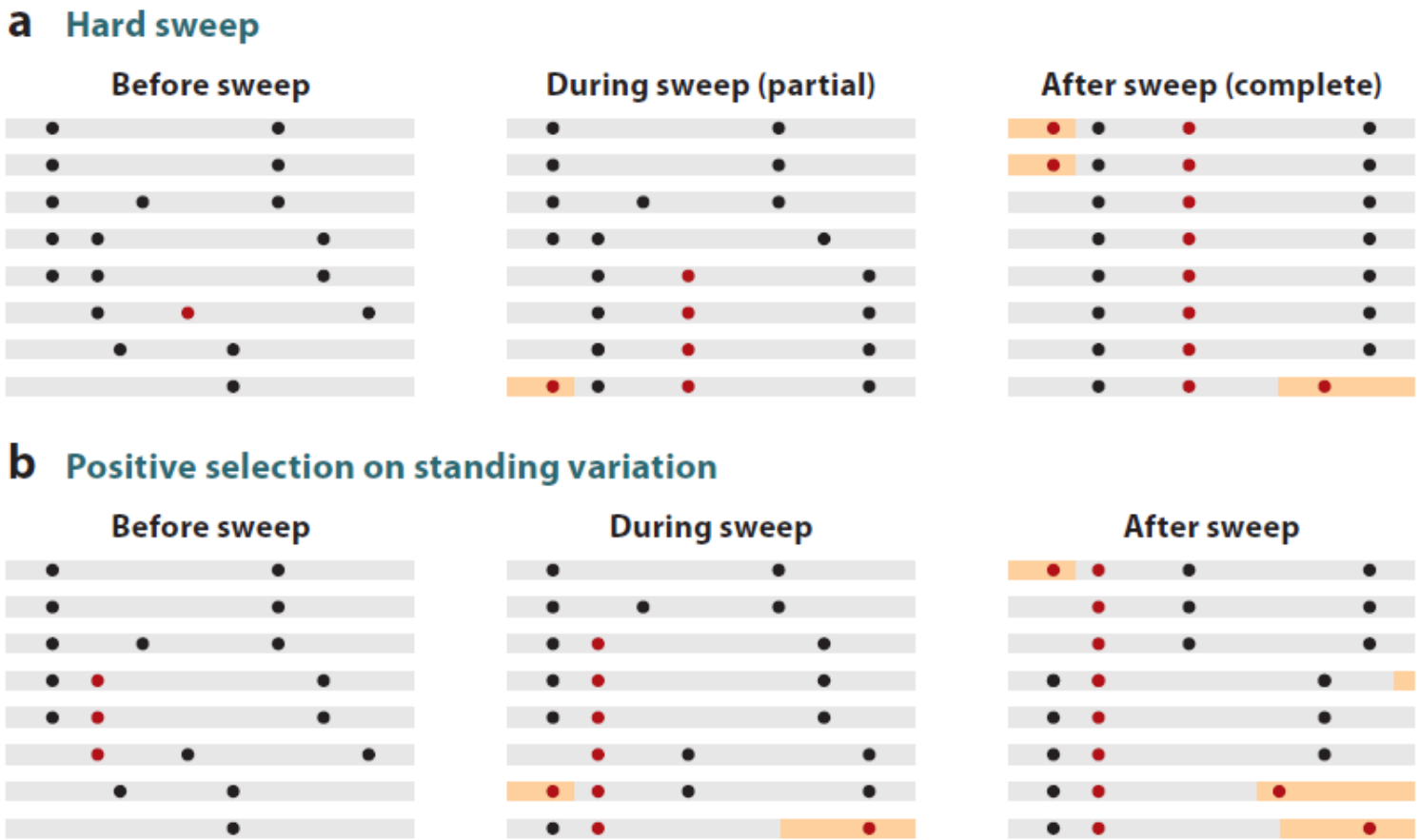
Pozitivní selekce zvyšuje mezidruhovou divergenci v expresi v testes

- Testes vykazuje zhruba třikrát vyšší poměr divergence a variability v expresi ve srovnání se somatickými tkáněmi (srdce, mozek, ledviny, játra).
- V testes často odlišná regulace transkripce ve srovnání se somatickými tkáněmi (alternativní sestřih či alternativní počátky transkripce).
- Důsledek vyšší míry pohlavního výběru v testes či meiotického tahu?



