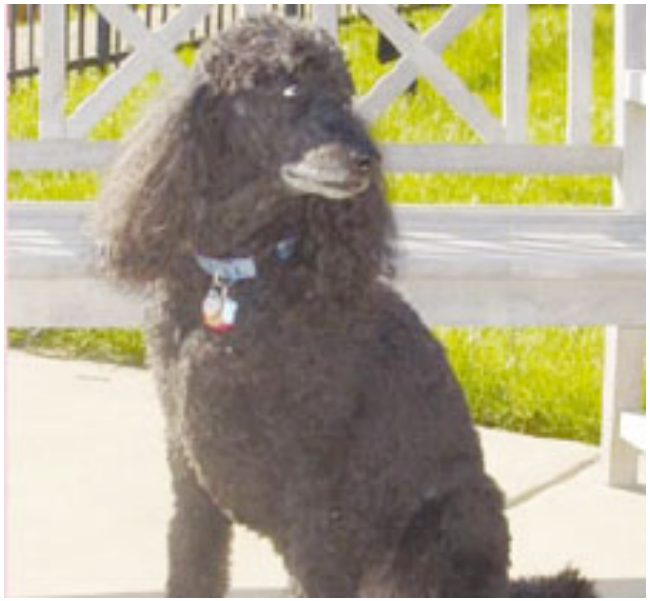
Engineering of the dog and its wild relatives

Tasha



Shadow

TALK OUTLINE

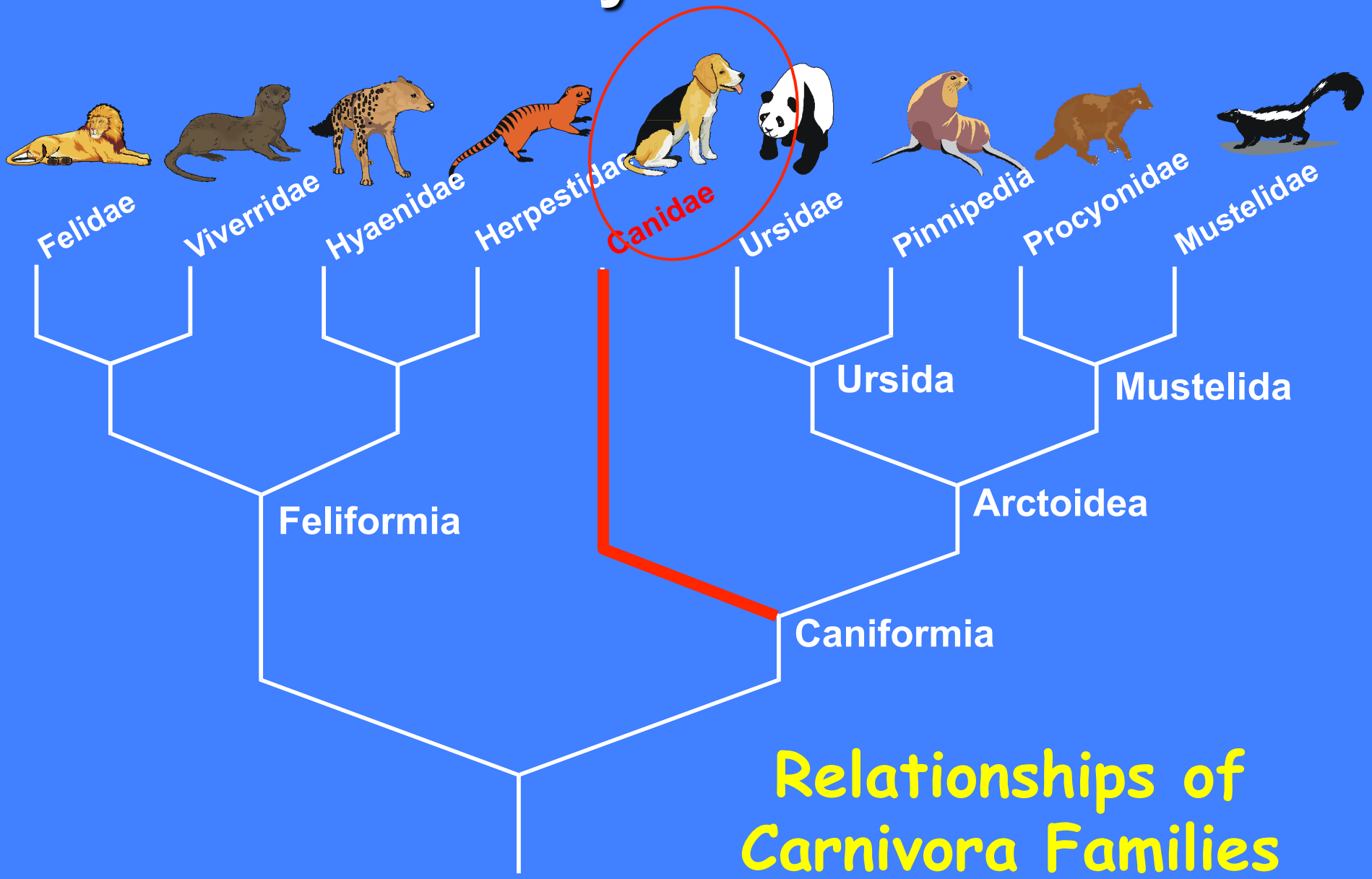
1) Introduction
   a) Evolutionary framework
   b) Origin and diversification
   c) The dog as a model species
   d) The canine genome and canmap projects

2) How to engineer and genotype-phenotype map a dog
   a) Three ways to build a dog
   b) Association and selective sweep mapping
   c) IGF1, the body size gene
   d) Legs and fur
   e) The number of genes
   f) Evolving the dog

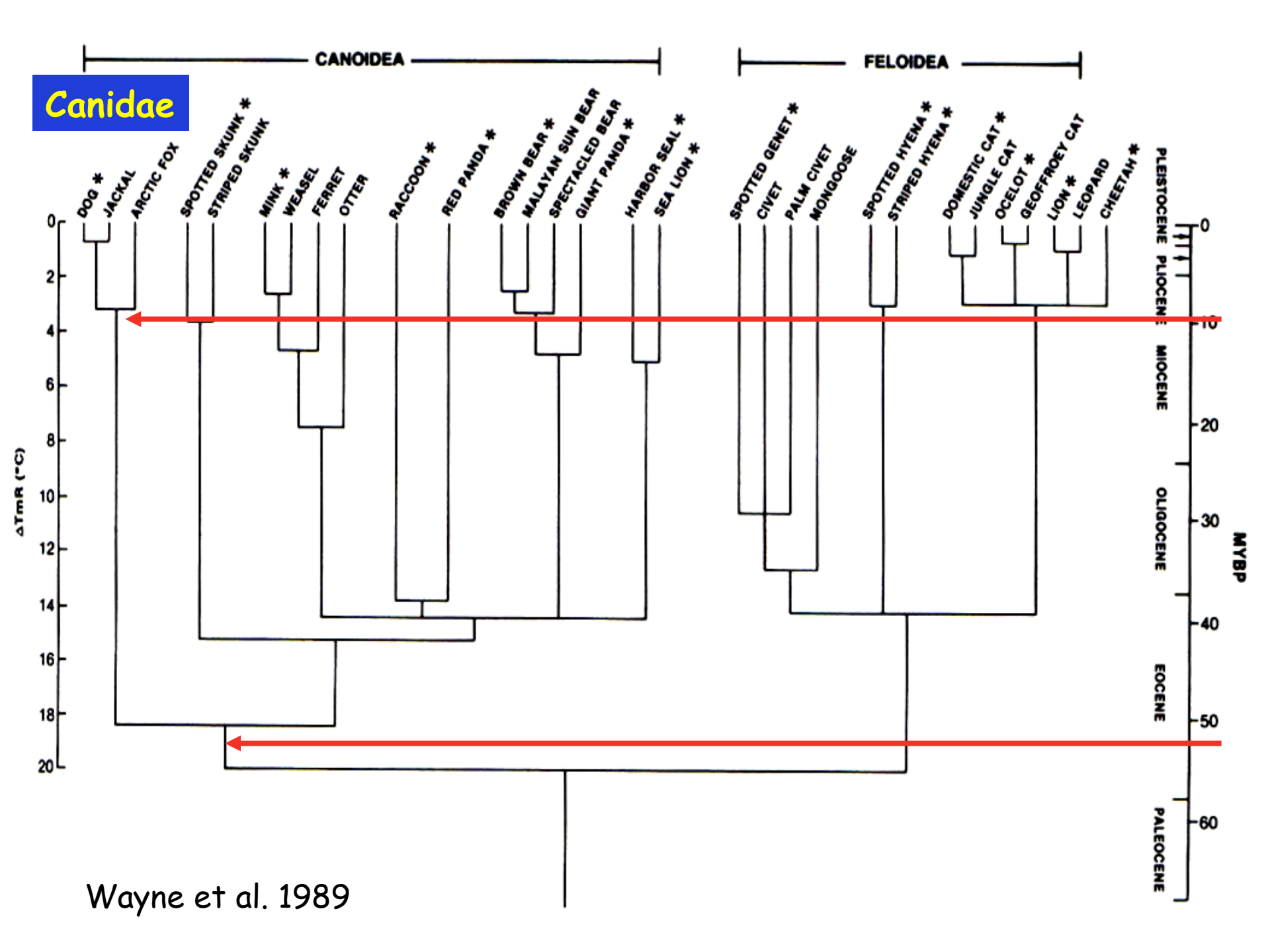
3) Engineering in the Wild
   a) Black and gray coat colors of wolves

4) The future
   a) Exome capture and complete genome sequencing

a. The evolutionary framework



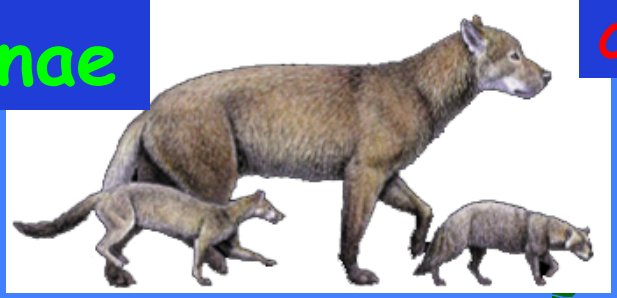
Relationships of Carnivora Families



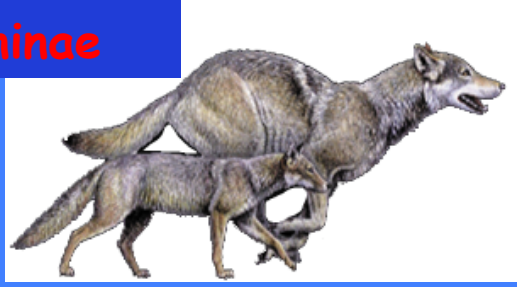
Wayne et al. 1989

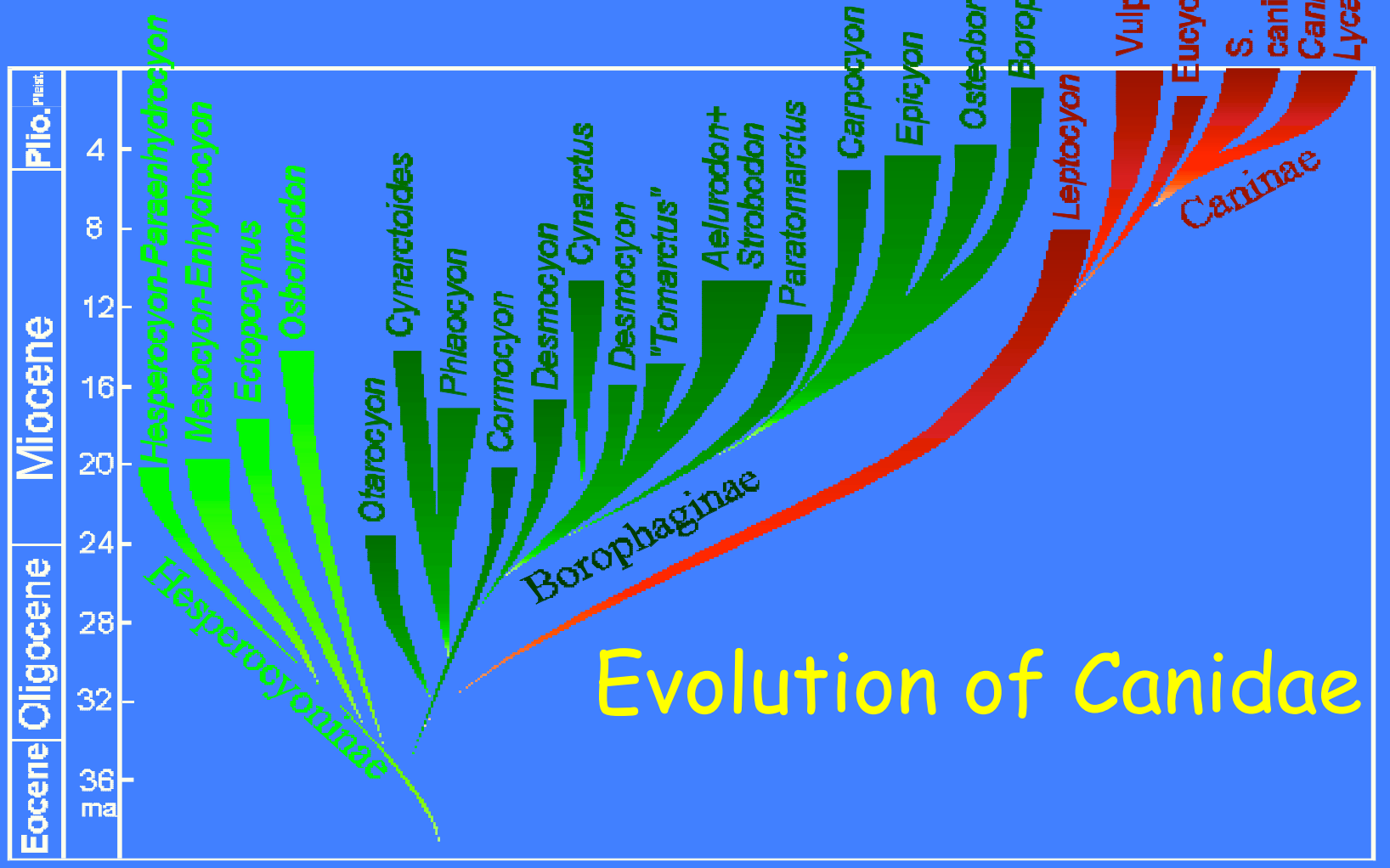
Borophaginae

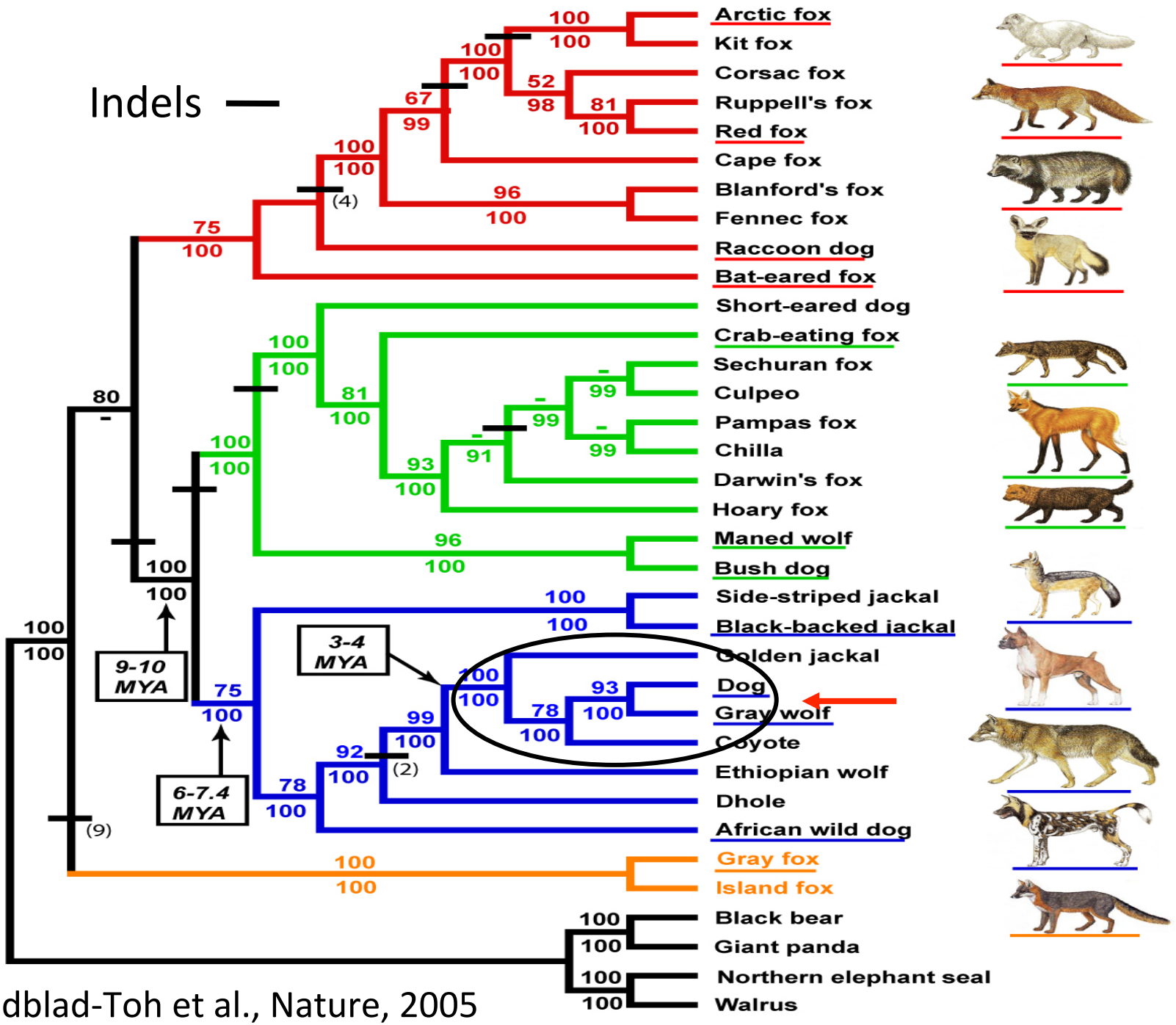
Hesperocyoninae



Caninae







Lindblad-Toh et al., Nature, 2005

CONCLUSIONS

1) The dog has an ancient divergence from other carnivores dating to about 55 million years ago.

2) The extant Canidae has a relatively recent origin about 10 million years ago and is divided into three primary groupings. The dog is assigned to the wolf-like group.

3) The dog is extremely close to gray wolves and shows no evidence of origin from other wolf-like canids.

