The genomic signature of dog domestication reveals adaptation to a starch-rich diet

Mikk Eelmets
Journal Club
25.03.2013

Canis lupus

Canis lupus familiaris

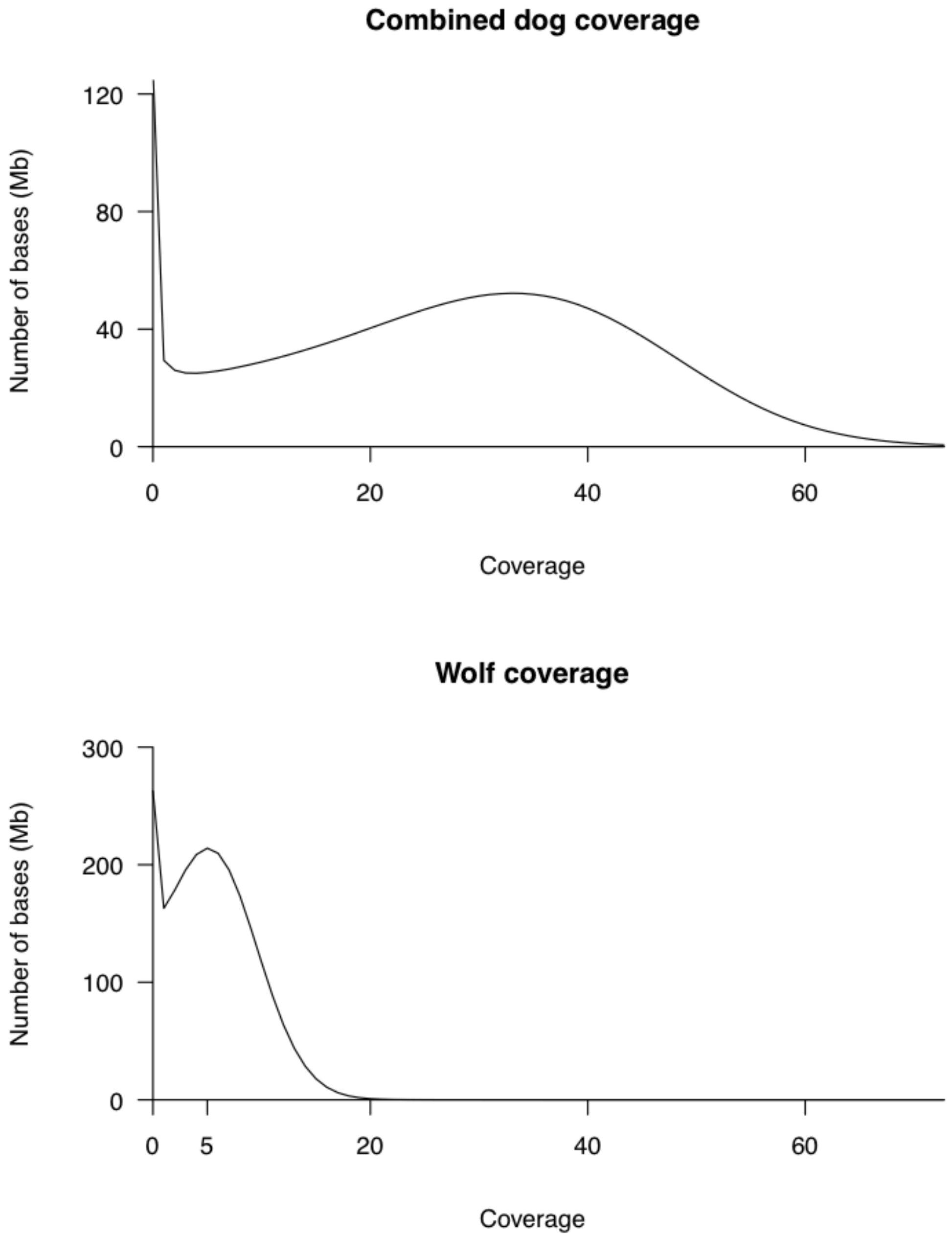
Fossilised dog remains
33,000-year-old doglike canid from the Altai Mountains in Siberia
12,000– 11,000 years BP found buried together with humans in Israel could represent the earliest verified dog remains

