

2024 | DANI DOKTORATA
BIOTEHNIČKOG PODRUČJA



Inbreeding and selection on the X chromosome in domestic animal populations

Mario Shihabi

University of Zagreb Faculty of Agriculture

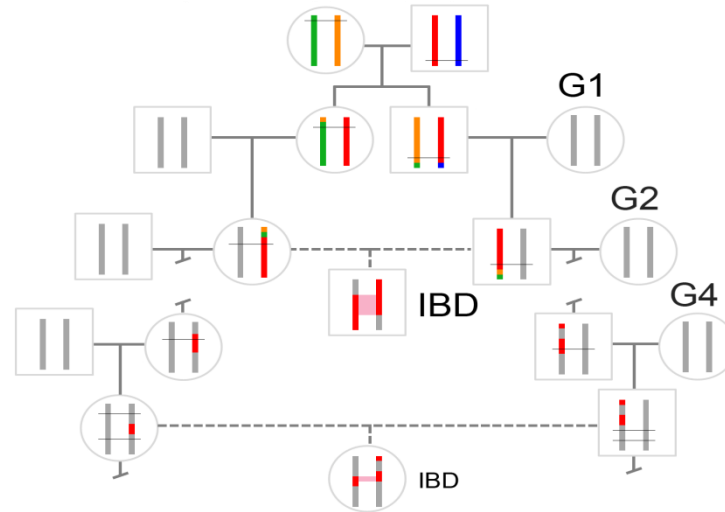
Inbreeding

- Mating of closely related individuals
- Unavoidable due to small N_e
- Used also for \uparrow productivity

\uparrow proportion of homozygosity

NEGATIVE CONSEQUENCES

- Inbreeding coefficient (F):
0 - 1 (or -1 to 1)



THROUGH HISTORY – Pedigree
MODERN AGE – **Genomic assessment**

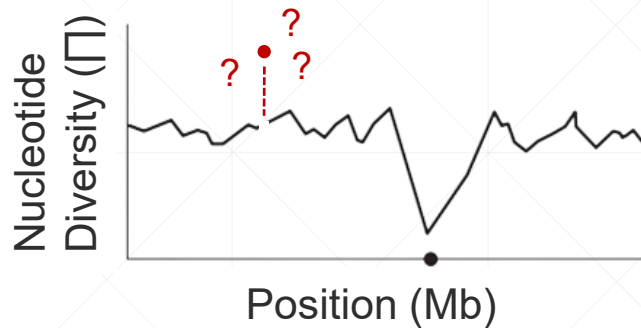


Selection

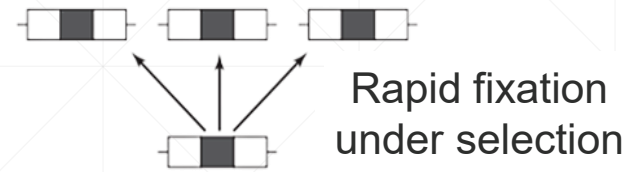
- Selection signals in the genome of the population
- Change in the frequency of certain alleles

a) Positive selection signals

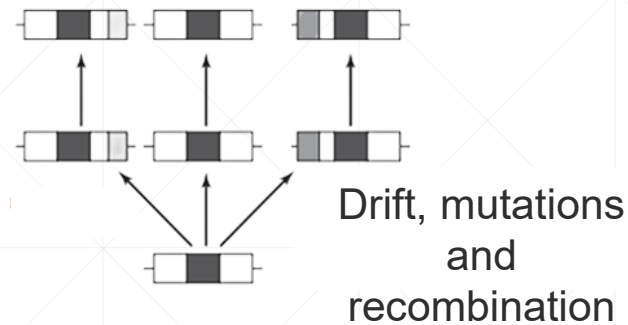
b) Balancing selection signals



a) *Hard sweep*: eROHi



b) *Soft sweep*: iHS, nSL



X chromosome

Sex ↔ Sex chromosomes

Two main systems of sex chromosomes in diploids:

XX / XY
♀ ♂

ZW / ZZ
♀ ♂



a) Pseudo-Autosomal Region (PAR)

b) nonPAR X chromosome

c) nonPAR Y chromosome

nonPAR X chromosome

- ↑ Genetic drift
- ↓ Recombination rate
- ↑ Linkage disequilibrium
- ↓ Mutation rate
- ↓ Ne
- ↓ Genetic diversity



