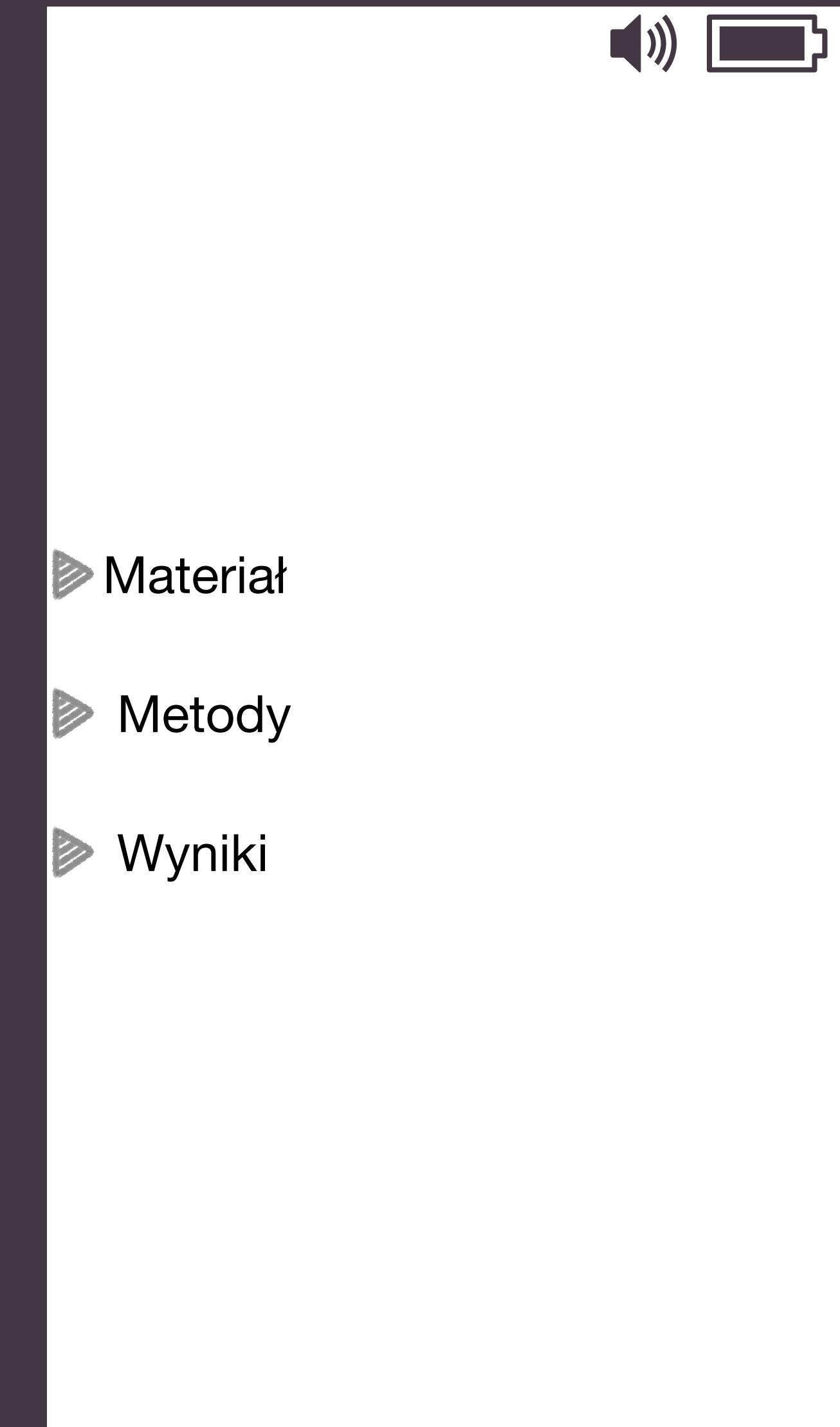


Analiza markerów genetycznych nowotworu listwy mlecznej u psa domowego *(Canis lupus familiaris)* w obrębie chromosomu 15

Daria Plewa

**Praca wykonana pod kierunkiem
dr inż. Anny Muchy**

Plan prezentacji





Materiał

- ▶ Dane dotyczą 118 suk (*Canis lupus familiaris*) w wieku 5 miesięcy – 16,5 roku, różne rasy i mieszańce.
- ▶ 36 samic z guzami listwy mlecznej (w tym 19 przypadków złośliwych).

