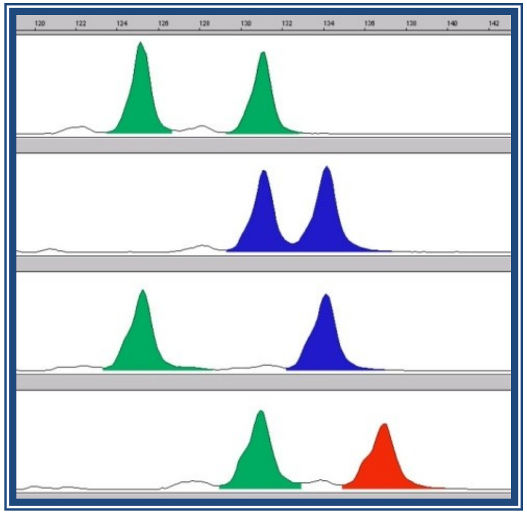
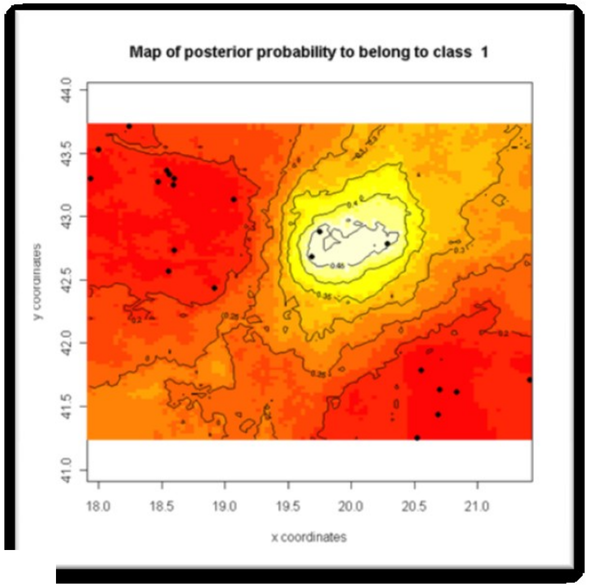
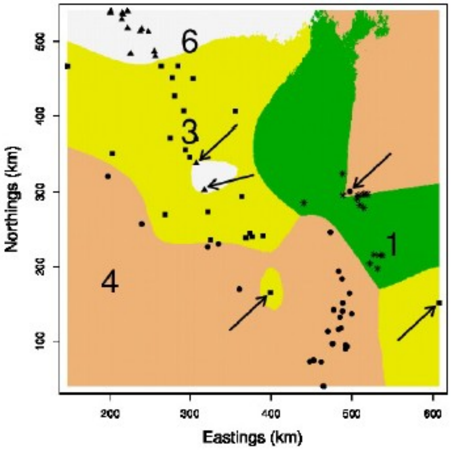
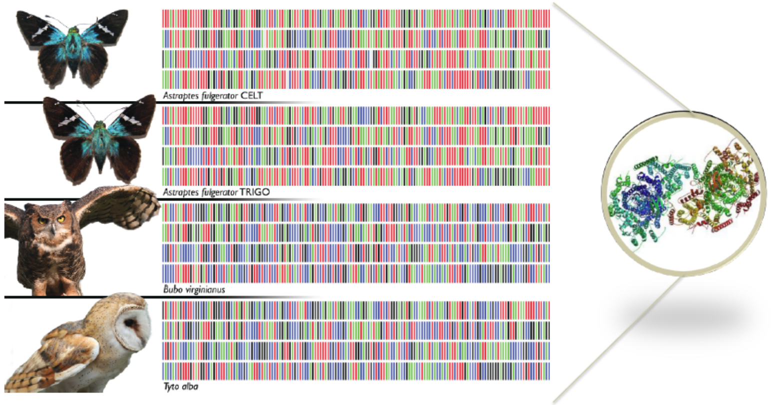
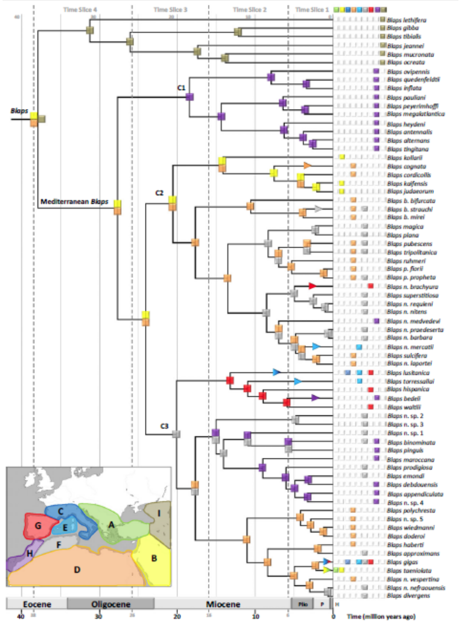








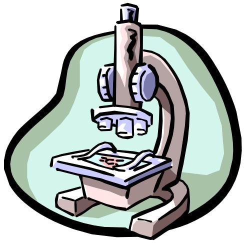


Molekulární ekologie

J. Bryja, M. Macholán, A. Konečný

Co je molekulární ekologie?

Uměle vytvořený obor vymezený technickým přístupem. Na ekologické a evoluční problémy hledá odpověď na základě molekulárně-genetických dat.



Klasické problémy a metody ekologie
(zoologie, evoluční biologie)

+

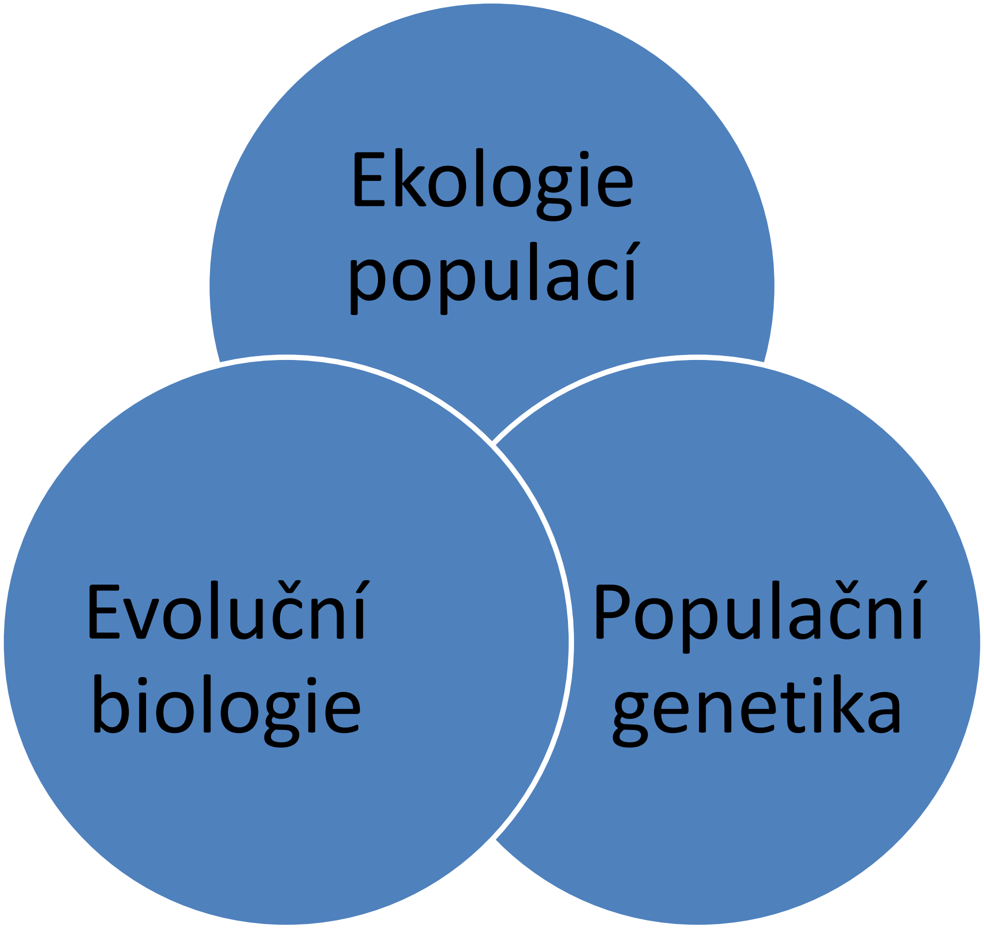
Molekulárně-genetická data a populačně-genetické analýzy

(Zoologové a botanici nakoupili cyklery a sekvenátory, snažili se je využít i k něčemu jinému než je fylogenetická analýza => vznikla molekulární ekologie)

Pracuje na různých úrovních variability DNA (genom, jedinec, populace, skupina populací, max. skupina blízce příbuzných druhů)

*Je to vlastně aplikovaná populační genetika – **analyzuje a interpretuje** získaná molekulárně-genetická data*

Mezioborová disciplína



Ekologie populací

Evoluční biologie

Populační genetika

Molekulární ekologie

Mechanismy mikroevoluce

(teorie)

+

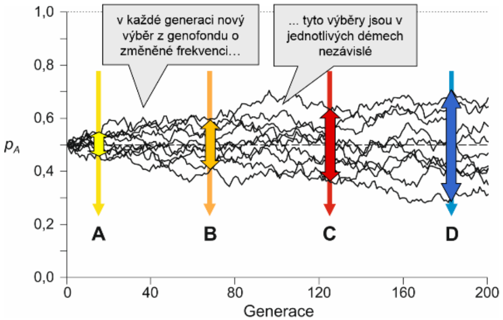
Genetické metody v zoologii

(praxe – data)

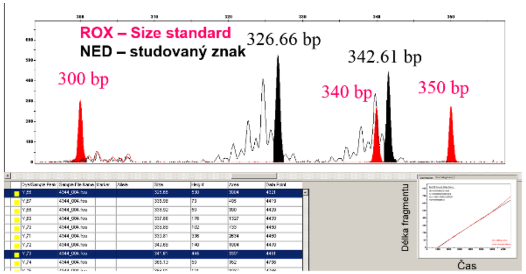
=

Molekulární ekologie

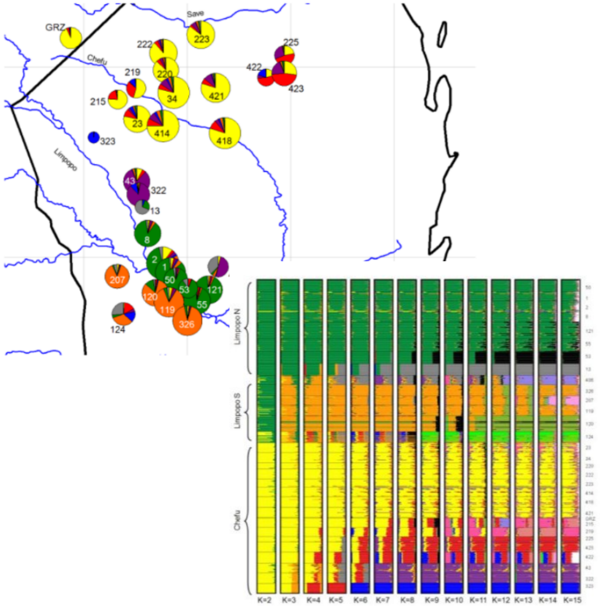
(praktické postupy a interpretace výsledků)



Závěr 9: Drift vede k divergenci mezi démy.



Genotyp mikrosatelitu na lokusu NED = 326/342 nebo 327/343
Programy: GeneMapper, Genotyper, Geneious, GeneMarker, ...



Příbuzné přednášky, tj. co se zde objeví jen okrajově?

- M. Macholán - Evoluční biologie + Mechanismy mikroevoluce
- J. Bryja, M. Macholán - Genetické metody v zoologii
- J. Zukal – Behaviorální ekologie
- S. Pekár – Ekologie populací
- aj. (molekulární ekologie „prorůstá všude“)

Její význam vzrůstá ...

- Je populární – časopis Molecular Ecology (od 1992) – dnes 24 čísel za rok
- Vyšly i její učebnice
- Na řešení velmi odlišných problémů používá obdobné metody

MOLECULAR ECOLOGY

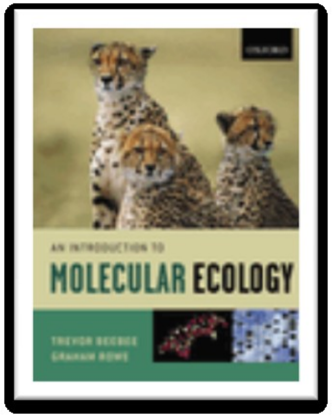
Editor-in-Chief: Loren Rieseberg

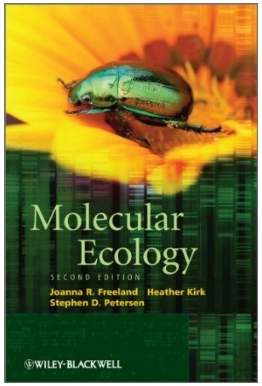
Impact factor: 6.185

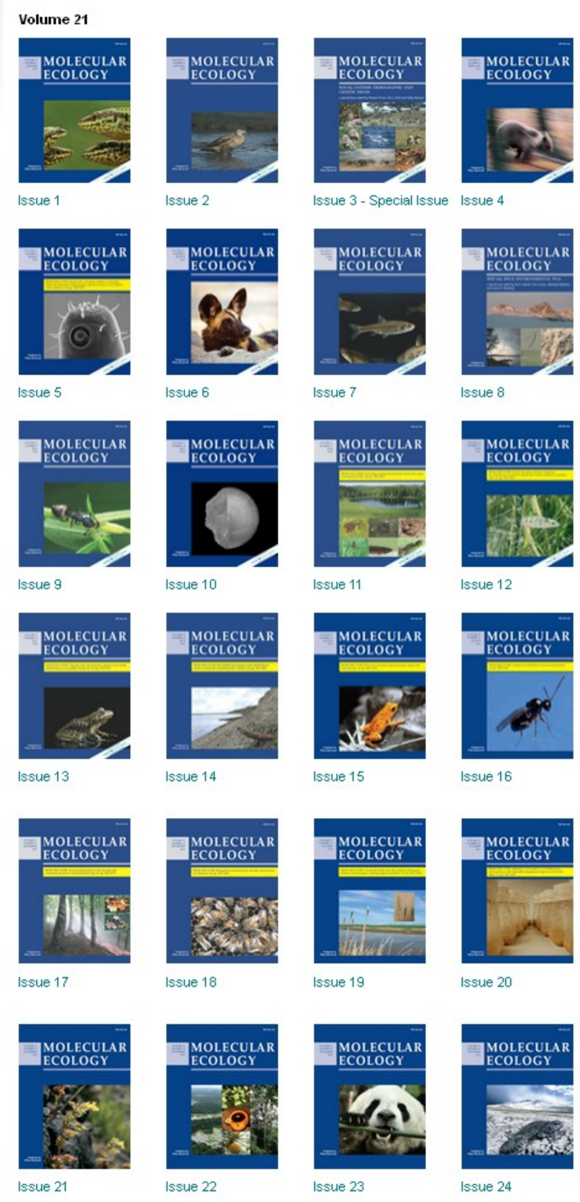
2020 Journal Citation Reports (Clarivate Analytics): 62/297 (Biochemistry & Molecular Biology) 16/166 (Ecology) 7/50 (Evolutionary Biology)