Problem

- Lack of research on inbreeding and selection on the X chromosome in domestic animal populations
- Influence on many economically and evolutionarily important traits
- Good potential for methodological progress (male -> hemizygous status)



Hypotheses & Objectives

Inbreeding

on

X chromosome

VS

Autosomes

1. Higher?
2. With greater variability?

Y X



Based on
haplotype richness
difference

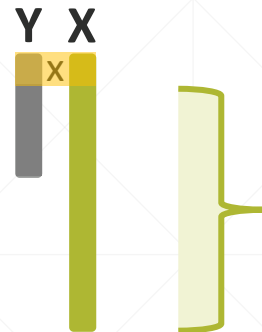
3. To identify positive
selection signals?



Hypotheses & Objectives

1. Estimate and compare Inbreeding on X chromosome *AND* Autosomes in domestic animal populations

2. Establish new method for identifying positive selection signals



Based on
haplotype richness
difference

3. Test it in native Croatian sheep breeds



Materials & Methods

Two distinct populations per:

Cattle

Illumina BovineHD BeadChip
(777 K SNPs)

Metapopulation of native Croatian breeds 79 ♀ ; 33 ♂
(ANAGRAMs project)

Nellore 1267 ♀ ; 973 ♂
(Brazil)

Dog

Illumina CanineHD BeadChip
(172 K SNPs)

Labrador Retriever 129 ♀ ; 145 ♂
(Binversie et al., 2020)

Patagonian Sheepdog 41 ♀ ; 105 ♂
(Barrios et al., 2022)

Sheep

Illumina OvineHD BeadChip
(606 K SNPs)

Metapopulation of native Croatian breeds 101 ♀ ; 101 ♂
(ANAGRAMs project)

Soay 102 ♀ ; 87 ♂
(Johnston et al., 2016)

Inbreeding

Selection

+ 10 mouflons
ancestral
information

Materials & Methods

Runs Of Homozygosity (ROH)

a) Empirical approach → GOLDEN HELIX Accelerating the Quest for Significance™

- Min. ROH = 1 Mb, Min. SNP = 15
- het. and missing per class (Ferenčaković *et al.*, 2013)

b) Statistical approach → RZooROH

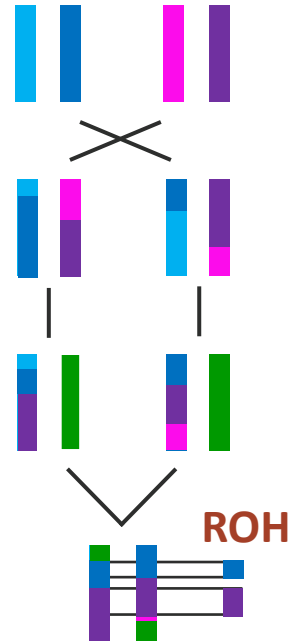
- 1 cM = 1 Mb (Druet and Gautier, 2017)

Inbreeding based on IBD status (F_{ROH}) → Spans from 0 to 1

In males and
females separately
&
In females only

- a) Autosomal (F_{ROH_SVS} & $F_{ROH_RZooROH}$)
- b) X chromosomal ($F_{ROH_SVS_X}$ & $F_{ROH_RZooROH_X}$)

$$F_{ROHk} = \frac{\sum_k \text{Length (ROH}_k\text{)}}{\text{Lgenome (info used)}}$$



Materials & Methods

Inbreeding based on IBS status (F_{LH1} , F_{VR1} , F_{YA2}) -> Spans from -1 to 1

F_{LH1} (Li & Horvitz, 1953)

- **PLINK** → F_{HOM}
- Based on deviations from HW equilibrium

F_{VR1} (VanRaden, 2008)

- **GCTA** → F_{HOM} with the option --make-grm-alg 1
- Based on correlation between actual and expected genetic values of individuals

F_{YA2} (Yang *et al.*, 2010)

- **GCTA** → F_{HAT3}
- Based on correlation between uniting gametes

In males and
females separately

&

In females only

- a) Autosomal (F_{LH1} , F_{VR1} & F_{YA2})
- b) X chromosomal (F_{LH1_X} , F_{VR1_X} & F_{YA2_X})