- ▶ Próbki: krew żylna i tkanki zdrowe oraz nowotworowe.
- ▶ Izolacja DNA: kit Sherlock AX (A&A Biotechnology).
- ▶ Analiza ilościowa: spektrometr NanoDrop2000P (Thermo Scientific).



**POLSKA FEDERACJA
HODOWCÓW BYDŁA
I PRODUCENTÓW MLEKA**

- ▶ 4791 SNP na chromosomie 15 (Illumina CanineHD BeadChip).
- ▶ Numer próbki i informacja o obecności nowotworu (złośliwego lub niezłośliwego).



Metody

Plink

- ▶ Selekcja SNP
- ▶ Analiza asocjacyjna SNP
- ▶ Analiza nierównowagi sprzężeń
- ▶ Konstrukcja bloków



Selekcja SNP

Etapy selekcji wraz z wykorzystanymi parametrami

etap selekcji	wykorzystane parametry
brakujące genotypy (GENO)	95%
usunięcie osobników (MIND)	90%
częstość występowania rzadzkiego allele (MAF)	5%
równowaga Hardy'ego-Weinberga (HWE)	1×10^{-6} oraz 1×10^{-10}



Analiza asocjacyjna

Modele dla krytych zostało
przeprowadzona analiza

modele
dominacyjny
kodominacyjny
recesywny
naddominacyjny
log-addytywny

Ogólna forma modelu analitycznego

$$\log\left(\frac{\pi}{1 - \pi}\right) = \alpha + \beta G + \gamma Z$$

Objaśnienie modelu genetycznego

π – prawdopodobieństwo przypadku (wystąpienia nowotworu listwy mlecznej)

α, β, γ – estymowane wartości parametrów modelu

G – polimorfizm kodujący 3 genotypy (homozygotę dominującą, heterozygotę lub

homozygotę reesywną dla konkretnego allelu)

Z – macierz czynników niegenetycznych wpływających na prawdopodobieństwo wystąpienia choroby.



Analiza nierównowagi sprzężeń

Wzór na współczynnik D'

$$D' = P_{AB} - P_A * P_B$$

Wzór na współczynnik korelacji r^2

$$r^2 = \frac{(P_{AB} - P_A * P_B)^2}{P_A * (1 - P_A) * P_B * (1 - P_B)}$$

Objaśnienia wykorzystanych wzorów

P_{AB} – to rzeczywista częstość występowania haplotypu składającego się z alleli A i B.

P_A – to częstość allelu A w populacji.

P_B – to częstość allelu B w populacji.



Konstrukcja bloków sprzężonych

- ▶ Wyznaczenie współczynników r^2
- ▶ Wybranie SNP dla których $r^2 > 0.9$
- ▶ Tworzenie bloków sprzężonych SNP



R

- ▶ Wizualizacja wyników analizy asocjacyjnej
- ▶ Konstrukcja haplotypów
- ▶ Estymacja frekwencji haplotypów
- ▶ Analiza asocjacyjna haplotypów