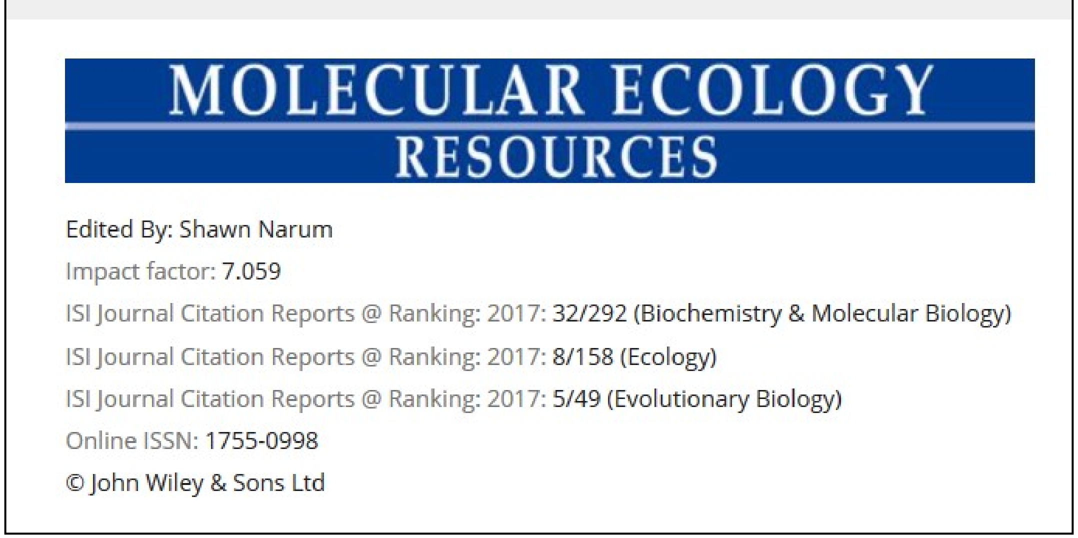
Online ISSN: 1365-294X

© John Wiley & Sons Ltd









MOLECULAR ECOLOGY
RESOURCES

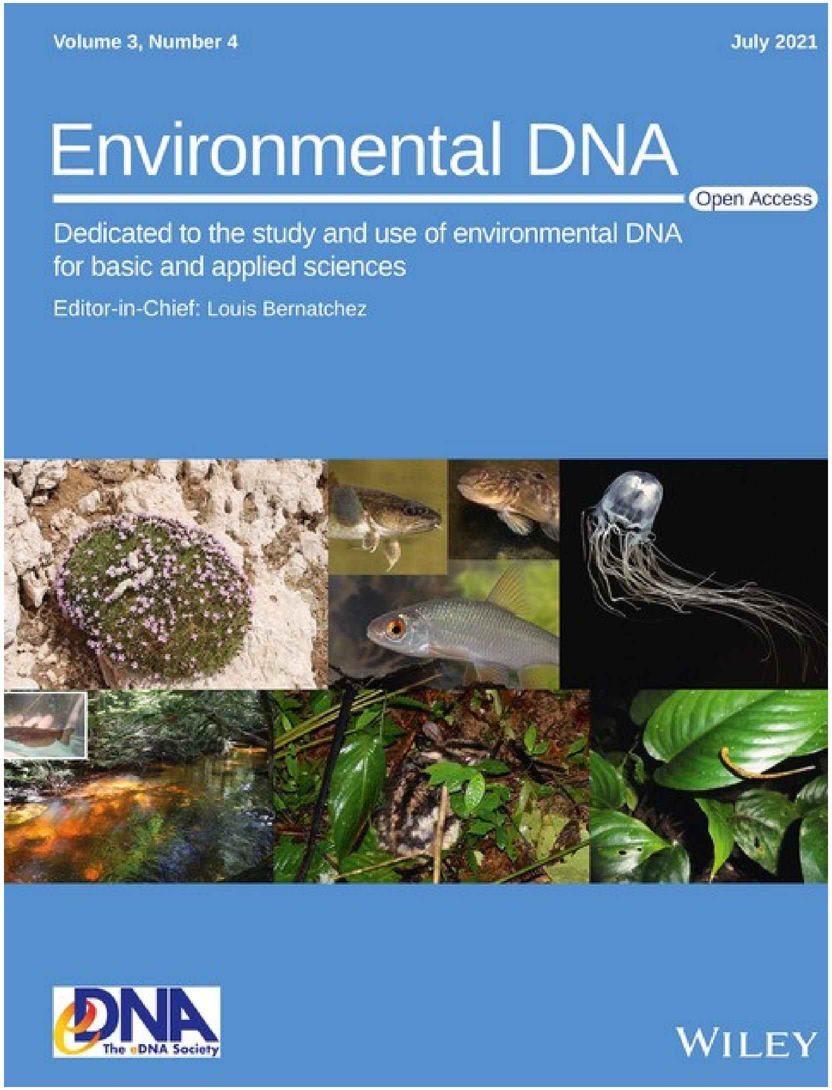
Edited By: Shawn Narum
Impact factor: 7.059
ISI Journal Citation Reports @ Ranking: 2017: 32/292 (Biochemistry & Molecular Biology)
ISI Journal Citation Reports @ Ranking: 2017: 8/158 (Ecology)
ISI Journal Citation Reports @ Ranking: 2017: 5/49 (Evolutionary Biology)
Online ISSN: 1755-0998
© John Wiley & Sons Ltd

Molecular and Statistical Advances
These papers primarily present **new techniques for collecting and analysing data for molecular ecology** studies. These submissions can also describe meaningful comparisons of statistical, computational or molecular methods, or alternatively demonstrate the existence of important problems with current procedures.

Computer Programs
These articles typically present **new computer software or substantial updates of existing programs**. Authors should clearly describe the need for the program and the rationale behind its design, as well as a summary of functions, usage and output. When relevant, authors should present an evaluation of a program's performance compared to existing software based on real or simulated data. Programmers are urged to remember that 'user-friendly' programs are more likely to be used by the community, and that it is helpful if programs require standard input/output file formats (e.g., Genepop or Arlequin). Submissions should include a user manual or README file with adequate guidance for new users. Software and supporting documentation should be accessible from a long-term server (e.g., github), but can additionally be made available at academic websites.

Permanent Genetic Resources
These articles describe the development of **significant genetic resources for application to evolutionary or ecological questions**. For example, these papers could describe NGS projects in which **sequenced transcriptomes, genome fractions or whole genomes** have been analyzed such that a readily usable resource is presented to the molecular ecology community. These articles may also present data on novel applications of the **standard DNA barcoding loci to a hundred or more species**, where the paper presents a readily usable resource.

+ **Invited technical reviews, Opinions**, etc.



Volume 3, Number 4

July 2021

Environmental DNA

Open Access

Dedicated to the study and use of environmental DNA for basic and applied sciences

Editor-in-Chief: Louis Bernatchez

The eDNA Society

WILEY

- **Experimental eDNA work:** Testing the impact of physico-chemical factors (e.g. natural biogeochemistry and PCR pollutants) on eDNA, degradation, transport, shedding and detection rate, comparing detectionand abundance estimate with conventional methods.
- **Trophic and community ecology:** Ecosystem dynamics, functional diversity, predator-prey interactions (e.g. diet analysis), host-associated microbiota.
- **Palaeo-environments:** Past species and community diversity and abundance measurements, inference in space and time.
- **Biomonitoring, conservation biology:** Single- and multi-species detection, comprehensive biodiversity at different scales, abundance estimates, detection of rare, cryptic and endangered species, non-invasive sampling, management (e.g. fisheries), occurrence and detection estimates.
- **Invasion biology:** Early species detection at low abundance, passive surveillance, impacts on ecosystems, vectors and pathways of dispersal.
- **Environmental assessment:** Impacts of pollutants and other environmental disturbance on species and communities, microbial source tracking (fecal bacteria or pathogens).
- **Physical eDNA properties:** Uptake and transformation based on geochemistry, particles, organic chemistry or microbial community.
- **Techniques and methods:** Engineering development, developing, testing and evaluating eDNA biotechnology and biostatistical approaches.
- **Applications in citizen science and biodiversity education**

Proč používat molekulární metody v ekologii?

- **Často nelze jinak**
- paternita – páření často skryté a nemusí vést k oplození
- identifikace z trusu, chlupů - pohyb jedinců skrytě žijících druhů
- izolace populací – nemusí být zřejmá
- počet migrantů – nelze sledovat naráz všechny jedince
- adaptace – nejsou zřejmé na první pohled (např. imunitní geny)

Proč používat molekulární metody v ekologii?

- ecological, evolutionary, and population genomics
- population structure and phylogeography
- landscape genomics
- community ecology and coevolution
- reproductive strategies
- relatedness and kin selection
- sex allocation
- population genetic theory
- analytical methods development
- conservation genetics
- speciation and hybridization
- microbial biodiversity
- evolutionary dynamics of QTLs
- ecological interactions
- molecular adaptation and environmental genomics
- impact of genetically modified organisms

MOLECULAR ECOLOGY

Editor-in-Chief: Loren Rieseberg

Impact factor: 6.185

2020 Journal Citation Reports (Clarivate Analytics): 62/297 (Biochemistry & Molecular Biology) 16/166 (Ecology) 7/50 (Evolutionary Biology)

Online ISSN: 1365-294X

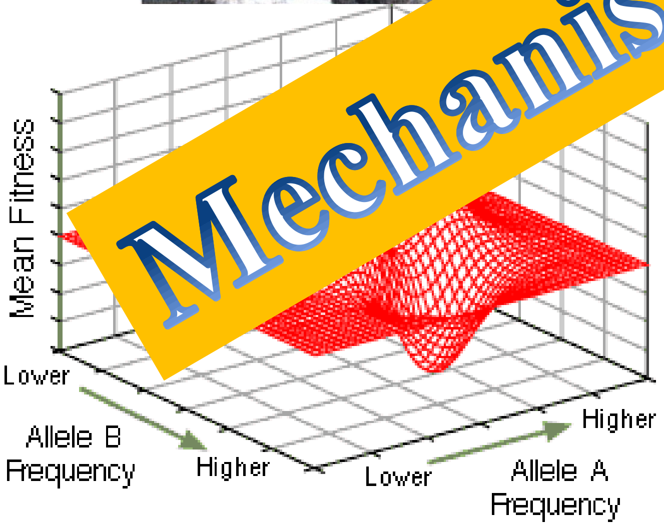
© John Wiley & Sons Ltd

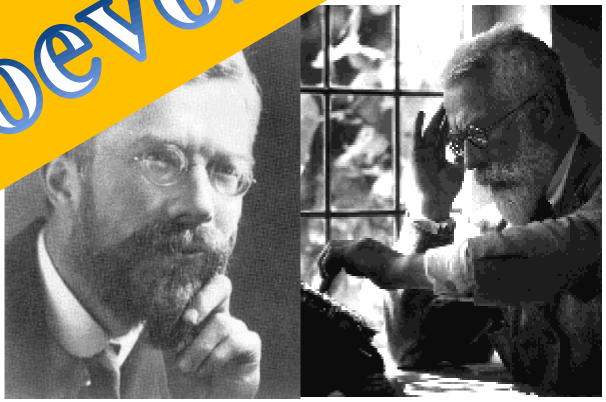
Vychází z populační genetiky

- Slavní zakladatelé moderní syntézy, třicátá léta
- Matematické modely spojující genetiku a evoluční teorii

