

Usporedba metoda za sastavljanje i anotaciju mitohondrijskih i jezgrinih genoma na primjeru filogenije divokoza (*Rupicapra spp.*)

Dan doktorata biotehničkog područja 2021.

17.09.2021



DNA kao dokaz o distribuciji i vitalnosti ugrožene Balkanske divokoze

STUDENT:
Toni Tešija, mag. ing. agr.

STUDIJSKI SAVJETNIK/CI:
doc. dr. sc. Toni Safner

Uvod

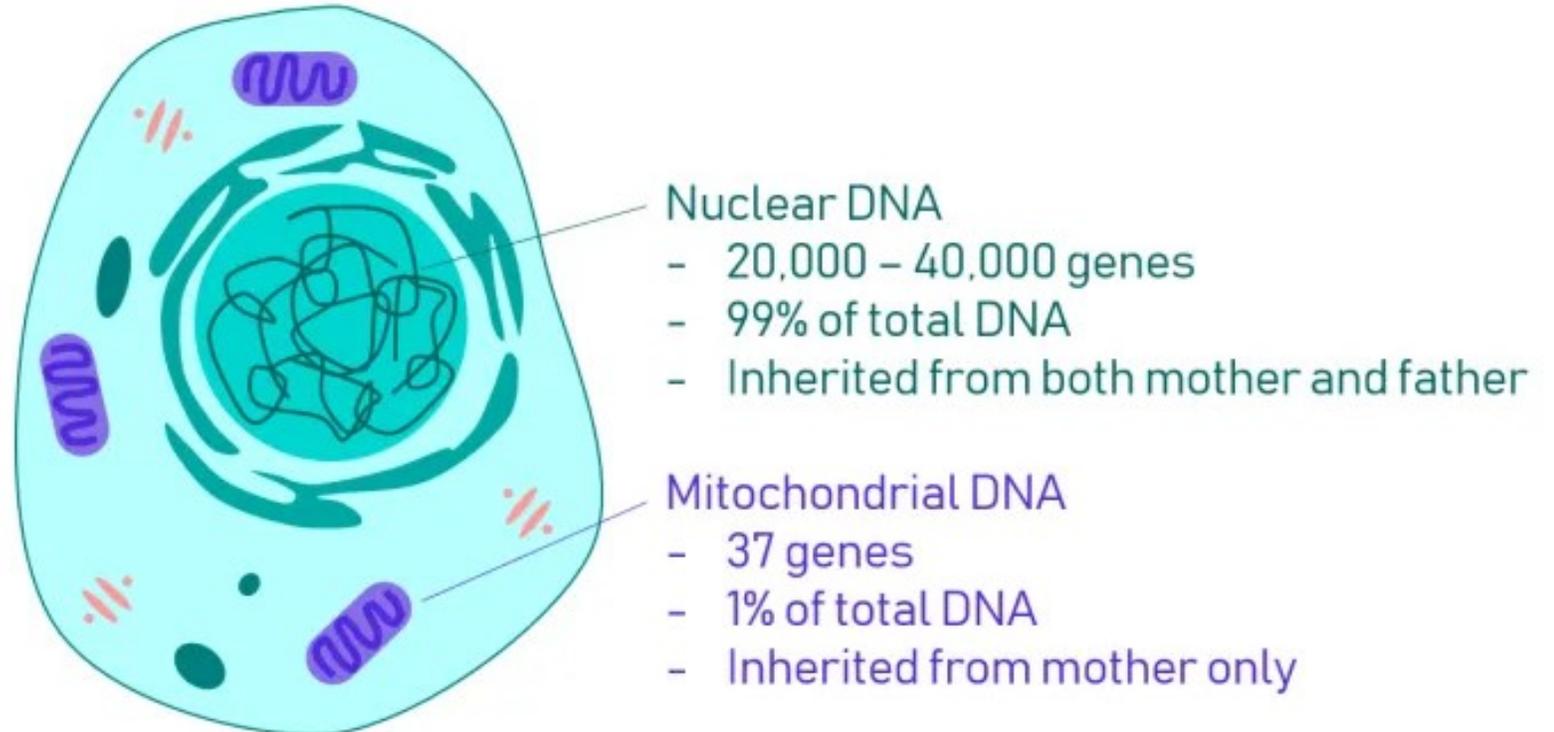
GENOM → skup svih genetičkih informacija nekog organizma