Khaitovich et al. 2006. Nat Rev Genet

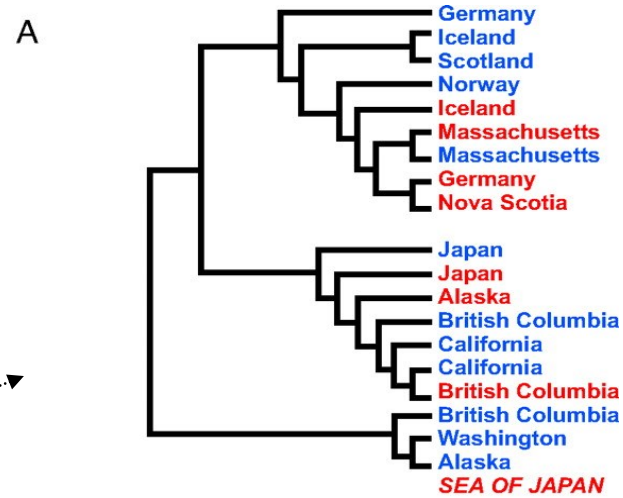
Adaptace pomocí *denovo* mutací vs. adaptace ze stávající genetické variability



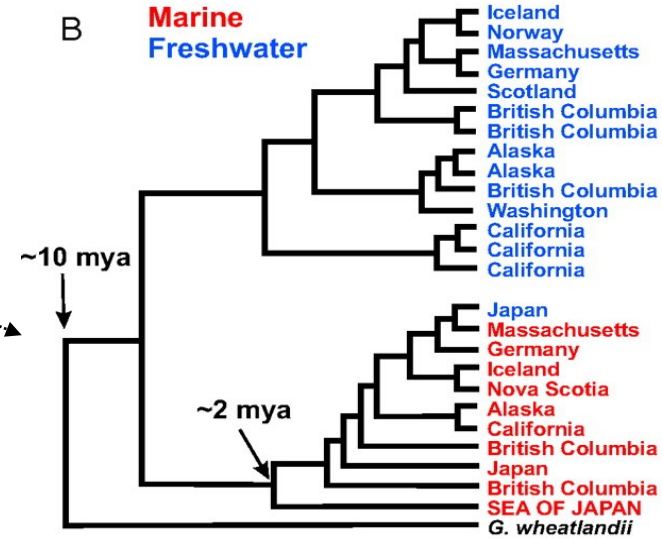
Fu and Akey, 2013

Adaptace ze stávající genetické variability a paralelní evoluce

koljuška tříostná
forma mořská (nahoře) a říční (dole)



- Říční forma vznikla mnohokrát nezávisle při kolonizaci řek.

