



# Bi5130 Základy práce s lidskou aDNA

Mgr. et Mgr. Kristýna Brzobohatá

[brzobohata@sci.muni.cz](mailto:brzobohata@sci.muni.cz)

Laboratoř biologické a molekulární antropologie,  
ÚEB, PřF, Mu



## Metody analýzy aDNA:

### Metody I. – 12. 10.

- 1) Výběr vhodného vzorku, očištění, separace tkáně
- 2) Izolace DNA - lyzace, dekalifikace, purifikace
- 3) Amplifikace aDNA

### Metody II. – 19. 10.

- 1) qPCR, klonování aDNA
- 2) Sangerova sekvenace
- 3) Masivní paralelní sekvenování
- 4) aDNA a epigenetika, aRNA

# Metody II.

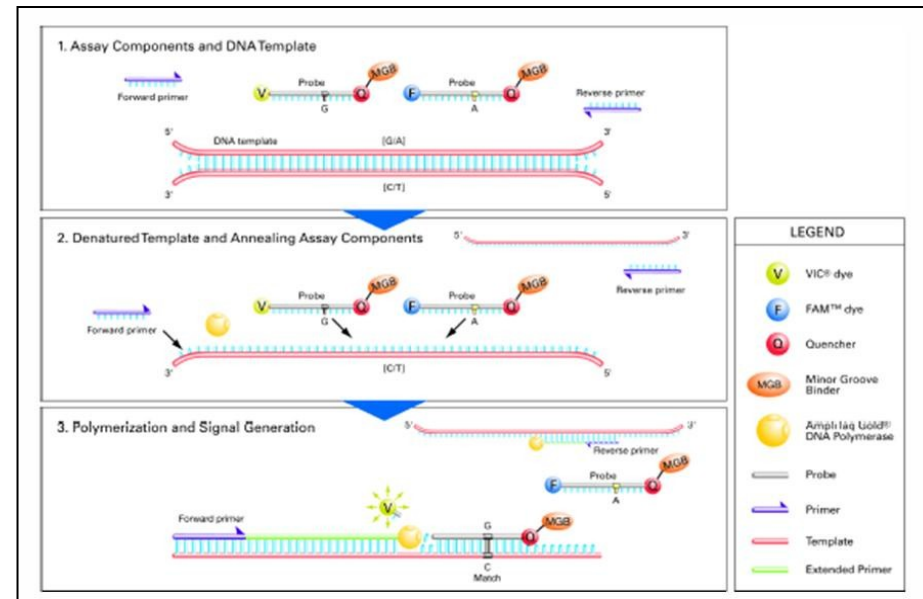
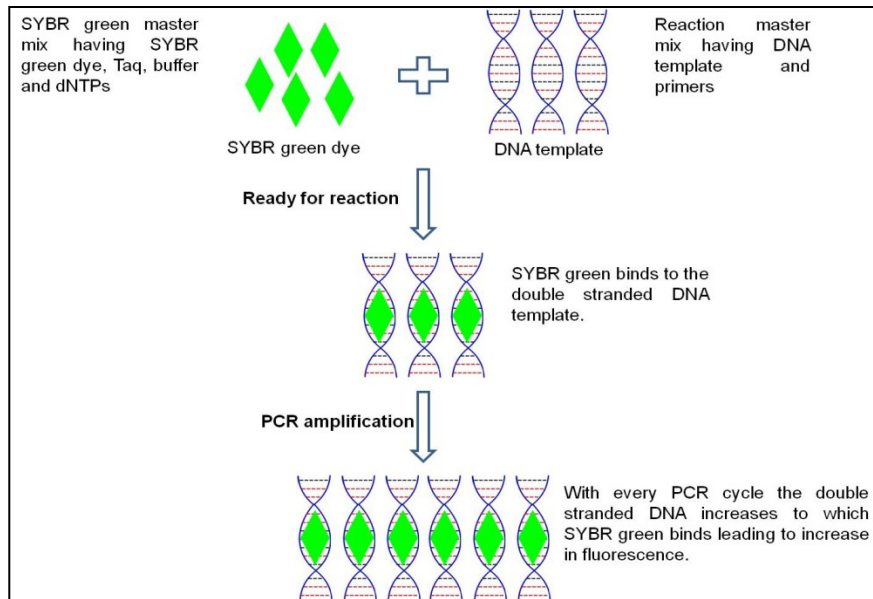


## Kvantitativní polymerázová řetězová reakce a její aplikace na aDNA

Princip qPCR:

<http://www.youtube.com/watch?v=QVeVIM1yRMU>

- použití interkalačních barviv ( SYBR Green, Eva Green)
- použití duálně značených sond (TaqMan®)