Materials & Methods

New method for identification of positive selection signals
Haplotype Richness Drop (HRiD)

♂ XY

→ Hemizygous status on the X chromosome (without PAR)

Effective number of alleles (haplotypes) = $1 / \sum p_i^2$ (Kimura & Crow, 1964)

No selection, value ≈ 1

Positive selection, value > 1

$$\text{HRiD}_{wi+1} = \frac{n_{h_{wi}} + n_{h_{wi+2}}}{2 n_{h_{wi+1}}}$$

w_1

w_i

w_{i+1}

w_{i+2}

w_n

w_1

w_i

w_{i+1}

w_{i+2}

w_n

Materials & Methods

New method for identification of positive selection signals

Haplotype Richness Drop (HRiD)

- Sliding window approach (size of 70 SNPs \approx 500 kb; slide of 35 SNPs \approx 250 kb)

windows with $-\log(P) \geq 3.3$

One-sided test

(HRiD = 2.8)

- Median-Joining Network (MJN)** = to illustrate the phylogenetic relationship between ancestral and derived haplotypes in each identified selection signal

ancestral haplotypes = mouflons

Results – Inbreeding

Inbreeding coefficients
on **autosomes** and **X**
chromosome in
females
(Means \pm SE)

Autosomes

X chromosome

Croatian
cattle breeds

Nellore

Labrador
Retriever

Patagonian
Sheepdog

Croatian
sheep breeds

Soay

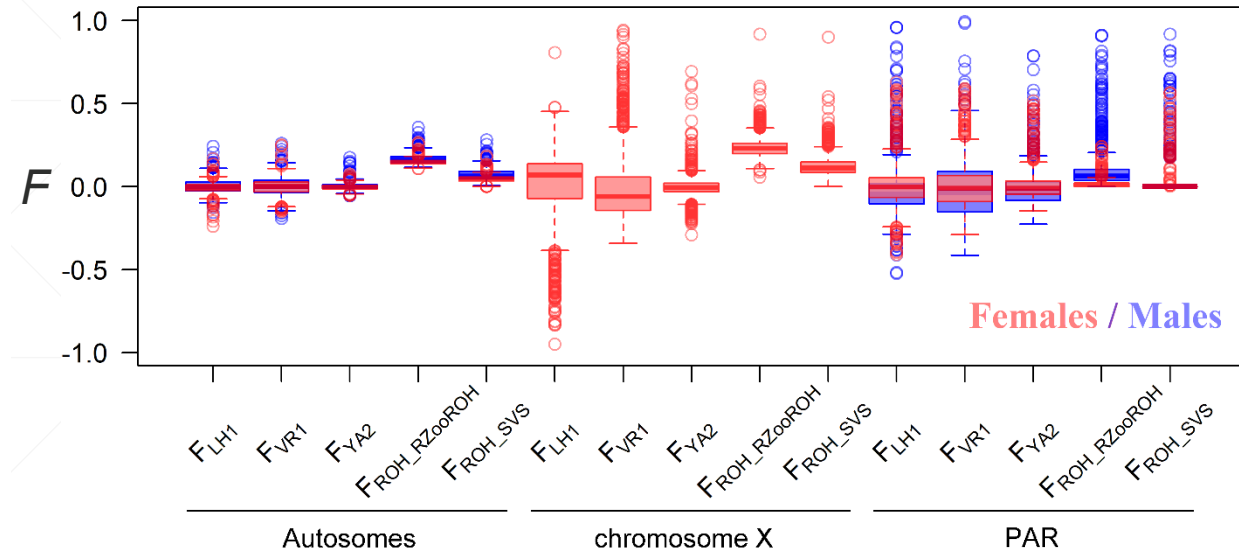