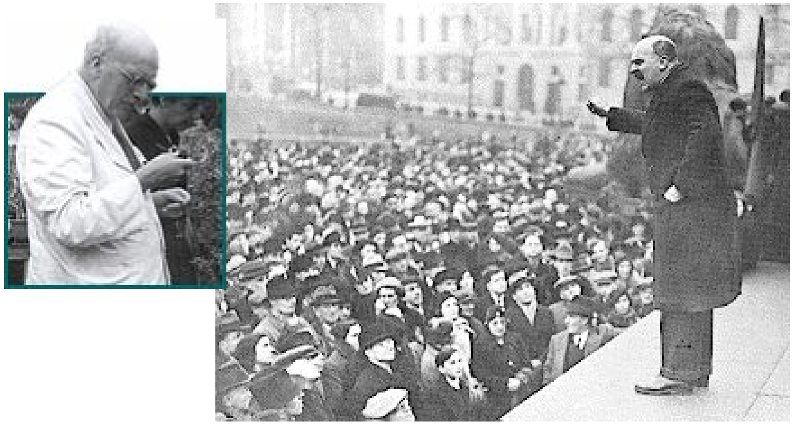
Mechanismy mikroevoluce

Sewall Wright
adaptivní krajina



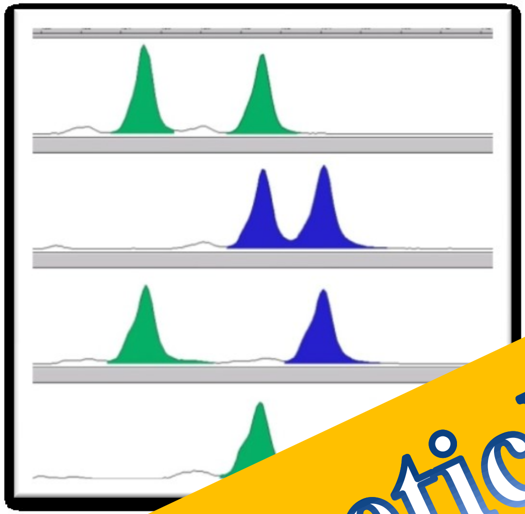


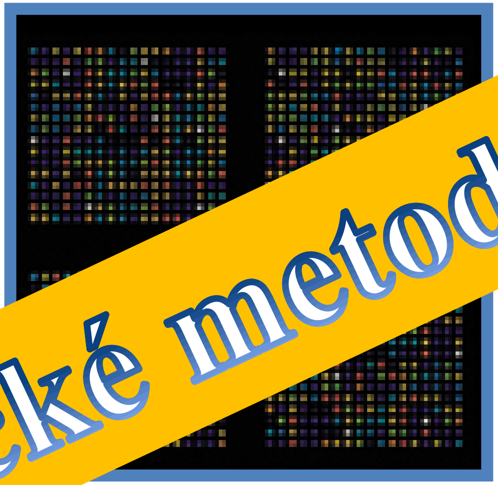
John B. S. Haldane



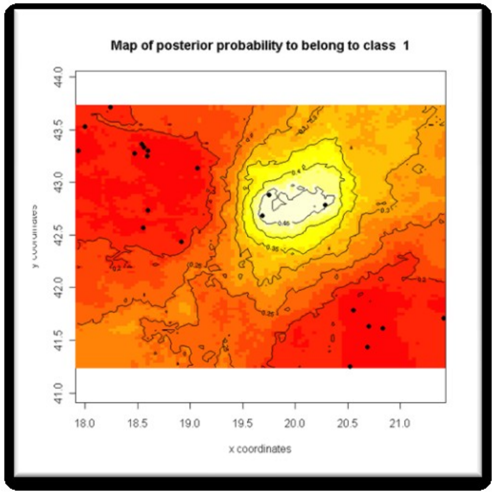
Genetické metody v zoologii

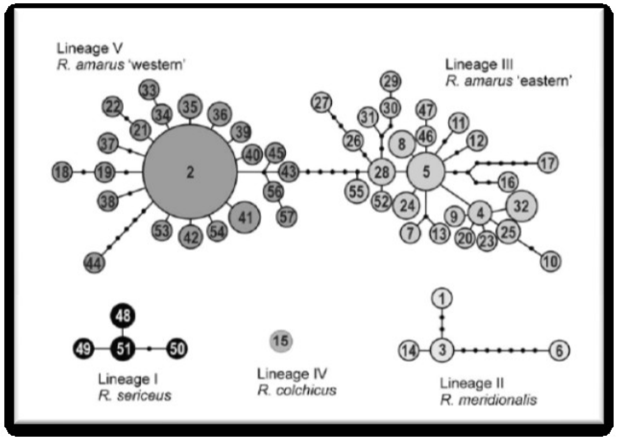
Technické výlety (omezen





Analýza dat



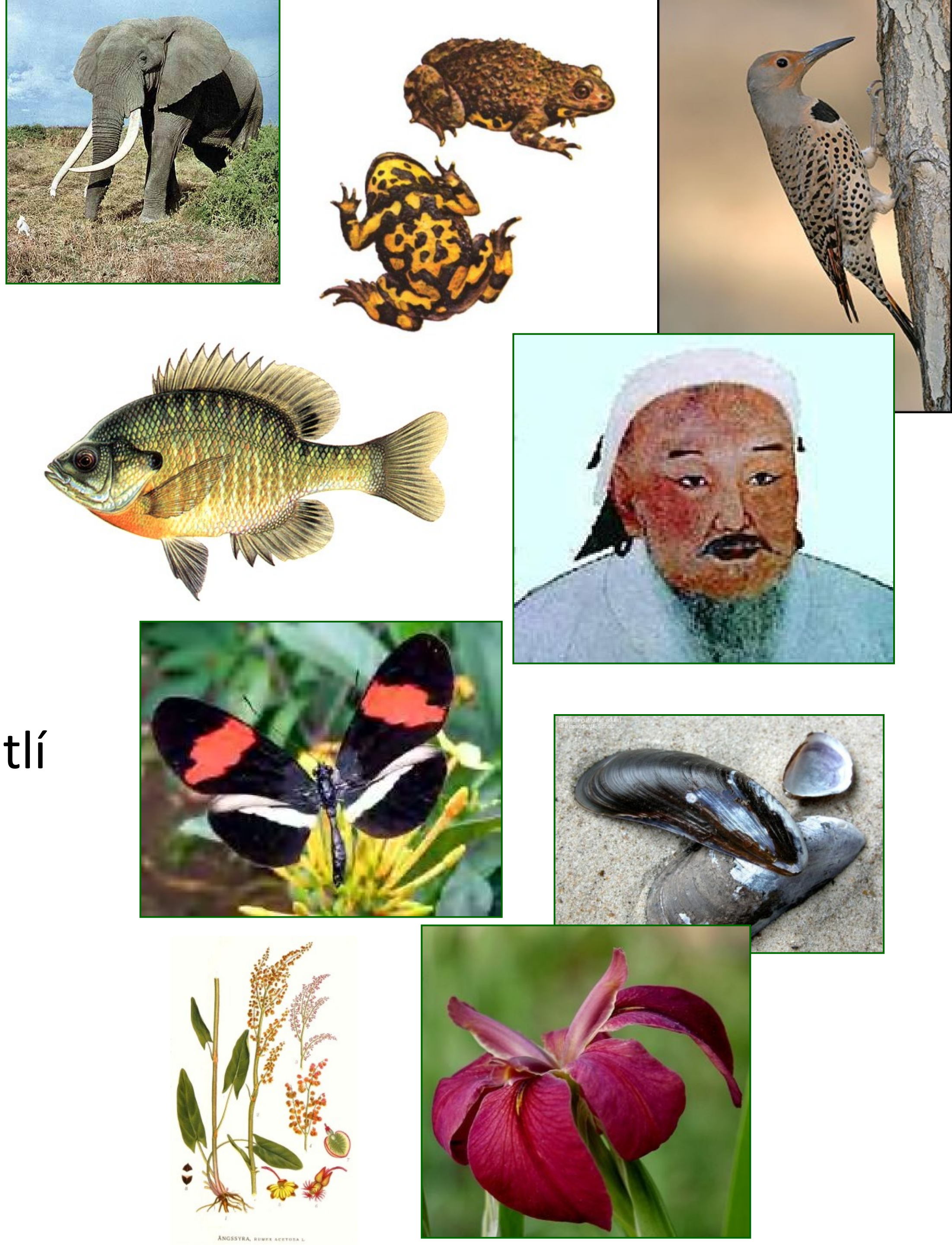


✓ Diploidní s pohlavním rozmnožováním

✓ Většinou obratlovci

✓ Budou ale i někteří bezobratlí

✓ Rostliny fungují často jinak!
Ale občas i o nich bude řeč.



Získání genetických dat – viz Genetické metody v zoologii

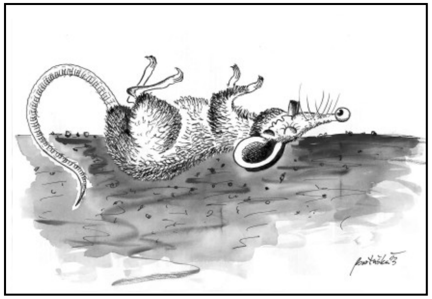
Genotypizace – analýza genotypu

- stanovení formy určitého úseku DNA (alely, haplotypu) - výběr daného znaku (= markeru) souvisí s úrovní genetické variability

1) izolace celkové DNA z tkání
2) amplifikace požadovaného úseku DNA (u PCR-based metod)
3) studium variability daného úseku (lokus)

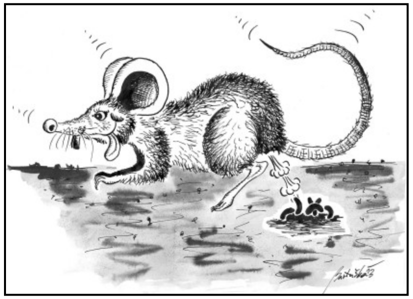
Způsoby získání DNA z volně žijících živočichů

1. **destrukční** – živočich je usmrcen kvůli získání tkání potřebných na genetické analýzy



2. **nedestrukční (invazivní)** – živočich je odchycen a je mu odebrán vzorek tkáně nebo krve

3. **neinvazivní** – zdroj DNA je „zanechán za živočichem“ a je získán bez potřeby odchytu, manipulace či dokonce pozorování



Izolace DNA

- rozmanitý biologický materiál – musí obsahovat buněčná **jádra nebo mitochondrie** s nedegradovanou DNA
- dnes většinou komerční kity
- velký vliv **fixace** vzorků

Genetické markery

- **Kódující DNA (geny)**
- Přepisované sekvence
- Genetický kód
- Ovlivňují fenotyp
- Podléhají přírodnímu výběru
- Narůstající význam v molekulární ekologii (transkriptomika)

- **Nekódující DNA**
- Nefunkční (neznámá funkce)
- Neutrální k přírodnímu výběru – větší variabilita
- Většina DNA u eukaryot
- Pseudogeny
- Repetitivní DNA

Typy genetických markerů

- *sekvence jaderné nebo organelové (mt, cp) DNA* – **Sangerovo sekvenování** nebo „**next-generation sequencing**“

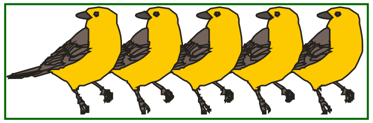
- *jaderné znaky*

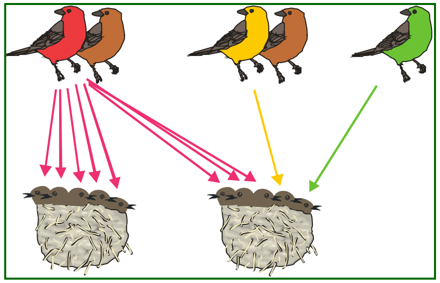
  \- **dominantní** (AFLP) – multi-locus markery

  \- **kodominantní** (mikrosatelity, SNPs) – single locus markery

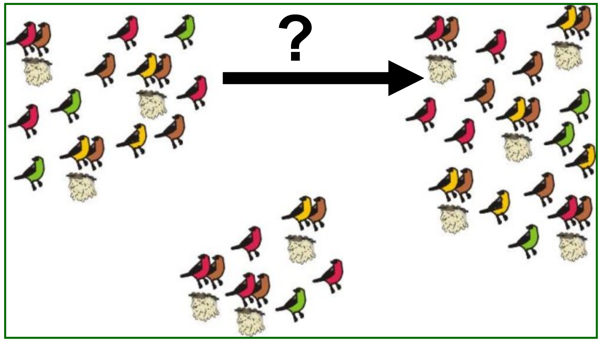
Různé otázky – různé přístupy

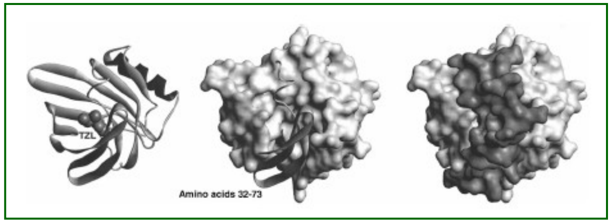
- Příbuznost (neutrální znaky)
  - identita (stopy stejného jedince, klony)
  - paternita, vzdálenější příbuzní
  - vztah populací (izolovanost, výměna migrantů)
  - fylogeografie (historie šíření)
  - hybridizace, hybridní zóny



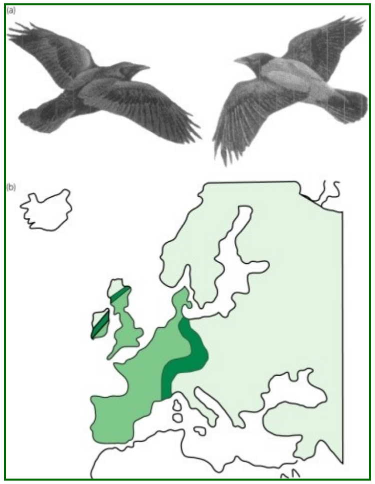


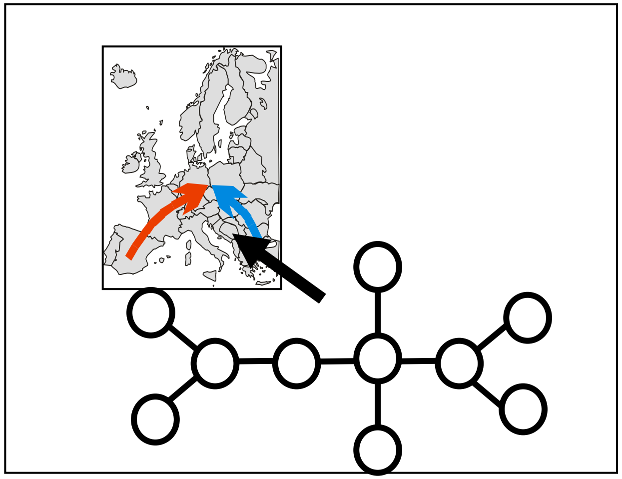
- Geny pod selekcí
  - MHC, MUP, ABP, reprodukční proteiny
  - geny pro zbarvení
  - detekce selekce

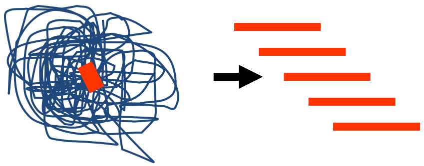




- Ochranářská genetika

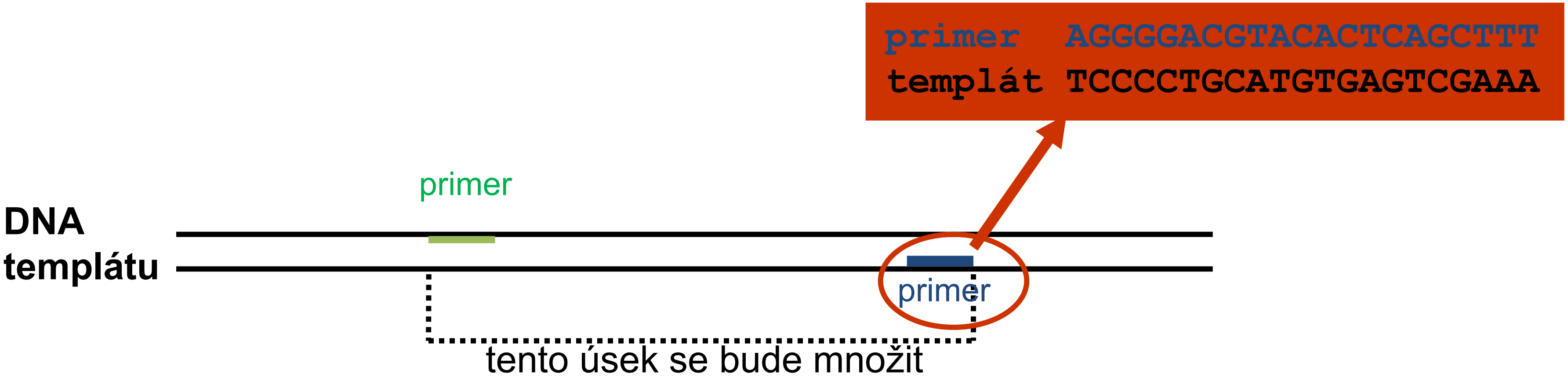






PCR

- Z celkové DNA si namnožíme jen úsek, který nás zajímá.
- Co se bude množit? To určí **primery**.
- **Primery** – krátké oligonukleotidy komplementární k úsekům ohraničujícím místo našeho zájmu.