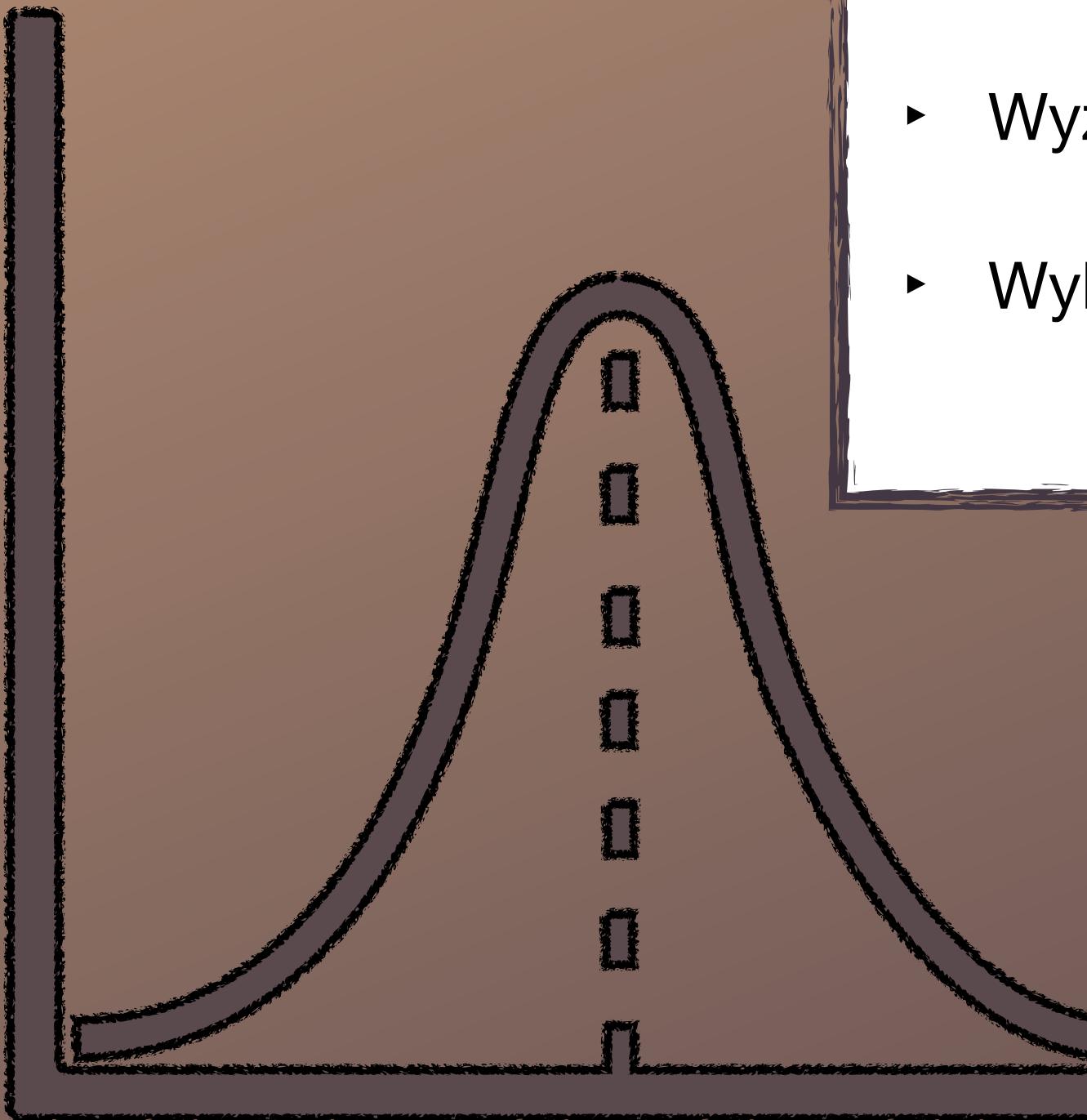
Konstrukcja haplotypów

- ▶ Inicjacja
- ▶ Oczekiwanie
- ▶ Maksymalizacja



Analiza asocjacyjna haplotypów

- Obliczenie ilorazu szans (OR, ang. odds ratio)
- Wyznaczenie haplotypu referencyjnego
- Wykorzystanie testu Walda do obliczenia istotności statystycznej