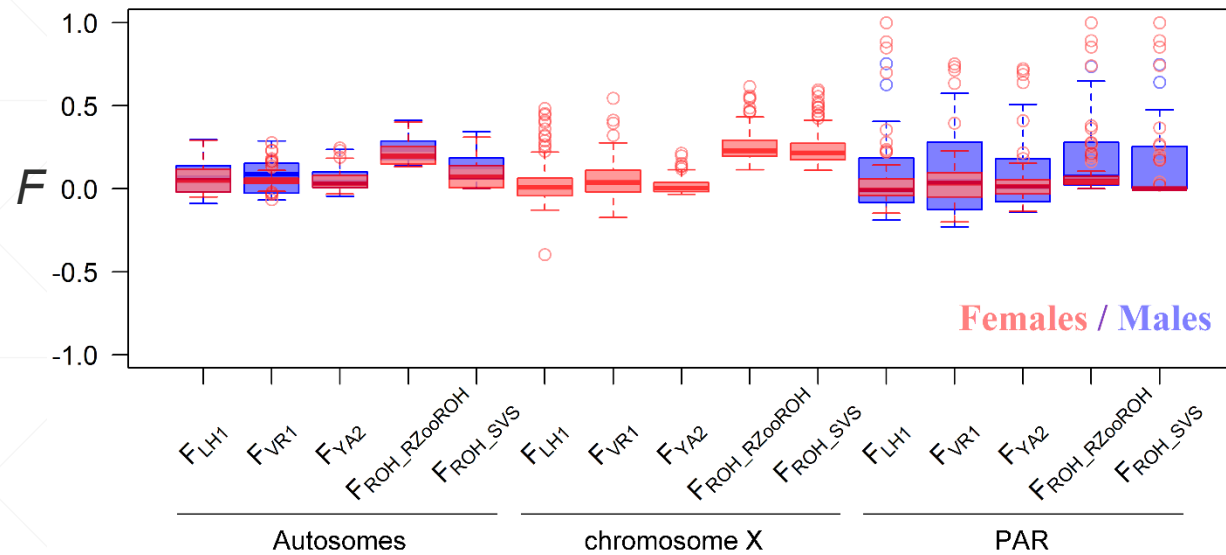
	F_{ROH_SVS}	$F_{ROH_RZooROH}$	F_{VR1}	F_{LH1}	F_{YA2}
	0.087 \pm 0.009	0.211 \pm 0.008	0.057 \pm 0.007	0.058 \pm 0.007	0.050 \pm 0.010
	0.255 \pm 0.013	0.264 \pm 0.012	0.053 \pm 0.015	0.053 \pm 0.006	0.020 \pm 0.018
	0.046 \pm 0.001	0.148 \pm 0.000	-0.007 \pm 0.001	-0.006 \pm 0.000	-0.003 \pm 0.001
	0.124 \pm 0.002	0.236 \pm 0.002	-0.009 \pm 0.003	0.003 \pm 0.002	-0.003 \pm 0.002
	0.213 \pm 0.006	0.220 \pm 0.005	0.048 \pm 0.006	0.048 \pm 0.006	0.044 \pm 0.007
	0.396 \pm 0.012	0.272 \pm 0.013	0.064 \pm 0.019	0.064 \pm 0.014	0.066 \pm 0.018
	0.109 \pm 0.009	0.101 \pm 0.009	0.026 \pm 0.009	0.026 \pm 0.008	0.024 \pm 0.010
	0.367 \pm 0.017	0.176 \pm 0.021	0.045 \pm 0.025	0.045 \pm 0.019	0.034 \pm 0.023
	0.062 \pm 0.007	0.137 \pm 0.006	0.060 \pm 0.007	0.060 \pm 0.006	0.058 \pm 0.007
	0.144 \pm 0.014	0.226 \pm 0.013	0.080 \pm 0.015	0.080 \pm 0.015	0.078 \pm 0.015
	0.255 \pm 0.002	0.321 \pm 0.002	-0.010 \pm 0.003	-0.010 \pm 0.002	-0.009 \pm 0.003
	0.422 \pm 0.009	0.453 \pm 0.009	-0.021 \pm 0.015	-0.021 \pm 0.009	-0.015 \pm 0.015
	X > Autosomes		X \approx Autosomes		

Results – Inbreeding

Distribution of inbreeding estimated on autosomes, X chromosome and PAR in **cattle**