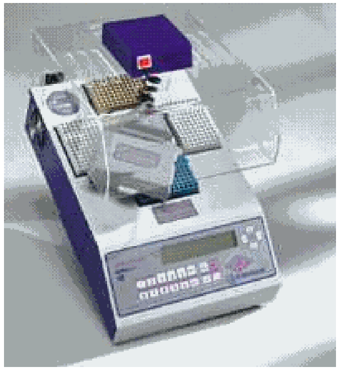


```
primer  AGGGGACGTACACTCAGCTTT
templát TCCCCTGCATGTGAGTCGAAA
```

PCR


Cycler MJ Research


Cycler Eppendorf


RoboCycler Stratagene

Cykly (obvykle 20-40):
denaturace (95°C)
nasednutí primerů (50-65°C)
elongace=polymerizace (72°C)

Nejprve však často prodlužená denaturace celkové DNA

Nakonec prodloužená elongace

Příklad programu

95 C 3 min

95 C 30 s
60 C 30 s
72 C 1 min
35x zpět

72 C 10 min

„Molekulárně-genetické“ metody

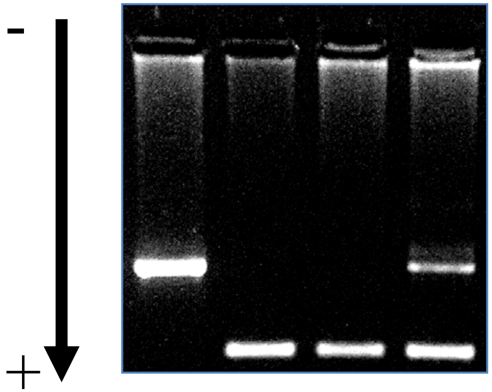
- analýza polymorfismu DNA
- délkový polymorfismus (princip mikrosatelitů)

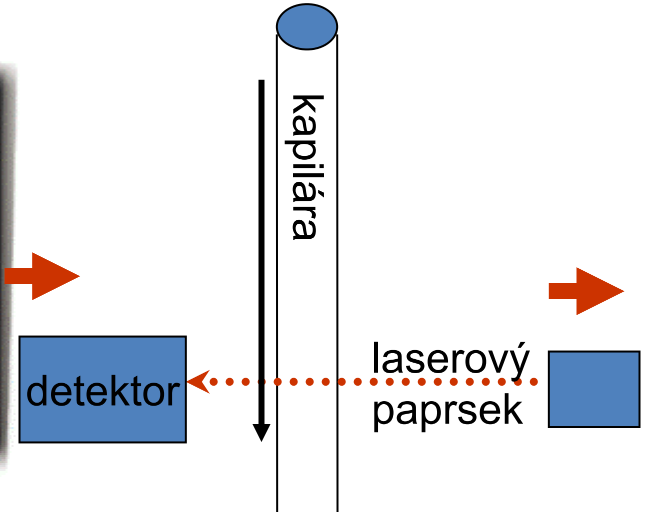
CGCACATCTCTAGCTTCGATTCAGGAA

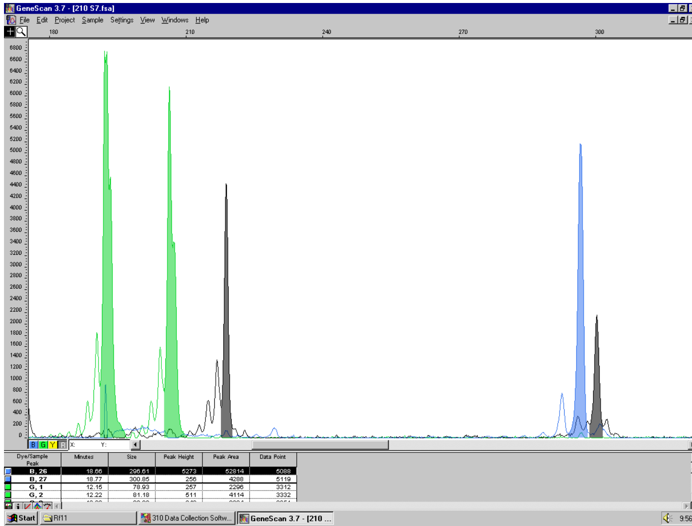
CGCATCTCTAGCTTTGATTCAGGAA

Rozdělení fragmentů DNA podle velikosti

- Agarosa - Hrubé rozdělení (do rozdílu 15 bp)
- Polyakrylamid – Přesnější rozdělení (4 bp)
- Sekvenátor, fragmentační analýza – nejpřesnější (fluorescenčně značené PCR fragmenty, např. značené primery)







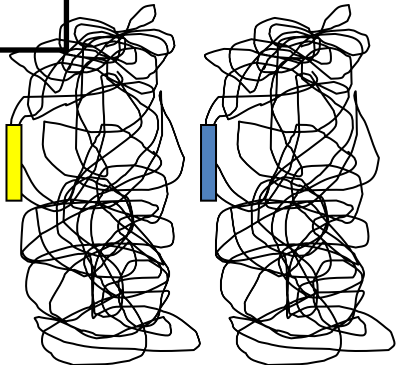
Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Point
B, 26	18.66	296.61	5273	52814	5088
B, 27	18.77	300.85	256	4288	5119
G, 1	12.15	78.93	257	2296	3312
G, 2	12.22	81.18	511	4114	3332

„Molekulárně-genetické“ metody

- analýza polymorfismu DNA
- sekvenční polymorfismus (princip SNPs):

CGCATCTCTAGCTTCGATTCAGGAA

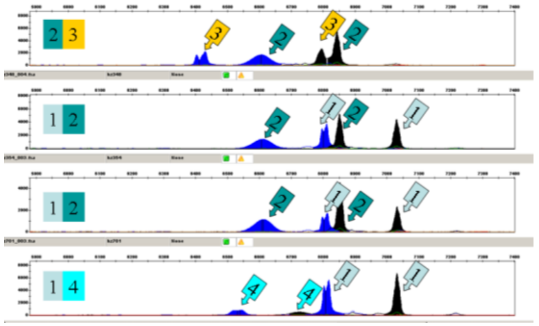
CGCATCTCTAGCTTTGATTCAGGAA

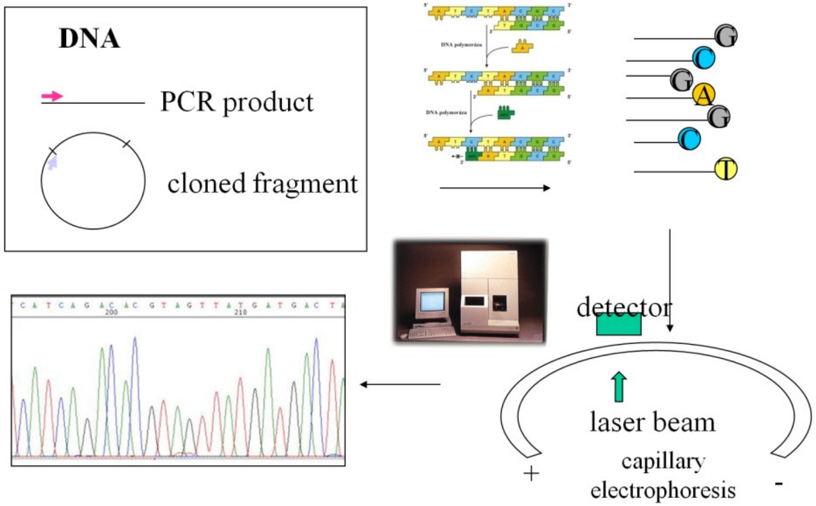


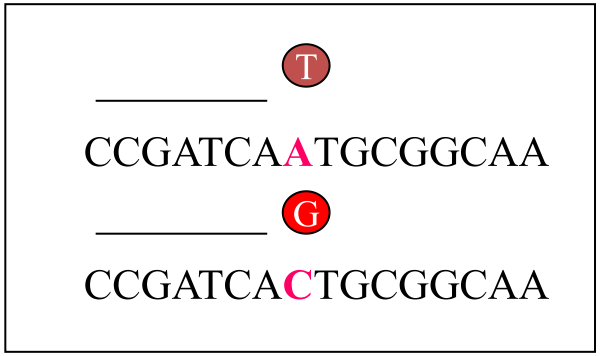
genotyp diploidního jedince: **C/T**

Studium variability DNA fragmentů, které se neliší délkou

- Sangerovo sekvencování (velmi dobré pro mtDNA, u nDNA problém s odlišením alel u heterozygotů)
- SNP („single nucleotide polymorphism“) analýza – např. RFLP, SSCP, microarrays – chips, atd.



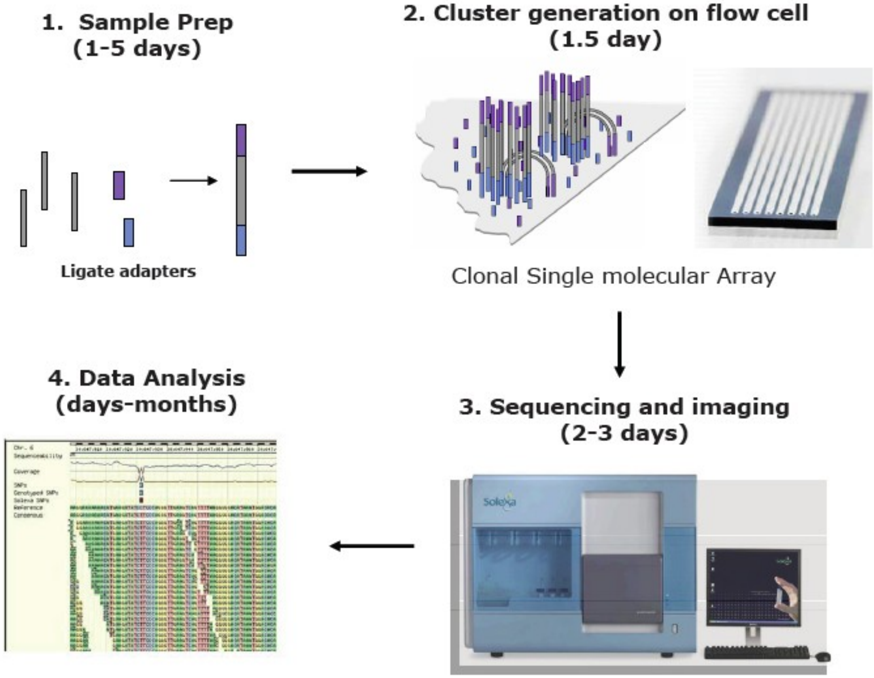




CCGATCAATGCGGCAA

CCGATCACTGCGGCAA

„High-throughput sequencing“



1. Sample Prep
(1-5 days)

Ligate adapters

2. Cluster generation on flow cell
(1.5 day)

Clonal Single molecular Array

3. Sequencing and imaging
(2-3 days)

4. Data Analysis
(days-months)

Illumina

MinION

Typy populačně-genetických dat

mikrosatelity

Jedinec	Marker 1	Marker 2
Ind_1	170/172	133/136
Ind_2	168/172	133/139
Ind_3	168/168	136/139