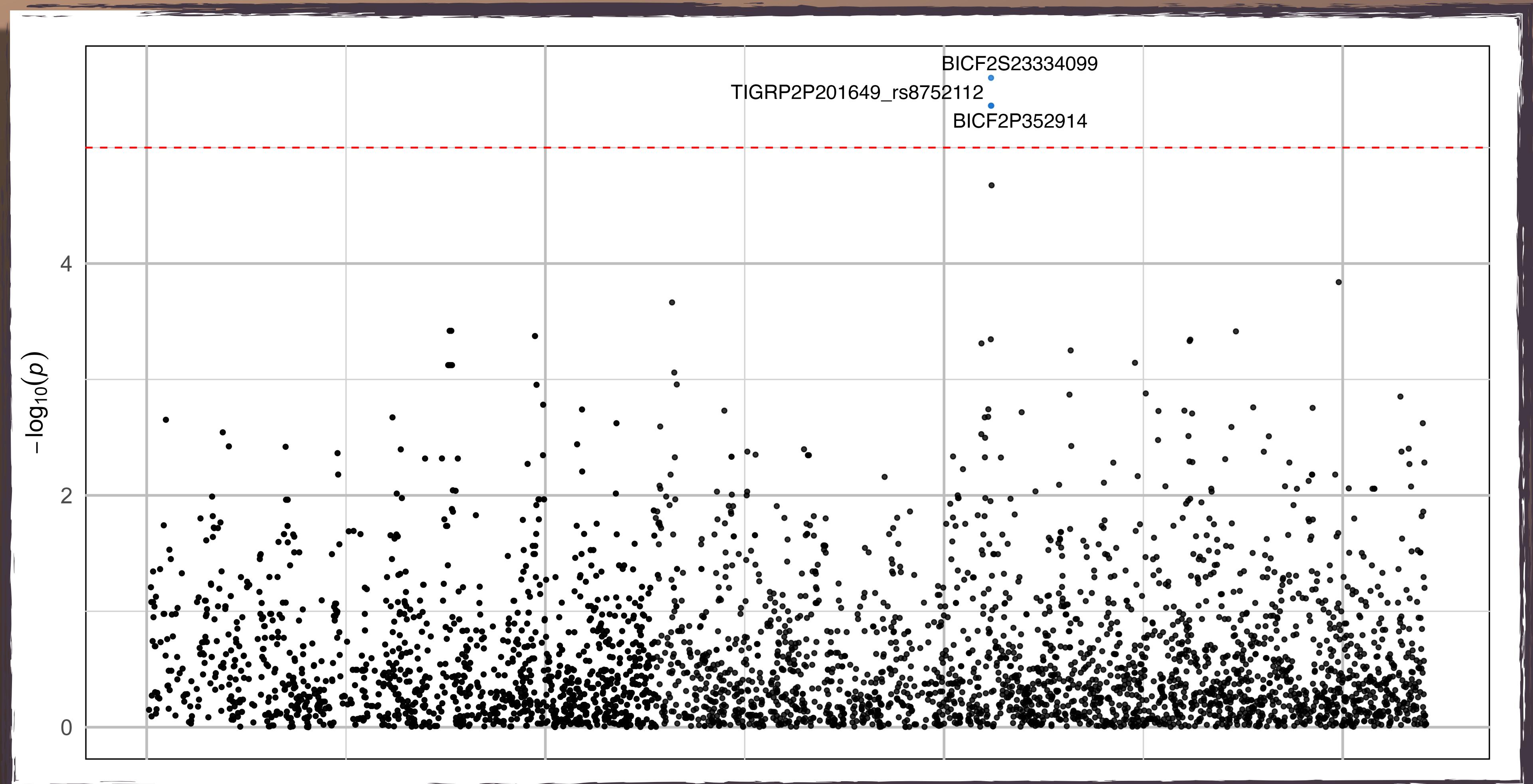
Wyniki

Selekcja SNP

etap selekcji	liczba usuniętych SNP	liczba pozostałych SNP
brakujące genotypy (GENO)	218	4573
częstość występowania rzadszego allelu (MAF)	828	3745
równowaga Hardy'ego-Weinberga (HWE)	86	3659



Analiza asocjacyjna SNP



Analiza asocjacyjna SNP

Markery uzyskane przez analizę asocjacyjną, wraz z przypisanym modelem dziedziczenia, lokalizacją i sąsiadującymi genami

marker	model	pozycja [pary zasad]	geny sąsiadujące
TIGRP2P201649_rs87521 12	dominacyjny	42346461	ASCL1 (Achaete-Scute Family BHLH Transcription Factor 1)
BICF2S23334099	dominacyjny	42352911	ASCL1 (Achaete-Scute Family BHLH Transcription Factor 1)
BICF2P352914	dominacyjny	42358113	ASCL1 (Achaete-Scute Family BHLH Transcription Factor 1)