### Kvantifikace vzorku aDNA

- Odhalení principů degradace a konzervace molekuly aDNA (Allentoft *et al.*, 2012; Campos *et al.*, 2012)
- Optimalizace purifikačního protokolu (Lee *et al.*, 2010)
- Zvolení nejvhodnějšího postupu následných analýz (King *et al.* 2010)
- Detekce přítomnosti inhibitorů PCR (Sustovic *et al.*, 2005)
- Způsob průkaznosti autenticity aDNA (Haak *et al.*, 2010)

### Optimalizované postupy nebo komerční forenzní soupravy

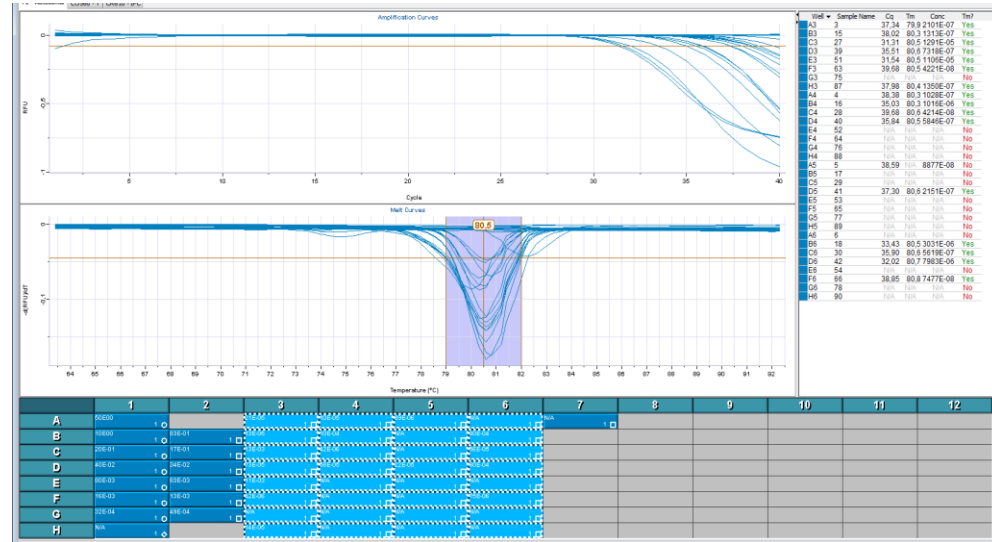
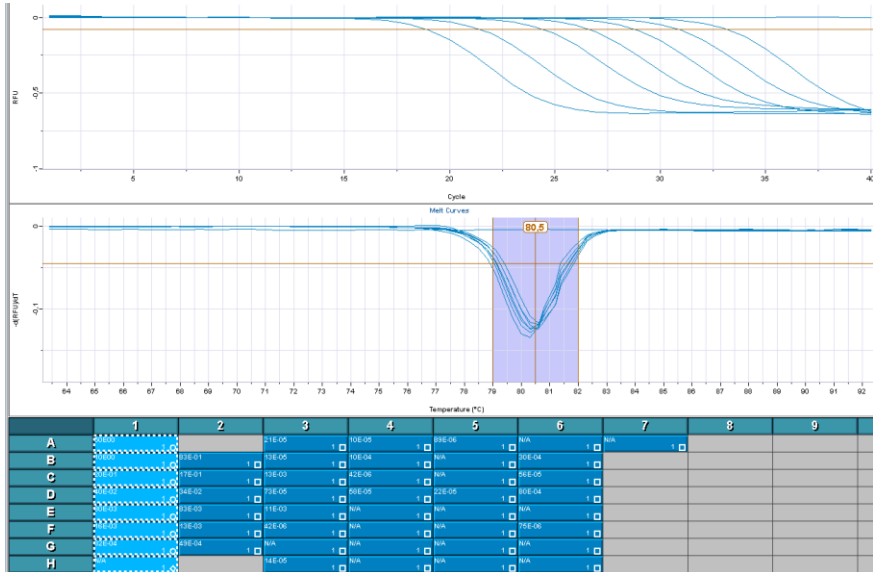
Vlastní design - mitochondriální DNA i DNA nukleární (Alonso *et al.*, 2004).

Forenzní soupravy – odolné inhibitorům; během jedné reakce je amplifikován úsek autozomální DNA i marker na chromozomu Y (Krenke *et al.*, 2008).