SNPs

Jedinec	Marker 1	Marker 2
Ind_1	A/T	C/T
Ind_2	A/T	T/T
Ind_3	T/T	C/T

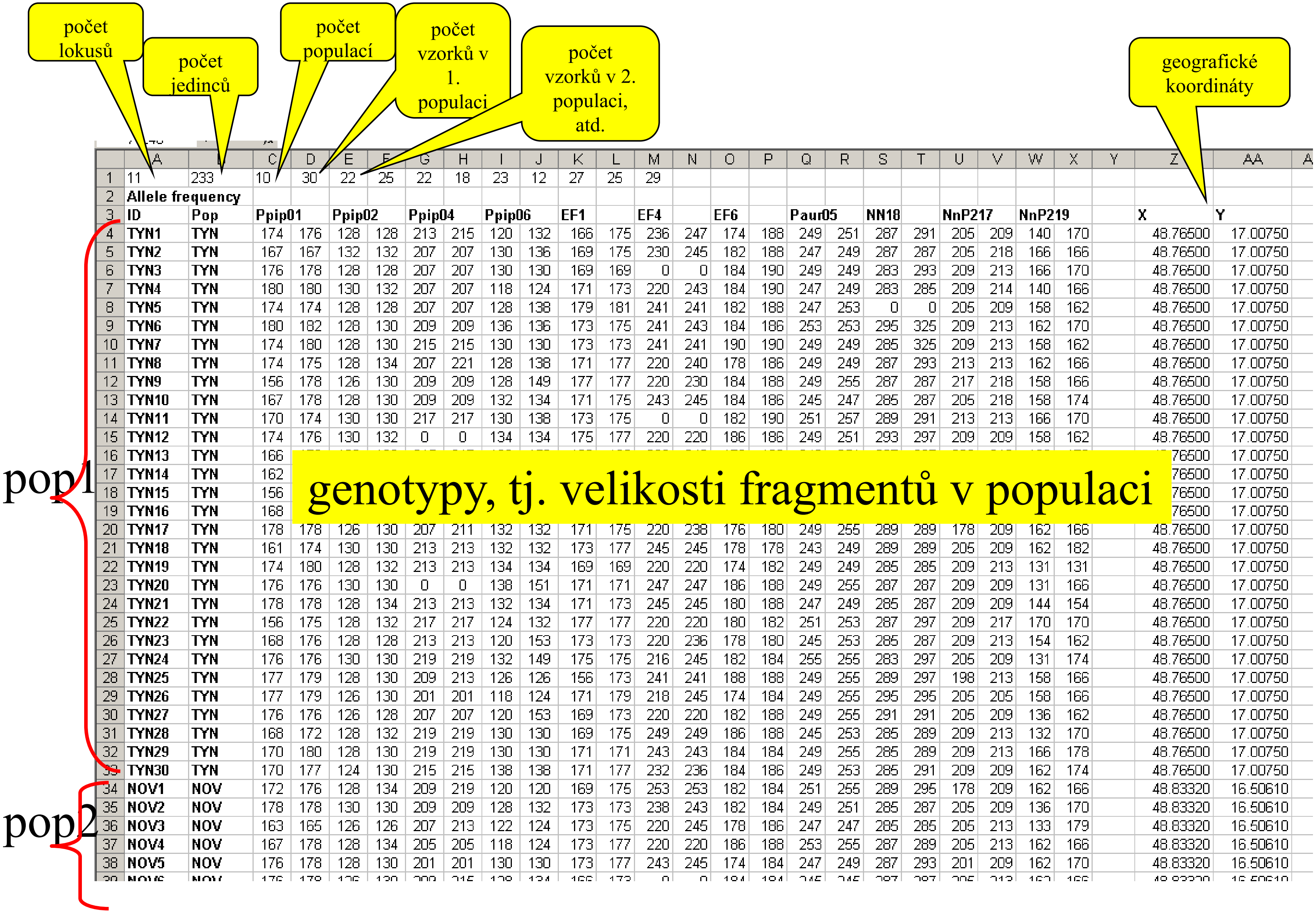
SINE

Jedinec	Marker 1	Marker 2
Ind_1	+/-	-/-
Ind_2	+/+	+/+
Ind_3	-/-	+/-

AFLP

Jedinec	Marker 1	Marker 2
Ind_1	+	-
Ind_2	+	+
Ind_3	-	-

Typy získaných dat – kodominantní znaky



počet lokusů

počet jedinců

počet populací

počet vzorků v 1. populaci

počet vzorků v 2. populaci, atd.

geografické koordináty

genotypy, tj. velikosti fragmentů v populaci

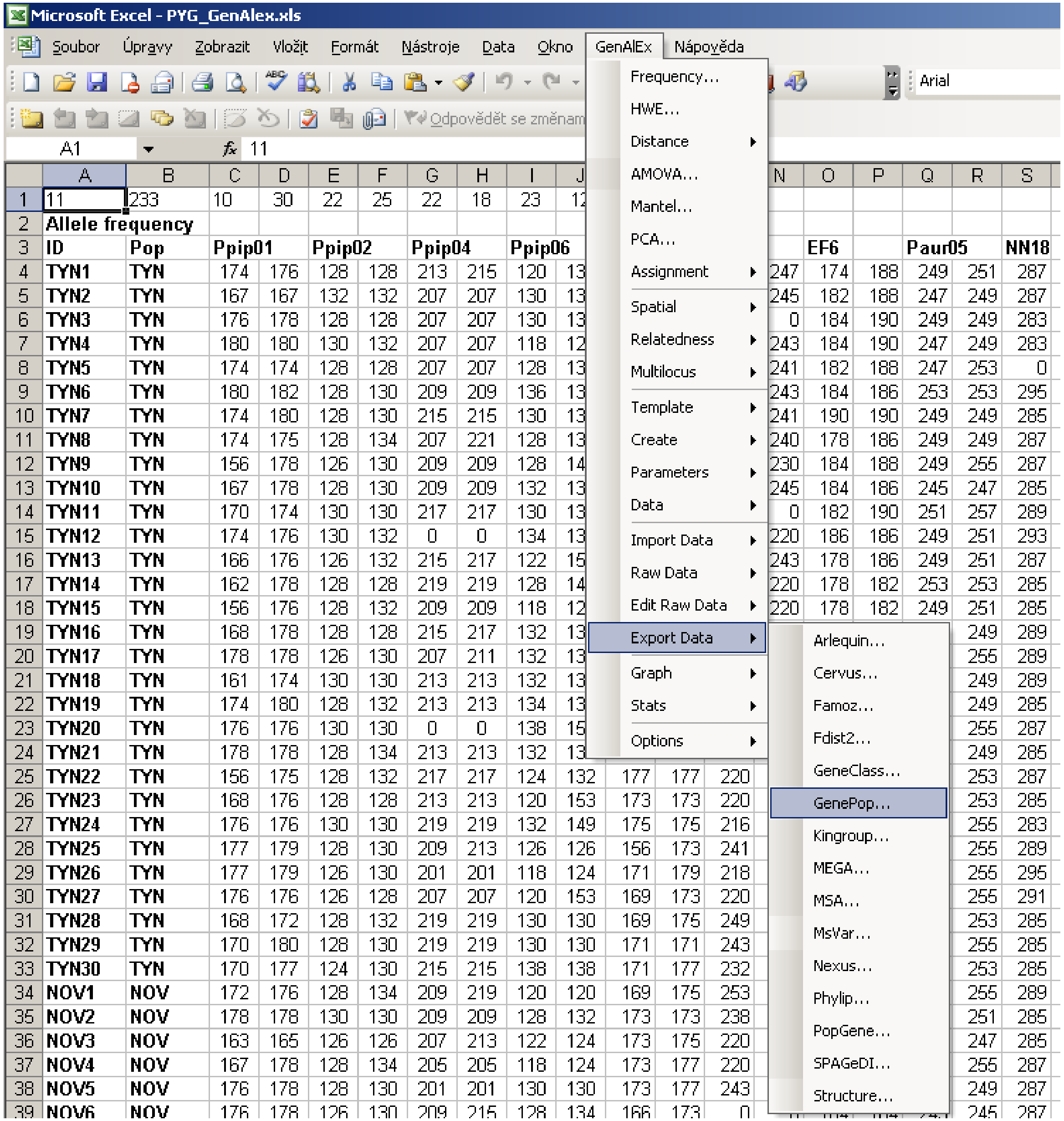
pop1

pop2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	11	233	10	30	22	25	22	18	23	12	27	25	29														
2	**Allele frequency**																										
3	**ID**	**Pop**	**Ppip01**		**Ppip02**		**Ppip04**		**Ppip06**		**EF1**		**EF4**		**EF6**		**Paur05**		**NN18**		**NnP217**		**NnP219**			**X**	**Y**
4	**TYN1**	**TYN**	174	176	128	128	213	215	120	132	166	175	236	247	174	188	249	251	287	291	205	209	140	170		48.76500	17.00750
5	**TYN2**	**TYN**	167	167	132	132	207	207	130	136	169	175	230	245	182	188	247	249	287	287	205	218	166	166		48.76500	17.00750
6	**TYN3**	**TYN**	176	178	128	128	207	207	130	130	169	169	0	0	184	190	249	249	283	293	209	213	166	170		48.76500	17.00750
7	**TYN4**	**TYN**	180	180	130	132	207	207	118	124	171	173	220	243	184	190	247	249	283	285	209	214	140	166		48.76500	17.00750
8	**TYN5**	**TYN**	174	174	128	128	207	207	128	138	179	181	241	241	182	188	247	253	0	0	205	209	158	162		48.76500	17.00750
9	**TYN6**	**TYN**	180	182	128	130	209	209	136	136	173	175	241	243	184	186	253	253	295	325	209	213	162	170		48.76500	17.00750
10	**TYN7**	**TYN**	174	180	128	130	215	215	130	130	173	173	241	241	190	190	249	249	285	325	209	213	158	162		48.76500	17.00750
11	**TYN8**	**TYN**	174	175	128	134	207	221	128	138	171	177	220	240	178	186	249	249	287	293	213	213	162	166		48.76500	17.00750
12	**TYN9**	**TYN**	156	178	126	130	209	209	128	149	177	177	220	230	184	188	249	255	287	287	217	218	158	166		48.76500	17.00750
13	**TYN10**	**TYN**	167	178	128	130	209	209	132	134	171	175	243	245	184	186	245	247	285	287	205	218	158	174		48.76500	17.00750
14	**TYN11**	**TYN**	170	174	130	130	217	217	130	138	173	175	0	0	182	190	251	257	289	291	213	213	166	170		48.76500	17.00750
15	**TYN12**	**TYN**	174	176	130	132	0	0	134	134	175	177	220	220	186	186	249	251	293	297	209	209	158	162		48.76500	17.00750
16	**TYN13**	**TYN**	166																							76500	17.00750
17	**TYN14**	**TYN**	162																							76500	17.00750
18	**TYN15**	**TYN**	156																							76500	17.00750
19	**TYN16**	**TYN**	168																							76500	17.00750
20	**TYN17**	**TYN**	178	178	126	130	207	211	132	132	171	175	220	238	176	180	249	255	289	289	178	209	162	166		48.76500	17.00750
21	**TYN18**	**TYN**	161	174	130	130	213	213	132	132	173	177	245	245	178	178	243	249	289	289	205	209	162	182		48.76500	17.00750
22	**TYN19**	**TYN**	174	180	128	132	213	213	134	134	169	169	220	220	174	182	249	249	285	285	209	213	131	131		48.76500	17.00750
23	**TYN20**	**TYN**	176	176	130	130	0	0	138	151	171	171	247	247	186	188	249	255	287	287	209	209	131	166		48.76500	17.00750
24	**TYN21**	**TYN**	178	178	128	134	213	213	132	134	171	173	245	245	180	188	247	249	285	287	209	209	144	154		48.76500	17.00750
25	**TYN22**	**TYN**	156	175	128	132	217	217	124	132	177	177	220	220	180	182	251	253	287	297	209	217	170	170		48.76500	17.00750
26	**TYN23**	**TYN**	168	176	128	128	213	213	120	153	173	173	220	236	178	180	245	253	285	287	209	213	154	162		48.76500	17.00750
27	**TYN24**	**TYN**	176	176	130	130	219	219	132	149	175	175	216	245	182	184	255	255	283	297	205	209	131	174		48.76500	17.00750
28	**TYN25**	**TYN**	177	179	128	130	209	213	126	126	156	173	241	241	188	188	249	255	289	297	198	213	158	166		48.76500	17.00750
29	**TYN26**	**TYN**	177	179	126	130	201	201	118	124	171	179	218	245	174	184	249	255	295	295	205	205	158	166		48.76500	17.00750
30	**TYN27**	**TYN**	176	176	126	128	207	207	120	153	169	173	220	220	182	188	249	255	291	291	205	209	136	162		48.76500	17.00750
31	**TYN28**	**TYN**	168	172	128	132	219	219	130	130	169	175	249	249	186	188	245	253	285	289	209	213	132	170		48.76500	17.00750
32	**TYN29**	**TYN**	170	180	128	130	219	219	130	130	171	171	243	243	184	184	249	255	285	289	209	213	166	178		48.76500	17.00750
33	**TYN30**	**TYN**	170	177	124	130	215	215	138	138	171	177	232	236	184	186	249	253	285	291	209	209	162	174		48.76500	17.00750
34	**NOV1**	**NOV**	172	176	128	134	209	219	120	120	169	175	253	253	182	184	251	255	289	295	178	209	162	166		48.83320	16.50610
35	**NOV2**	**NOV**	178	178	130	130	209	209	128	132	173	173	238	243	182	184	249	251	285	287	205	209	136	170		48.83320	16.50610
36	**NOV3**	**NOV**	163	165	126	126	207	213	122	124	173	175	220	245	178	186	247	247	285	285	205	213	133	179		48.83320	16.50610
37	**NOV4**	**NOV**	167	178	128	134	205	205	118	124	173	177	220	220	186	188	253	255	287	289	205	213	162	166		48.83320	16.50610
38	**NOV5**	**NOV**	176	178	128	130	201	201	130	130	173	177	243	245	174	184	247	249	287	293	201	209	162	170		48.83320	16.50610

formát GenAlex

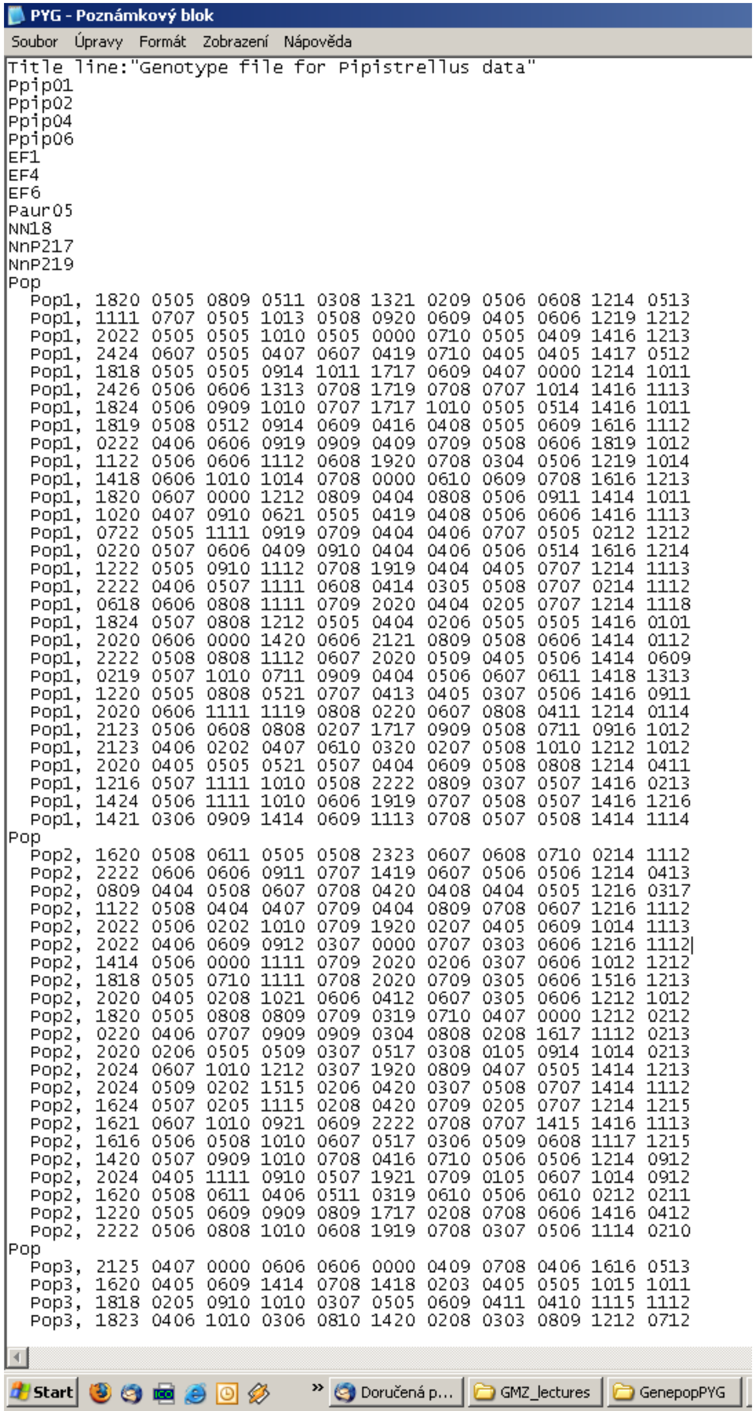
GenAlEx
Genetic Analysis in Excel
Version 6.1