- Jezgrin genom ~ 3 Gb
- Mitohondrijski genom (mtDNA) ~ 16 kb

ACTG

Rekonstrukcija genoma:

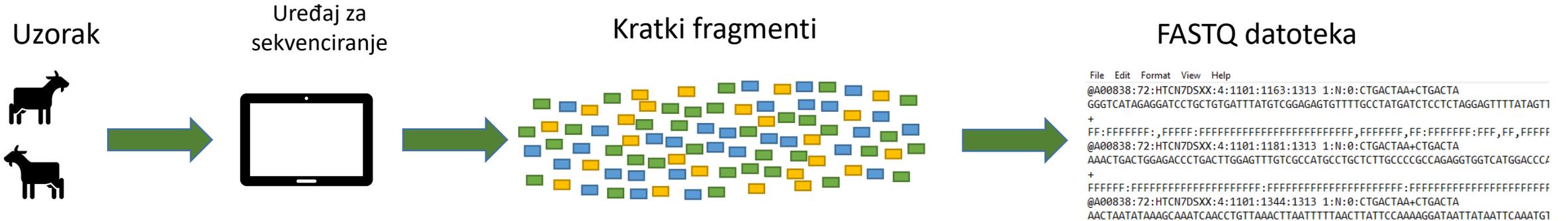
- Sekvenciranje
- Sastavljanje
- Anotacija



Izvor: www.harward.edu

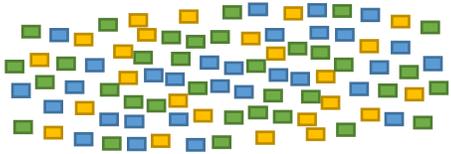
Sekvenciranje DNA

- Uređaj za sekvenciranje lomi i sekvencira stotine tisuća kratkih fragmenata DNA
- FASTQ datoteka → informacija o svakom fragmentu: kvaliteta svake baze, redoslijed i broj nukleotida, orijentacija (4 linije po fragmentu)