CHIHUAHUA

HOMO SAPIENS

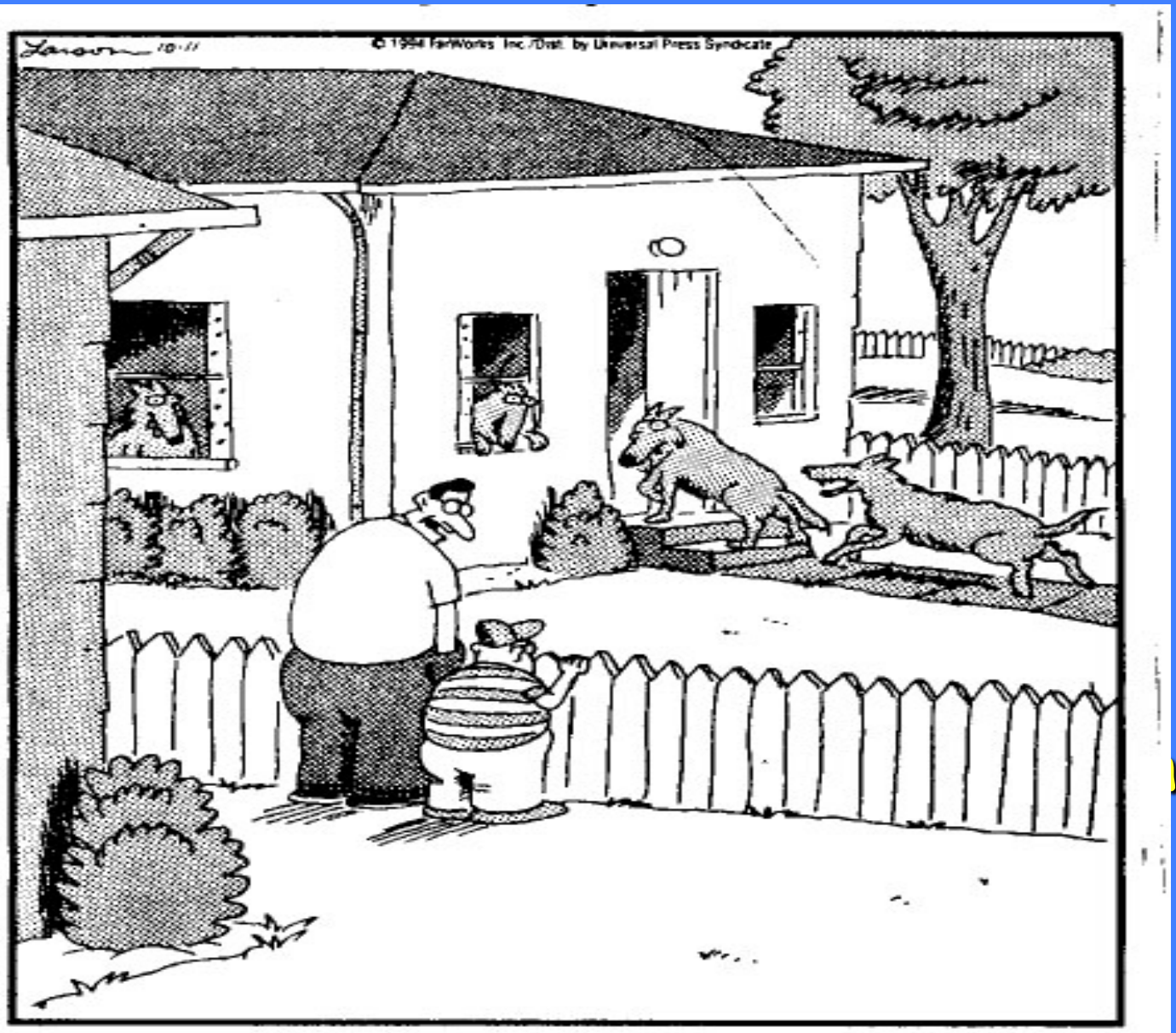
GREAT DANE

b. ORIGIN AND DIVERSIFICATION

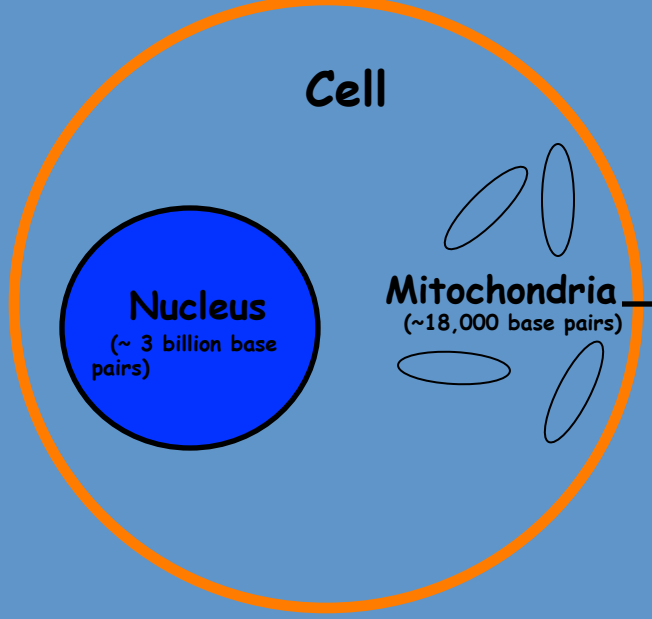
1) W

2) W
(fro

3) H
of d



"I know you miss the Wainwrights, Bobby, but they were weak and stupid people — and that's why we have wolves and other large predators."



Mitochondrial DNA (mtDNA)

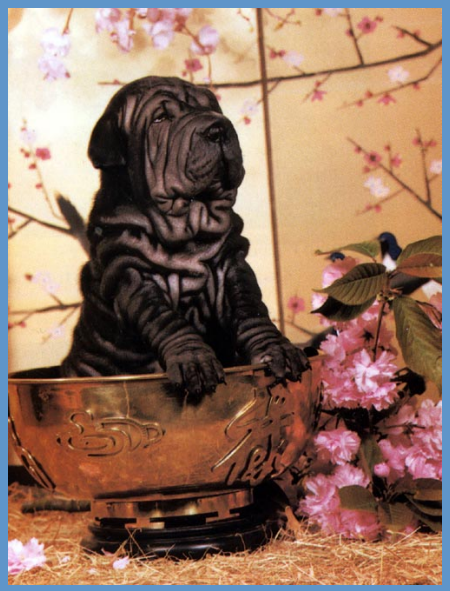
- One linked array of genes
- One locus
- Control region sequence
- Maternally inherited

...ATCATCCC... **Specimen 1**

...ATCTTCCC... **Specimen 2**

140 dogs from 67 breeds

162 wolves from 27 populations

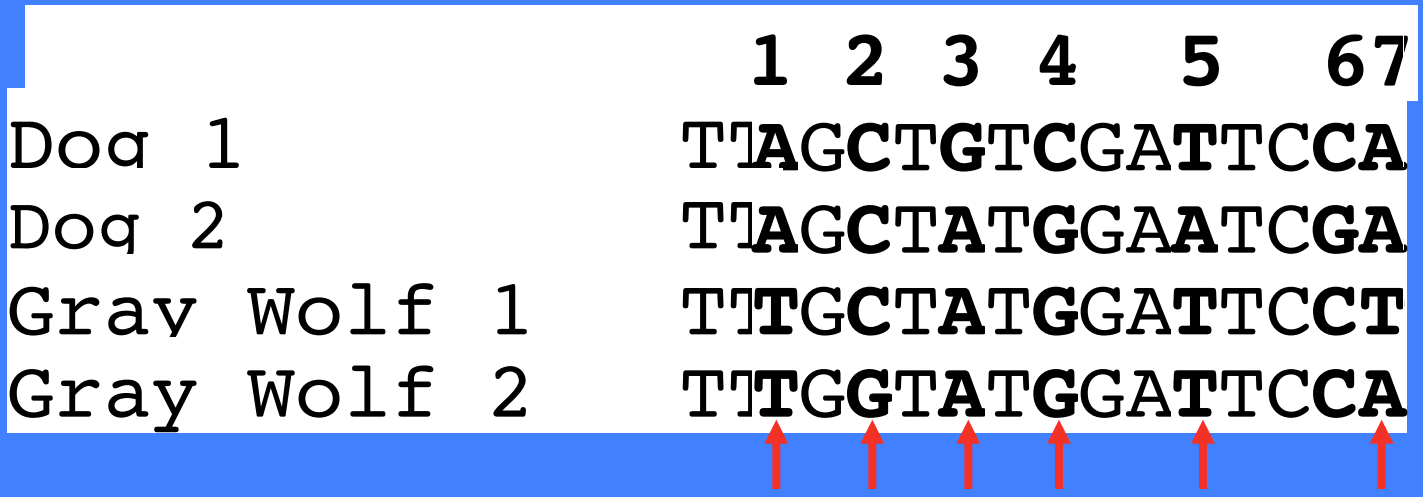


Vilà et al., Science, 1997

The sequence diversity within dogs is similar to that within wolves

Average sequence divergence in dogs = 2.06% ± 0.07%
26 haplotypes in 140 dogs from 67 breeds

Average sequence divergence in wolves = 2. 1% ± 0.04%
27 haplotypes in 162 wolves from 27 populations



Tabulation of DNA sequence differences among individuals

TTGCAGCTCTCC
TTGCAGCTCTCC

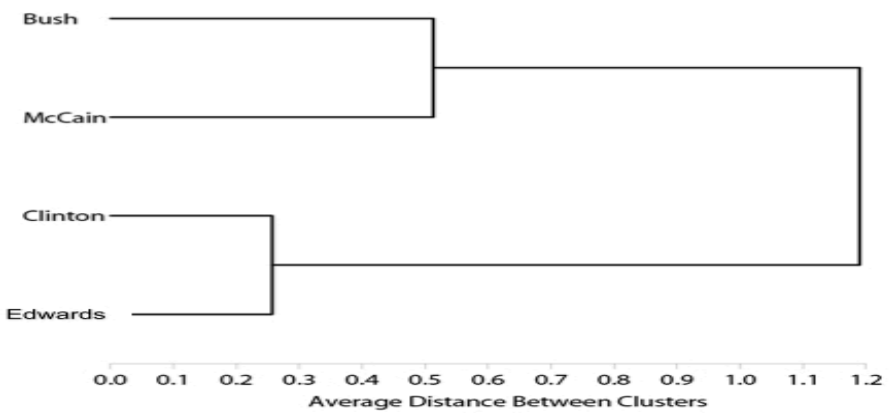
TTGCAGCTCTCC
ATGCAGCTCTCG

ATGCAGCTCTCG
ATGCTGCTCTCG

ATGCTGCTCTCG
ATGCTGCTCTCG

	Bush	McCain	Clinton	Edwards
Bush	0	.	.	.
McCain	2	0	.	.
Clinton	5	3	0	.
Edwards	6	4	1	0

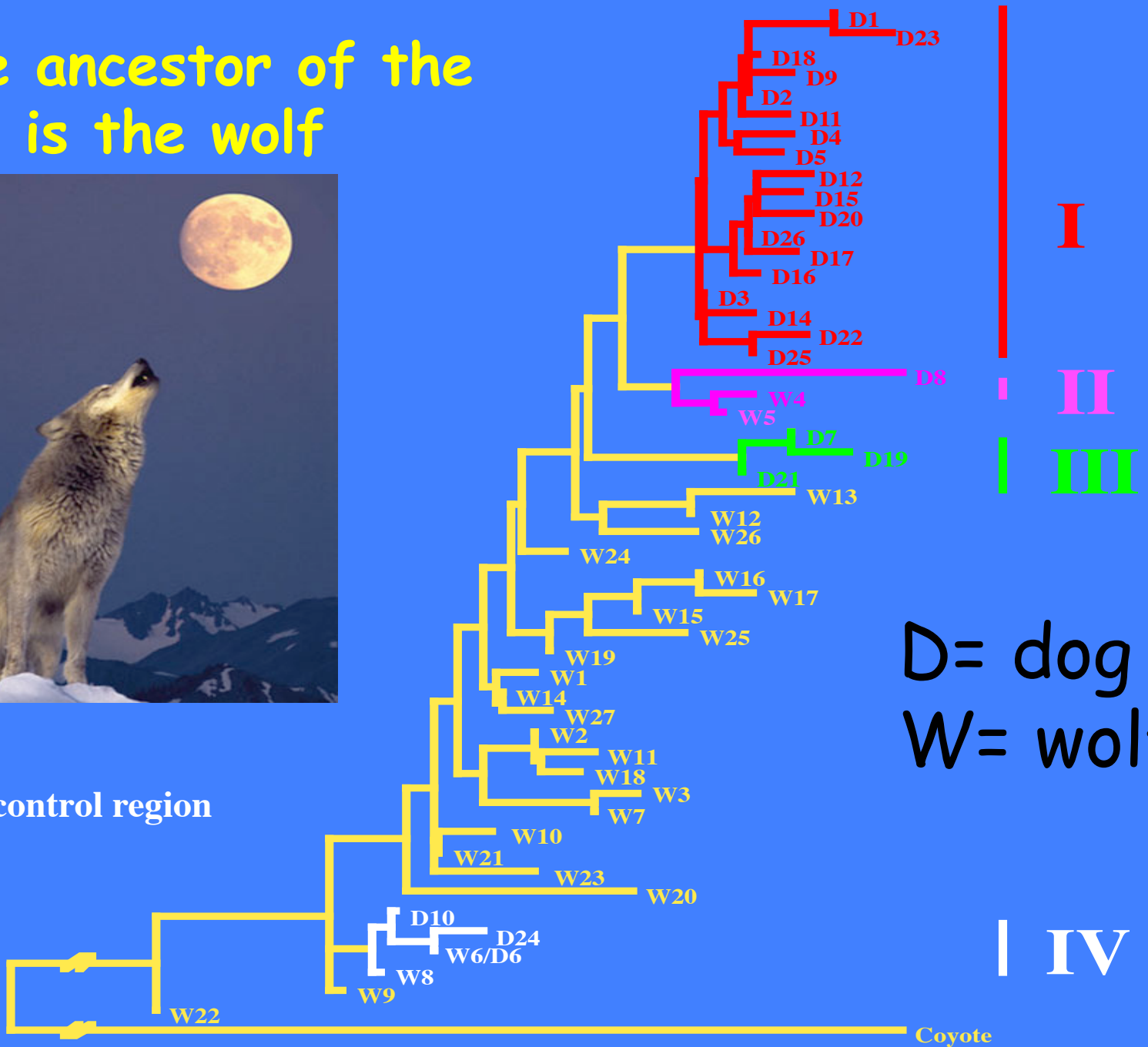
DNA differences can be summarized in a "tree"