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	11	233	10	30	22	25	22	18	23	12									
2	**Allele frequency**																		
3	**ID**	**Pop**	**Ppip01**		**Ppip02**		**Ppip04**		**Ppip06**						**EF6**		**Paur05**		**NN18**
4	**TYN1**	**TYN**	174	176	128	128	213	215	120	13				247	174	188	249	251	287
5	**TYN2**	**TYN**	167	167	132	132	207	207	130	13				245	182	188	247	249	287
6	**TYN3**	**TYN**	176	178	128	128	207	207	130	13				0	184	190	249	249	283
7	**TYN4**	**TYN**	180	180	130	132	207	207	118	12				243	184	190	247	249	283
8	**TYN5**	**TYN**	174	174	128	128	207	207	128	13				241	182	188	247	253	0
9	**TYN6**	**TYN**	180	182	128	130	209	209	136	13				243	184	186	253	253	295
10	**TYN7**	**TYN**	174	180	128	130	215	215	130	13				241	190	190	249	249	285
11	**TYN8**	**TYN**	174	175	128	134	207	221	128	13				240	178	186	249	249	287
12	**TYN9**	**TYN**	156	178	126	130	209	209	128	14				230	184	188	249	255	287
13	**TYN10**	**TYN**	167	178	128	130	209	209	132	13				245	184	186	245	247	285
14	**TYN11**	**TYN**	170	174	130	130	217	217	130	13				0	182	190	251	257	289
15	**TYN12**	**TYN**	174	176	130	132	0	0	134	13				220	186	186	249	251	293
16	**TYN13**	**TYN**	166	176	126	132	215	217	122	15				243	178	186	249	251	287
17	**TYN14**	**TYN**	162	178	128	128	219	219	128	14				220	178	182	253	253	285
18	**TYN15**	**TYN**	156	176	128	132	209	209	118	12				220	178	182	249	251	285
19	**TYN16**	**TYN**	168	178	128	128	215	217	132	13								249	289
20	**TYN17**	**TYN**	178	178	126	130	207	211	132	13								255	289
21	**TYN18**	**TYN**	161	174	130	130	213	213	132	13								249	289
22	**TYN19**	**TYN**	174	180	128	132	213	213	134	13								249	285
23	**TYN20**	**TYN**	176	176	130	130	0	0	138	15								255	287
24	**TYN21**	**TYN**	178	178	128	134	213	213	132	13								249	285
25	**TYN22**	**TYN**	156	175	128	132	217	217	124	132	177	177	220					253	287
26	**TYN23**	**TYN**	168	176	128	128	213	213	120	153	173	173	220					253	285
27	**TYN24**	**TYN**	176	176	130	130	219	219	132	149	175	175	216					255	283
28	**TYN25**	**TYN**	177	179	128	130	209	213	126	126	156	173	241					255	289
29	**TYN26**	**TYN**	177	179	126	130	201	201	118	124	171	179	218					255	295
30	**TYN27**	**TYN**	176	176	126	128	207	207	120	153	169	173	220					255	291
31	**TYN28**	**TYN**	168	172	128	132	219	219	130	130	169	175	249					253	285
32	**TYN29**	**TYN**	170	180	128	130	219	219	130	130	171	171	243					255	285
33	**TYN30**	**TYN**	170	177	124	130	215	215	138	138	171	177	232					253	285
34	**NOV1**	**NOV**	172	176	128	134	209	219	120	120	169	175	253					255	289
35	**NOV2**	**NOV**	178	178	130	130	209	209	128	132	173	173	238					251	285
36	**NOV3**	**NOV**	163	165	126	126	207	213	122	124	173	175	220					247	285
37	**NOV4**	**NOV**	167	178	128	134	205	205	118	124	173	177	220					255	287
38	**NOV5**	**NOV**	176	178	128	130	201	201	130	130	173	177	243					249	287
39	**NOV6**	**NOV**	176	178	126	130	209	215	128	134	166	173	0					245	287

http://www.anu.edu.au/BoZo/GenAlEx/

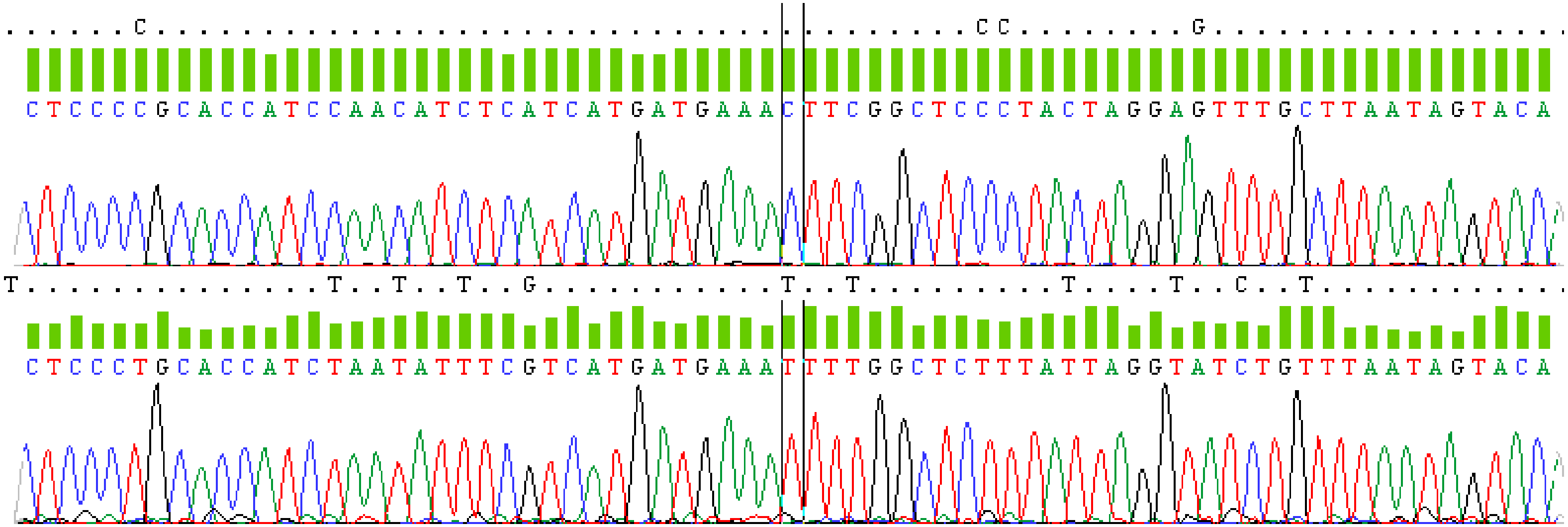
Genepop file format – jednoduchý ASCI kód (.txt)



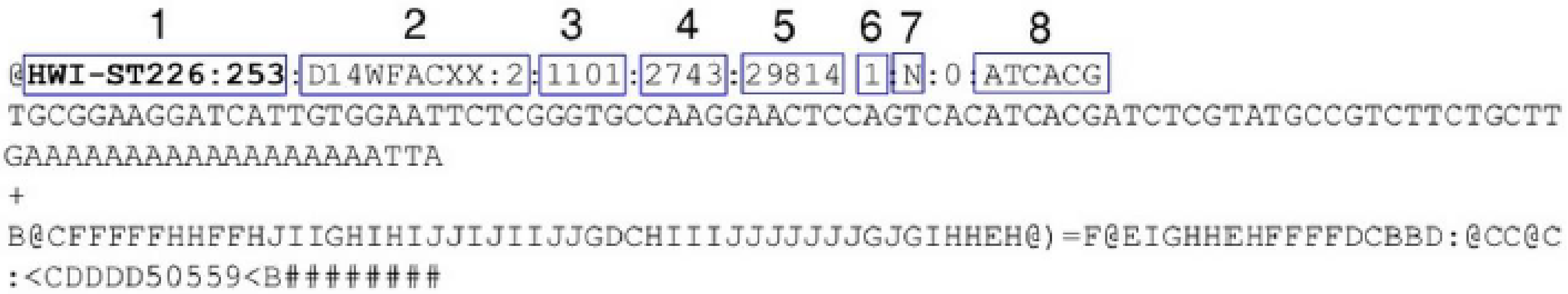
```
Title line:"Genotype file for Pipistrellus data"
Ppip01
Ppip02
Ppip04
Ppip06
EF1
EF4
EF6
Paur05
NN18
NnP217
NnP219
Pop
   Pop1, 1820 0505 0809 0511 0308 1321 0209 0506 0608 1214 0513
   Pop1, 1111 0707 0505 1013 0508 0920 0609 0405 0606 1219 1212
   Pop1, 2022 0505 0505 1010 0505 0000 0710 0505 0409 1416 1213
   Pop1, 2424 0607 0505 0407 0607 0419 0710 0405 0405 1417 0512
   Pop1, 1818 0505 0505 0914 1011 1717 0609 0407 0000 1214 1011
   Pop1, 2426 0506 0606 1313 0708 1719 0708 0707 1014 1416 1113
   Pop1, 1824 0506 0909 1010 0707 1717 1010 0505 0514 1416 1011
   Pop1, 1819 0508 0512 0914 0609 0416 0408 0505 0609 1616 1112
   Pop1, 0222 0406 0606 0919 0909 0409 0709 0508 0606 1819 1012
   Pop1, 1122 0506 0606 1112 0608 1920 0708 0304 0506 1219 1014
   Pop1, 1418 0606 1010 1014 0708 0000 0610 0609 0708 1616 1213
   Pop1, 1820 0607 0000 1212 0809 0404 0808 0506 0911 1414 1011
   Pop1, 1020 0407 0910 0621 0505 0419 0408 0506 0606 1416 1113
   Pop1, 0722 0505 1111 0919 0709 0404 0406 0707 0505 0212 1212
   Pop1, 0220 0507 0606 0409 0910 0404 0406 0506 0514 1616 1214
   Pop1, 1222 0505 0910 1112 0708 1919 0404 0405 0707 1214 1113
   Pop1, 2222 0406 0507 1111 0608 0414 0305 0508 0707 0214 1112
   Pop1, 0618 0606 0808 1111 0709 2020 0404 0205 0707 1214 1118
   Pop1, 1824 0507 0808 1212 0505 0404 0206 0505 0505 1416 0101
   Pop1, 2020 0606 0000 1420 0606 2121 0809 0508 0606 1414 0112
   Pop1, 2222 0508 0808 1112 0607 2020 0509 0405 0506 1414 0609
   Pop1, 0219 0507 1010 0711 0909 0404 0506 0607 0611 1418 1313
   Pop1, 1220 0505 0808 0521 0707 0413 0405 0307 0506 1416 0911
   Pop1, 2020 0606 1111 1119 0808 0220 0607 0808 0411 1214 0114
   Pop1, 2123 0506 0608 0808 0207 1717 0909 0508 0711 0916 1012
   Pop1, 2123 0406 0202 0407 0610 0320 0207 0508 1010 1212 1012
   Pop1, 2020 0405 0505 0521 0507 0404 0609 0508 0808 1214 0411
   Pop1, 1216 0507 1111 1010 0508 2222 0809 0307 0507 1416 0213
   Pop1, 1424 0506 1111 1010 0606 1919 0707 0508 0507 1416 1216
   Pop1, 1421 0306 0909 1414 0609 1113 0708 0507 0508 1414 1114
Pop
   Pop2, 1620 0508 0611 0505 0508 2323 0607 0608 0710 0214 1112
   Pop2, 2222 0606 0606 0911 0707 1419 0607 0506 0506 1214 0413
   Pop2, 0809 0404 0508 0607 0708 0420 0408 0404 0505 1216 0317
   Pop2, 1122 0508 0404 0407 0709 0404 0809 0708 0607 1216 1112
   Pop2, 2022 0506 0202 1010 0709 1920 0207 0405 0609 1014 1113
   Pop2, 2022 0406 0609 0912 0307 0000 0707 0303 0606 1216 1112
   Pop2, 1414 0506 0000 1111 0709 2020 0206 0307 0606 1012 1212
   Pop2, 1818 0505 0710 1111 0708 2020 0709 0305 0606 1516 1213
   Pop2, 2020 0405 0208 1021 0606 0412 0607 0305 0606 1212 1012
   Pop2, 1820 0505 0808 0809 0709 0319 0710 0407 0000 1212 0212
   Pop2, 0220 0406 0707 0909 0909 0304 0808 0208 1617 1112 0213
   Pop2, 2020 0206 0505 0509 0307 0517 0308 0105 0914 1014 0213
   Pop2, 2024 0607 1010 1212 0307 1920 0809 0407 0505 1414 1213
   Pop2, 2024 0509 0202 1515 0206 0420 0307 0508 0707 1414 1112
   Pop2, 1624 0507 0205 1115 0208 0420 0709 0205 0707 1214 1215
   Pop2, 1621 0607 1010 0921 0609 2222 0708 0707 1415 1416 1113
   Pop2, 1616 0506 0508 1010 0607 0517 0306 0509 0608 1117 1215
   Pop2, 1420 0507 0909 1010 0708 0416 0710 0506 0506 1214 0912
   Pop2, 2024 0405 1111 0910 0507 1921 0709 0105 0607 1014 0912
   Pop2, 1620 0508 0611 0406 0511 0319 0610 0506 0610 0212 0211
   Pop2, 1220 0505 0609 0909 0809 1717 0208 0708 0606 1416 0412
   Pop2, 2222 0506 0808 1010 0608 1919 0708 0307 0506 1114 0210
Pop
   Pop3, 2125 0407 0000 0606 0606 0000 0409 0708 0406 1616 0513
   Pop3, 1620 0405 0609 1414 0708 1418 0203 0405 0505 1015 1011
   Pop3, 1818 0205 0910 1010 0307 0505 0609 0411 0410 1115 1112
   Pop3, 1823 0406 1010 0306 0810 1420 0208 0303 0809 1212 0712
```