Patterns of genomic variation indicate that dog domestication started at least 10,000 years BP in southern East Asia or the Middle East

Sequencing

- 12 wolves, 60 dogs (14 diverse breeds)
- 12 individuals per pool
- AB SOLiD ver3 (mate-pair)
- Aligned to CanFam 2.0 reference
- Bioscope 1.1
- 3.79M SNP, 0.51M Indels, 26619 CNVs
- Unique for dogs 1.77M SNPs, for wolf 0.14M SNPs

	WOLVES	**n**	**sequence cov.**	**assembly cov. (%)**
Pool 1	Sweden	1		
	Spain	3		
	Russia	3		
	Belarus	2		
	Bulgaria	1		
	USA	1		
	Canada	1		
	Total	**12**	**6.2x**	**89**
	DOGS			
Pool 2	Smålandsstövare	3		
	Norwegian Elkhound	3		
	Swedish Elkhound	3		
	Finnish Lapphund	3	5.9x	89
Pool 3	Cocker Spaniel	3		
	Springer Spaniel	3		
	Golden Retriever	3		
	Labrador Retriever	3	5.3x	87
Pool 4	Drever	12	6.8x	90
Pool 5	Belgian Tervueren	12	6.8x	90
Pool 6	Bearded Collie	3		
	Hovawart	3		
	Giant Schnauzer	3		
	German Shephard	3	6.7x	90
	Total	**60**	**29.8x**	**94**



Distribution of sequence coverage for the five dog pools combined and a single wolf pool

Selection signals

- pooled heterozygosity (H_p)
- fixation index (F_{ST})