Courtesy of Lynn Jorde

The ancestor of the dog is the wolf

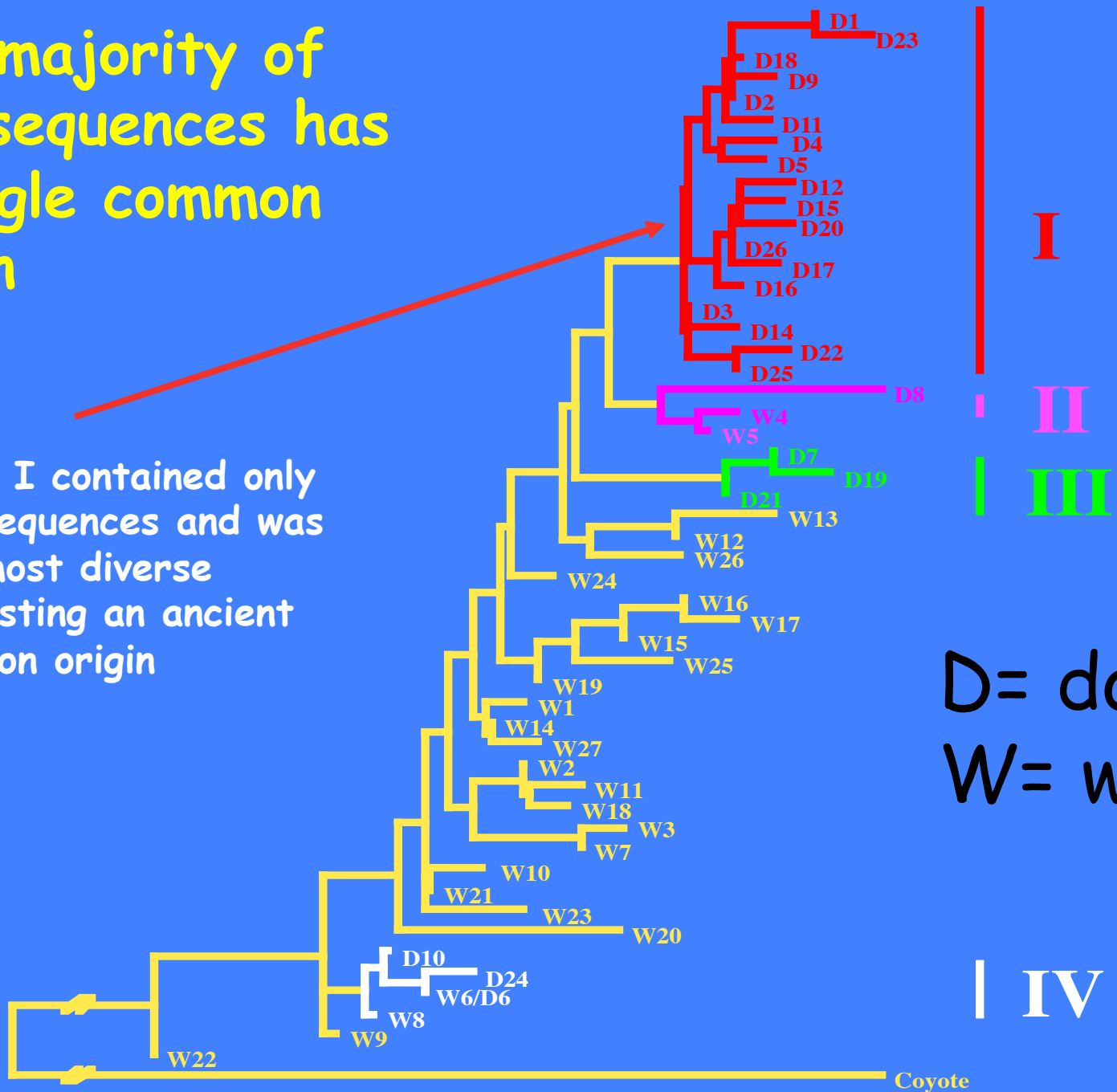
261bp control region



D= dog
W= wolf

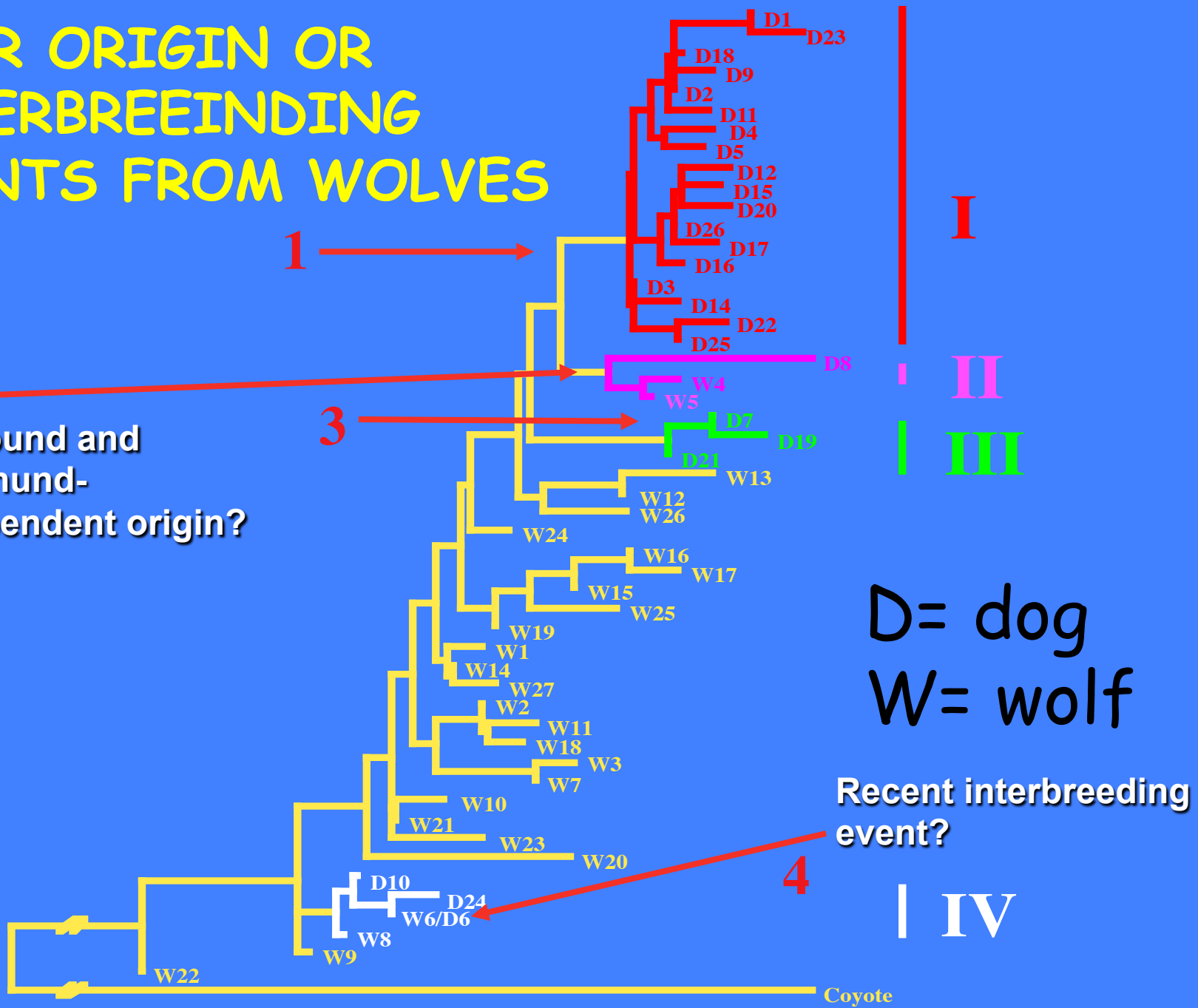
The majority of dog sequences has a single common origin

Clade I contained only dog sequences and was the most diverse suggesting an ancient common origin



D= dog
W= wolf

FOUR ORIGIN OR INTERBREEINDING EVENTS FROM WOLVES



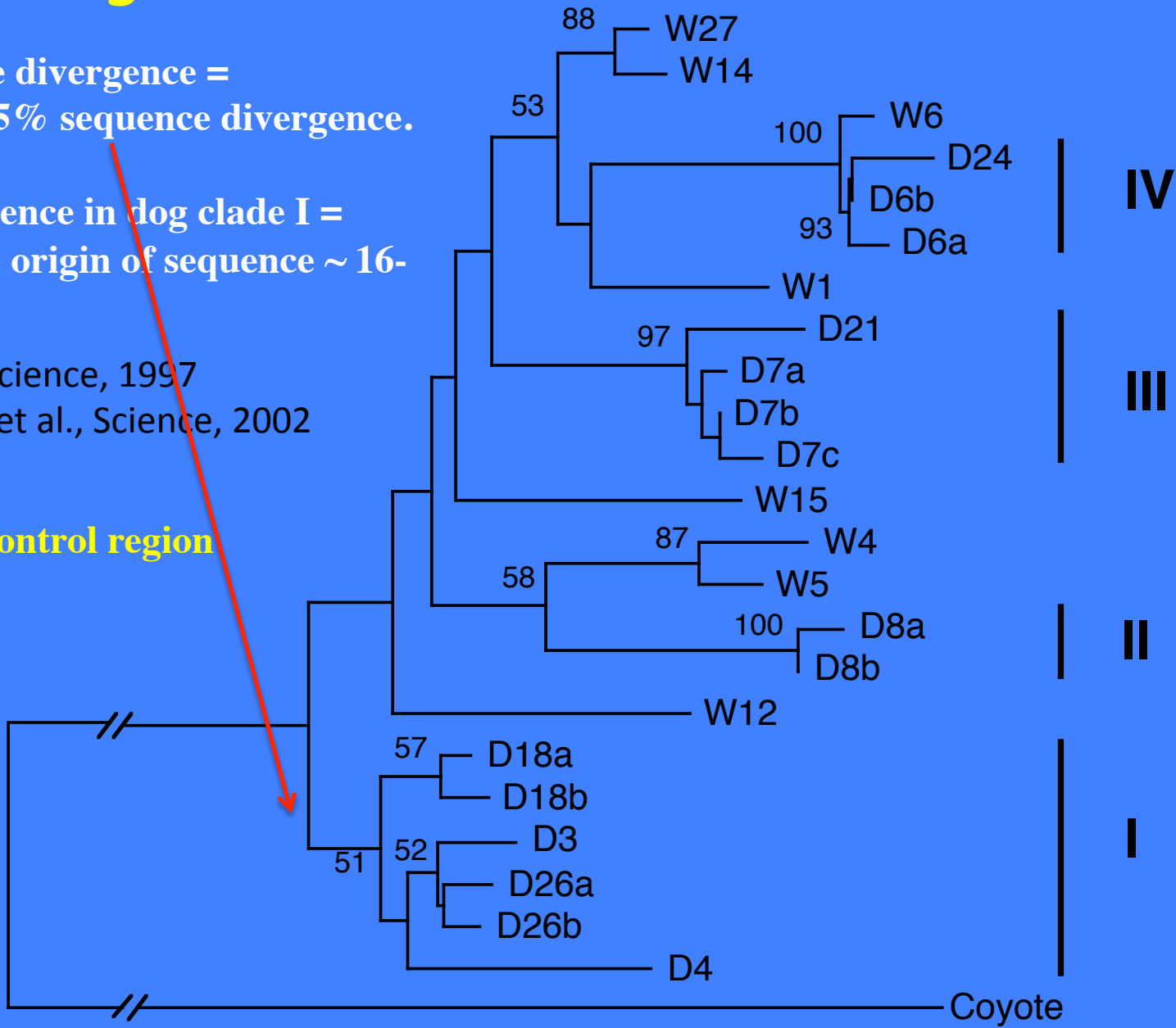
The Timing of Domestication

Wolf/coyote divergence =
1 mybp, 7.5% sequence divergence.

Max divergence in dog clade I =
1.0% hence origin of sequence ~ 16-100Kya

Vila et al., Science, 1997
Savolainen et al., Science, 2002

1030bp control region



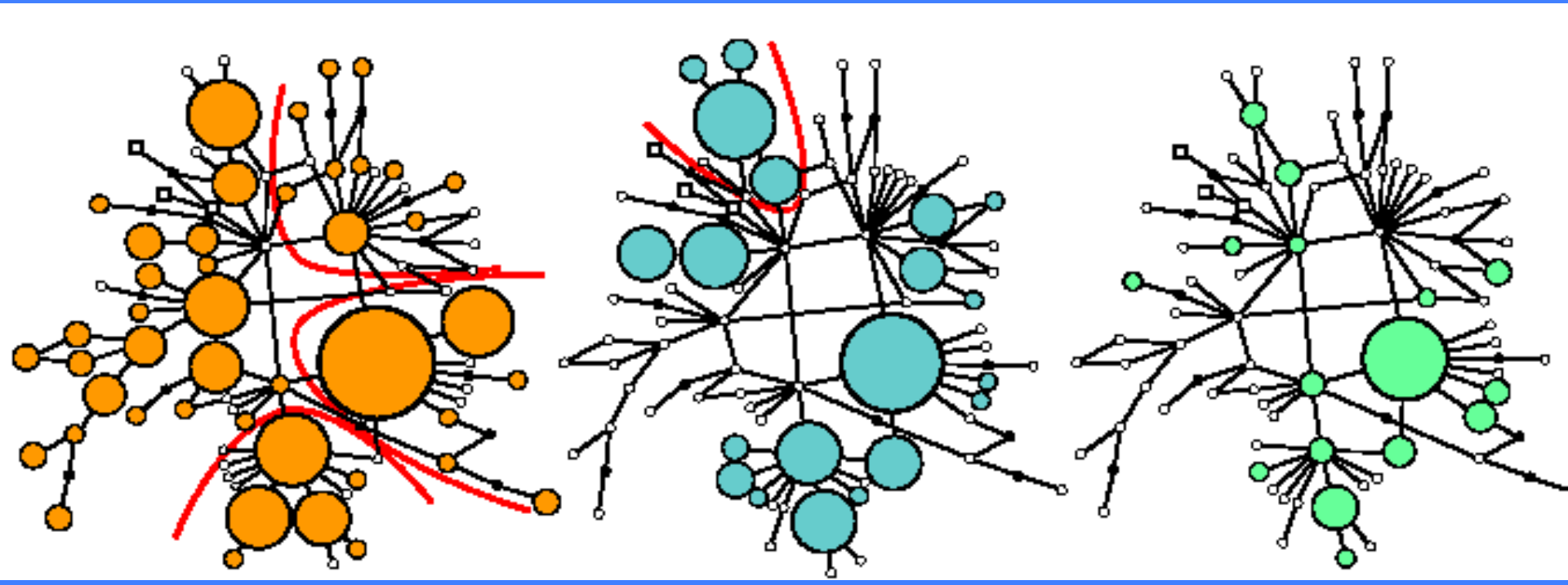
Dogs breeds were founded by multiple matralines or interbred

	I	II	III	IV
Chow Chow (n= 3)	D1, D2, D3			
Border Collie (n= 3)	D1, D5			
Wirehaired Dachshund (n= 3)	D5			D10
Australian Dingo (n= 4)	D18			
Norwegian Elkhound (n= 9)	D3	D8		
German Shepherd (n= 8)	D4, D5		D7, D19	D6
Afghan Hound (n= 3)				D6
Siberian Husky (n= 3)	D3, D18		D7	
Jämthund (n= 3)		D8	D7	
Flat-coated Retriever (n= 3)	D4			D10
Golden Retriever (n= 6)	D4, D15			D6, D24
Labrador Retriever (n= 6)	D4, D12			
Samoyed (n= 3)	D1, D4, D5			
Giant Schnauzer (n=3)	D4		D7	
English Setter (n= 4)	D3, D5			
Irish Setter (n= 3)	D1, D9			
Mexican Hairless (n= 19)	D1, D3, D4, D26		D7, D21	D6