- jednotlivé alely pro daný lokus jsou seřazeny podle velikosti a očíslovány
- tj. např. 128/130 je převedeno na 10/11

Typy dat: Sangerovo sekvenování



Illumina *fastq* format

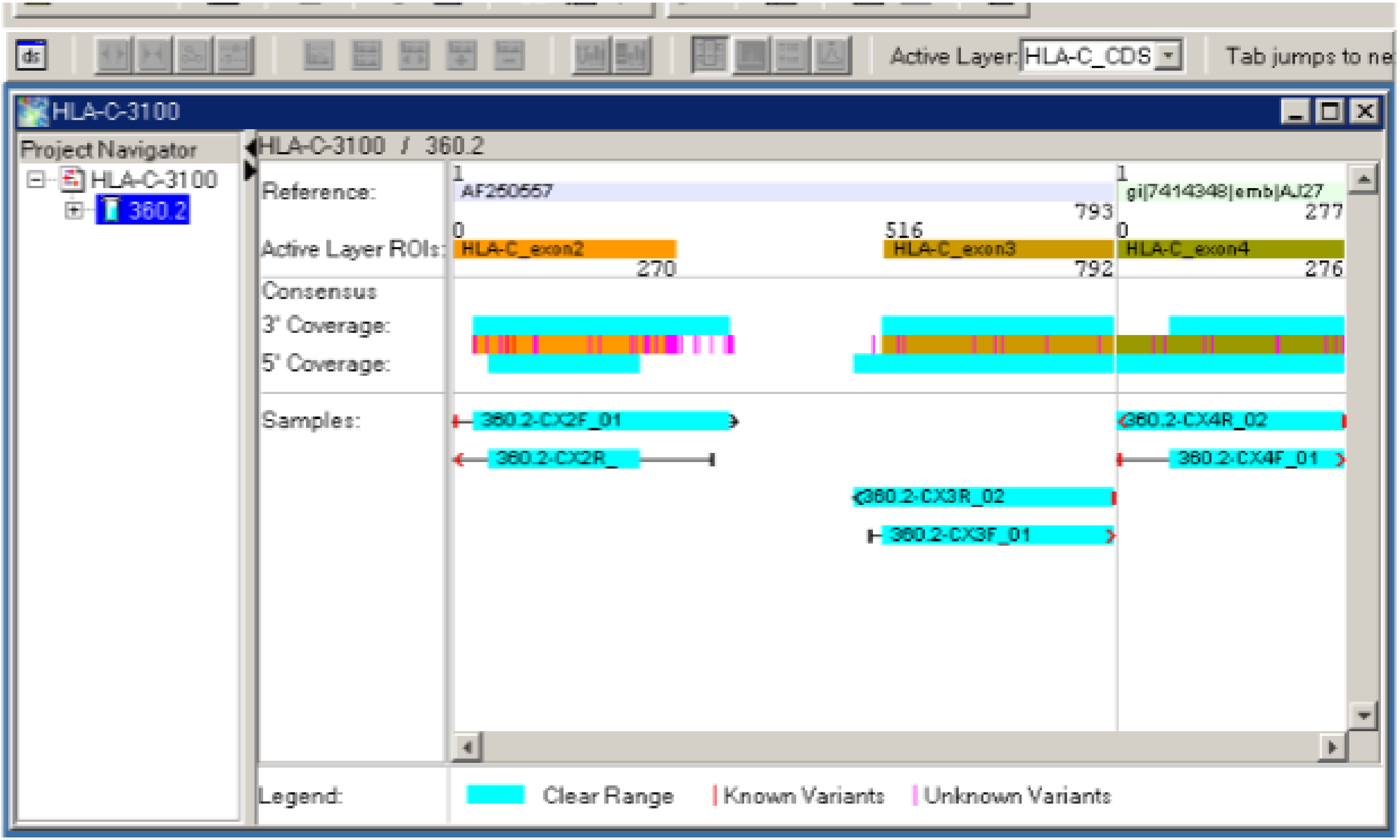


```
@HWI-ST226:253:D14WFACXX:2:1101:2743:29814 1:N:0:ATCACG
TGCGGAAGGATCATTGTGGAATTCTCGGGTGCCAAGGAACTCCAGTCACATCACGATCTCGTATGCCGTCTTCTGCTT
GAAAAAAAAAAAAAAAAAATTA
+
B@CFFFFFHHFFHJIIGHIHIJJIJIIJJGDCHIIIJJJJJJGJGIHHEH@)=F@EIGHHEHFFFFDCBBD:@CC@C
:<CDDDD50559<B########
```

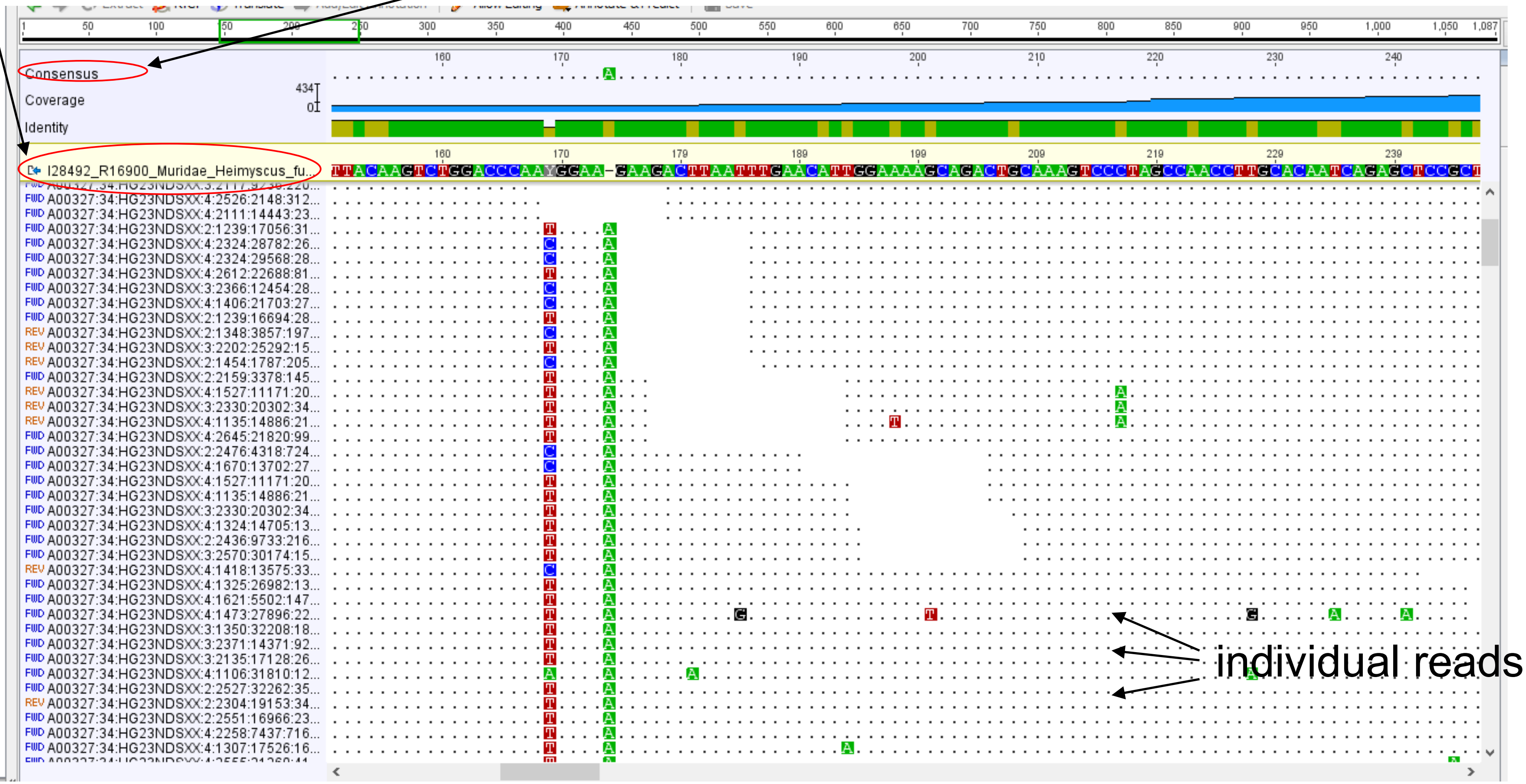
1. unique instrument ID and run ID
2. Flow cell ID and lane
3. tile number within the flow cell lane
4. 'x'-coordinate of the cluster within the tile
5. 'y'-coordinate of the cluster within the tile
6. the member of a pair, /1 or /2 *(paired-end or mate-pair reads only)*
7. N if the read passes filter, Y if read fails filter otherwise
8. Index sequence

Typy získaných dat – sekvence DNA

„Alignment“ → contig (ze stejného jedince)



Contigs from NGS



reference (in resequencing)

consensus

coverage

individual reads

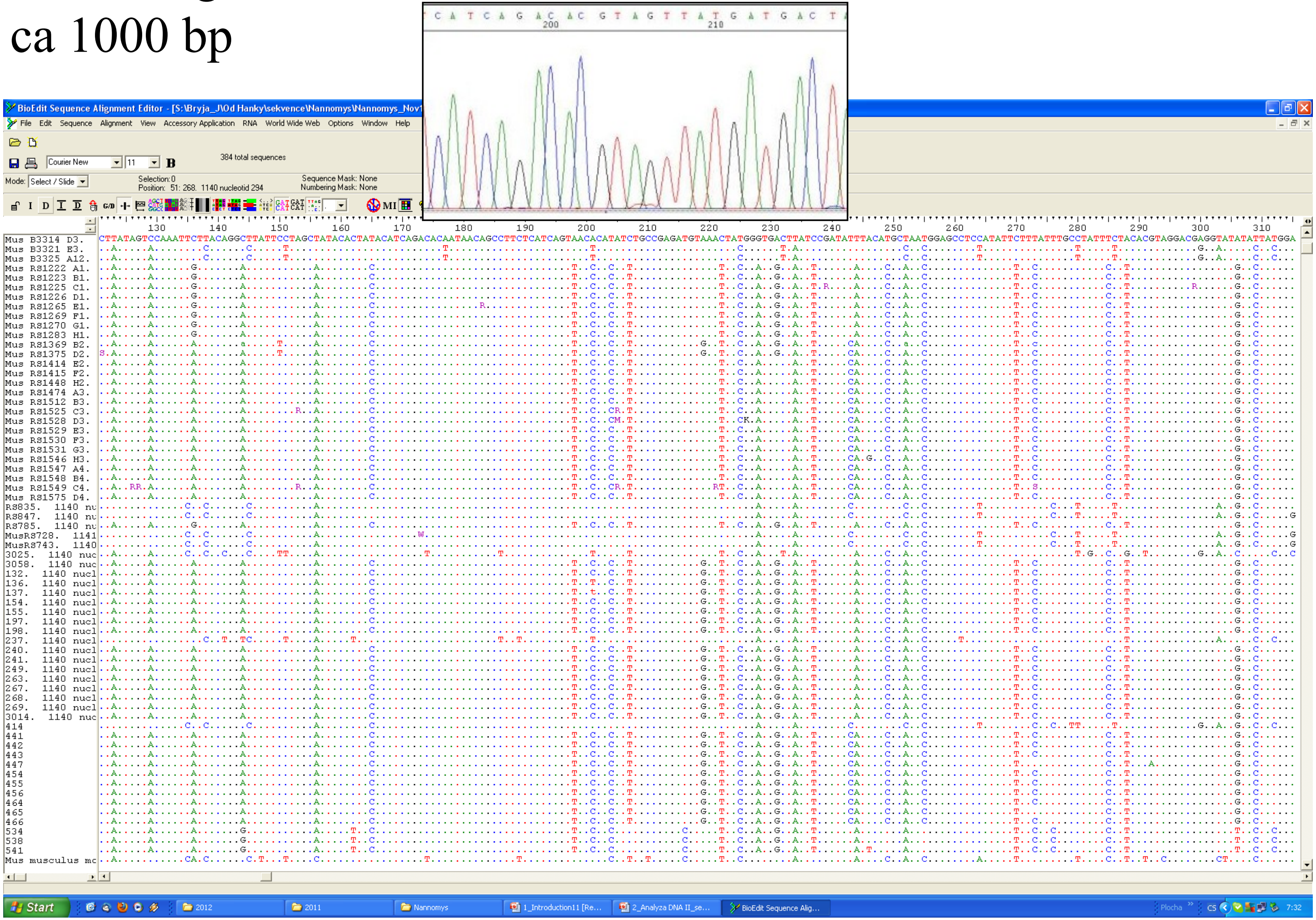
I28492_R16900_Muridae_Heimyscus_fu...