# Metody II.



## Kvantifikace vzorku aDNA – forenzní kit Plexor HY



Během jedné reakce jsou amplifikovány standardy se známou sekvencí a neznámé vzorky. Duálně značené sondy nasedají **lokus na chromozomu Y (SR)** a **na nukleární DNA**. Narůst fluorescence ( $\Delta R_n$ ) u standardů je porovnáváno s nárůstem  $\Delta R_n$  u neznámých vzorků. Z měřené míry fluorescence u standardů je dopočítána koncentrace u vzorků.

[promega.com/resources/protocols/technical-manuals/0/plexor-hy-system-protocols](http://promega.com/resources/protocols/technical-manuals/0/plexor-hy-system-protocols)

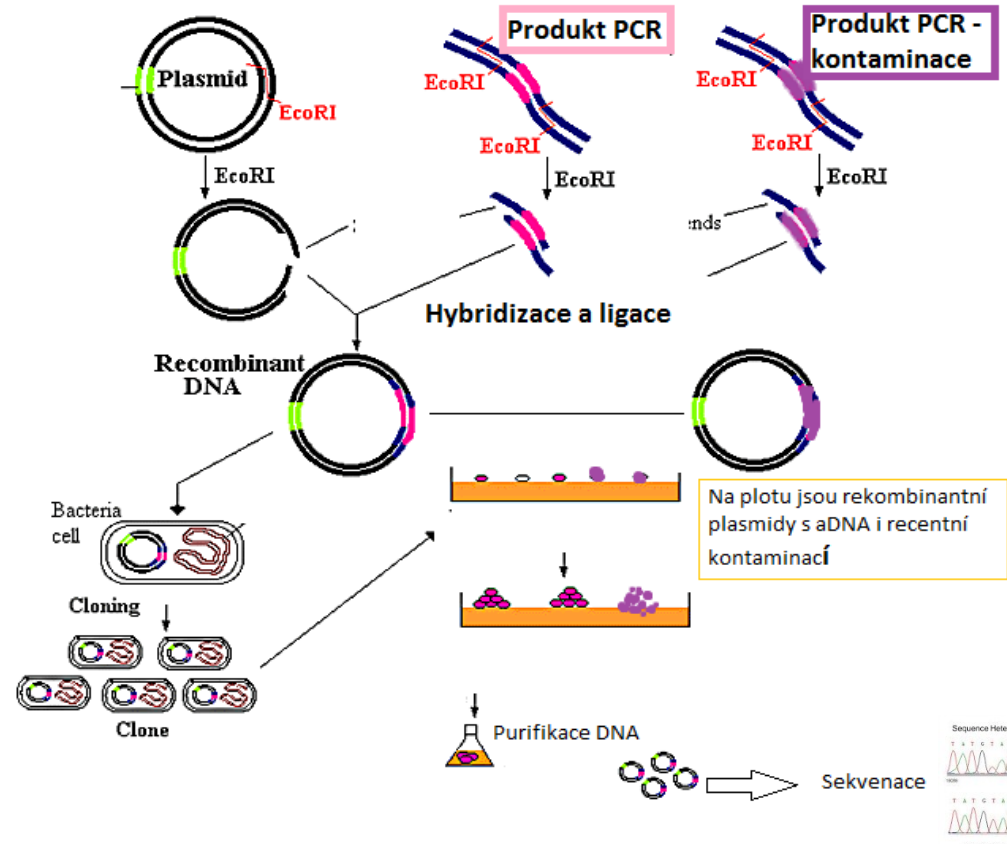
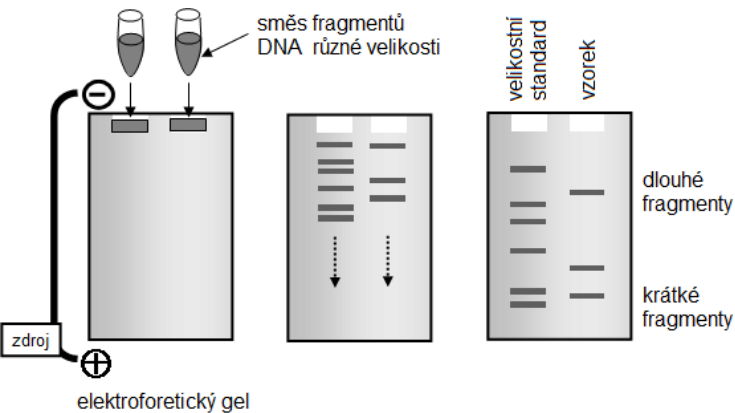


## Klonování aDNA

Po konvenční amplifikaci jsou vzorky vizualizovány pomocí gelové elektroforézy. Pokud byla PCR úspěšná v drtivé většině případů následuje sekvenování.