Origin in East Asia

1. High diversity in East Asia suggests an origin there.

2. Founding by multiple matralines is consistent w/ origin 15 to 40k



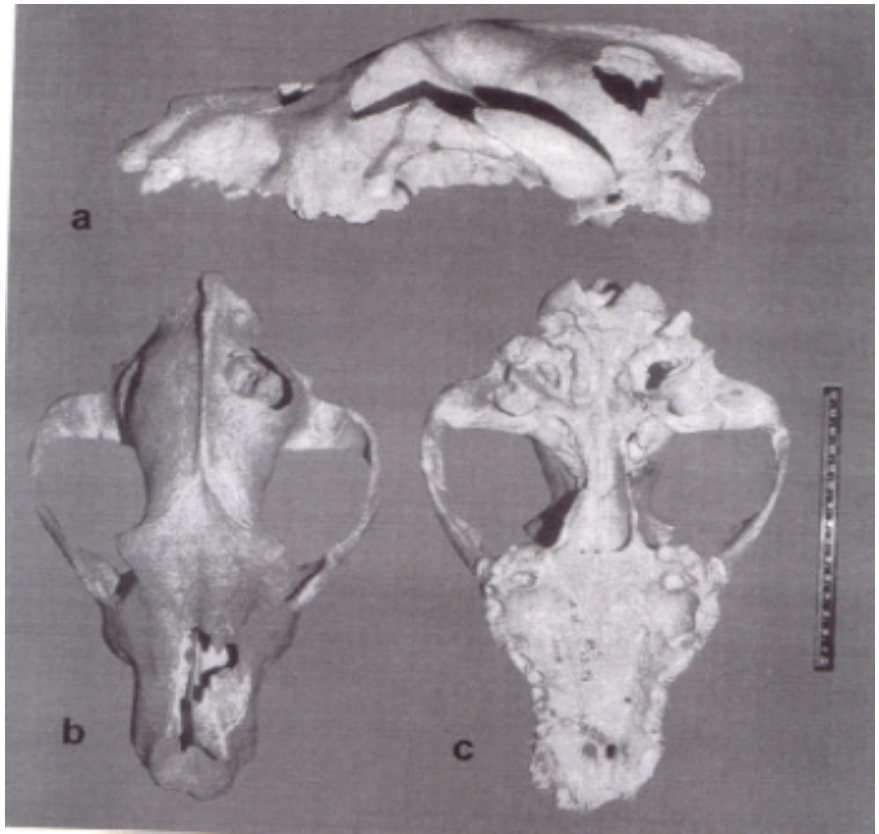
East Asia

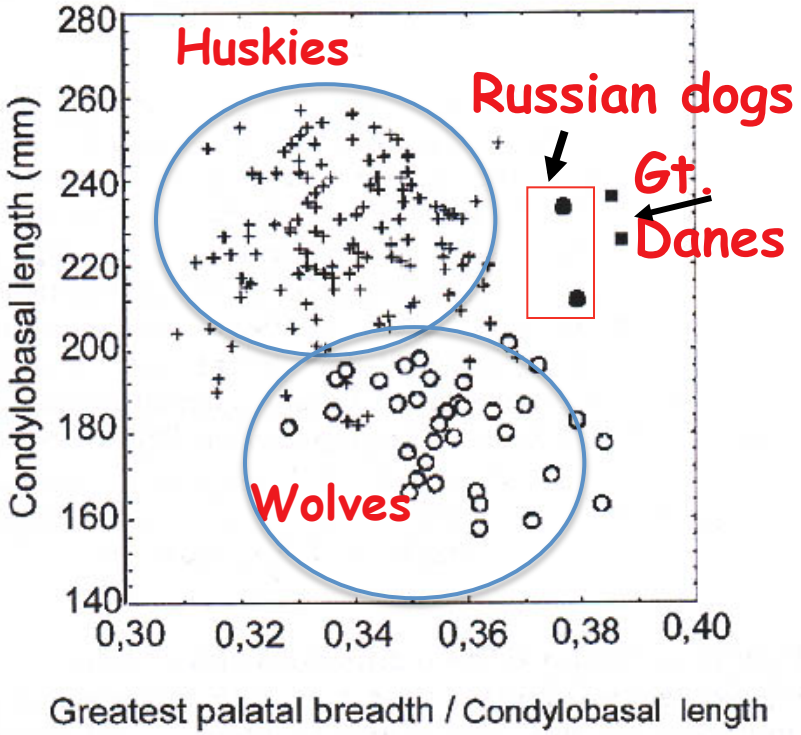
Europe

Southwest Asia

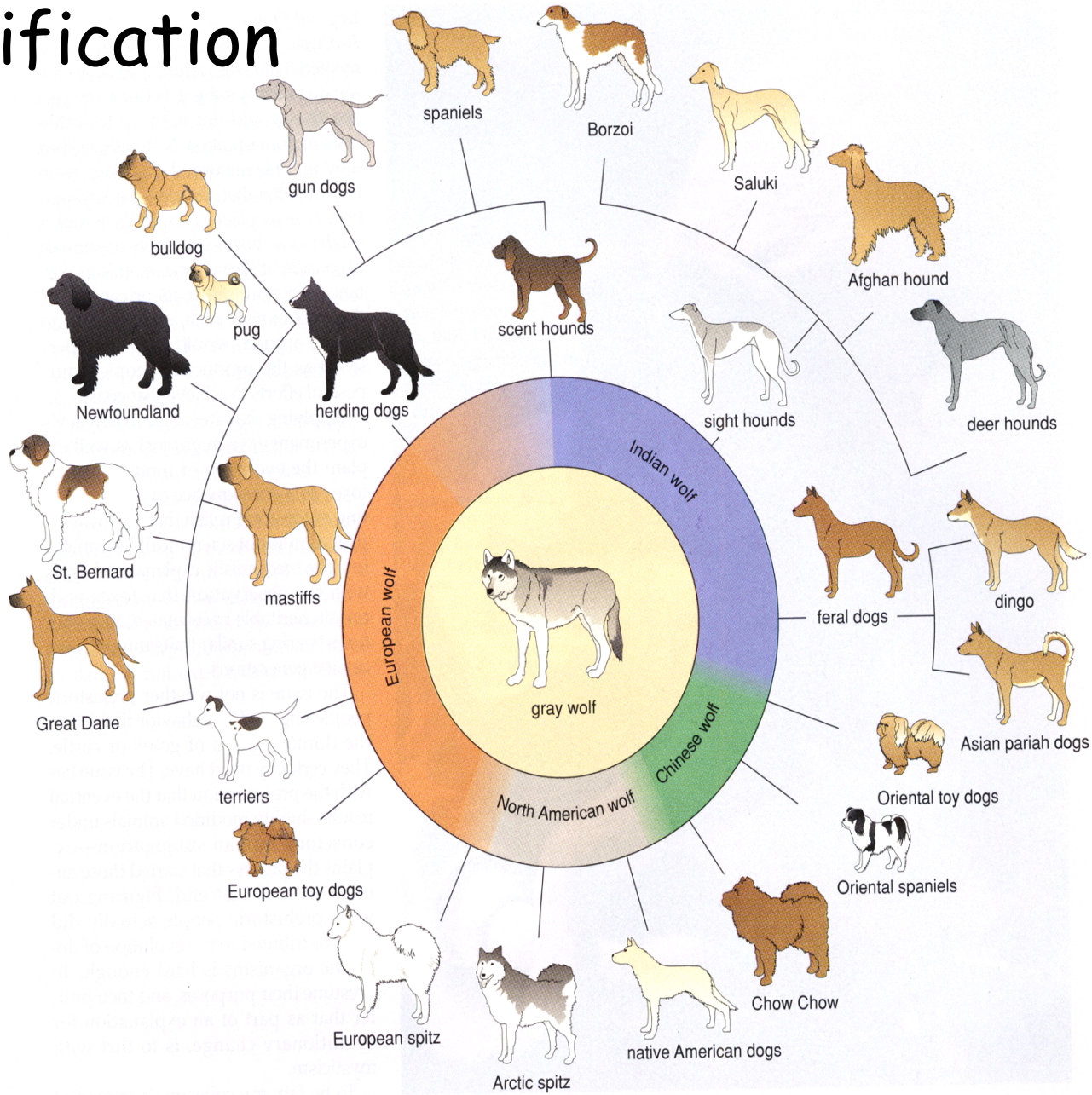
Savolainen et al., Science, 2002

The Oldest dogs: 15-31K years ago

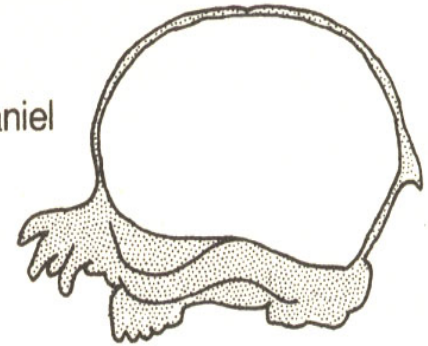




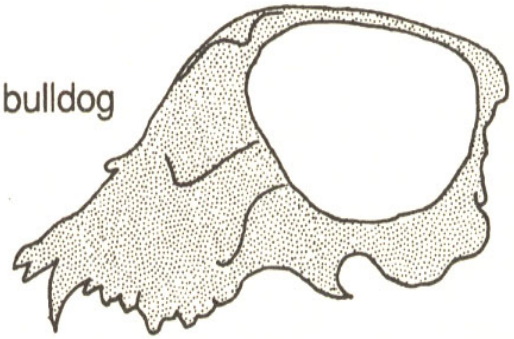
Diversification



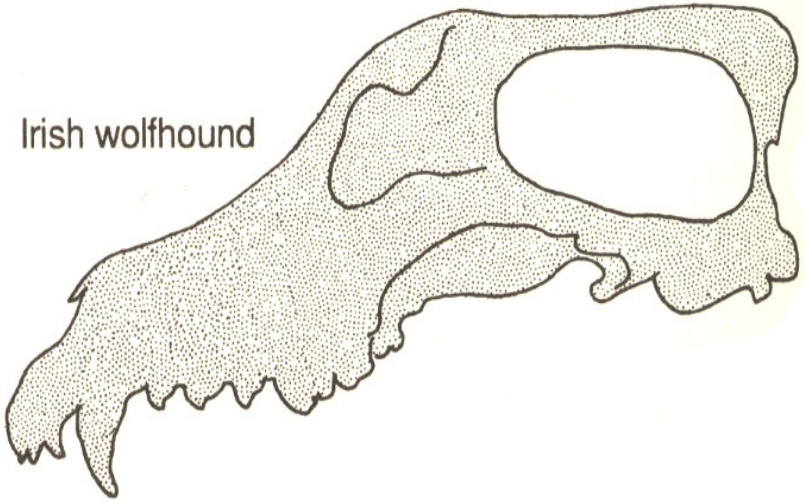
King Charles spaniel



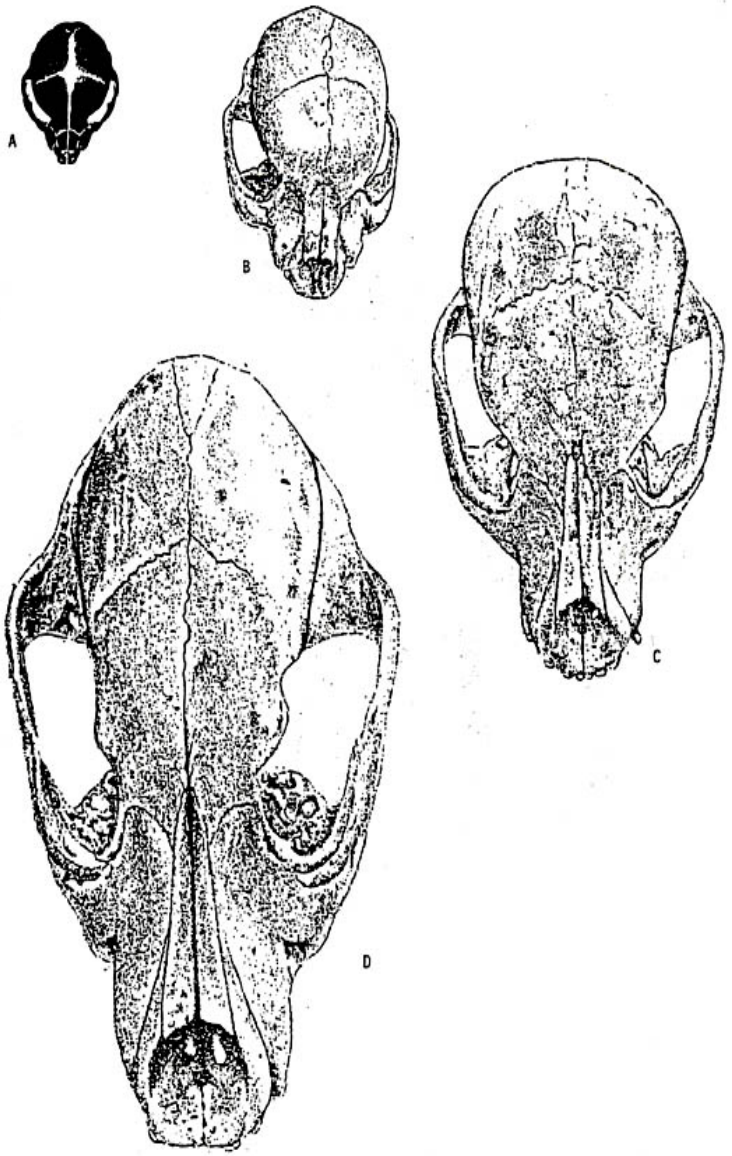
English bulldog



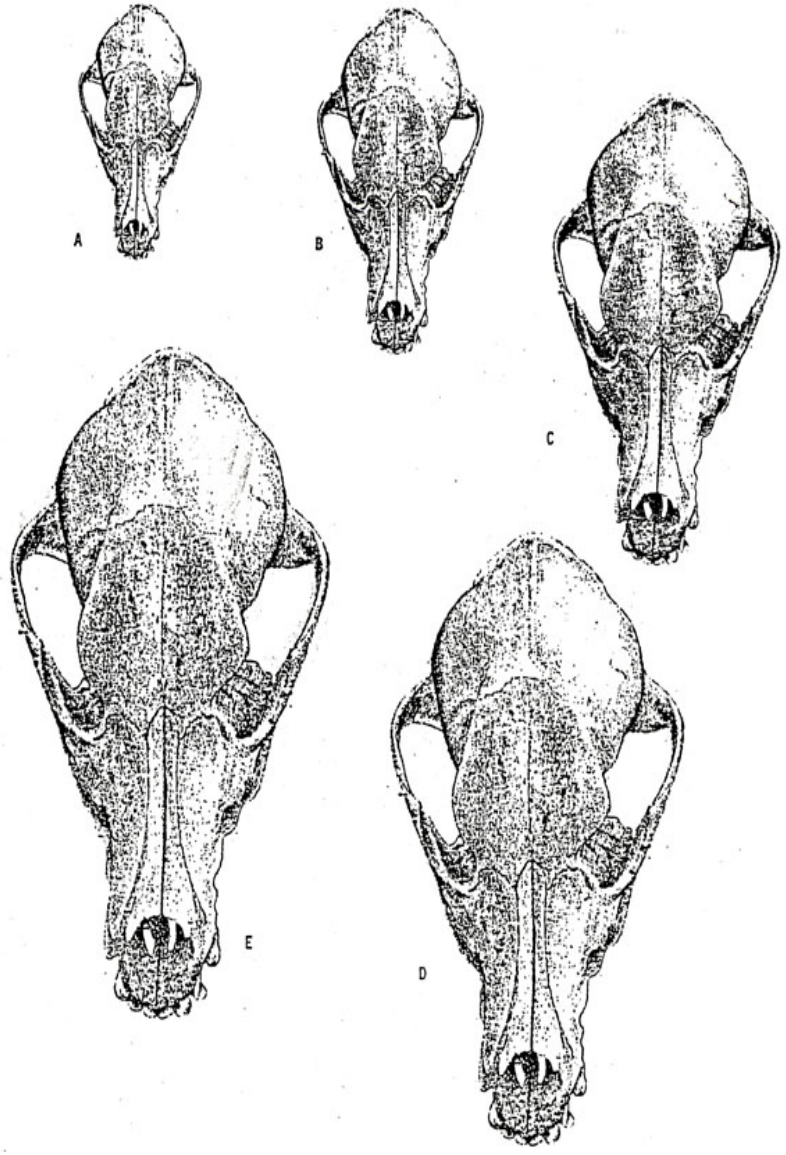
Irish wolfhound

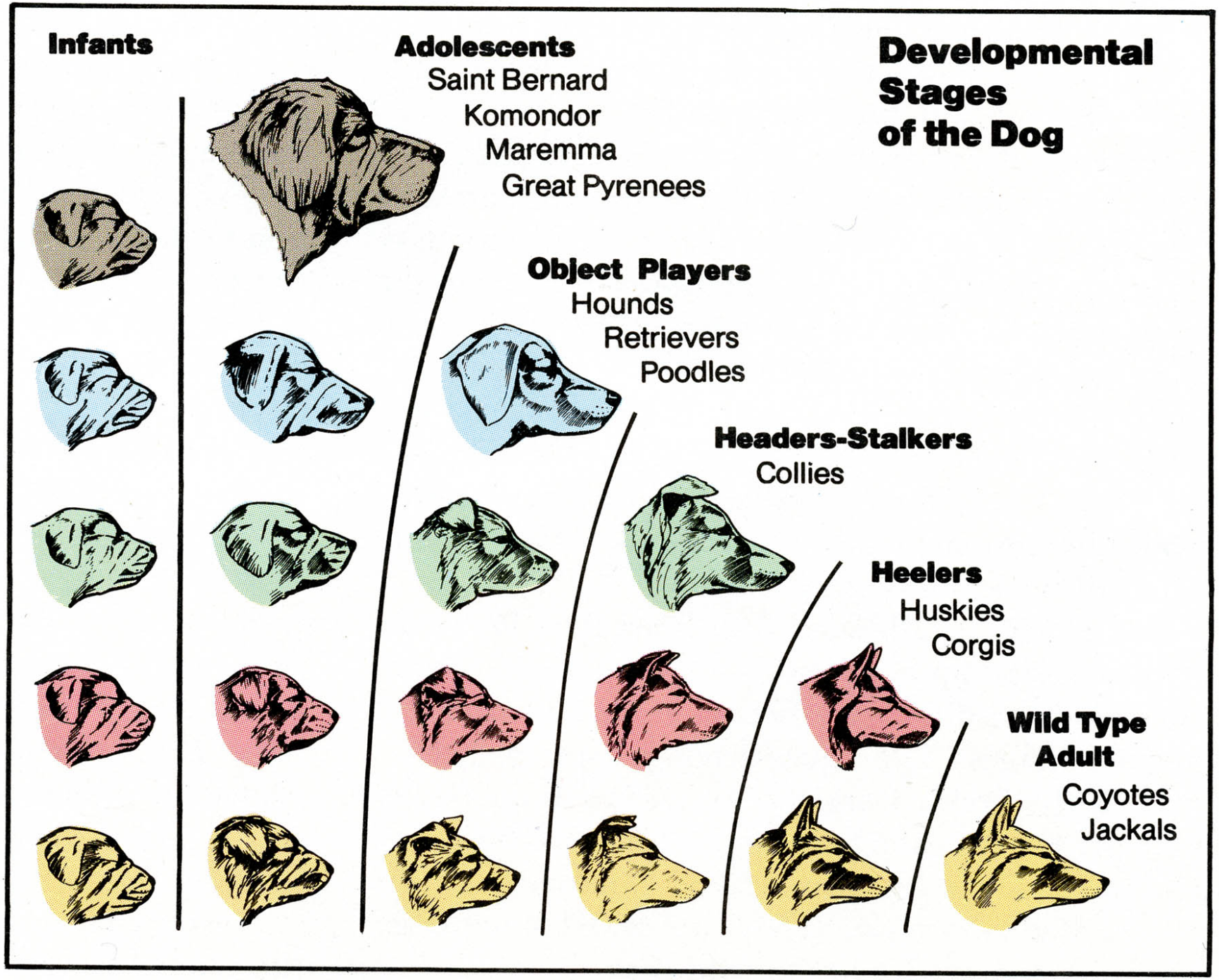


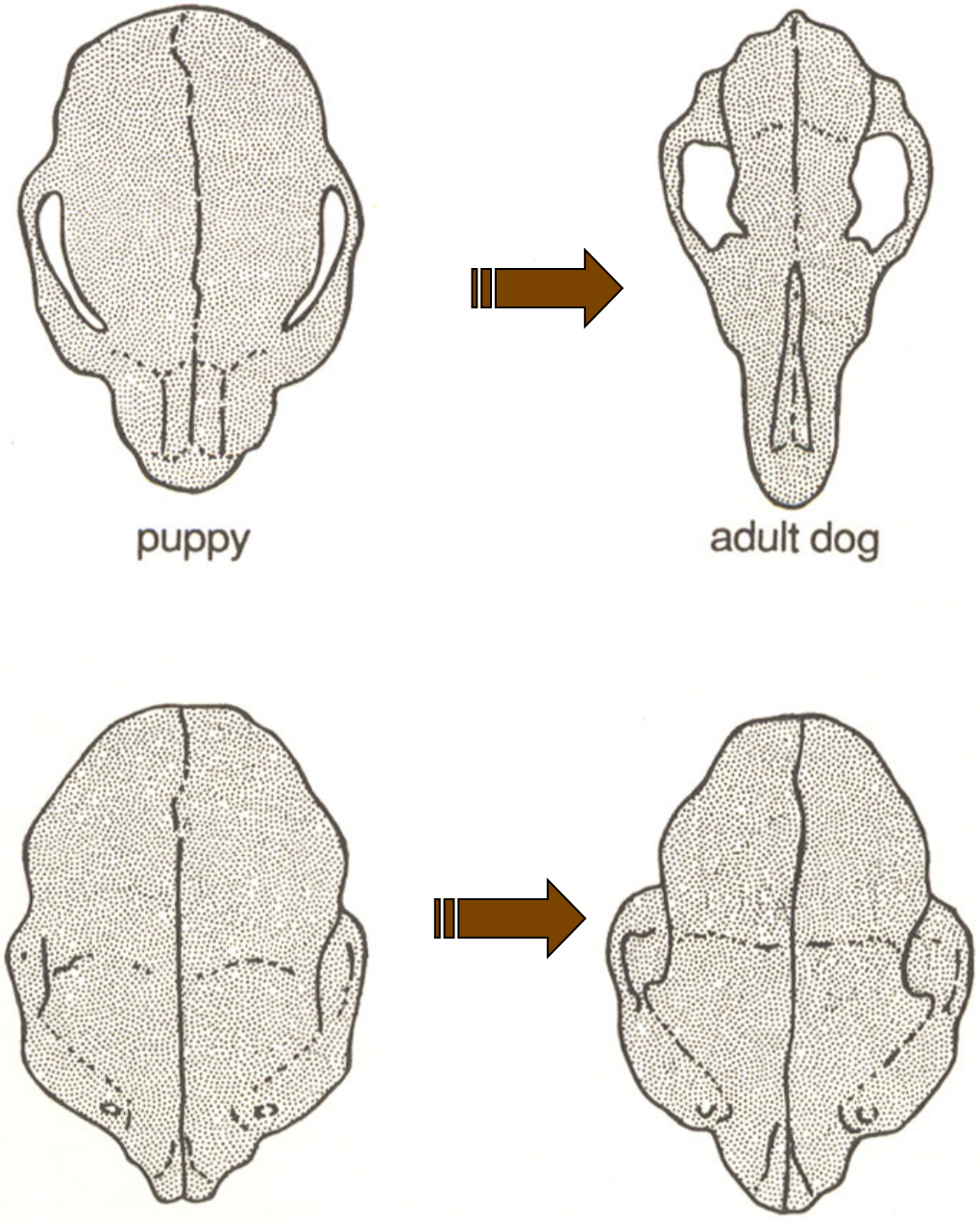
Allometric Growth



Proportionate Growth

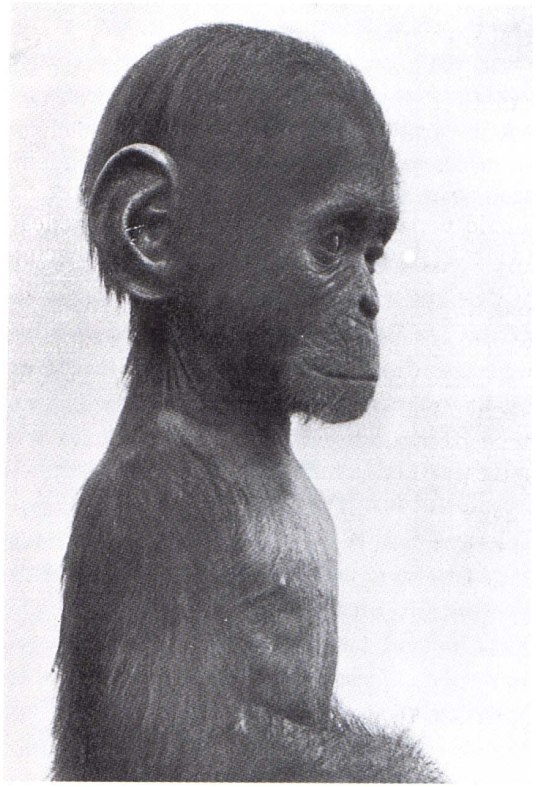



Infants
Adolescents
Saint Bernard
Komondor
Maremma
Great Pyrenees
Developmental Stages of the Dog
Object Players
Hounds
Retrievers
Poodles
Headers-Stalkers
Collies
Heelers
Huskies
Corgis
Wild Type Adult
Coyotes
Jackals

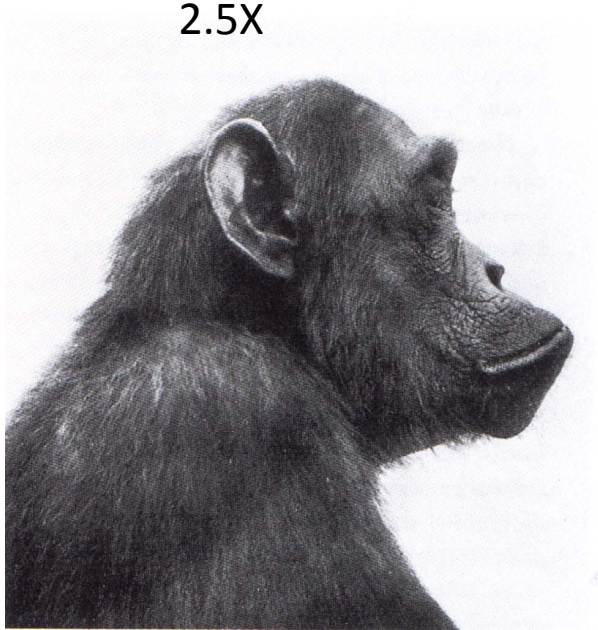


Wayne, Evolution, 1986

Neoteny (juvenilization)



2.5X



3.3X

CONCLUSIONS

1) The dog has high mitochondrial DNA sequence diversity suggesting a diverse origin from multiple wolf matralines from a wide region.