Sastavljanje genoma mapiranjem

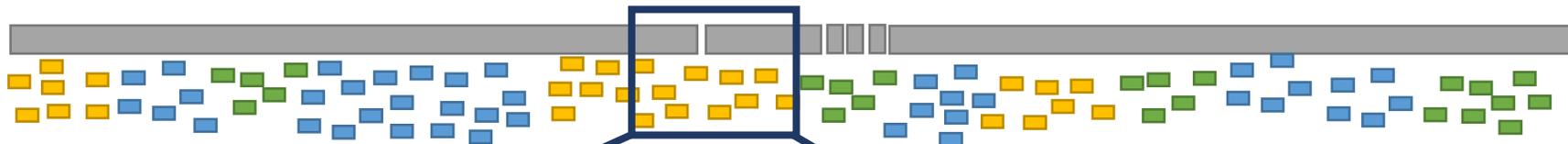
1. KRATKI FRAGMENTI



2. REFERENTNI GENOM



3. MAPIRANJE DNA FRAGMENTATA



ATCGTGAAGCCTGACGTGAGCTGAGCGTGATTAACGCAC

ATAA TAAG

Fragment 1

CTAA CGTGG

Fragment 2

TGACGTGG

Fragment 3

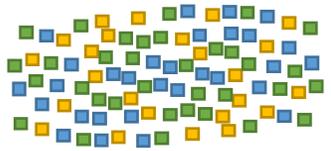
AACGCTG

Fragment 4

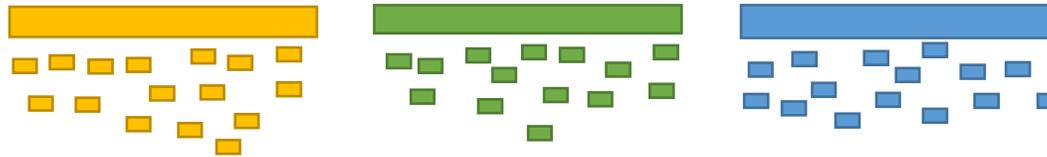
SNPs
INDELS
SNVs
CSVs

Sastavljanje genoma *de novo*

1. KRATKI FRAGMENTI



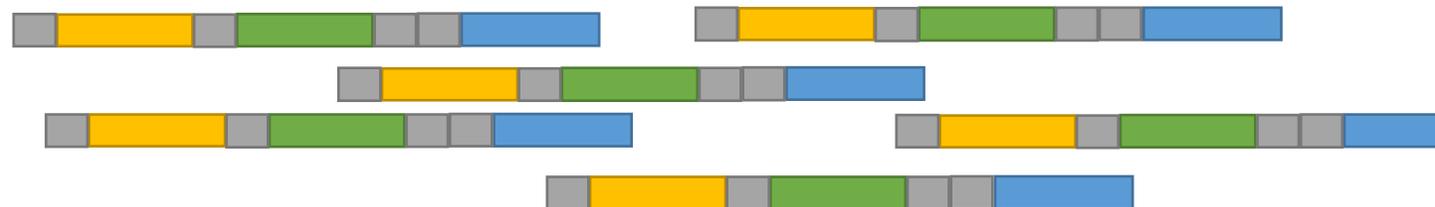
2. KONTIG (eng. *contigs*)



3. SKAFOLD (eng. *scaffolds*)



4. POPUNJAVANJE PRAZNINA (eng. *gap filling*)



3. METODA
HIBRIDNA
METODA
SASTAVLJANJA

Anotacija genoma

Anotacija - **biološke** informacije povezuju se s dijelovima genoma

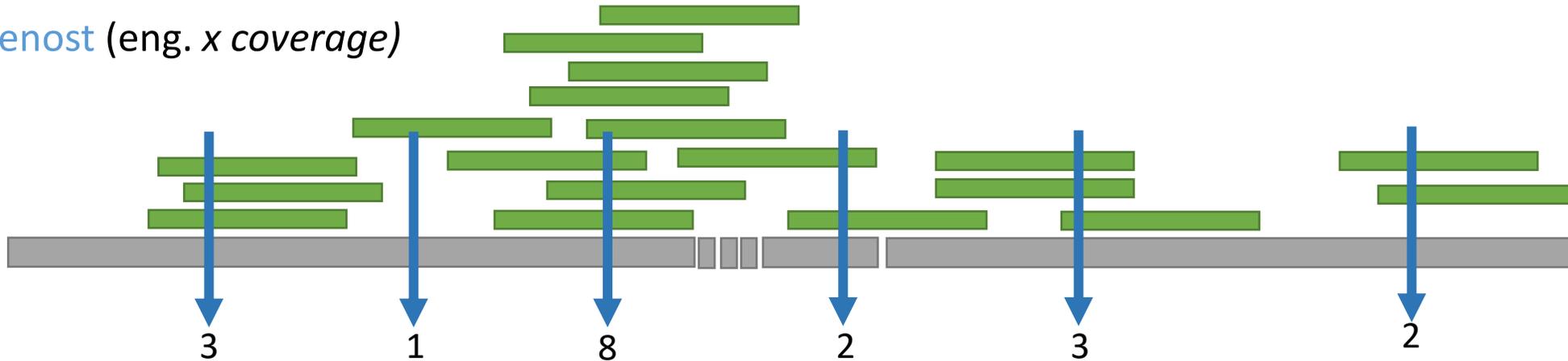
- **Strukturna** anotacija - identifikacija elemenata u genomu (kodirajuća područja i genska struktura)
- **Funkcionalna** anotacija - vezanje biološke informacije na genomske elemente
- Programski alati analiziraju **strukturu** i **sastav** sekvence te je uspoređuju s poznatim genomom srodne vrste