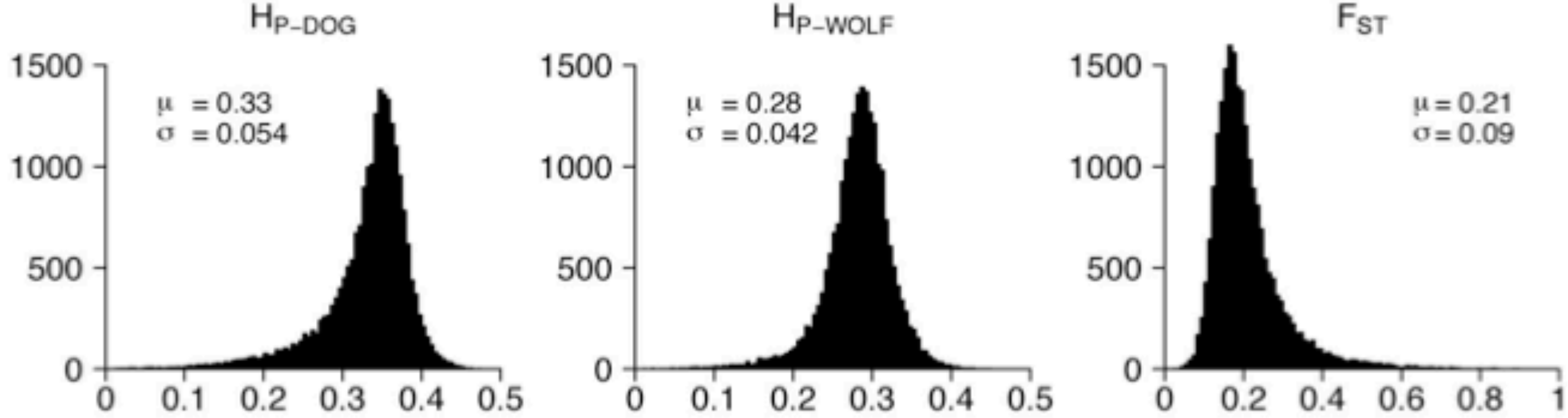
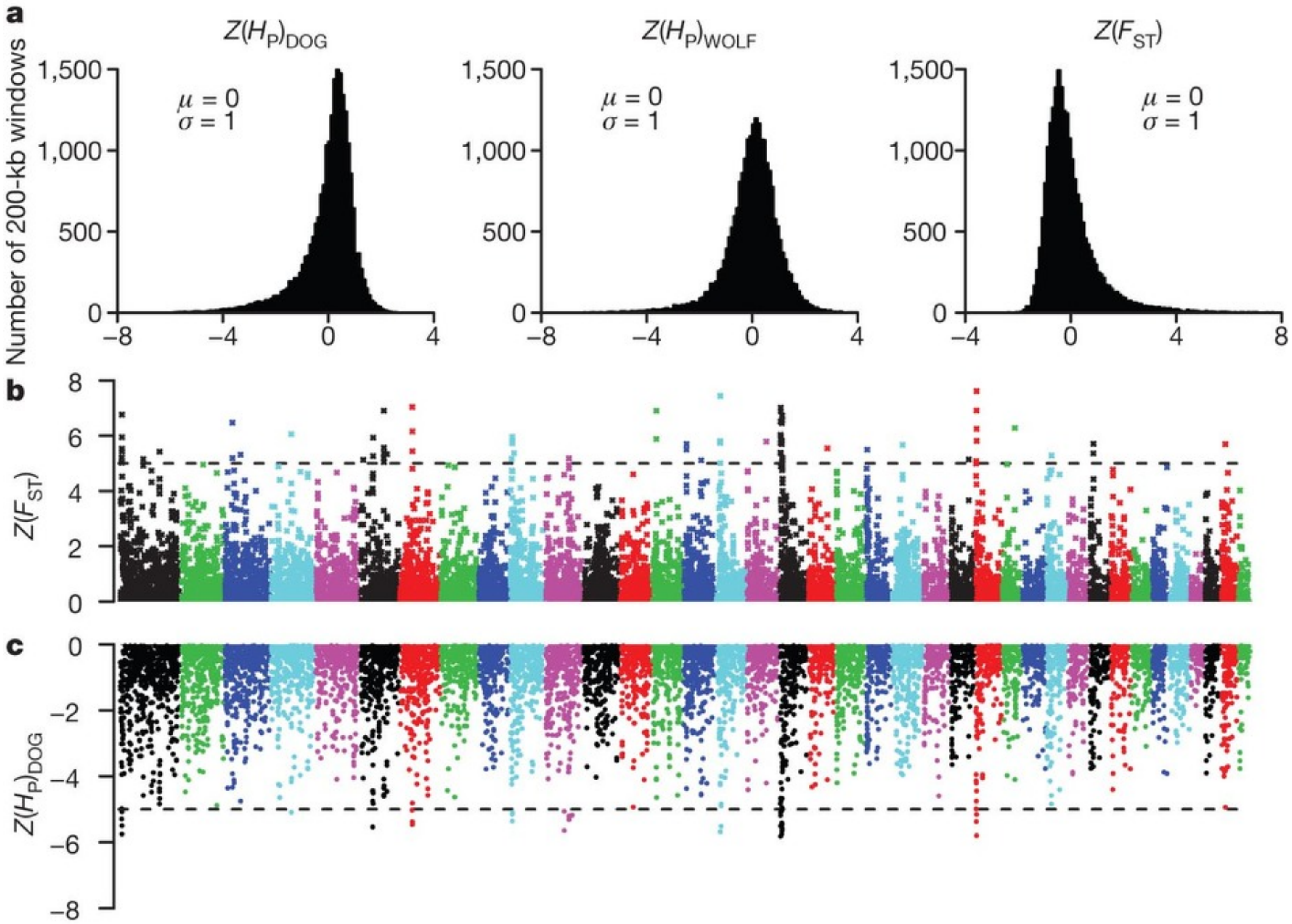


Figure S2. Distribution of heterozygosity and fixation index. Distribution of average pooled heterozygosity in dog ($H_{\text{P-DOG}}$) and wolf ($H_{\text{P-WOLF}}$) respectively, as well as average fixation index (F_{ST}), for autosomal 200 Kb windows (σ, standard deviation; μ, average).