2) However, the higher diversity of dog mtDNA in East Asia suggest that dogs may have had a long history there.

3) Dogs are the oldest domesticated animals, originating 15-100,000 years ago.

4) Neoteny has been a common mode of change, affecting behavior and morphology.

Questions

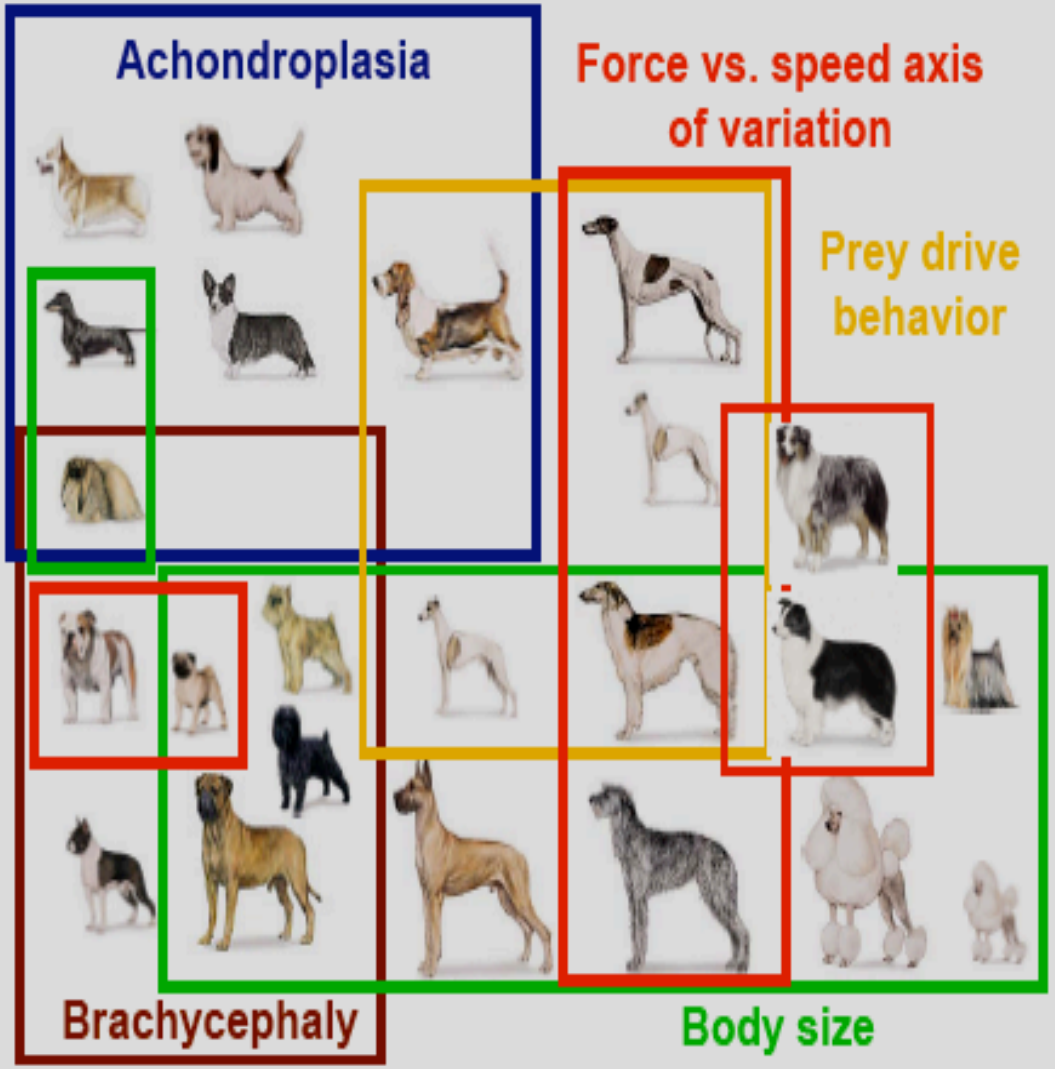
1) The reason dogs appear to have evolved in East Asia is based on:
   a. The fossil record
   b. Historical record
   c. Morphological studies
   d. Higher mtDNA diversity
   e. Lower mtDNA diversity

2) The apparent phenotypic diversity in skull shape in dogs as compared to cats reflects what general difference between the two species:
   a. Selection for bite strength
   b. Selection for bigger brain size
   c. Occlusal and breathing limitations
   d. Differences in developmental gradients

3) The dog is derived from:
   a. an extinct species
   b. a diversity of wolf-like species
   c. gray wolves
   d. none of the above

C. The Dog as a Novel Model Species

- Each breed fixed for specific traits
- Creation of breeds recent and rapid
  - Closed breeding
  - Popular sires
  - Population bottlenecks
- Small number of genes responsible for many complex traits



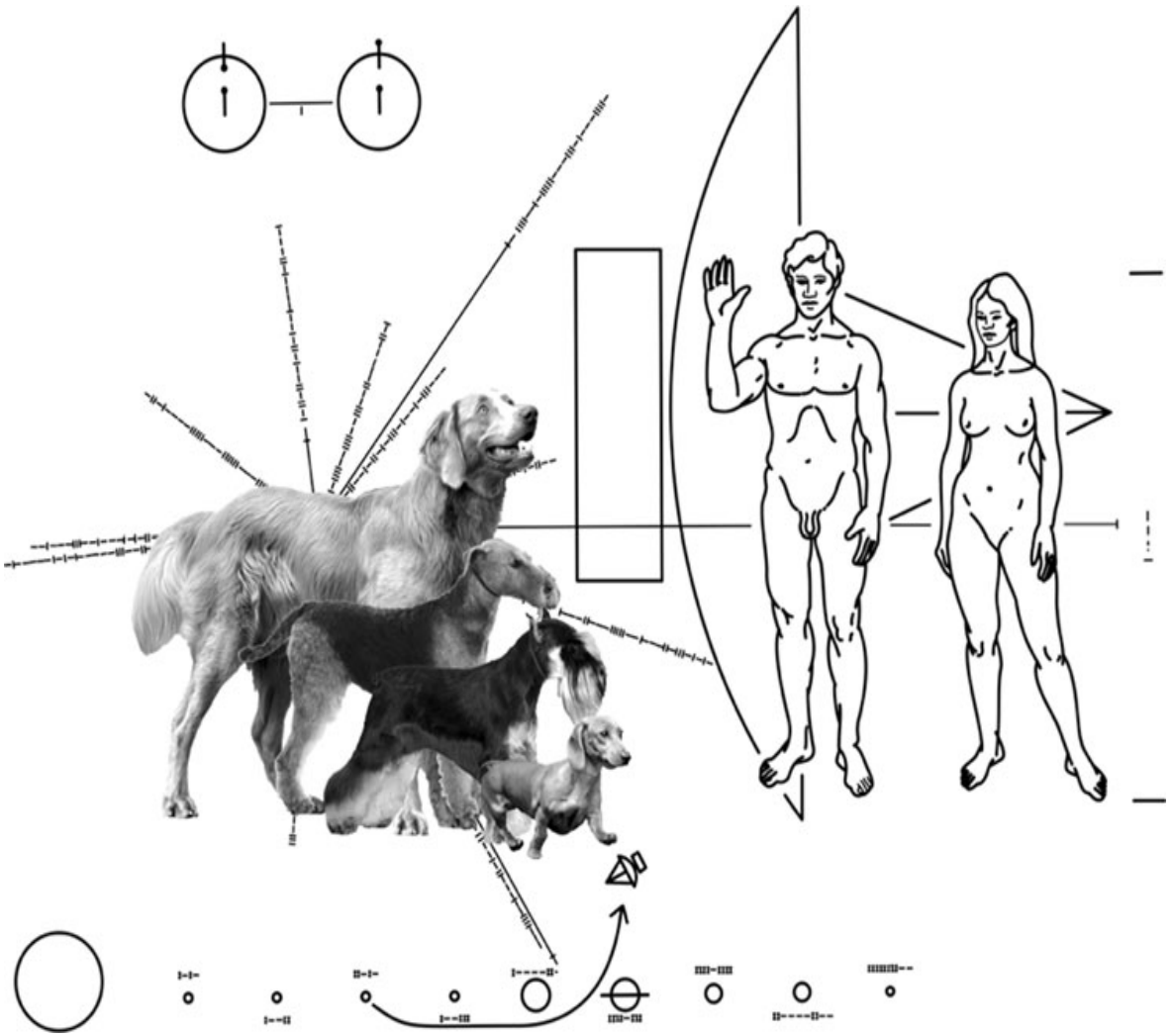


Fig. 1 Dog breed diversity helps to cure humans. The photo of four dog breeds, meant to represent canine diversity, is superimposed on the famous pictorial message sent on the Pioneer 10 spacecraft. The Pioneer plaque shows a man and a woman, a hyperfine transition of neutral hydrogen (*circles top left*), the relative position of the Sun with respect to the center of the galaxy and 14 pulsars (*lines*), the solar system (*circles bottom*), and silhouette of the spacecraft. It aimed to represent how the human species portrays itself and its place in the galaxy

Derrien et al. Mammalian Genome, 2012

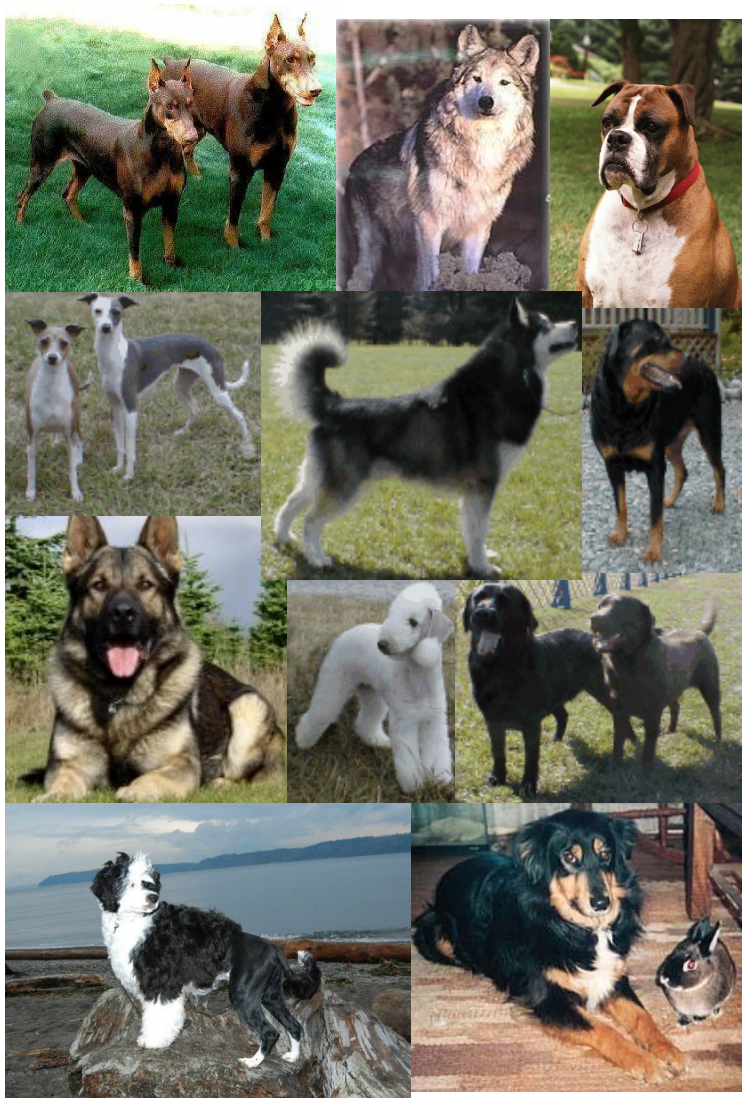
D. The Sequencing of the Dog

Dog Genome Resources

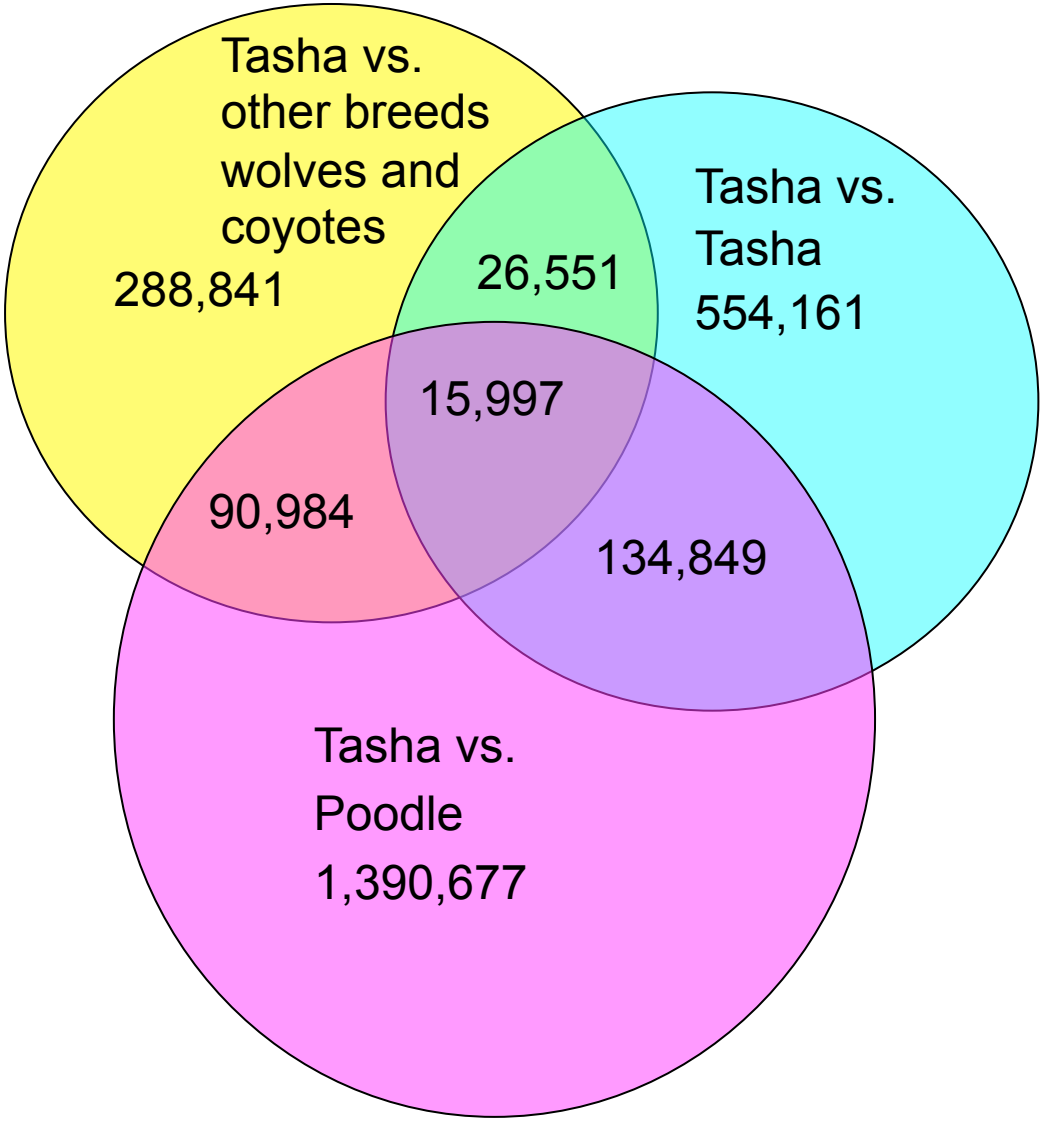
- Complete genome sequenced - 2.4 Gb
- Boxer – 7.5X coverage (Linblad-Toh et al., Nature, 2005)
- Poodle – 1.5X coverage (Kirkness et al., Science 2004)
- 19,120 Dog genes predicted, nearly all with human homologs
- 2.1-2.6 Million SNPs generated from Boxer and Poodle complete genomes, 10 other dog breeds (100,000 sequence reads each), and 4 wolves and one coyote (20,000 sequence reads each)
- 127,000 Dog SNP microarray chip
- Searchable annotated genome (e.g. UCSC Genome Browser, Ensembl) e.g. allows identification of new microsatellite or SNP loci in regions of interest
- Syntenic map of Dog, Human, Mouse, Rat