A) *Native Croatian breeds*

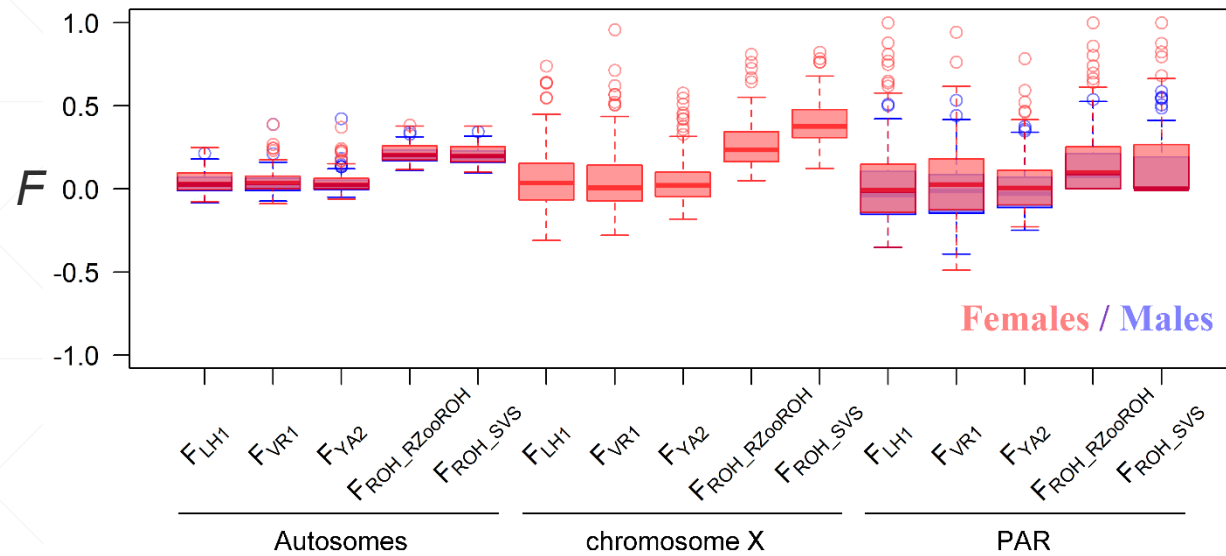
B) *Nellore*



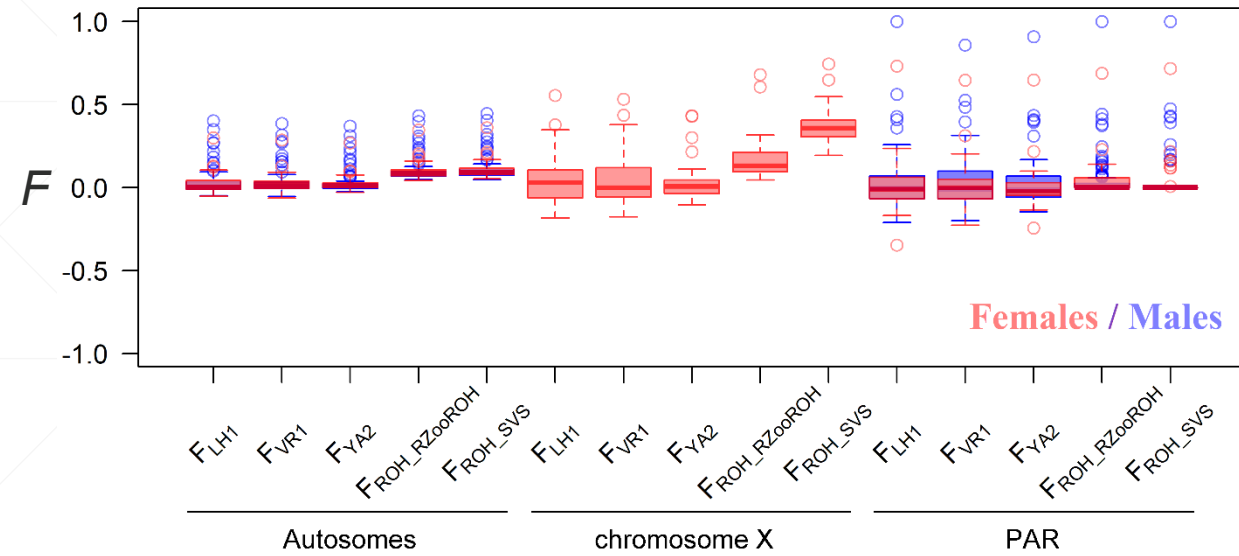
Results – Inbreeding

Distribution of inbreeding estimated on autosomes, X chromosome and PAR in **dog**