Analiza asocjacyjna SNP

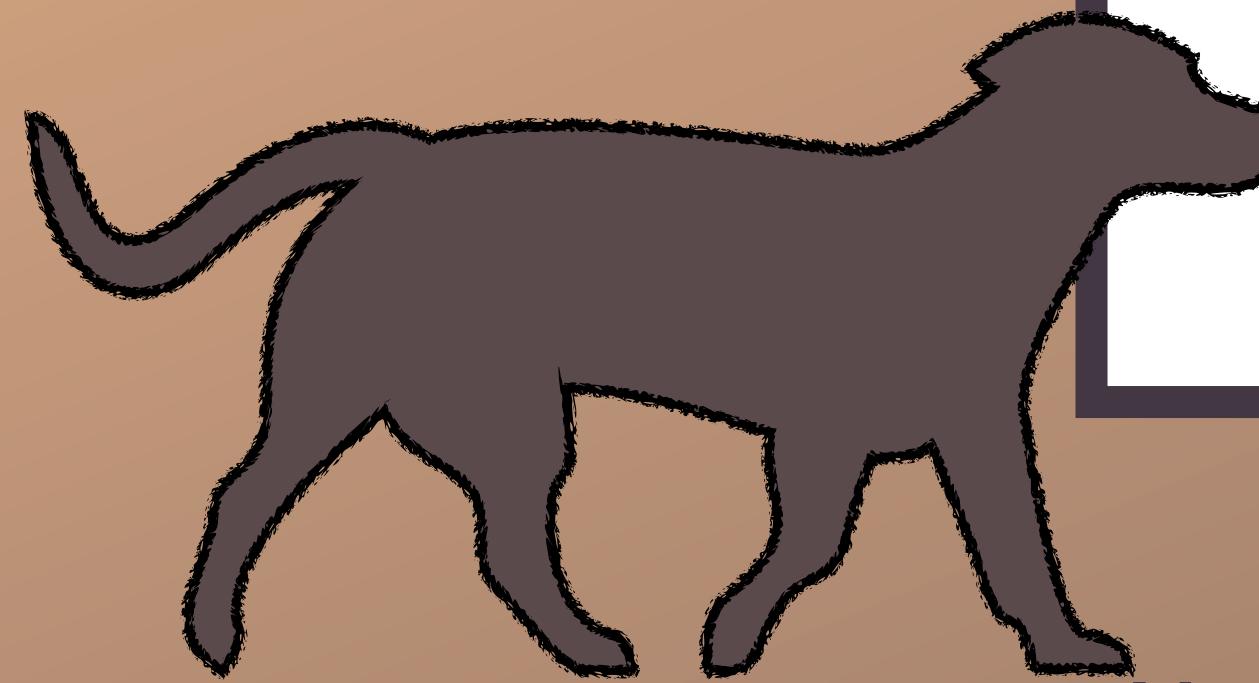
Liczba osobników chorych i zdrowych w zależności od genotypu w obrębie
TIGRP2P201649_rs8752112

genotyp	próba badawcza	próba kontrolna	OR	95% przedział ufności dla OR	p-wartość
A/A	32	38	1.00	-	4.35×10^{-6}
G/A-G/G	4	44	0.11	0.03; 0.33	

Analiza asocjacyjna SNP

Liczba osobników chorych i zdrowych w zależności od genotypu w obrębie
BICF2S23334099

genotyp	próba badawcza	próba kontrolna	OR	95% przedział ufności dla OR	p-wartość
G/G	32	37	1.00	-	2.50×10^{-6}
A/G-A/A	4	45	0.10	0.03; 0.32	



Analiza asocjacyjna SNP

Liczba osobników chorych i zdrowych w zależności od genotypu w obrębie
BICF2P352914

genotyp	próba badawcza	próba kontrolna	OR	95% przedział ufności dla OR	p-wartość
G/G	32	38	1.00	-	4.35×10^{-6}
A/G-A/A	4	44	0.11	0.03; 0.33	