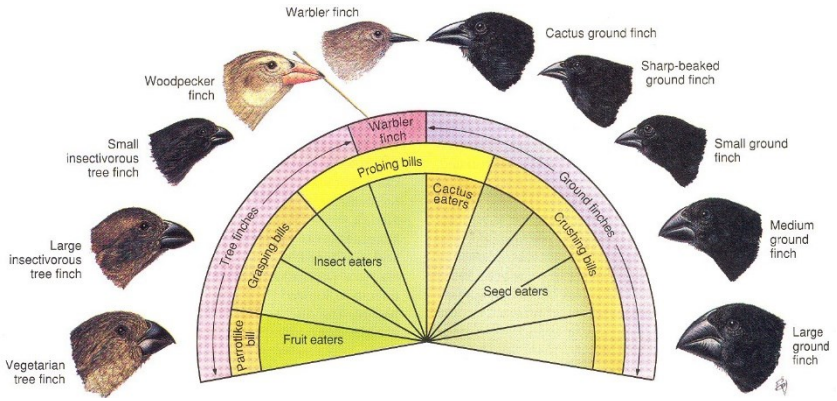


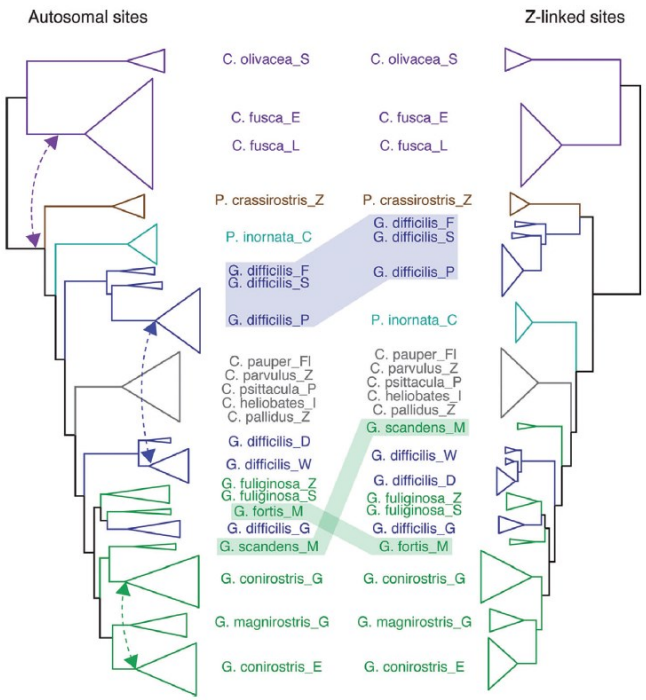
- Fylogeneze založená na genu *Ectodisplasin* (*Eda*), který odpovídá za určité rozdíly mezi mořskou a říční formou. Ukazuje, že alely způsobující říční fenotyp jsou staré a do řek se musely dostat už při kolonizaci mořskou formou.
- Platí i pro další geny odpovědné za říční fenotyp. Tyto geny často v inverzích.

Jones et al. Nature. 2012

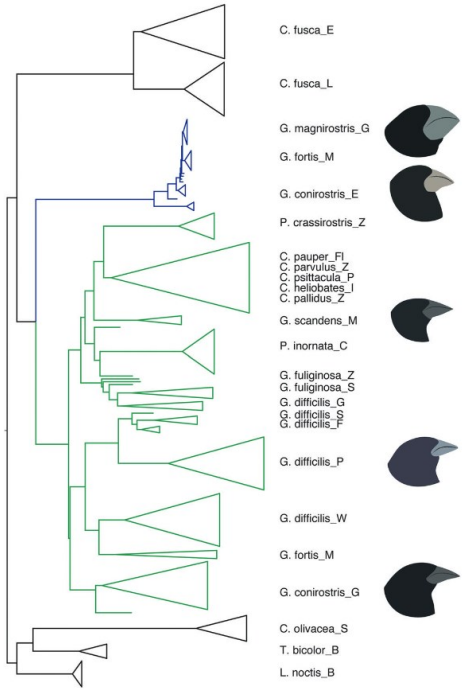
Darwinovy pěnkavy

- Adaptivní radiace začala před cca 900 000 lety.
- Hybridizace mezi druhy.
- Adaptivní introgrese genu pro velikost zobáku (ALX1)?