Je vhodné, aby samotnému sekvenování předcházelo klonování produktů PCR v plasmidech. Tento krok odhalí případné post mortem mutace nebo kontaminaci tím, že **výsledky sekvenování jednotlivých kolonií nebudou totožné.**

**!VERIFIKACE ZÍSKANÝCH VÝSLEDKŮ!**

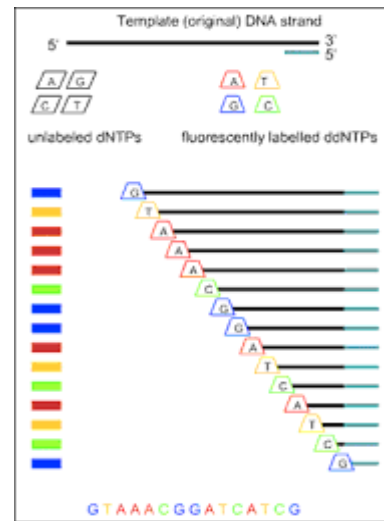
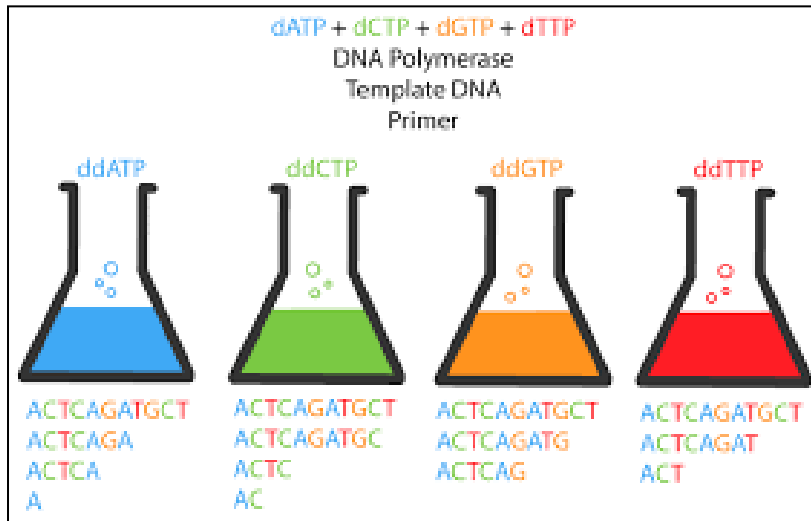


## Metody II.



### Sangerova sekvenace

- použití modifikované dideoxy-terminátorové metody
- použití fluorescenčních barviv pro detekci řetězců DNA
- elektroforetická separace produktů v kapiláře
- použití fotobuňky k detekci fluorescence barviv při jejich průchodu gelem nebo kapilárou
- přímý přenos výstupu fotobuňky do počítače – automatické analyzování výsledků inkorporace terminačních fluorescenčně značených dideoxynukleotidů (ddNTP)





### Sekvenování nové generace (Next gene sequencing)

- Známe také jako masivní paralelní sekvenování, sekvenování druhé generace
- Bez ohledu na použitou platformu NGS zahrnuje následující kroky:
  1. Příprava genomové knihovny
  2. Sekvencování a čtení
  3. Bioinformatická analýza

Platformy – Solid, Roche 454

- Illumina
- Ion Torrent (Thermofisher)