- Cost: 46 million dollars

The Dog Genome Project



2.6 Million SNPs Total-SNP Every 5 Kb



E. The CanMap Project

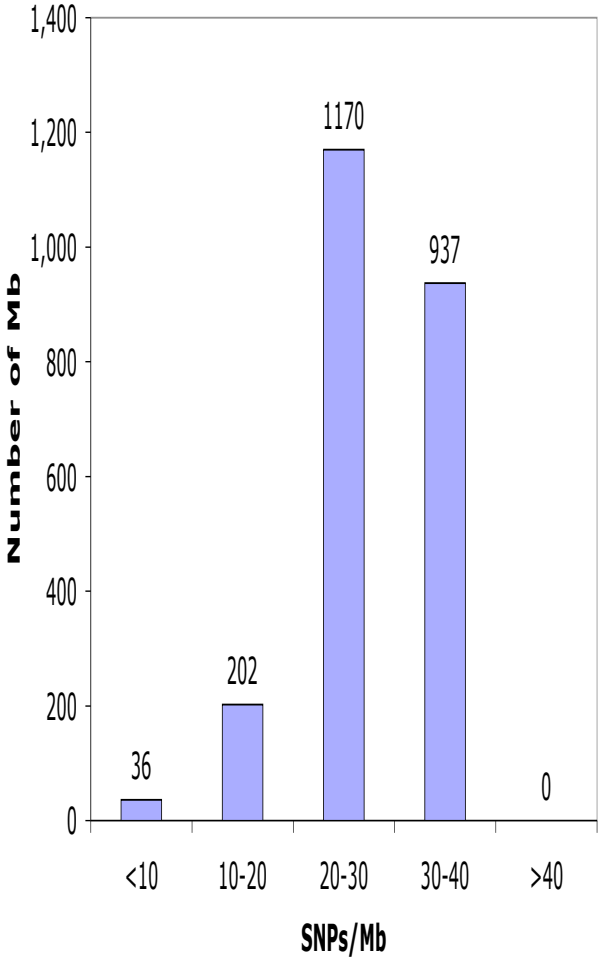
(like the human hapmap project aims to characterize genomic variation)

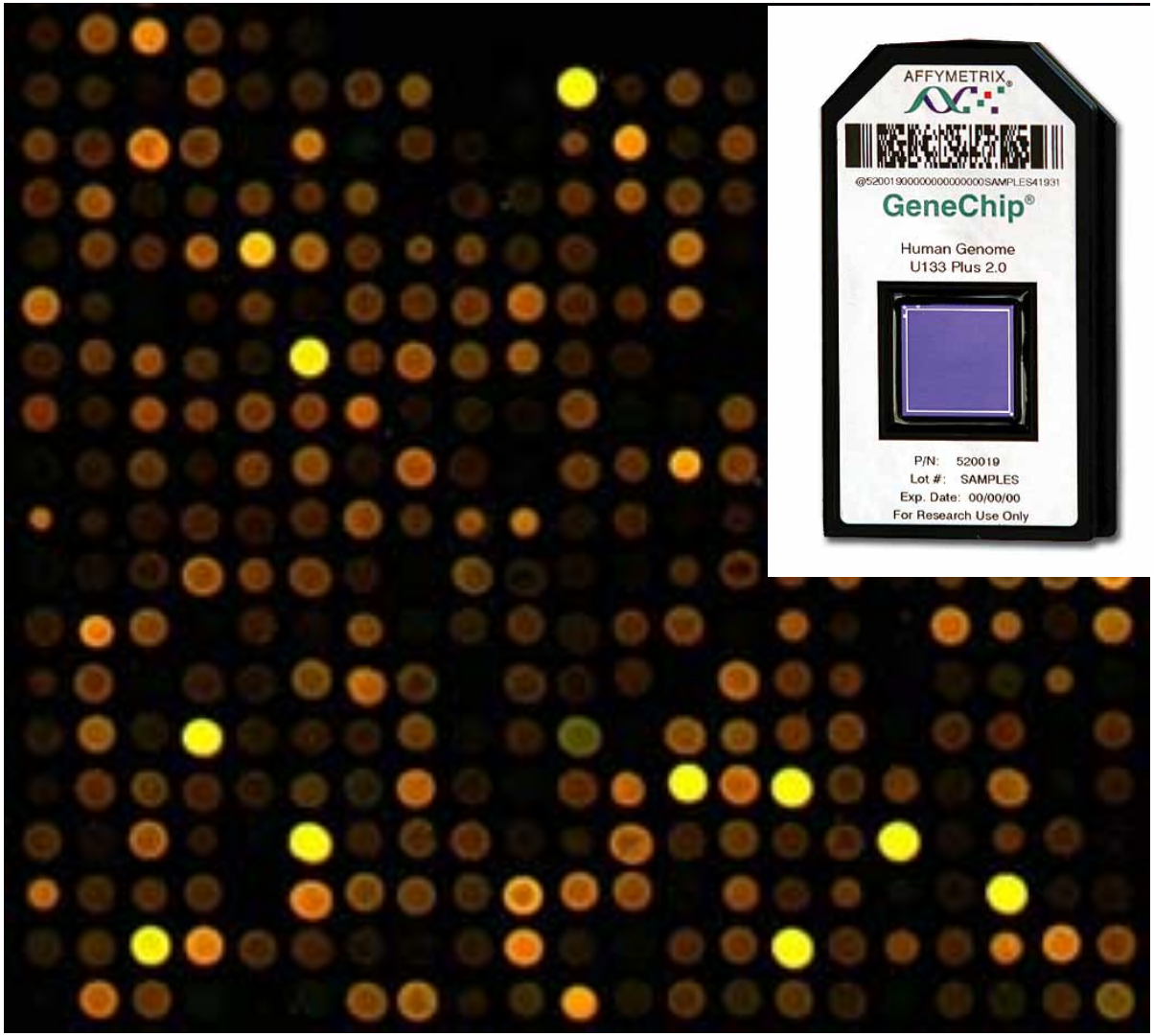
- CanMap Project:
  - Dr. Elaine Ostrander (NHGRI/NIH - CGB)
  - Dr. Robert Wayne (UCLA EEB)
  - Dr. Carlos Bustamante (Cornell/Stanford)
  - Affymetrix and Michele Cargill
- 127,000 SNPs (all Perfect Match oligos)
  - Magic genotyping algorithm- 50-70K
- Quality assessment for standardizing across batch
- Excluding SNPs:
  - When sample call rate is <95%
  - When SNP MAF < 1%
  - When SNP heterozygosity > 60

Samples

Collect DNA from 5-10 unrelated dogs from 85 breeds and gray wolves (the ancestor of dogs)

The Canine SNP genotyping microarray





The CanMap Project

Mapping Genes Underlying Phenotypic Diversity among Domestic Dog Breeds

About
Genetic Diversity
Breed History
Phenotypes
Software
Upload data for Analysis

The domestic dog is a phenotypically diverse species in which discrete phenotypes segregate along breed boundaries. We aim to characterize genetic variation in the domestic dog and identify genomic regions associated with skeletal and coat color differences among dog breeds. Our research design consists of three distinct components:

(1) A genome-wide survey of SNP variation.

(2) Screening the dog genome for footprints of selection.

(3) Genome-wide association mapping of breed-specific traits.

2. How to build a dog

2. How to build a dog

2. How to build a dog

B. Crossing distinct breeds

X

2. How to build a dog

C. Progressive "idea" driven selection

X

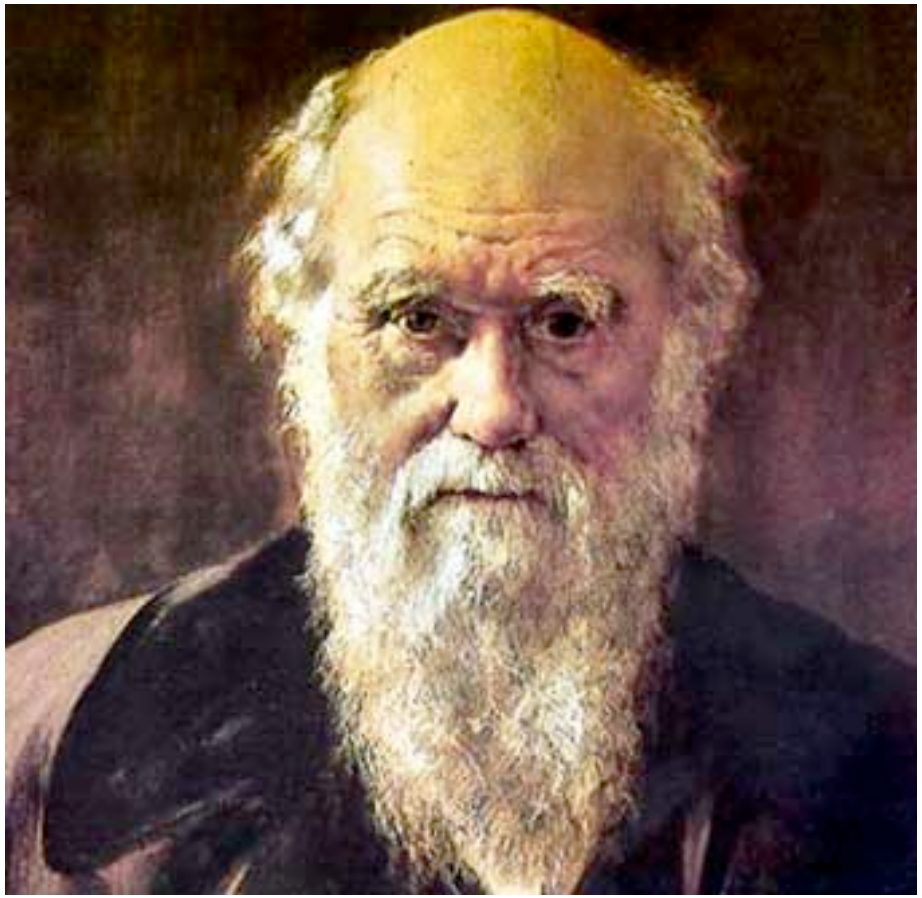
X

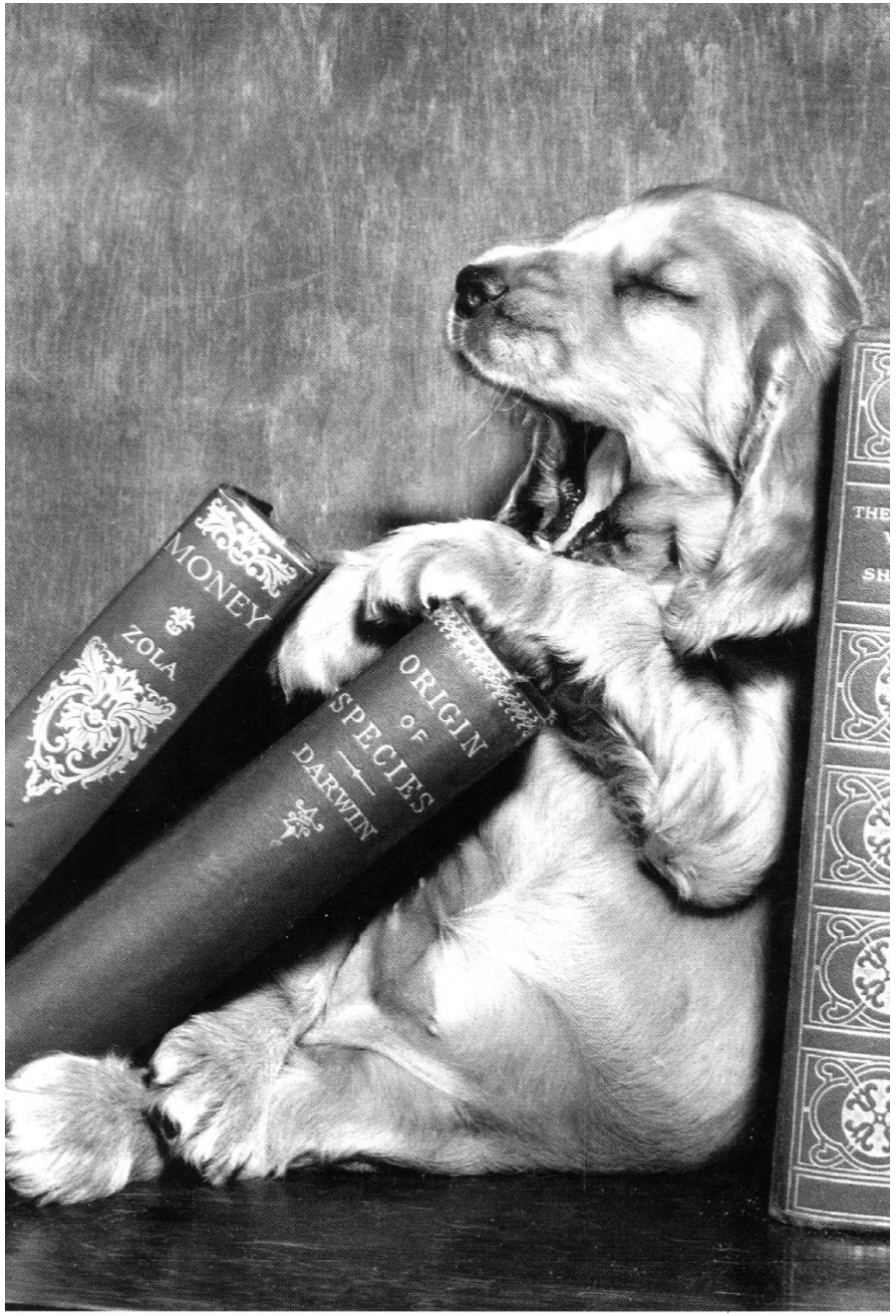
X

X

X

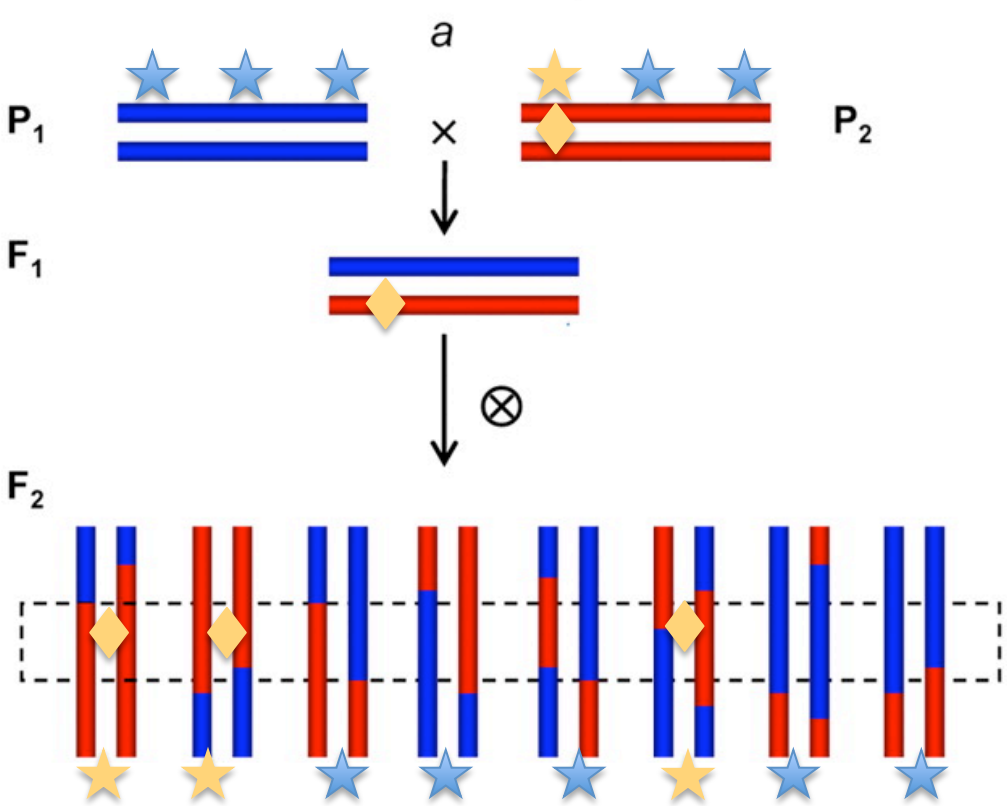
Shorter legs

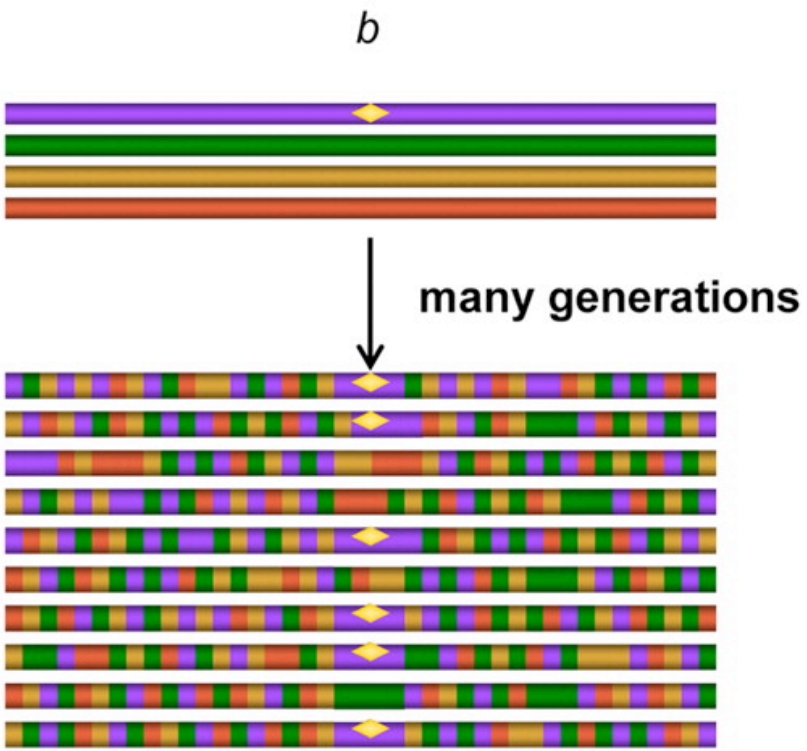




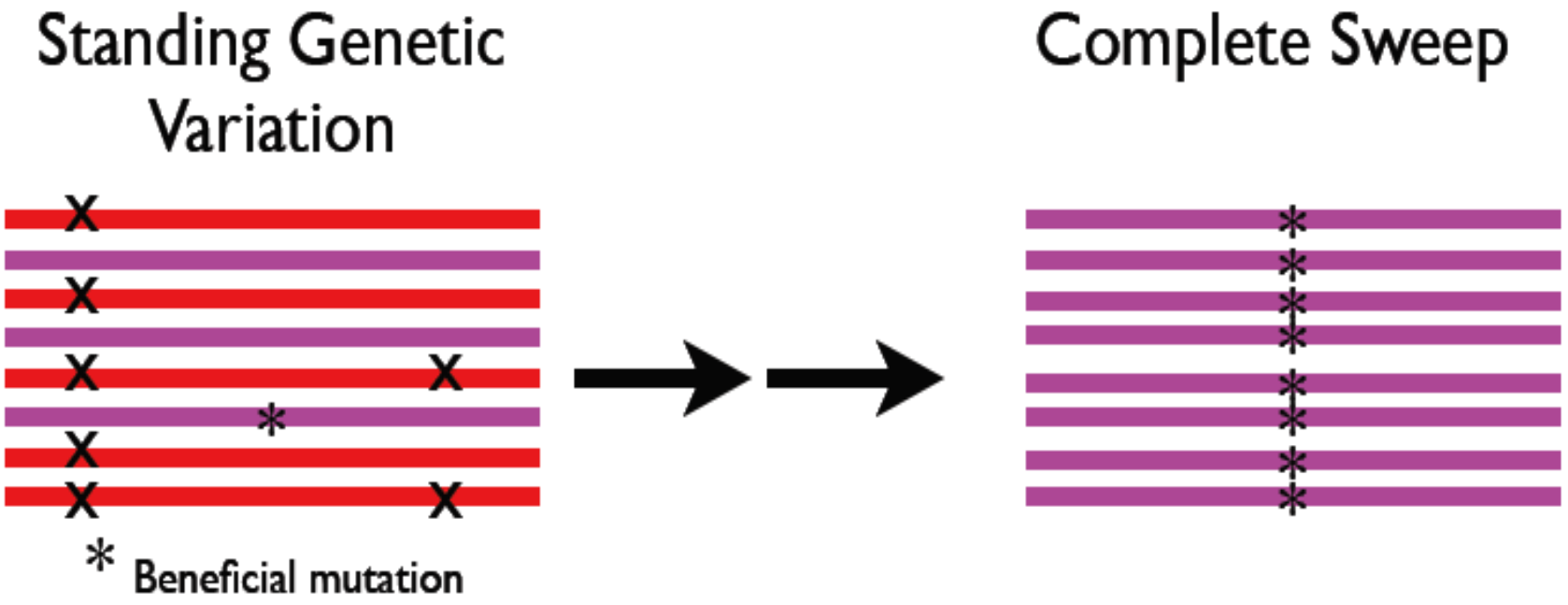
Darwin felt progressive artificial selection was analogous to natural selection