Selection signals



- a. Distribution of *Z-transformed average pooled heterozygosity in dog (Z(Hp)DOG) and wolf (Z(Hp)WOLF) respectively, as well as average fixation index (Z(FST)), for autosomal 200 kb windows (σ, standard deviation; μ, average)*
- b. The positive end of the *Z(Fst) distribution plotted along dog autosomes 1–38 (chromosomes are separated by colour). A dashed horizontal line indicates the cut-off (Z > 5) used for extracting outliers.*
- c. The negative end of the *Z(Hp) distribution plotted along dog autosomes 1–38. A dashed horizontal line indicates the cut-off (Z < −5) used for extracting outliers.*

Selection signals

- 36 unique autosomal candidate domestication region (CDR)
- 122 genes
- Overrepresented gene ontology terms in 25 categories, representing several groups of interrelated terms (nervous system development, sperm-egg recognition, digestion, regulation of molecular function)

Gene ontology term	P_{FDR} value	Gene count
Regulation of neuron differentiation	0.005	3 (26)
Multicellular organismal process	0.005	21 (3,822)
Digestion	0.008	4 (95)
Neuron differentiation	0.010	5 (210)
Regulation of molecular function	0.011	8 (671)
Central nervous system development	0.013	5 (235)
Regulation of developmental process	0.013	5 (236)
Generation of neurons	0.013	5 (242)
Nervous system development	0.013	8 (716)
Binding of sperm to zona pellucida	0.015	2 (12)
Sperm–egg recognition	0.015	2 (12)
Neurogenesis	0.015	5 (262)
Cell–cell recognition	0.019	2 (14)
Regulation of catalytic activity	0.020	7 (605)
Regulation of hydrolase activity	0.026	5 (307)
Fatty acid metabolic process	0.031	4 (191)
System development	0.034	11 (1,605)
Regulation of GTPase activity	0.039	4 (211)
Anatomical structure development	0.039	12 (2,005)
Intramembranous ossification	0.039	1 (1)
Quinolinate metabolic process	0.039	1 (1)
Starch metabolic process	0.039	1 (1)
Starch catabolic process	0.039	1 (1)
Glucocorticoid catabolic process	0.039	1 (1)
Cell development	0.039	9 (1,242)

Enriched terms are colour-coded to reflect relatedness in the ontology or functional proximity. Blue, nervous system development; green, sperm–egg recognition; grey, regulation of molecular function; orange, digestion. For each term, gene count shows number of genes in CDRs relative to total number of annotated genes (in parentheses).