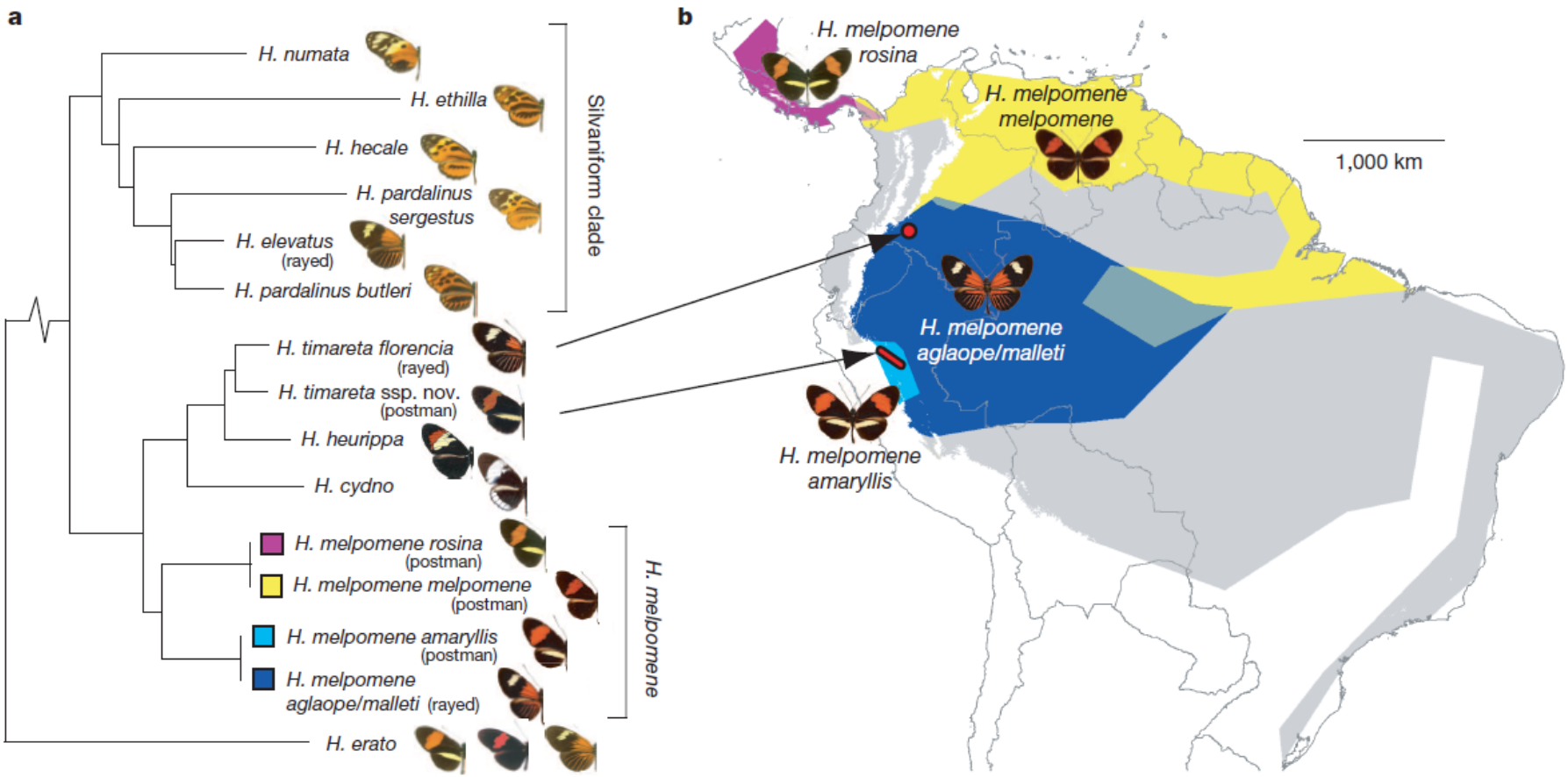
Fylogeneze Darwinových pěnkav



Haplotypový strom založený na genu ALX1

Lamichhaney et al. 2015

Adaptivní introgrese



Mezidruhová introgrese genů podmiňujících Mullerovské mimikry u motýlů rodu Heliconius

The Heliconius Genome Consorcium. Nature. 2012

Adaptivní introgrese: Co jsme zdědili od neandrtálců?