Ion PGM™ Sequencer



Ion Proton™ Sequencer



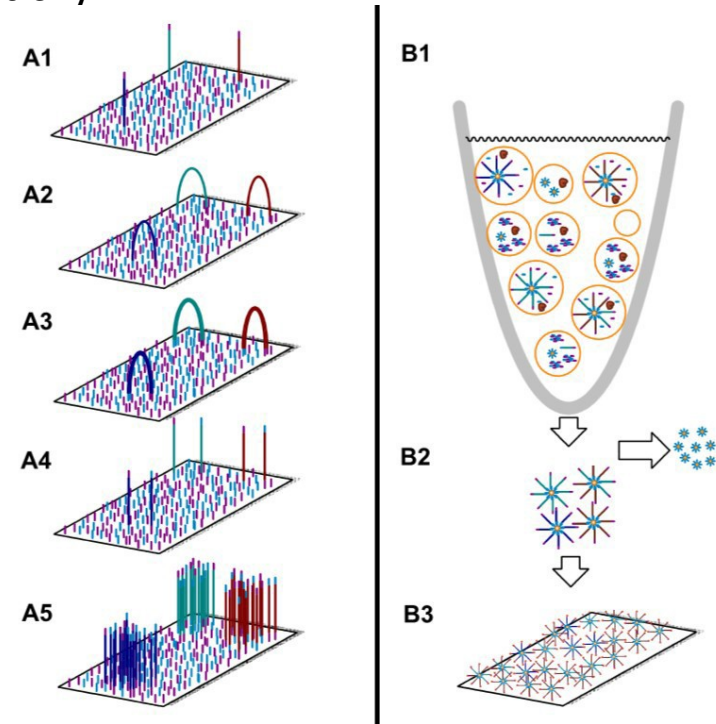
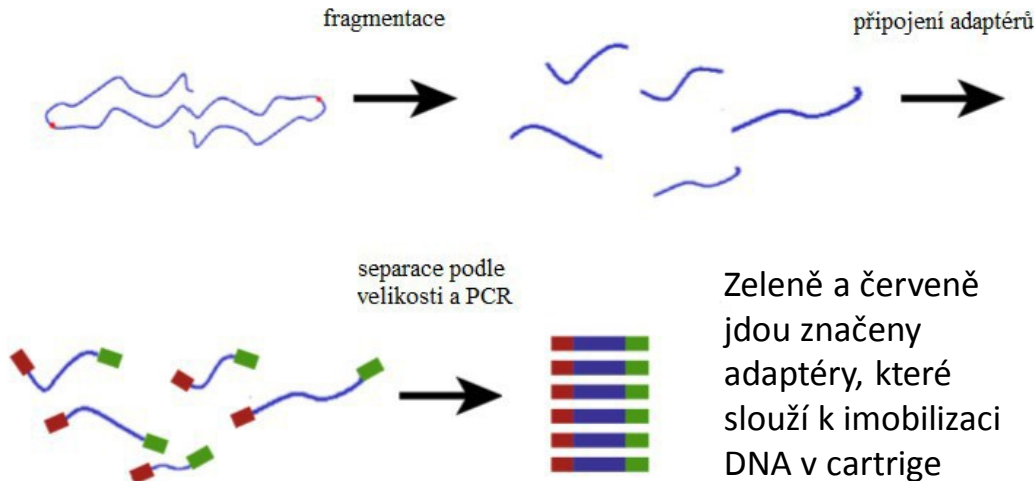


# Metody II.



## Příprava genomové knihovny – workflow (Illumina, Ion Torrent)

1. Fragmentace molekuly DNA – sonikace nebo enzymaticky
2. Připojení adaptérů
3. Imobilizace fragmentů DNA
4. Amplifikace – můstková (A) nebo emulzní (B)





## Sekvenace syntézou

Nukleotidy jsou značeny fluoroforem, při nasedání na aDNA jsou štěpeny a emituje záření. Čipem proudí tekutina, která přináší a odplavuje nové nukleotidy. Produkty PCR tvoří clustery.

**Illumina** – sekvenuje se jen z přímého řetězce.

<https://www.youtube.com/watch?v=womKfikWlxM>

**Ion Torrent** – na principu měření pH při začleňování nukleotidů.

<https://www.youtube.com/watch?v=WYBzbxIfuKs>

Data jsou v reálném čase online streamována na BaseSpace, odkud jsou dostupné na další analýzy.

# Metody II.



## Metagenomické analýzy aDNA - shotgun vs. target sequencing

Shotgun sequencing = celometagenomové sekvenování = „vše co je ve vzorku“

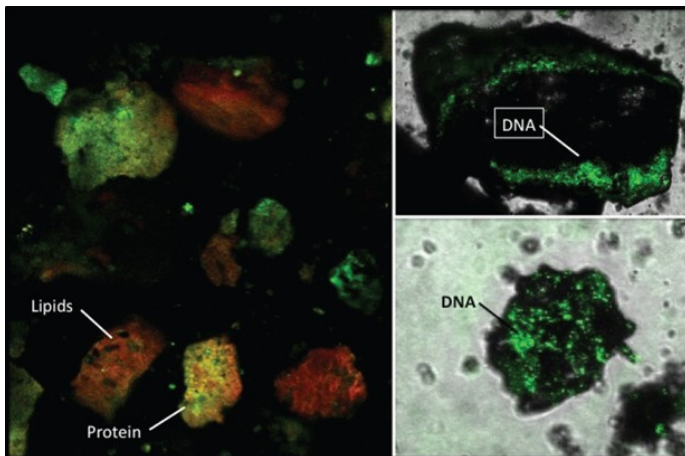
RESEARCH ARTICLE

### Retroviral DNA Sequences as a Means for Determining Ancient Diets

Jessica I. Rivera-Perez<sup>1\*</sup>, Raul J. Cano<sup>2</sup>, Yvonne Narganes-Storde<sup>3</sup>, Luis Chanlatte-Baik<sup>3</sup>, Gary A. Toranzos<sup>1</sup>