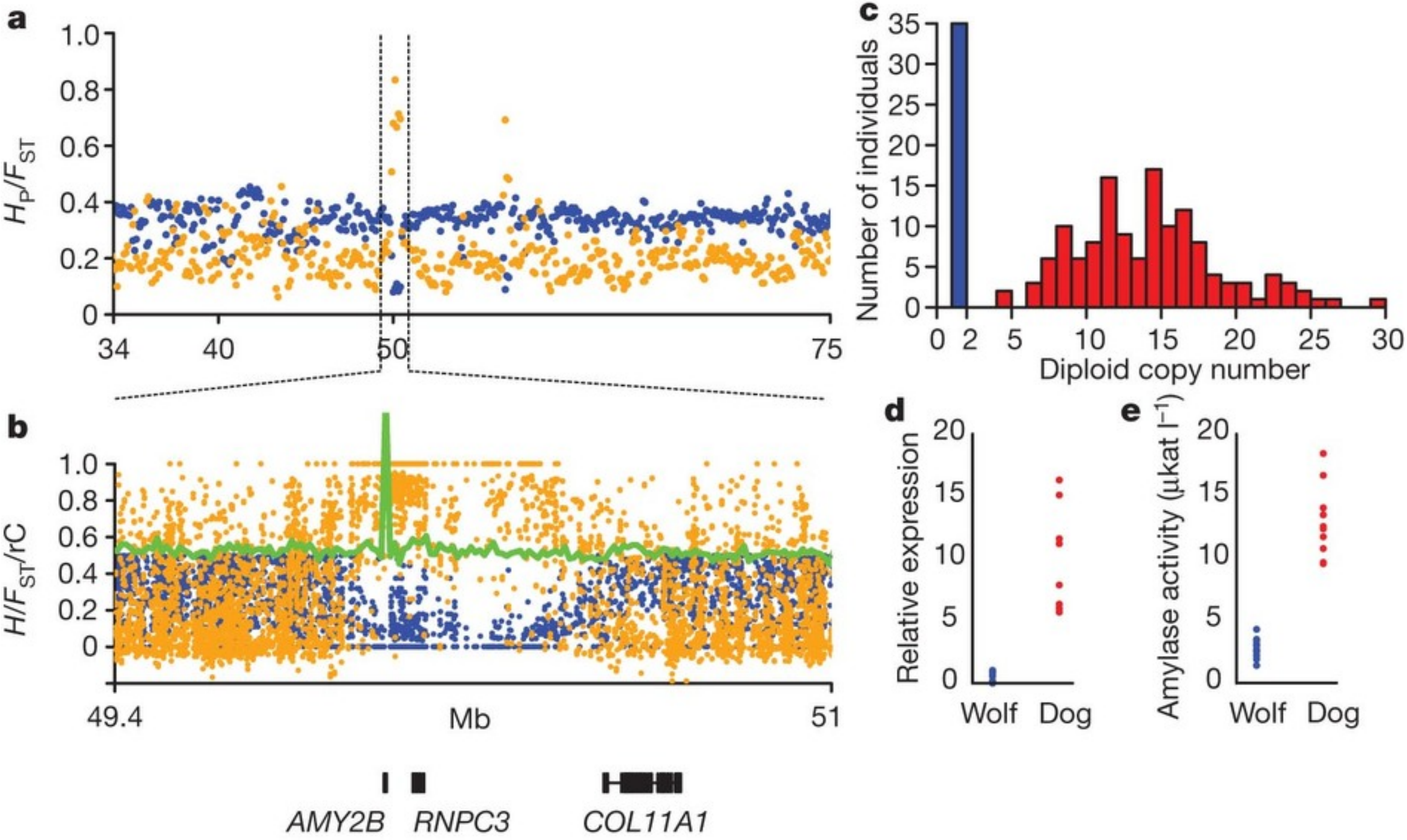
Starch metabolism in dogs

- **Three stages:**

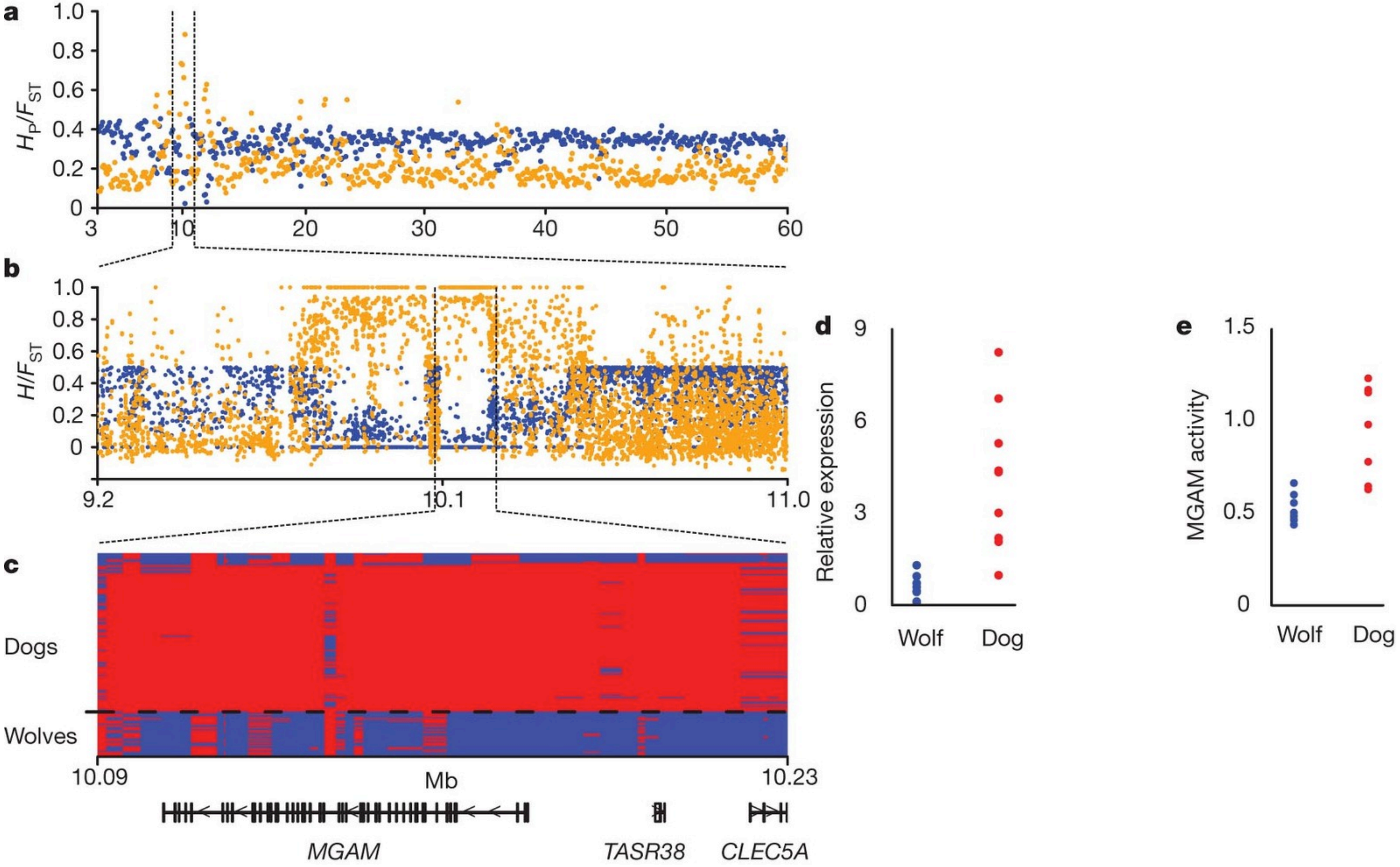
  1. starch is first cleaved to maltose and other oligosaccharides by **alpha-amylase** in the intestine;