A) *Labrador Retriever*



B) *Patagonian Sheepdog*

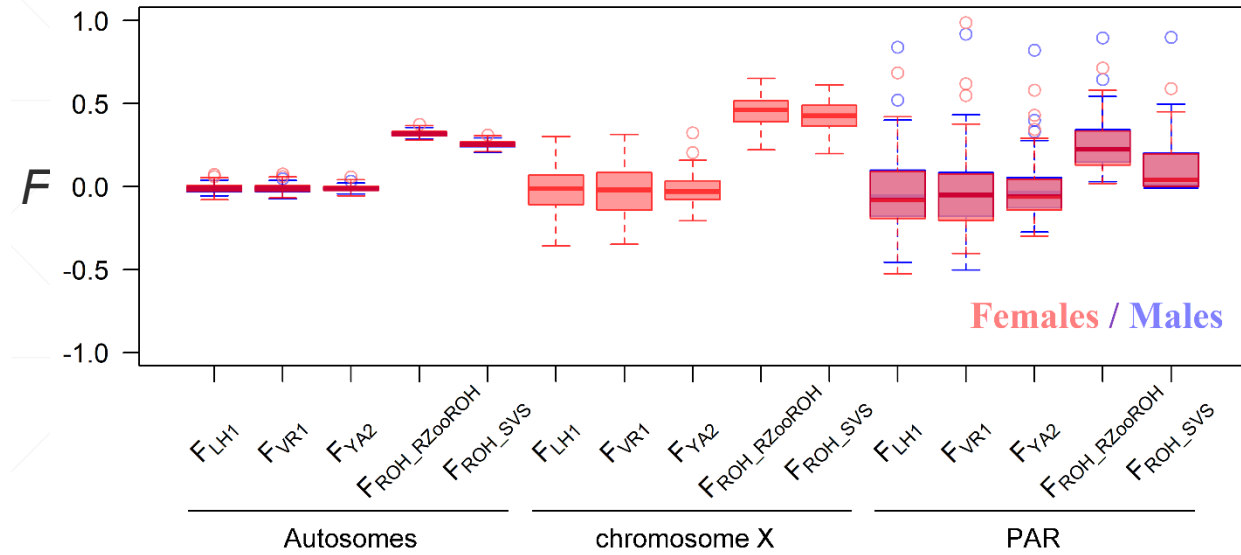
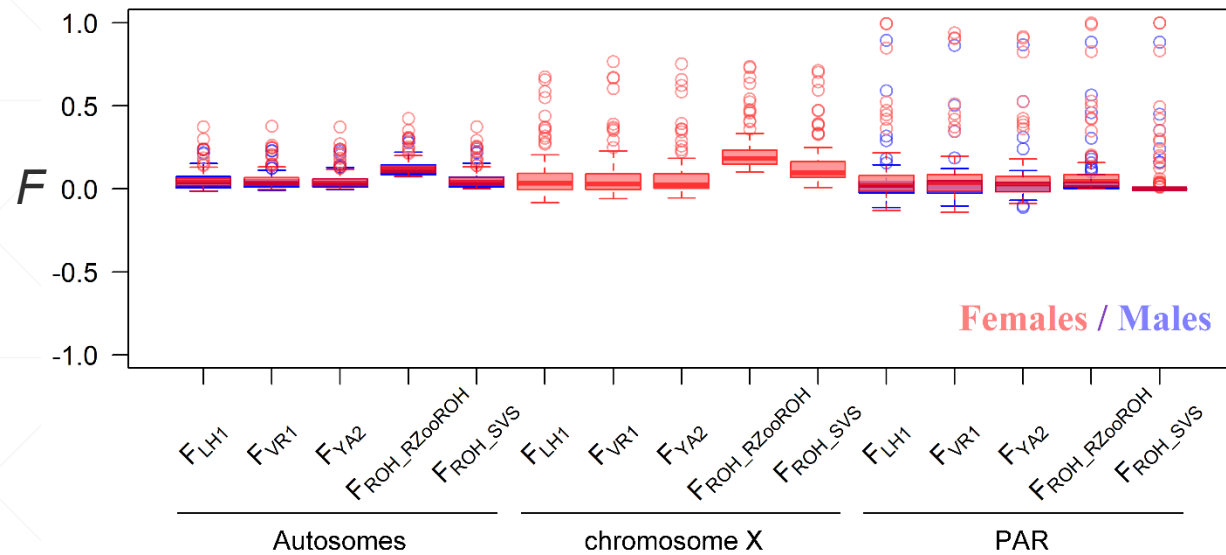