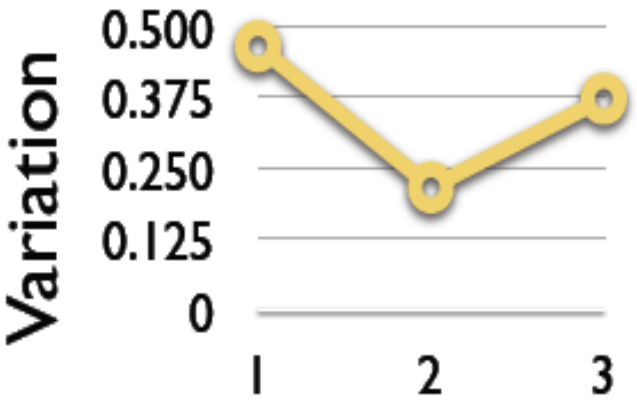
b. Association and selective sweep mapping

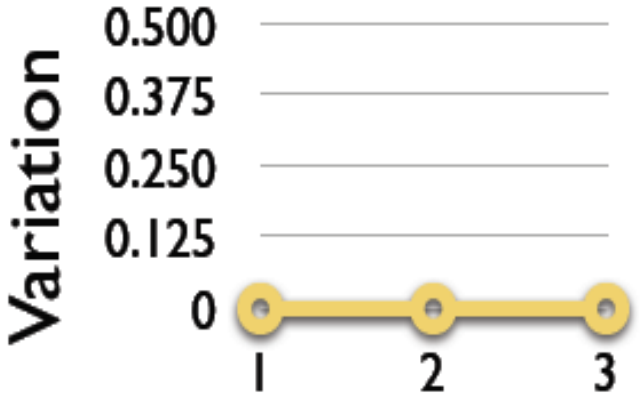


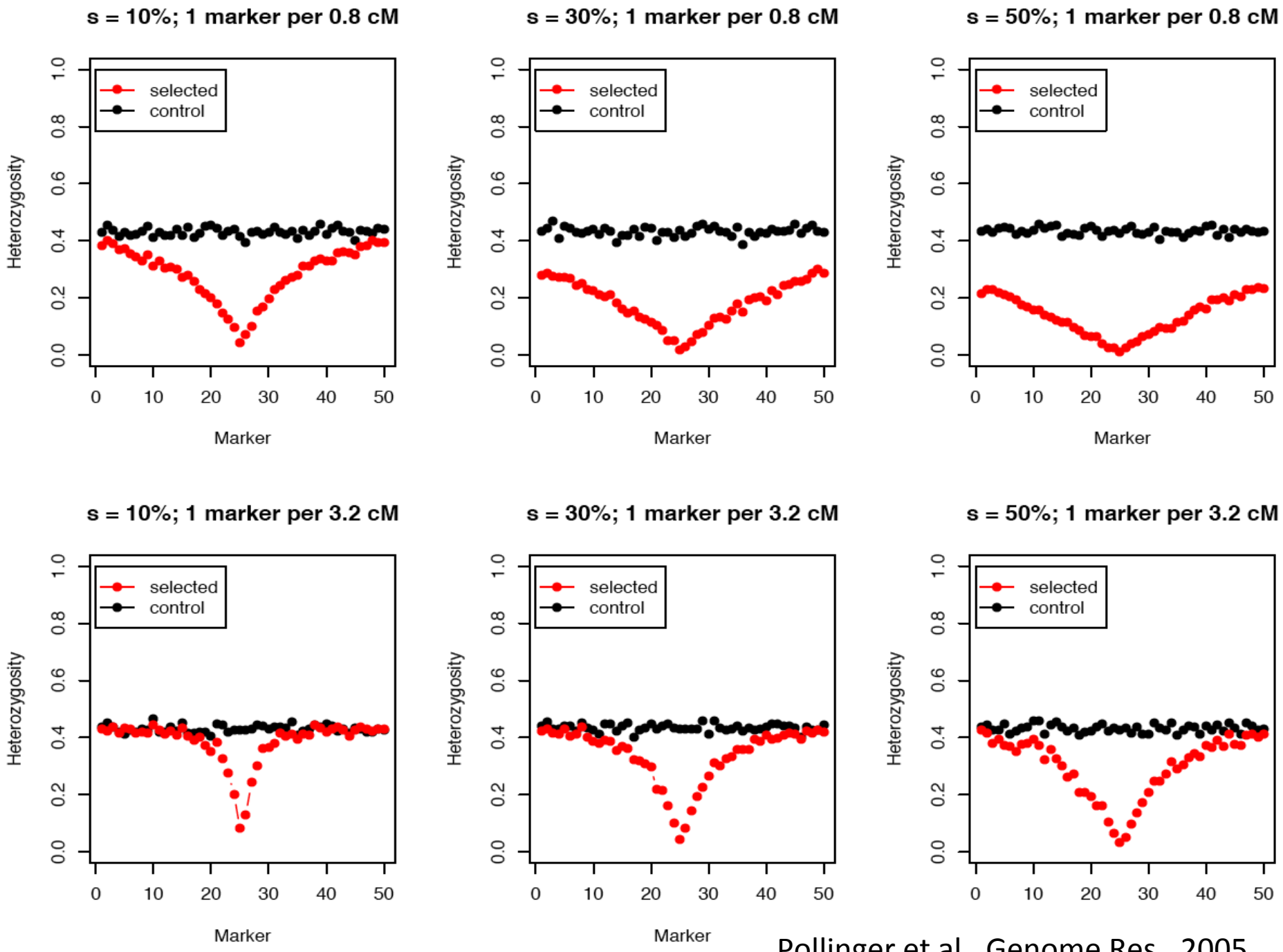


Selective Sweep Mapping



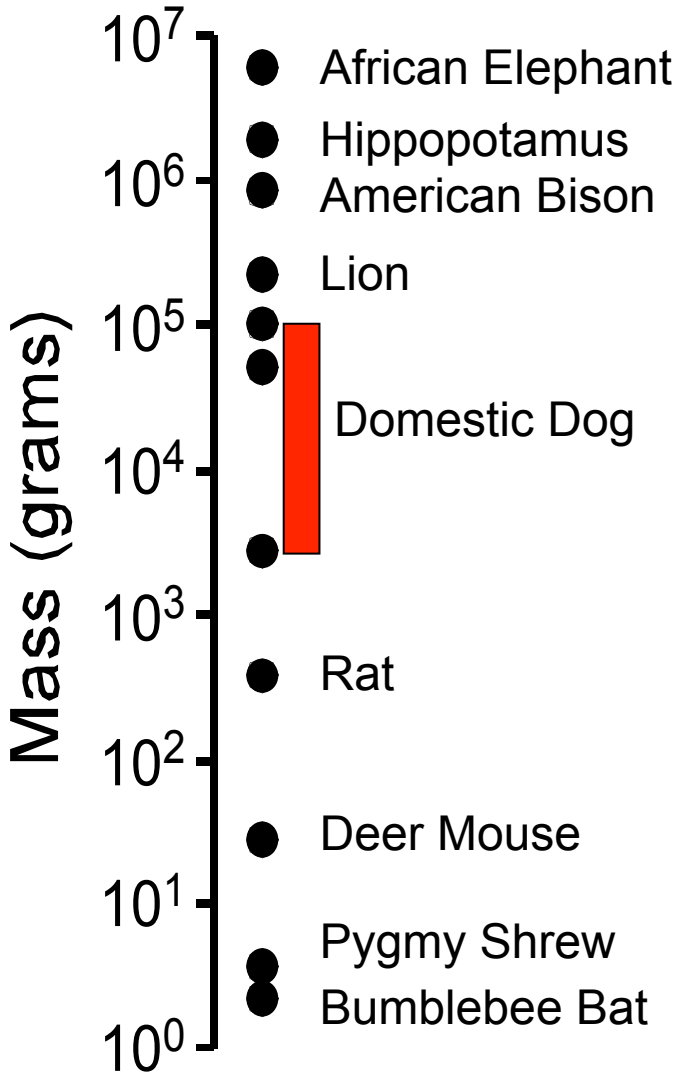






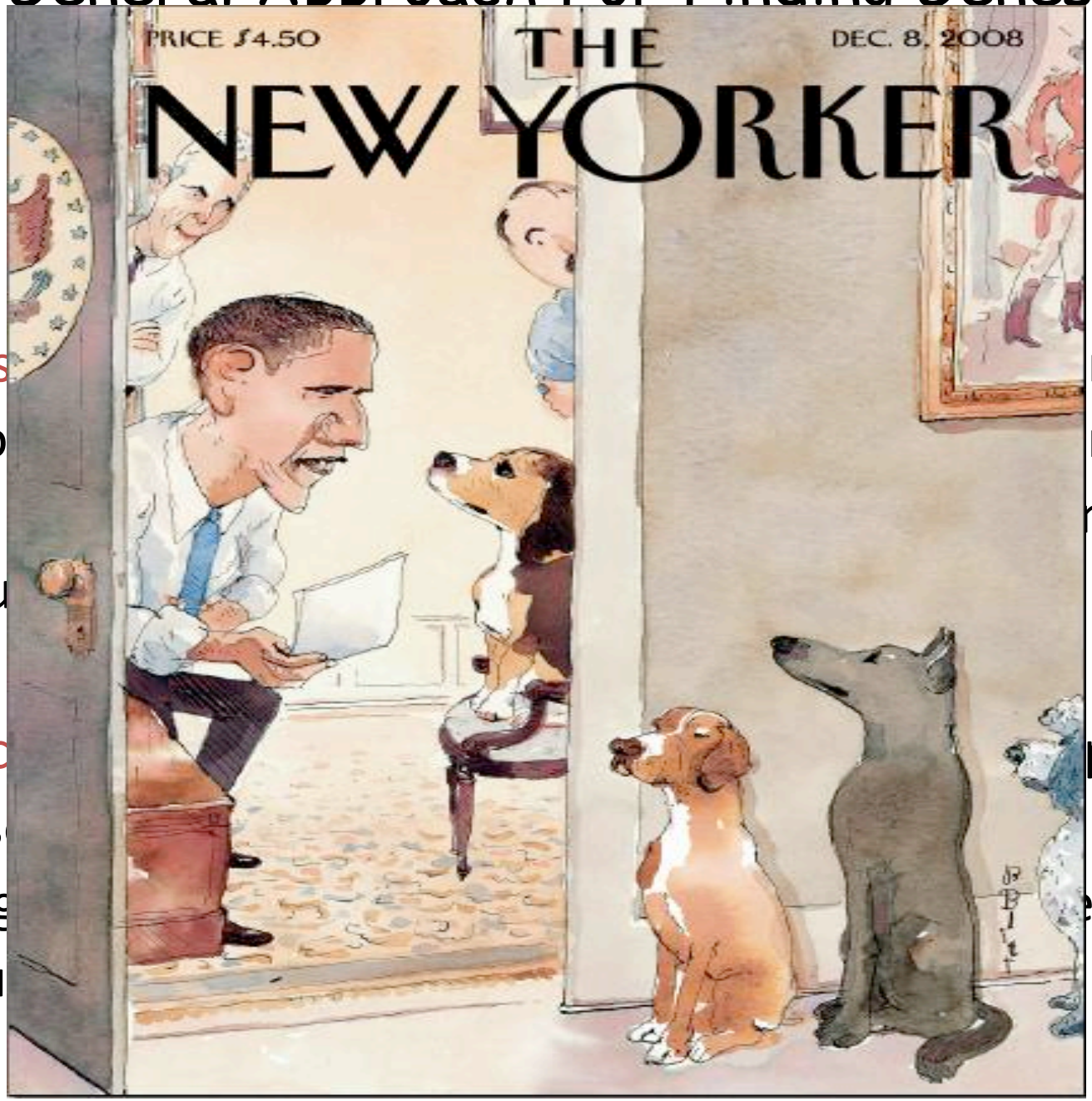
Pollinger et al., Genome Res., 2005

B. IGF1, the body size gene



Sutter et al., Science, 2007

A General Approach For Finding Genes Under



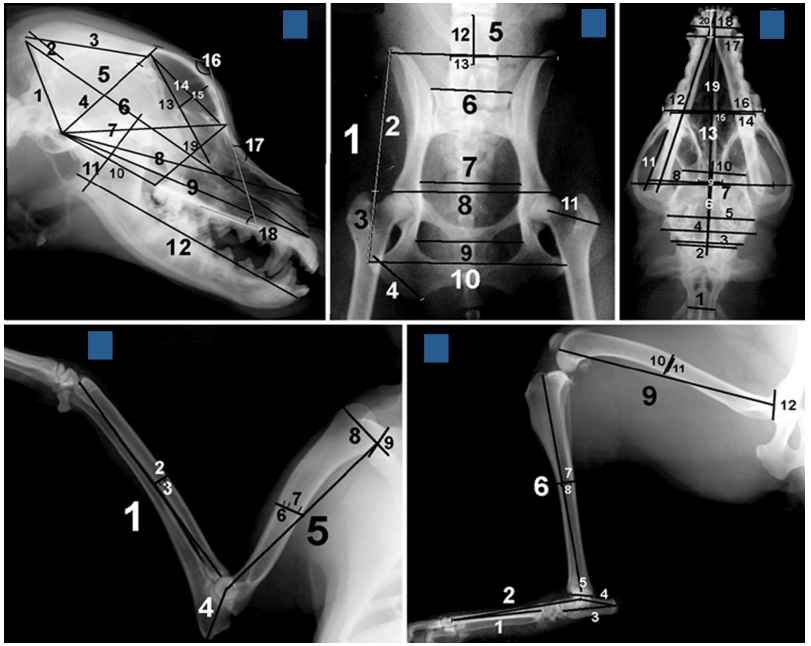
- Firs reed
(Po rphs to find
ma ns containing
cau

- Sec p and
ass and small
dog e causative
mu

Step 1: Genome-wide scan

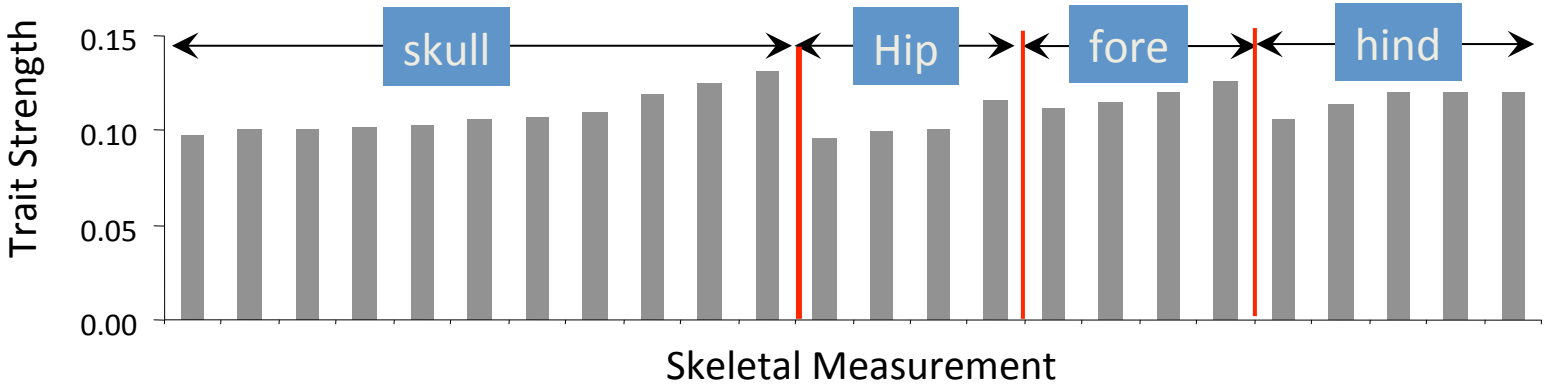
Genotype, Measure Skeletal Phenotypes, and Do Principal Component Analysis

- Focus on the Portuguese water dogs that have small and large morphs
- Genotype 500+ pedigreed dogs at 500 microsatellite markers
- Collect 91 skeletal measurements from x-ray images



- Do principal component analysis, which classifies phenotypic variation into independent systems of correlated traits
- Principal components are phenotypes so we can search for markers associated with them