<sup>1</sup> Department of Biology, University of Puerto Rico, Rio Piedras Campus, San Juan, Puerto Rico, <sup>2</sup> Center for Applications in Biotechnology, Biological Sciences Department, California Polytechnic State University, San Luis Obispo, California, United States of America, <sup>3</sup> Center for Archaeological Research, University of Puerto Rico, Rio Piedras Campus, San Juan, Puerto Rico

\* [jessica.i.rivera68@gmail.com](mailto:jessica.i.rivera68@gmail.com)



	Proviral sequence detected	Corresponding gene identification	Virus host organism	Previous osteological, archaeofaunal and botanical results
<b>Vertebrates<sup>a</sup></b>	Moloney murine leukemia virus	reverse transcriptase	<i>Mus musculus</i> (rodent)	<i>Isolobodon portoricensis</i> (hutia)
	Murine endogenous retrovirus	retrotransposable element ORF2	<i>Mus musculus</i>	<i>Heteropsomys insulans</i> (spiny rat)
	Bat endogenous retrovirus	polymerase polyprotein	<i>Pteropus alecto</i>	ND
	Tiger frog virus	thymidylate synthase	Frog	ND
	Grouper iridovirus	unknown protein	<i>Epinephelinae</i> (fish)	<i>Epinephelinae</i>
	Avian pox virus	HAL3 domain	Fowl, parrots, and other birds	<i>Pelicanus occidentalis</i> (among many other birds)
	Monkey endogenous retrovirus	H element-like protein	Marmosets	ND
<b>Invertebrates<sup>a</sup></b>	ND <sup>c</sup>			<i>Cardisoma guanhumii</i> (crab)
	ND			<i>Gecarcinus spp</i> (crab)
	ND			<i>Donax denticulatus</i> (oyster)
	ND			<i>Cittarium pica</i> (gastropod)
	ND			<i>Pelurodonte caracolla</i> (land snail)
	Mulberry endogenous retrovirus	polymerase polyprotein, transposon TNT 1–94	<i>Morus notabilis</i> (berry fruit)	ND
	ND			<i>Psidium guajava</i> (guava fruit)
<b>Plants<sup>b</sup></b>	ND			<i>Zea mays</i> (corn) <sup>d</sup>
	ND			<i>Ipomoea batatas</i> (sweet potato) <sup>d</sup>
	ND			<i>Cassava spp</i> (yucca)
	ND			<i>Capsicum spp</i> (pepper) <sup>d</sup>
	ND			

Viral sequences detected in this study are compared to diet components previously suggested by osteological, archaeofaunal and botanical studies of pre-Columbian cultures in Puerto Rico and Vieques.

<sup>a</sup>Previously suggested diet includes observations by the Spanish during their interactions with Caribbean indigenous cultures [20,21], as well as osteological and archaeofaunal analyses of conch shells, teeth and bones obtained from archaeological deposits in Sorcé, Vieques and Puerto Rico [22–25].

<sup>b</sup>Plants suggested as diet components include those proposed by archaeologists as well as plant remains recently identified by paleobotanists through microscopy analysis [26,27].

<sup>c</sup>ND none detected.

<sup>d</sup>Although these diet components were not detected in this shotgun metagenomic analysis, previous *18S rRNA* studies determined their presence in these coprolites [8].