- Neafrická lidská populace 2-4% genomu původem od neandrtálců.
- V Tibetské populaci mutace v genu EPAS1 usnadňuje život ve vysokých nadmořských výškách. Původem z neandrtálců.
- U inuitů fixované alely genů WARS2 a TBX15 původem z Denisovanů. Hrají roli v adaptaci na chladné podnebí (vytváření adipocytů a tepla).
- Adaptivní introgrese se také týká genů ovlivňujících kvalitu a zbarvení kůže a srsti a dále genů imunitního systému.
- Některé alely získané od neandrtálců způsobují choroby (např. obezitu, deprese).

Referát: Důsledky mezidruhové introgrese do lidského genomu

genes

MDPI

Review

Evolutionary and Medical Consequences of Archaic Introgression into Modern Human Genomes

Olga Dolgova * and Oscar Lao *

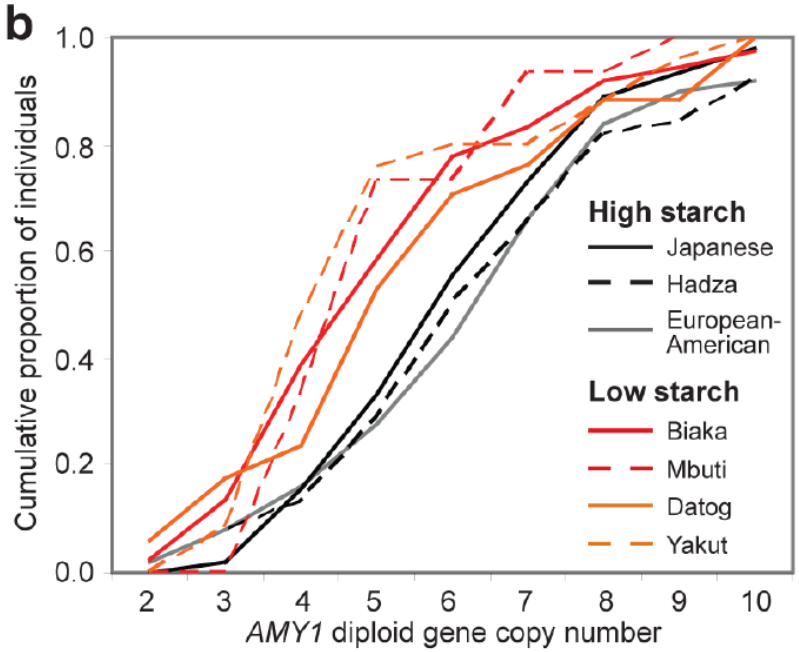
Population Genomics Group, Centre Nacional d'Anàlisi Genòmica, Centre de Regulació Genòmica (CRG-CNAG), Parc Científic de Barcelona, Baldiri Reixac 4, 08028 Barcelona, Catalonia, Spain

* Correspondence: olga.dolgova@cnag.crg.eu (O.D.); oscar.lao@cnag.crg.eu (O.L.); Tel.: +34-934-037-276 (O.D. & O.L.)