NAD4L CYTB Dloop
NAD5 12S RNA
NAD2 NAD4 NAD6 16SRNA
NAD3 NAD1
ATP6 COX3 COX1 ATP8
COX2



Rekonstrukcija genoma - IZAZOVI

1. Pokrivenost (eng. *x coverage*)



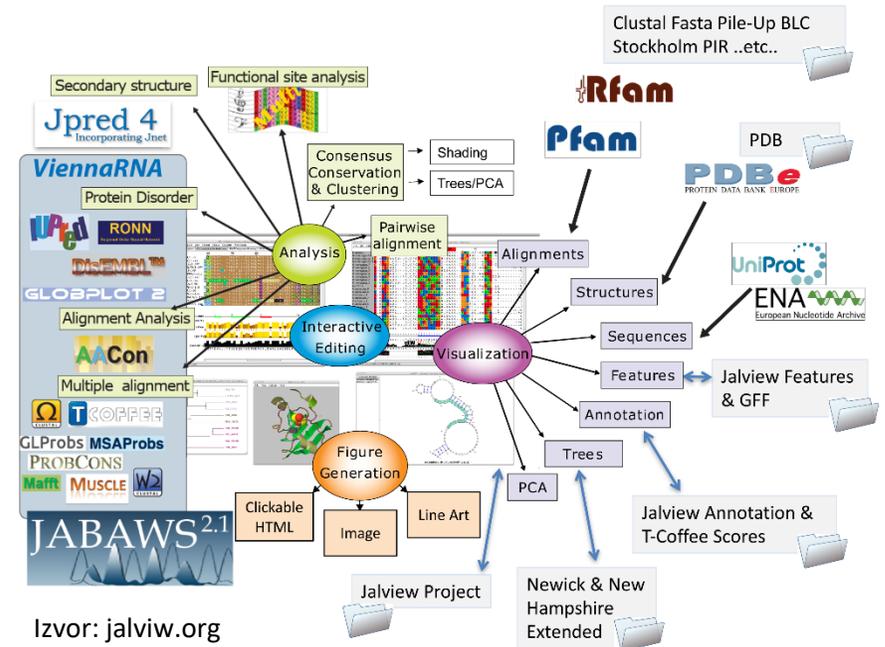
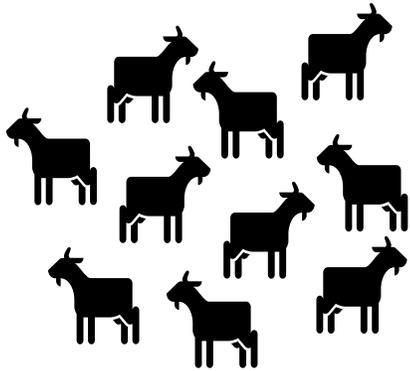
2. Veličina genomskih podataka (~50 GB po uzorku)