doi:10.1371/journal.pone.0144951.t003

## Metody II.

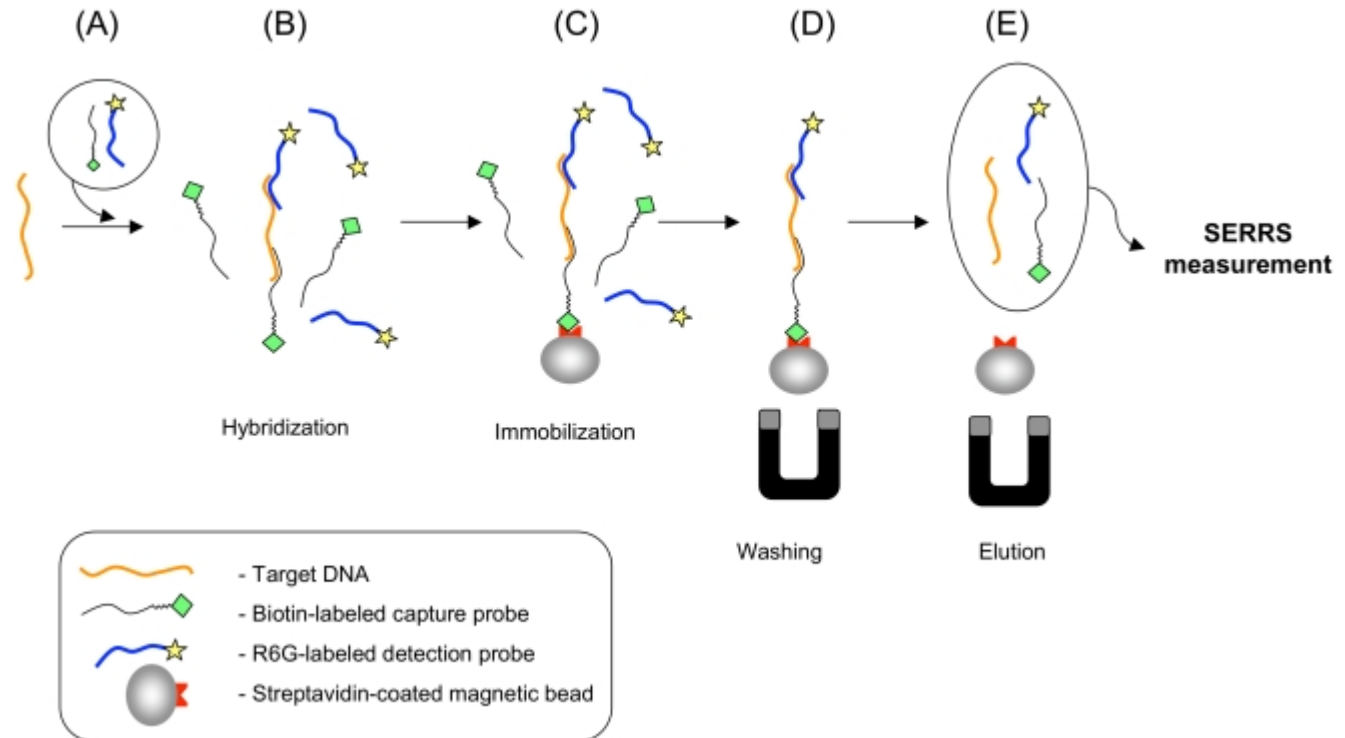


### Target sequencing

Target sequencing = enrichment genomů / genů zájmu

Označení může být provedeno metodou **capture hybridization** nebo jako **PCR enrichment**

### Capture hybridization



## Metody II.

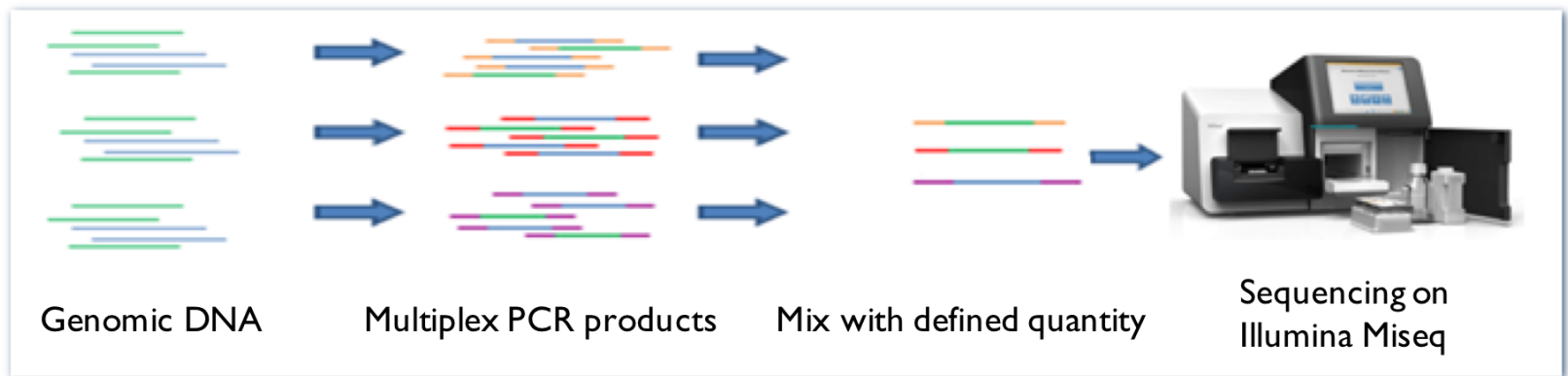


### Target sequencing

Target sequencing = enrichment genomů / genů zájmu

Označení může být provedeno metodou **capture hybridization** nebo jako **PCR enrichment**