TTACAAGTCTGGACCCAAYGGAA-GAAGACTTAATTTGAACATTGGAAAAGCAGACTGCAAAGTCCCTAGCCAACCTTGCACAATCAGAGCTCCGCT

FWD A00327:34:HG23NDSXX:4:2526:2148:312...
FWD A00327:34:HG23NDSXX:4:2111:14443:23...
FWD A00327:34:HG23NDSXX:2:1239:17056:31...
FWD A00327:34:HG23NDSXX:4:2324:28782:26...
FWD A00327:34:HG23NDSXX:4:2324:29568:28...
FWD A00327:34:HG23NDSXX:4:2612:22688:81...
FWD A00327:34:HG23NDSXX:3:2366:12454:28...
FWD A00327:34:HG23NDSXX:4:1406:21703:27...
FWD A00327:34:HG23NDSXX:2:1239:16694:28...
REV A00327:34:HG23NDSXX:2:1348:3857:197...
REV A00327:34:HG23NDSXX:3:2202:25292:15...
REV A00327:34:HG23NDSXX:2:1454:1787:205...
FWD A00327:34:HG23NDSXX:2:2159:3378:145...
REV A00327:34:HG23NDSXX:4:1527:11171:20...
REV A00327:34:HG23NDSXX:3:2330:20302:34...
REV A00327:34:HG23NDSXX:4:1135:14886:21...
FWD A00327:34:HG23NDSXX:4:2645:21820:99...
FWD A00327:34:HG23NDSXX:2:2476:4318:724...
FWD A00327:34:HG23NDSXX:4:1670:13702:27...
FWD A00327:34:HG23NDSXX:4:1527:11171:20...
FWD A00327:34:HG23NDSXX:4:1135:14886:21...
FWD A00327:34:HG23NDSXX:3:2330:20302:34...
FWD A00327:34:HG23NDSXX:4:1324:14705:13...
FWD A00327:34:HG23NDSXX:2:2436:9733:216...
FWD A00327:34:HG23NDSXX:3:2570:30174:15...
REV A00327:34:HG23NDSXX:4:1418:13575:33...
FWD A00327:34:HG23NDSXX:4:1325:26982:13...
FWD A00327:34:HG23NDSXX:4:1621:5502:147...
FWD A00327:34:HG23NDSXX:4:1473:27896:22...
FWD A00327:34:HG23NDSXX:3:1350:32208:18...
FWD A00327:34:HG23NDSXX:3:2371:14371:92...
FWD A00327:34:HG23NDSXX:3:2135:17128:26...
FWD A00327:34:HG23NDSXX:4:1106:31810:12...
FWD A00327:34:HG23NDSXX:2:2527:32262:35...
REV A00327:34:HG23NDSXX:2:2304:19153:34...
FWD A00327:34:HG23NDSXX:2:2551:16966:23...
FWD A00327:34:HG23NDSXX:4:2258:7437:716...
FWD A00327:34:HG23NDSXX:4:1307:17526:16...

Typy získaných dat – srovnání sekvencí mezi jedinci

PCR + Sanger
= ca 1000 bp



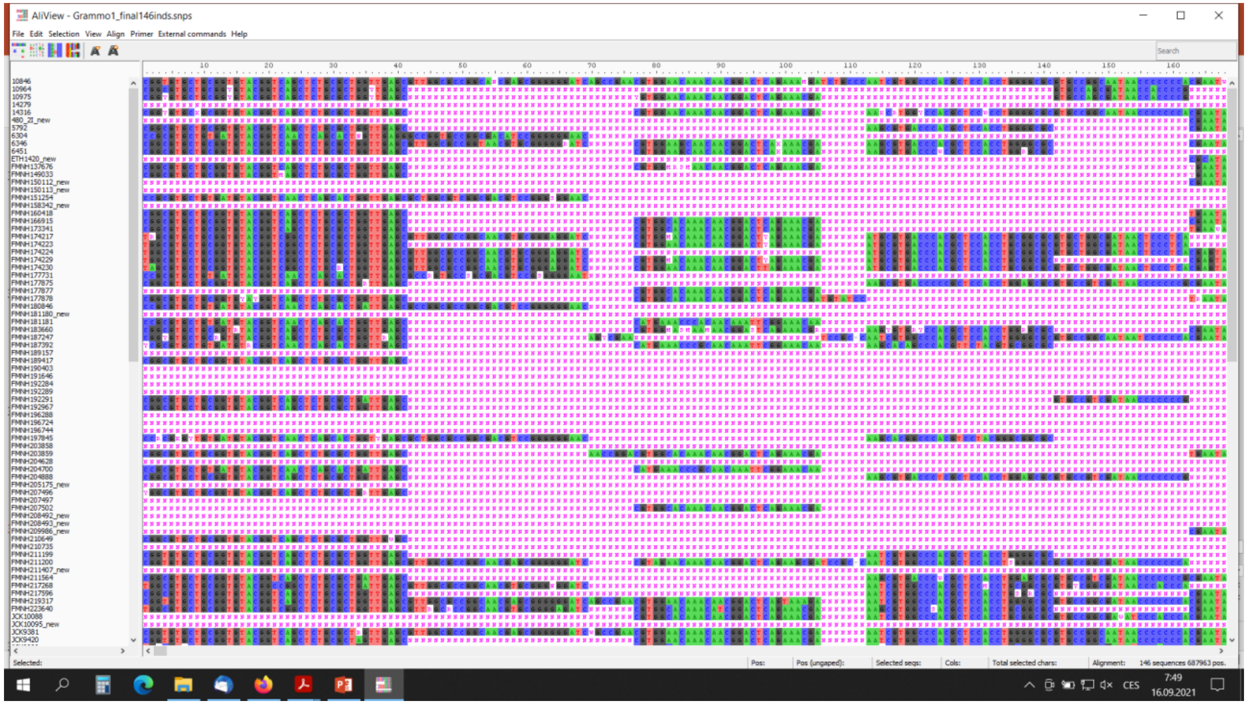
384 total sequences

Selection: 0
Position: 51: 268. 1140 nucleotid 294

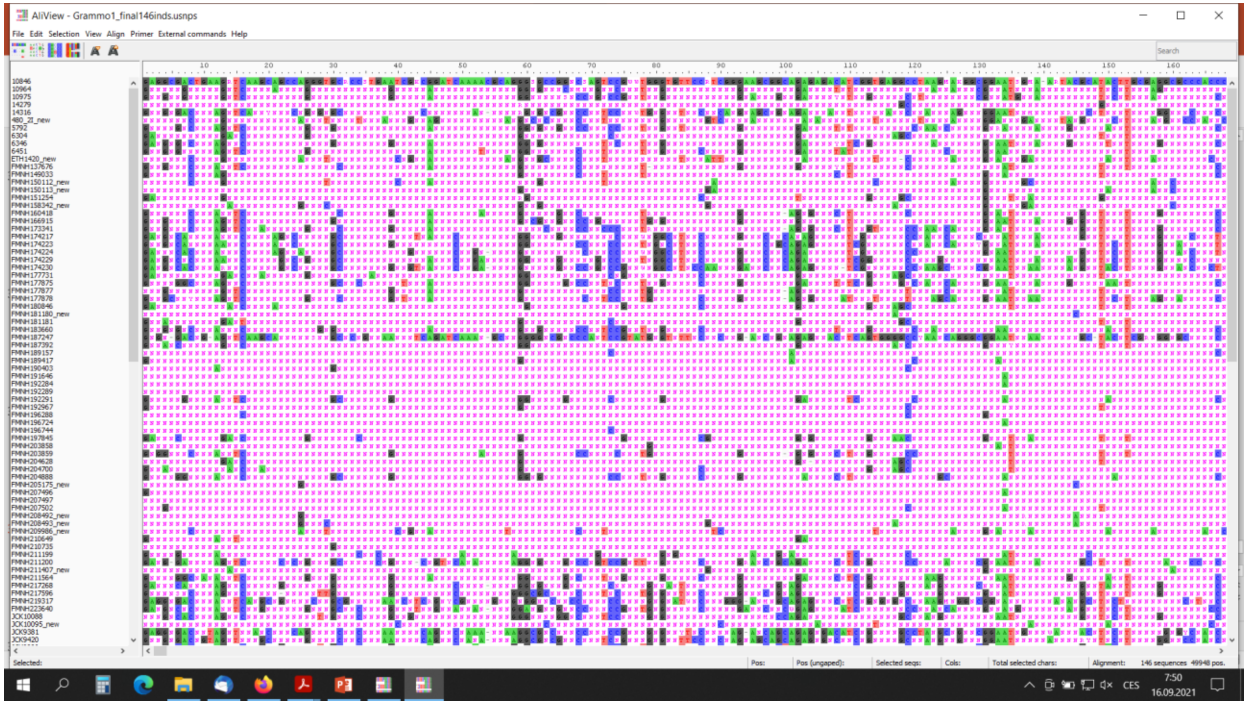
Sequence Mask: None
Numbering Mask: None

```
CTTATAGTCCAAATTCTTACAGGCTTATTCCTAGCTATACACTATACATCAGACACAATAACAGCCTTCTCATCAGTAACACATATCTGCCGAGATGTAAACTATGGGTGACTTATCCGATATTTACATGCTAATGGAGCCTCCATATTCTTTATTTGCCTATTTCTACACGTAGGACGAGGTATATATTATGGA
```

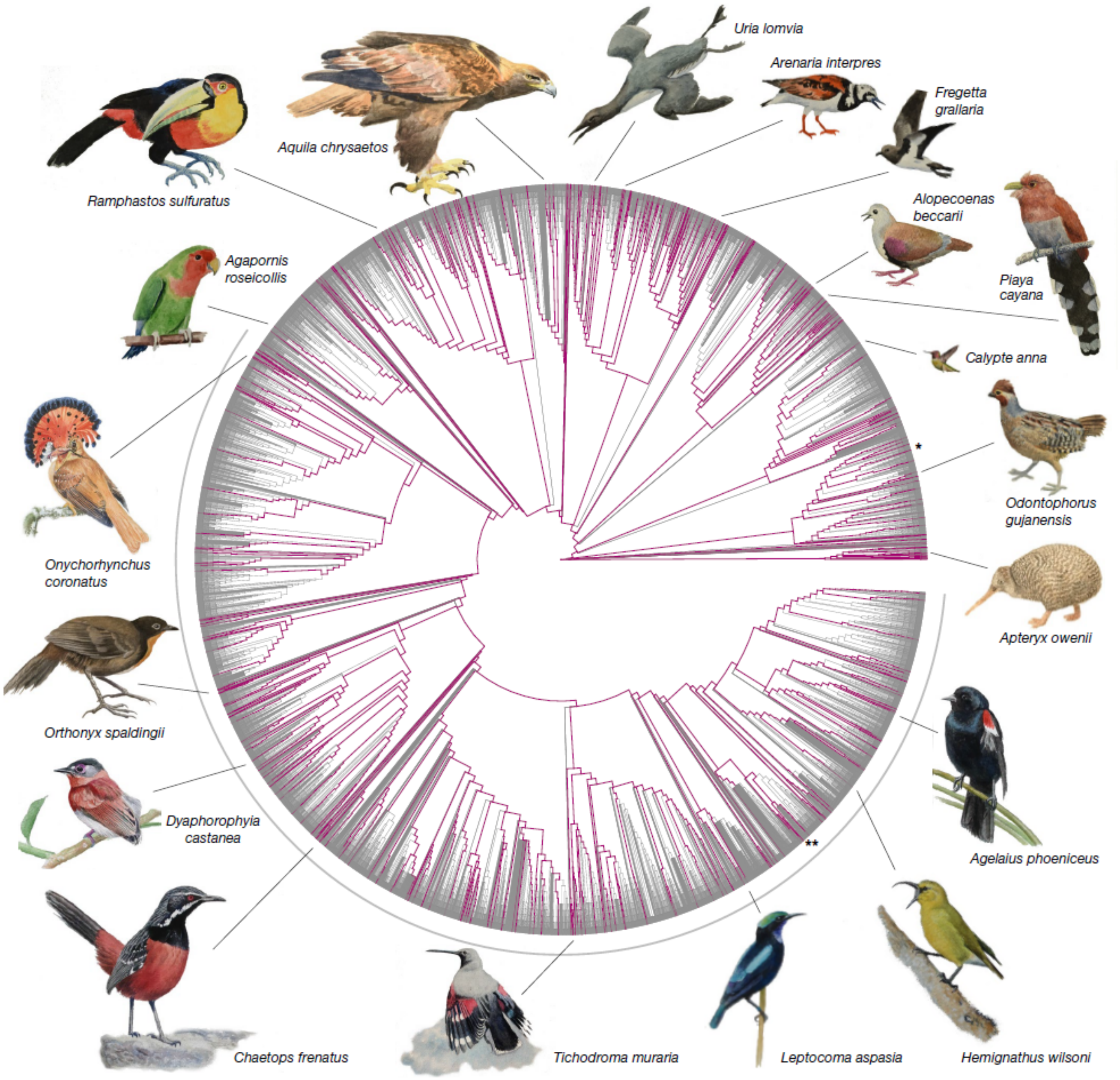
```
Mus B3314 D3.
Mus B3321 E3.
Mus B3325 A12.
Mus RS1222 A1.
Mus RS1223 B1.
Mus RS1225 C1.
Mus RS1226 D1.
Mus RS1265 E1.
Mus RS1269 F1.
Mus RS1270 G1.
Mus RS1283 H1.
Mus RS1369 B2.
Mus RS1375 D2.
Mus RS1414 E2.
Mus RS1415 F2.
Mus RS1448 H2.
Mus RS1474 A3.
Mus RS1512 B3.
Mus RS1525 C3.
Mus RS1528 D3.
Mus RS1529 E3.
Mus RS1530 F3.
Mus RS1531 G3.
Mus RS1546 H3.
Mus RS1547 A4.
Mus RS1548 B4.
Mus RS1549 C4.
Mus RS1575 D4.
RS835.  1140 nu
RS847.  1140 nu
RS785.  1140 nu
MusRS728.  1141
MusRS743.  1140
3025.  1140 nuc
3058.  1140 nuc
132.  1140 nucl
136.  1140 nucl
137.  1140 nucl
154.  1140 nucl
155.  1140 nucl
197.  1140 nucl
198.  1140 nucl
237.  1140 nucl
240.  1140 nucl
241.  1140 nucl
249.  1140 nucl
263.  1140 nucl
267.  1140 nucl
268.  1140 nucl
269.  1140 nucl
3014.  1140 nuc
414
441
442
443
447
454
455
456
464
465
466
534
538
541
Mus musculus mc
```



ddRAD
(„snps“ file)
= 500 000 bp
(depending on
sequencing intensity
etc.)



ddRAD
(„usnps“ file)
= 1 SNP/locus
= 50 000 SNPs
(depending on
sequencing intensity
etc.)



kompletní genomy (Feng et al. 2020 Nature)

Tak, a co teď s těmi daty ...