Konstrukcja i analiza asocjacyjna haplotypów

Otrzymane haplotypy z wartościami OR, p-wartością oraz najbliższymi genami lub *loci*

nr bloku	SNP	haplotypy	frekwencja	OR	p-wartość	najbliższe geny/ <i>loci</i>
94	BICF2G630433442 BICF2G630433437 BICF2G630433436 BICF2G630433434	TAAG AGGA	0.47 0.43	1.00 2.12	--- 0.01	LOC111098932, LOC111098894
143	BICF2S23539059 TIGRP2P201469_rs8530904 BICF2S23719375	AGG GAA	0.62 0.35	1.00 1.99	--- 0.01	PAH (Phenylalanine Hydroxylase)
144	TIGRP2P201494_rs8845947 BICF2P856505 BICF2P1114641	GAG AGA	0.65 0.33	1.00 2.44	--- 1×10^{-3}	ASCL1 (Achaete-Scute Family BHLH Transcription Factor 1), LOC102152557

Konstrukcja i analiza asocjacyjna haplotypów

Otrzymane haplotypy z wartościami OR, p-wartością oraz najbliższymi genami lub *loci*

nr bloku	SNP	haplotypy	frekwencja	OR	p-wartość	najbliższe geny/ <i>loci</i>
25	BICF2G630443763 TIGRP2P195538_rs8668139 BICF2G630443747 BICF2G630443732 BICF2G630443683 BICF2G630443676 BICF2G630443671 BICF2G630443668 BICF2G630443653 BICF2G630443628	CGAGAGGTAA GAGAGAAATC	0.55 0.38	1.00 0.49	--- 0.02	NRDC (Nardilysin Convertase)
111	TIGRP2P199588_rs9079514 BICF2S2395276	GG AA	0.56 0.43	1.00 0.59	--- 0.05	EEA1 (Early Endosome Antigen 1)
128	BICF2P692606 BICF2P830923	GA AG	0.58 0.41	1.00 0.56	--- 0.05	LOC111090052

Konstrukcja i analiza asocjacyjna haplotypów

Otrzymane haplotypy z wartościami OR, p-wartością oraz najbliższymi genami lub *loci*

nr bloku	SNP	haplotypy	frekwencja	OR	p-wartość	najbliższe geny/ <i>loci</i>
138	BICF2P686421 BICF2P204366	GA AG	0.64 0.35	1.00 0.55	--- 0.05	NR1H4 (Nuclear Receptor Subfamily 1 Group H Member 4)
142	BICF2P732671 BICF2P732670	GA AG	0.53 0.45	1.00 0.54	--- 0.02	PAH (Phenylalanine Hydroxylase)
146	IGRP2P201649_r s8752112 BICF2S23334099 BICF2P352914	GGA AAG	0.73 0.27	1.00 0.23	--- 1×10^{-3}	ENSCA FG00805 019903
172	BICF2G630426043 BICF2G630426028	AA GG	0.54 0.45	1.00 0.50	--- 0.02	LRAT (Lecithin Retinol Acyltransferase)



Podsumowanie

- Wszystkie wykryte SNP mają dominujący model dziedziczenia
- Obecność SNP oraz haplotypu zlokalizowanych w pobliżu ASCL1 sugeruje, że gen ten może być interesującym czynnikiem w kontekście nowotworu listwy mlecznej u psów.
- Geny związane z metabolizmem (PAH, NR1H4, LRAT) mogą wpływać na podatność na nowotwór poprzez zaburzenia metaboliczne.
- Geny związane z transportem wewnętrzkomórkowym i inwazją komórek (NRDC, EEA1) mogą uczestniczyć w mechanizmach progresji nowotworowej, co sugeruje ich potencjalną rolę w przerzutowaniu nowotworów listwy mlecznej.
- Warto przeprowadzić dalsze badania funkcjonalne, aby potwierdzić wpływ zidentyfikowanych genów na rozwój nowotworu.
- Możliwe jest także wykorzystanie tych markerów w selekcji hodowlanej, aby zmniejszyć ryzyko występowania nowotworów listwy mlecznej u określonych ras psów.

Dziękuje za uwagę!

Daria Plewa