  2. the oligosaccharides are subsequently hydrolysed by **maltase-glucoamylase**, sucrase and isomaltase to form glucose;

  3. glucose is transported across the plasma membrane by brush border protein **SGLT1**



a. Pooled heterozygosity, *Hp (blue), and average fixation index, Fst (orange), plotted for 200-kb windows across a chromosome 6 region harbouring AMY2B*
b. Heterozygosity, *Hp (blue), and fixation index, Fst (orange), for single SNPs in the selected region. Dog relative to wolf coverage, rC (green line), indicates increase in AMY2B copy number in dog. Genes in the region are shown below panel*
c. Histogram showing the distribution of diploid amylase copy number in wolf (*n = 35) (blue) and dog (n = 136) (red)*
d. Amylase messenger RNA expression levels in pancreas of wolf (*n = 12) and dog (n = 9).*
e. Amylase activity in serum from wolf (*n = 13) and dog (n = 12).*



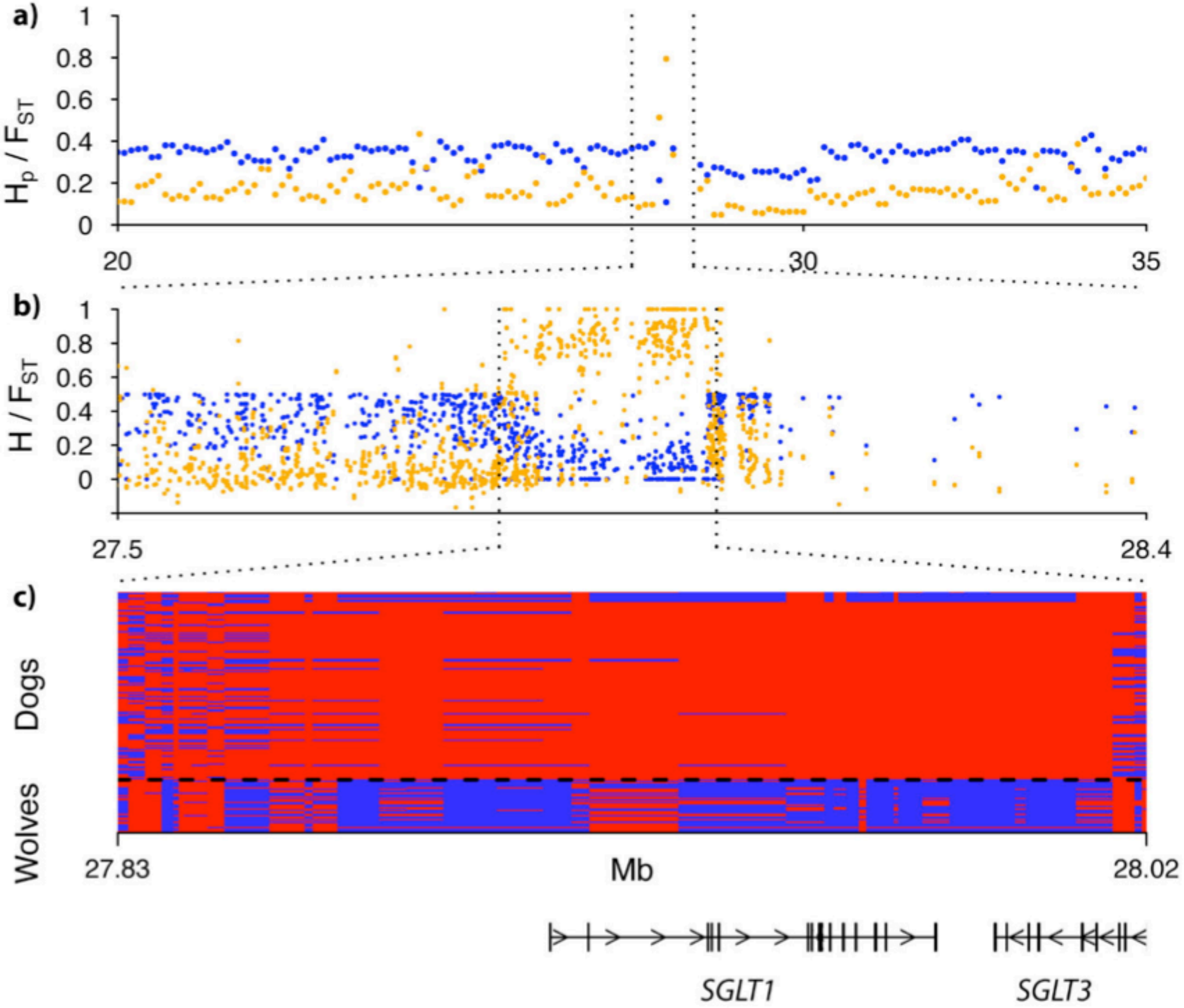
a. Pooled heterozygosity, Hp (blue), and average fixation index, Fst (orange), plotted for 200-kb windows across a chromosome 16 region harbouring MGAM

b. Heterozygosity, Hp (blue), and fixation index, Fst (orange), for single SNPs in the selected region

c. Haplotypes inferred from genotyping of 47 SNPs across the MGAM locus in 71 dogs and 19 wolves (red and blue colour are major and minor dog allele, respectively). Genes in the genotyped region are shown below panel.

d. MGAM mRNA expression levels in pancreas of wolf (n = 8) and dog (n = 9).

e. MGAM activity in serum from wolf (n = 8) and dog (n = 7).



a. Pooled heterozygosity, Hp (blue dots), and average fixation index, Fst (orange dots) plotted for 200 Kb windows across a region spanning 20-35 Mb on chromosome 26. Dashed vertical lines indicate the location of the selected region harbouring SGLT1.
b. Magnification of the region affected by selection showing heterozygosity, H (blue dots) and fixation index, FST (orange dots) estimated for single SNPs. Dashed horizontal lines delineate genotyped region shown in panel c
c. Haplotypes inferred from genotyping of 48 SNPs in 71 dogs and 19 wolves, shows the location of a 50.5 Kb region, spanning approximately 27.96-28.01 Mb, that is nearly fixed in all dogs. Red colour represents the major dog allele, while blue is minor dog allele. Genes residing in the genotyped region are shown below panel c.

Conclusion

- Dog domestication was accompanied by selection at three genes with key roles in starch digestion: *AMY2B, MGAM and SGLT1*
- This may suggest that a change of ecological niche (adaptations to thrive on a diet rich in starch) could have been the driving force behind the domestication process

Reference

E. Axelsson et al.

" The genomic signature of dog domestication reveals adaptation to a starch-rich diet. “

Nature 495, 360–364 (2013)

THANK YOU