Adaptivní evoluce genovou duplikací

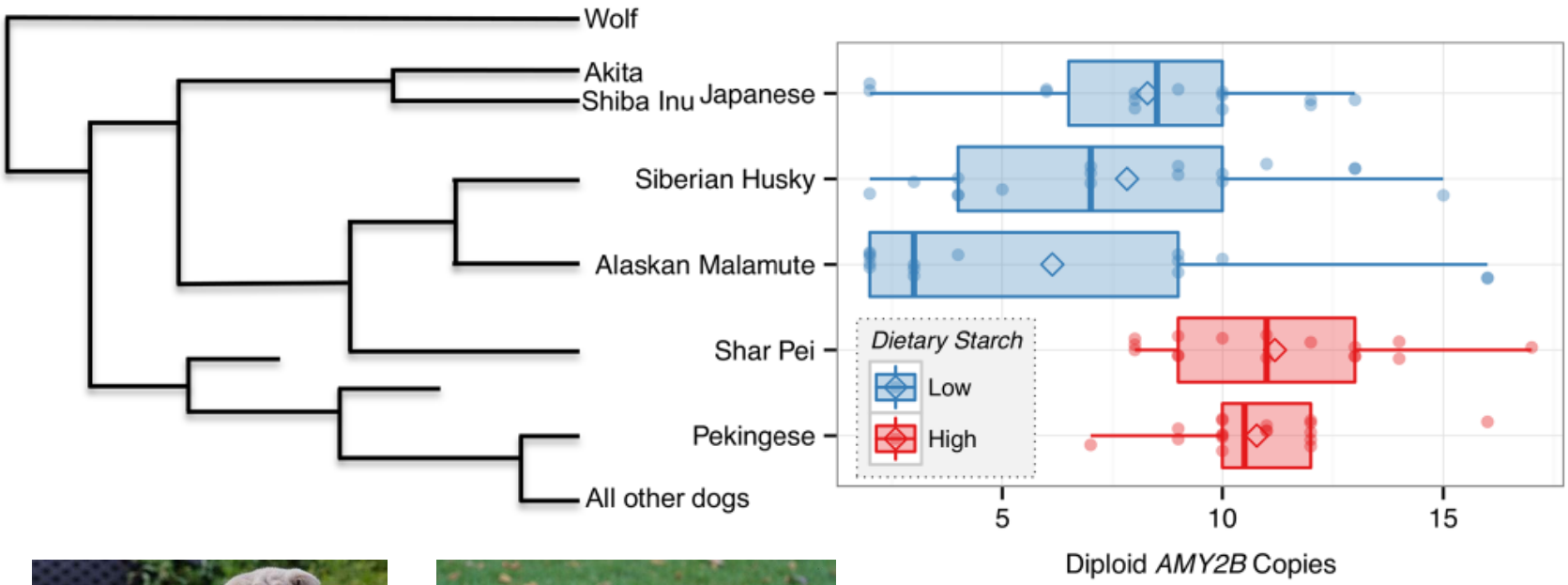
Salivary amylase gene (*AMY1*)

- Odpovědný za štěpení škrobu.
- Jeden z nejvariabilnějších genů co do počtu kopií v lidském genomu.
- Počet kopií AMY1 genu v genomu koreluje s množstvím enzymu ve slinách.
- Populace s vysokým množstvím škrobu v potravě (tj. živící se především zemědělstvím) mají více kopií AMY1 genu



Perry et al. 2007, Nature Genetics

Počet kopií genu pro amylázu (AMY2B) souvisí s typem potravy i u domestikovaných psů



Reiter et al. 2016

Epigenetické změny a adaptace

Referát: Jak tetry slepé přišly o oči

nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0569-4

An epigenetic mechanism for cavefish eye degeneration

Aniket V. Gore[1*], Kelly A. Tomins[1], James Iben[2], Li Ma[3], Daniel Castranova[1], Andrew E. Davis[1], Amy Parkhurst[1], William R. Jeffery[3] and Brant M. Weinstein[1*]

Coding and non-coding mutations in DNA contribute significantly to phenotypic variability during evolution. However, less is known about the role of epigenetics in this process. Although previous studies have identified eye development genes associated with the loss-of-eyes phenotype in the Pachón blind cave morph of the Mexican tetra *Astyanax mexicanus*, no inactivating mutations have been found in any of these genes. Here, we show that excess DNA methylation-based epigenetic silencing promotes eye degeneration in blind cave *A. mexicanus*. By performing parallel analyses in *A. mexicanus* cave and surface morphs, and in the zebrafish *Danio rerio*, we have discovered that DNA methylation mediates eye-specific gene repression and globally regulates early eye development. The most significantly hypermethylated and downregulated genes in the cave morph are also linked to human eye disorders, suggesting that the function of these genes is conserved across vertebrates. Our results show that changes in DNA methylation-based gene repression can serve as an important molecular mechanism generating phenotypic diversity during development and evolution.

Referáty:

Genomické ostrovy diferenciace

Důsledky mezidruhové introgrese do lidského genomu