Results – Inbreeding

Distribution of inbreeding estimated on autosomes, X chromosome and PAR in **sheep**

A) *Native Croatian breeds*

B) *Soay*



Results – Selection

Description of mapping statistics and annotation of genes inside positive selection signals identified by new HRiD method

Signal name	Position (Mb)	n_a	n_h	HRiD	$-\log(P)$	Candidate genes under selection
HRiD_w1	13.04–13.62	42	5.4	9.6	56.5	TMEM27, CDC42, CA5B, ZRSR2, AP1S2, GRPR
HRiD_w2	115.30–115.73	36	13.3	4.2	8.7	AMOT, LHFPL1
HRiD_w3	73.90–74.54	13	4.3	3.2	4.5	DACH2
HRiD_w4	73.57–74.20	10	1.9	3.1	4.1	CHM, DACH2
HRiD_w5	56.64–58.09	33	6.9	3.1	4.0	AR, OPHN1, YIPF6

- n_a = total number of unique alleles (haplotypes)
- n_h = effective number of alleles (haplotypes)
- Genes additionally identified as positive selection candidates by other methods (eROHI, iHS or nSL) are **bolded**
- Genes identified by all four methods are highlighted in **orange**

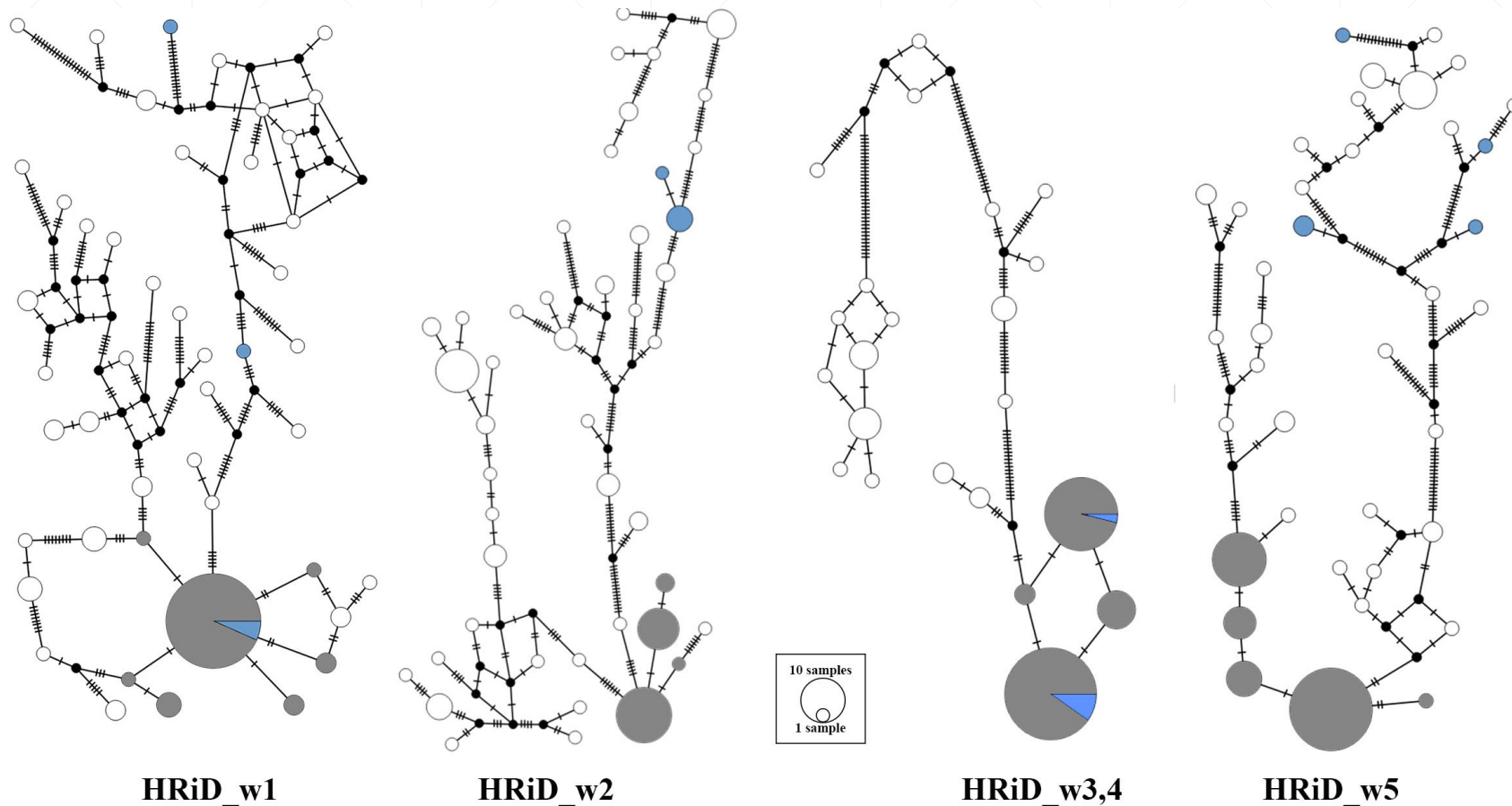
Results – Selection

Comparison of mapped HRiD signals with other sheep studies

Signal name	Position (Mb)	Overlapped signal (Mb)	Candidate genes	Reference
HRiD_w1	13.04–13.62	13.20–13.60	<i>CA5B, ZRSR2, AP1S2, GRPR</i>	Chen <i>et al.</i> (2018)
HRiD_w2	115.30–115.73	/		
HRiD_w3,4	73.57–74.54	73.69–74.09	<i>CHM, DACH2</i>	Zhu <i>et al.</i> (2015)
		73.44–73.94	<i>CHM, DACH2</i>	Zhu <i>et al.</i> (2020)
HRiD_w5	56.64–58.09	57.40–58.00	<i>AR, OPHN1, YIPF6</i>	Chen <i>et al.</i> (2018)
		56.50–58.00	<i>AR, OPHN1, YIPF6</i>	Liu <i>et al.</i> (2016)
		55.78–58.33	<i>AR, OPHN1, YIPF6</i>	Cesarani <i>et al.</i> (2022)