Prosječan broj koji prikazuje koliko je puta pojedina baza sekvencirana

3. Pohrana genomskih podataka i rezultata analiza

4. Vrijeme računanja

Materijali i metode



Dosadašnji rezultati

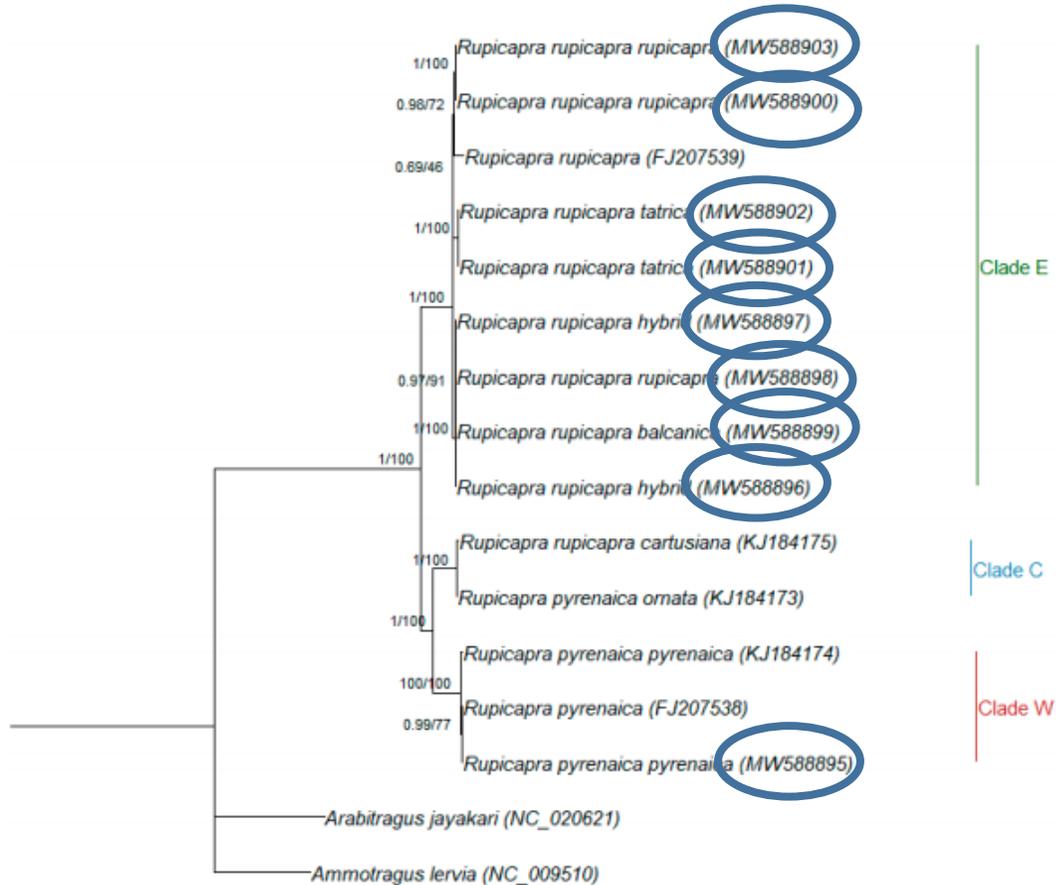


Figure 3. Rooted phylogenetic tree obtained by Bayesian inference for *Rupicapra* only. Nodal supports are indicated above the nodes (posterior probability for Bayesian inference and bootstrap percentages for maximum likelihood, respectively). GenBank Accession numbers are presented in parentheses.

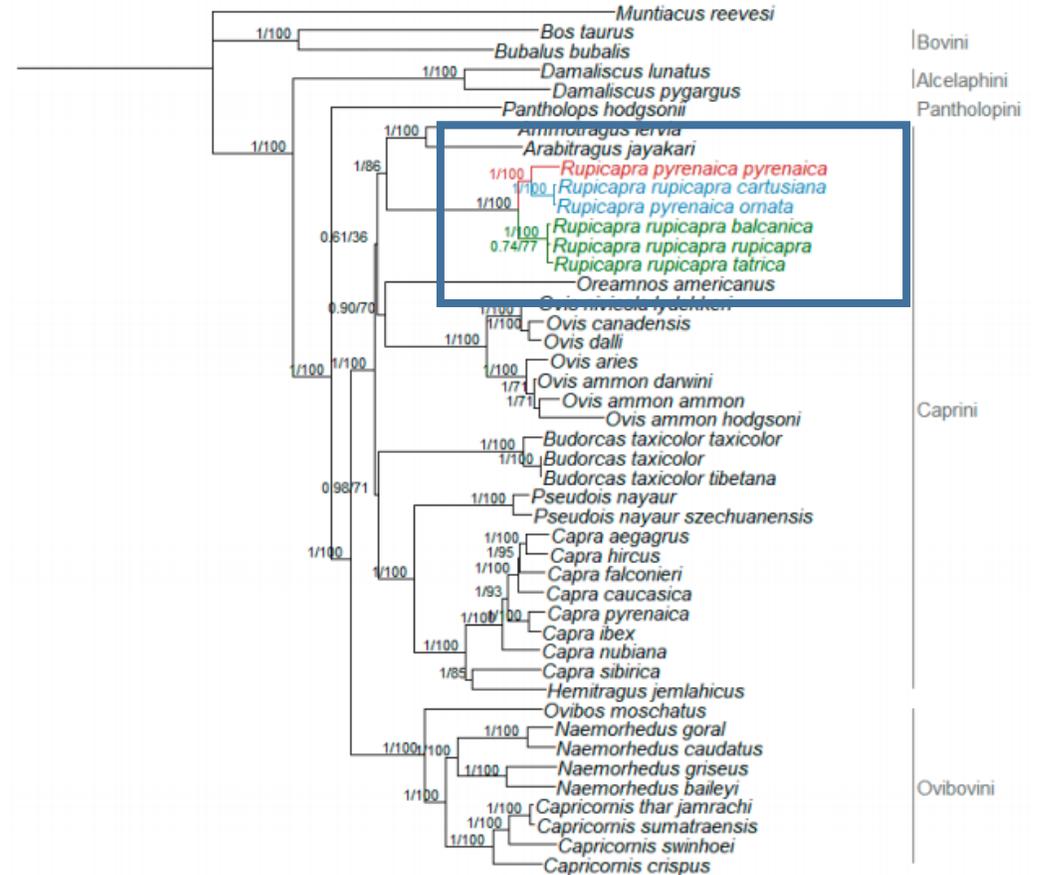


Figure 2. Rooted phylogenetic tree obtained by Bayesian inference for the whole *Caprinae*. Nodal supports are indicated above the nodes (posterior probability for Bayesian inference and bootstrap for maximum likelihood, respectively). Colored branches represent *Rupicapra* clades: red—clade W; blue—clade C; and green—clade E.

Hvala na pažnji!