### PCR Enrichment

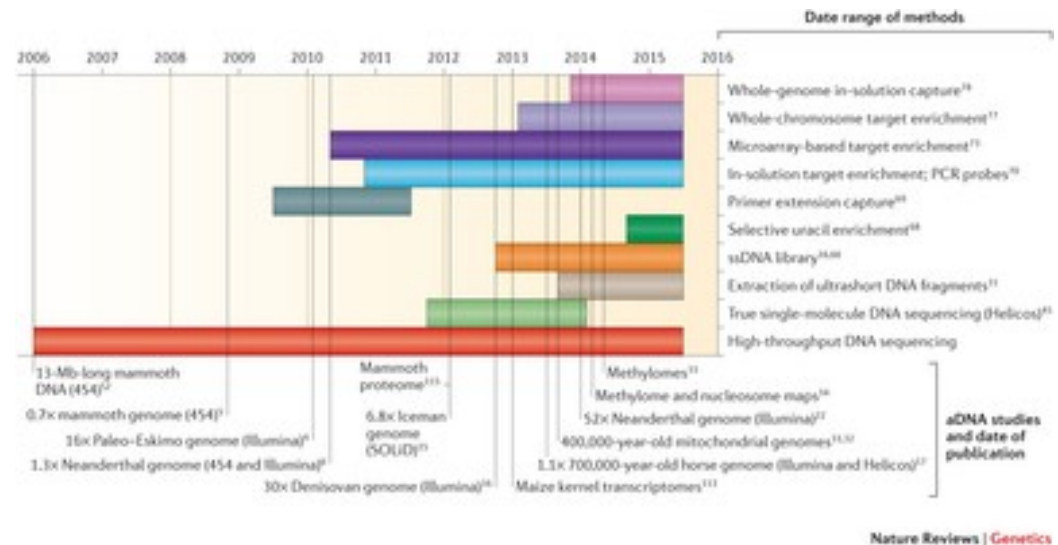


### **Schematic Workflow for FastTarget**



## Metodický přístup NGS a problémy vyvstávající z charakteristiky aDNA

- Metagenomika?
- Capture hybridization
- Enrichment – multiplex PCR
- Optimalizace použité polymerázy
- Barcoding - poolování vzorků
- Double read i single read
- **Hodnocení NGS dat!**



<http://www.nature.com/nrg/journal/v16/n7/full/nrg3935.html>



## Hodnocení výsledků NGS

**PALEOMIX** (University of Copenhagen, L. Orlando a kol.)

- Soubor implementovaných procedur na analýzu dat získaných z NGS
- Pipeline – zřetězované zpracování – překrývání strojových instrukcí, sled na sebe navazujících skriptů, které pracují paralelně vedle sebe a které umožňují analýzu velkého množství dat
- Vysoké nároky na server a uživatele (not really user friendly; prostředí Linux, iOS)
- 16-jádrový procesor; 800 GB operačního místa; 32 GB RAM; vysokorychlostní připojení na internet
- Možnost použití na moderní i starobylou DNA
- Identifikace poškození aDNA, oprava poškozené části
- Taxonomické zařazení do fylogenetického stromu i v případě vyhynulého druhu
- Analýza dat získaných ze shotgun i z target-enrichment sekvenování
- Fylogenomická a metagenomická analýza
- Umožnil charakterizovat úplný mitochondriální genom 29 vyhynulých druhů



## Hodnocení výsledků NGS

### **PALEOMIX**

**1) Genome alignment** = spojování jednotlivých osekvenovaných částí do řetězce, který bude odpovídat referenčnímu souboru knihovny tak aby bylo možno je porovnat.

**2) Resequencing** = reskvenování – porovnávání už osekvenovaného genomu s referenční knihovnou ; identifikace odlišností na úrovni genomu – SNP, příbuznost s referenčním souborem, zařazení do podskupin, případně vyloučení ze skupiny

**3) Authenticating aDNA data + Estimating contamination levels**

mapDamage, pmdtools – testování známých post-mortem poškození DNA

Deaminace cytosinu





## Hodnocení výsledků NGS

**Genome completion and error rates** = zabránění špatné interpretaci dat, aby se zabránilo špatnému fylogenetické zařazení, používá se třístupňové seřazení -

Three-way alignments:

Porovnávání s jedincem stejného druhu ze stejně starého vzorku (ancient DNA)

Porovnávání s jedincem stejného druhu ze současnosti (recent DNA)

Porovnávání s jedincem z taxonomicky blízkého druhu (šimpanz)

Vlastní chyby jsou však často závislé na rozsahu poškození samotné DNA a teda často neovlivnitelné

### **Další využití PALEOMIXU:**

Fylogenetická rekonstrukce (propojení evoluce a genomiky)

Rekonstrukce populační historie - např. Admixture

Sestrojení genome-wide epigenetických map – na základě míry metylace jednotlivých genů, resp. lokusů