		55.79–58.35	<i>AR, OPHN1, YIPF6</i>	Manzari <i>et al.</i> (2019)

Results – Selection

MJNs between ancestral and derived haplotypes within selection signals identified by HRiD



Most common haplotypes
with adjacent haplotypes
no more than
3 mutations apart

Ancestral haplotypes (mouflons)

HRiD_w1
(70 SNPs from 13.4 to 13.62 Mb)
HRiD_w2
(70 SNPs from 115.30 to 115.73)
HRiD_w3 and HRiD_w4
(105 SNPs from 73.57 and 74.54)
HRiD_w5
(70 SNPs from 56.64 and 58.09)

Conclusions

- Higher inbreeding was found on X chromosome compared to autosomes in all populations using F_{ROH_SVS} and $F_{ROH_RZooROH}$, while no differences were found using F_{LH1} , F_{VR1} and F_{YA2}
- Greater variability in inbreeding was found on X chromosome compared to autosomes using all five coefficients in all populations, with specific patterns observed across X chromosome
- High accuracy and reliability of new HRiD method were demonstrated:
 - a) most significant signal, was also most significant one using all 3 classical methods
 - b) 3 of 4 identified signals were consistently validated

Conclusions

- **Results emphasize the importance of:**

including X chromosome in inbreeding estimation and selection identification in domestic animal populations,

while new HRiD method opens up new possibilities in identifying positive selection signals using heterogametic sex haplotypes

Thank you for your attention!

The dissertation was conducted within the framework of the project:
ANAGRAMs, Project IP-2018-01-8708
("Application of NGS in assessment of genomic variability in ruminants")