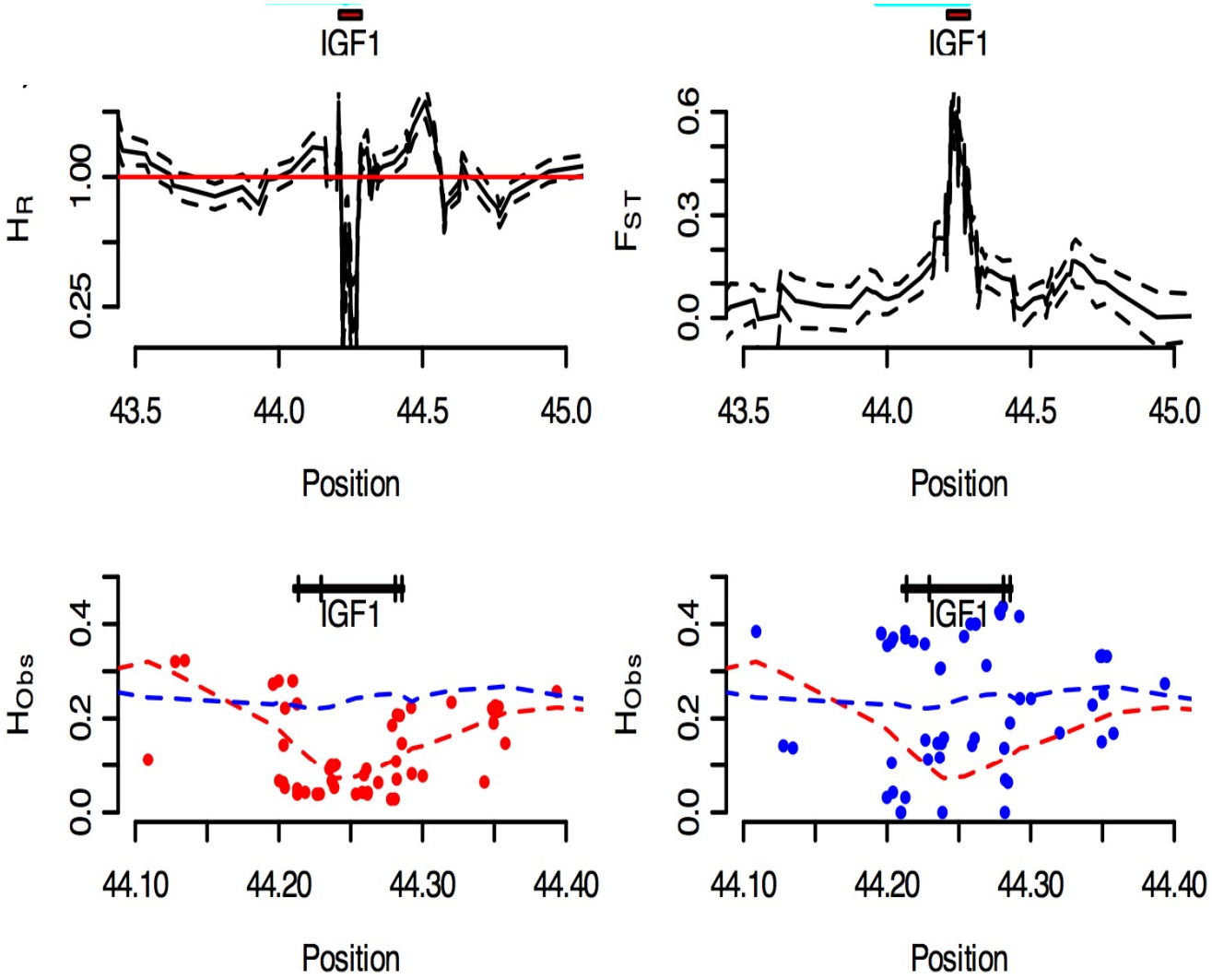
Principal Component 1 (Skeletal Size)



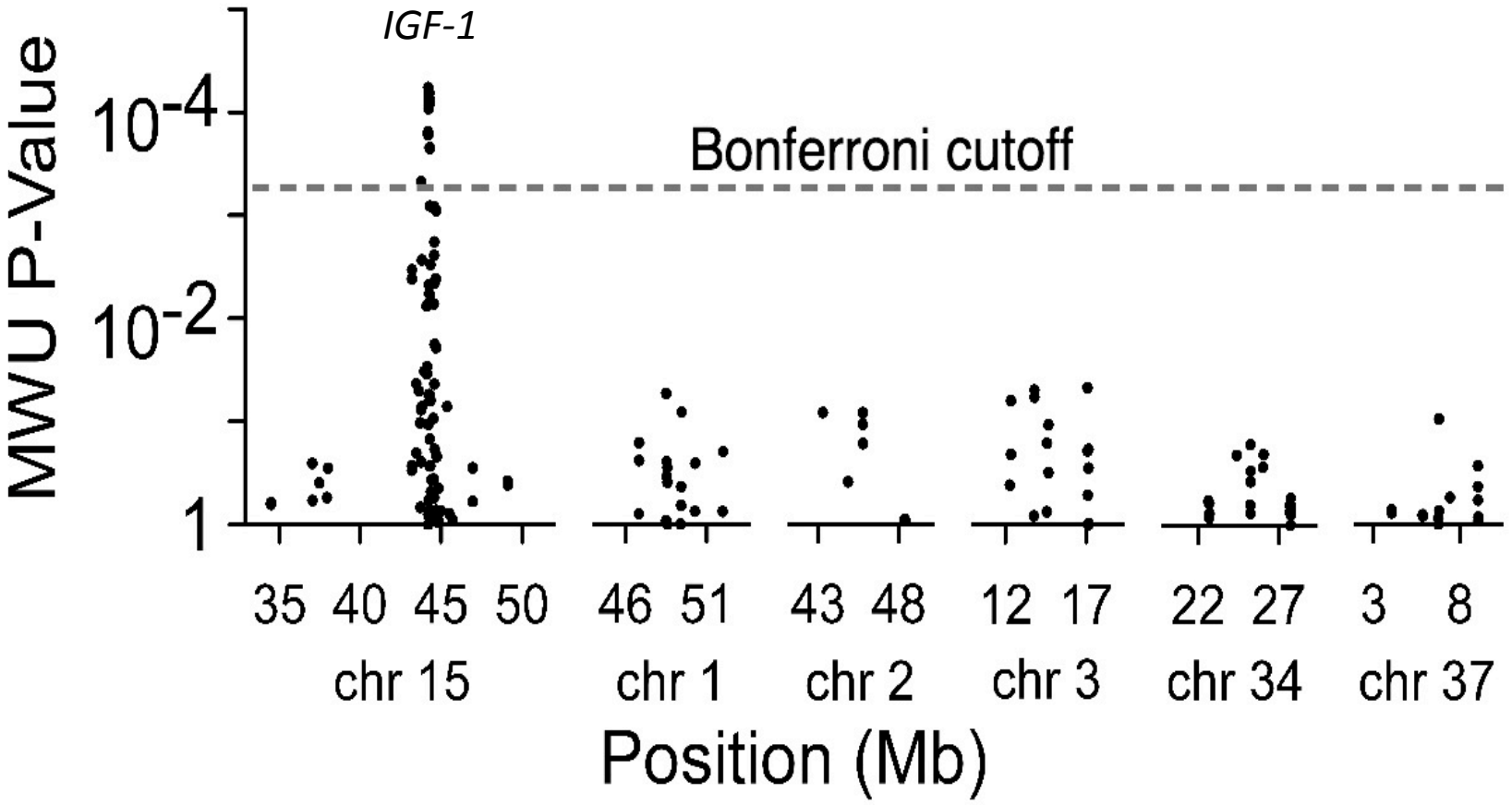
- PC1 is based on 42 metrics from pelvis, fore and hind limbs
- Highly heritable (h^2 0.45)
- 7 QTLs affect size
- A locus of interest: **FH2017 on CFA15 is linked to *IGF1!***
- Extensive re-sequencing near *IGF1* led to SNP discovery

Step 2: Selective and association mapping

(kg)	Breed
2	Chihuahua
2	Toy Fox Terrier
2	Pomeranian
3	Yorkshire Terrier
3	Japanese Chin
3	Chinese Crested
4	Italian Greyhound
4	Pekingese
5	Shih Tzu
6	Cav. King Charles Spar
6	Border Terrier
7	Miniature Schnauzer
7	Jack Russell Terrier
8	Boston Terrier
32	Giant Schnauzer
44	Akita
45	Bernese Mountain Dog
49	Great Pyrenees
54	Bullmastiff
54	Irish Wolfhound
59	Saint Bernard
	Great Dane
	Mastiff
<9	All Small Dogs
>31	All Giant Dogs



Step 2: Association mapping

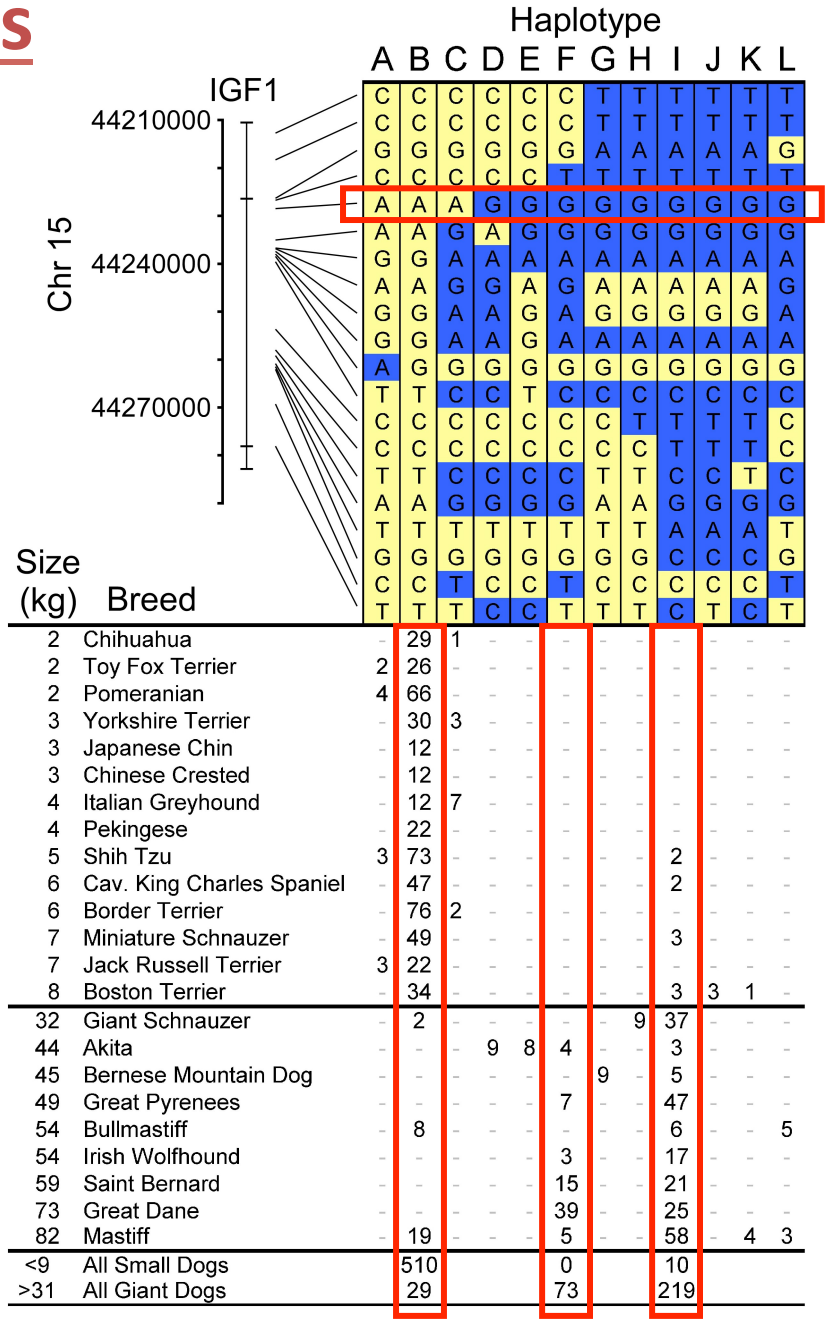


Step 2: Haplotype analysis

A single unique haplotype is associated with small size

- …Suggests a single ancient mutation for small size

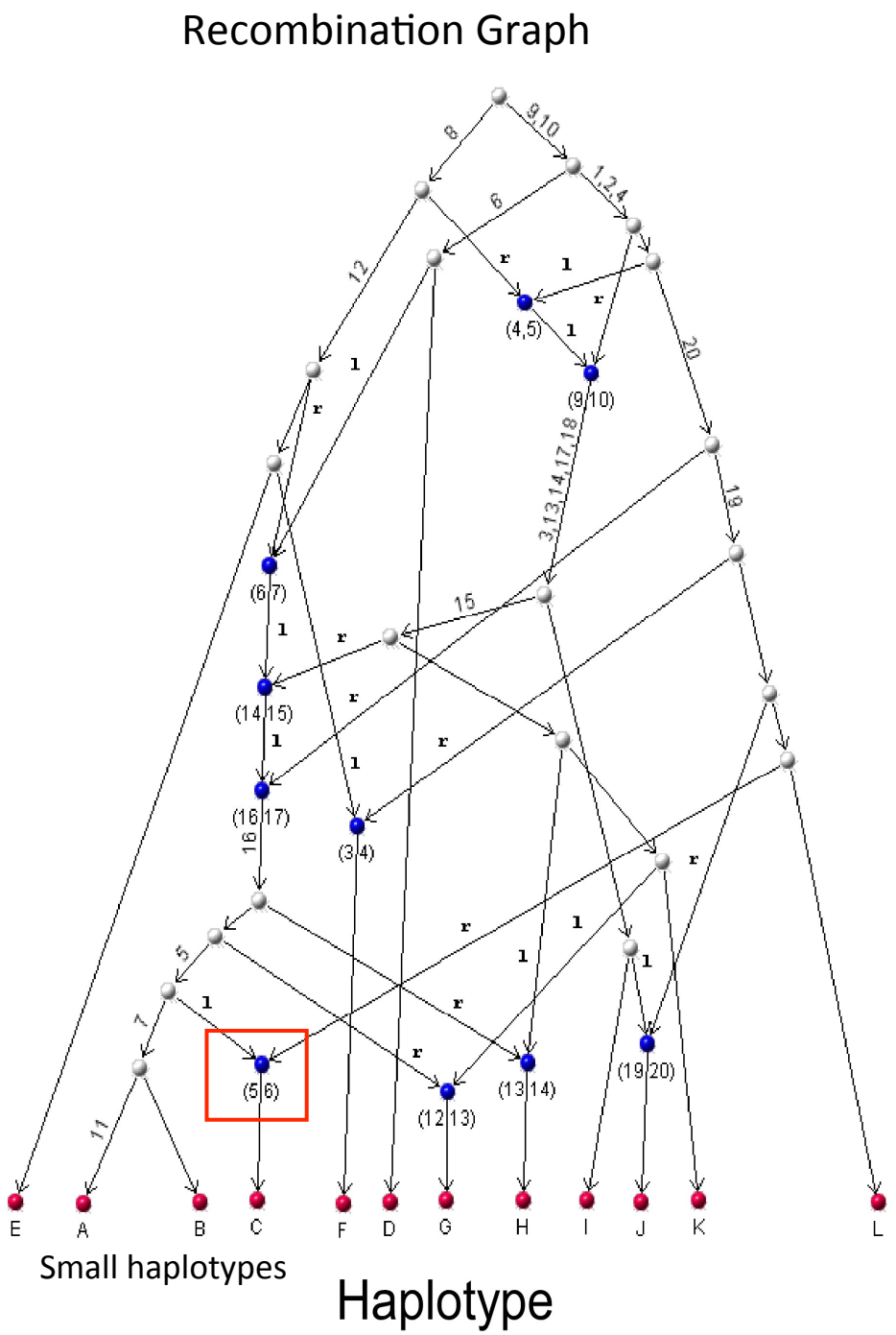
- Giant breeds carry two haplotypes rare in small breeds



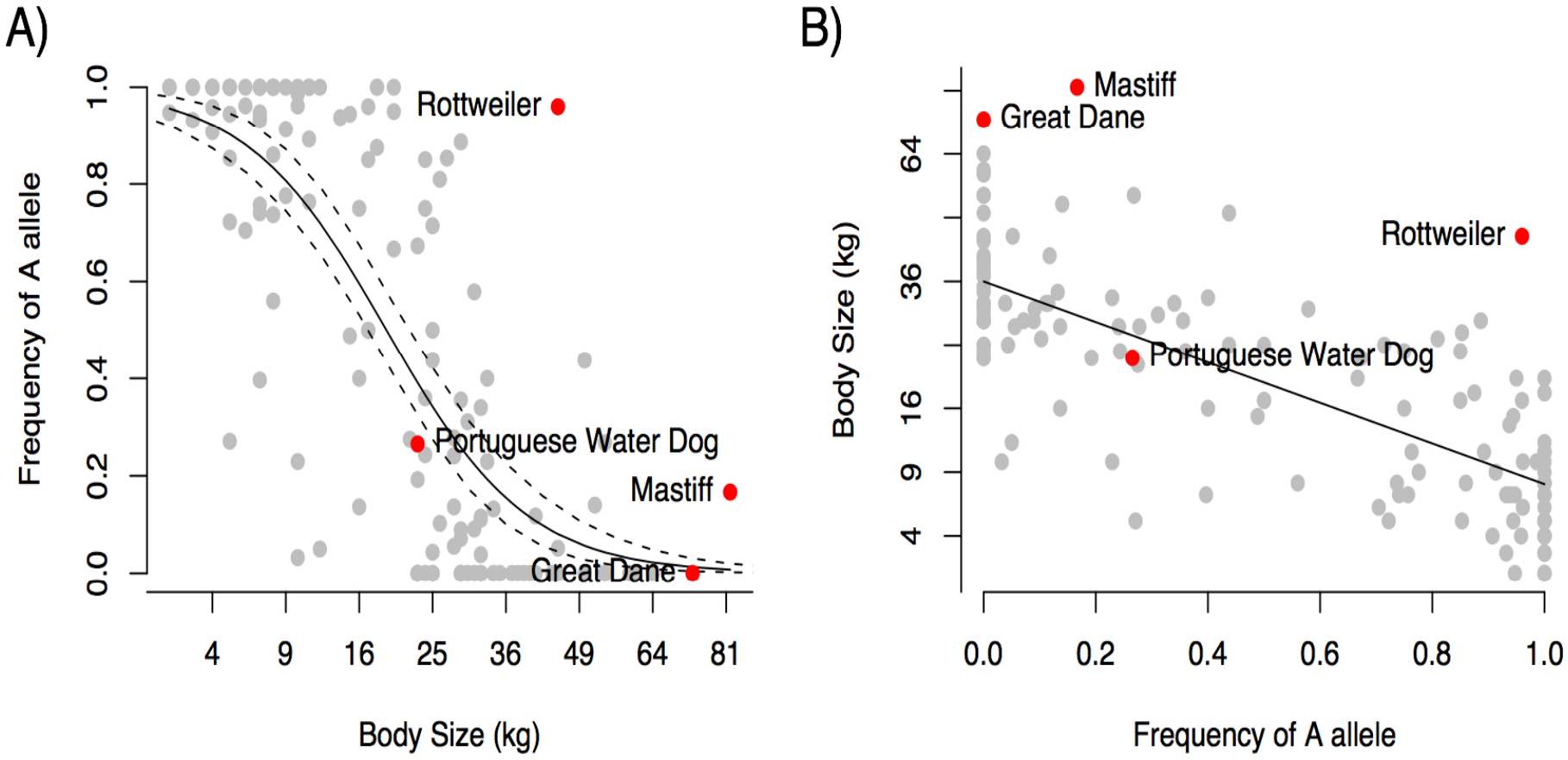
Step 3:
Association mapping

Core region of 8.7 kb contains only two possible causative mutations

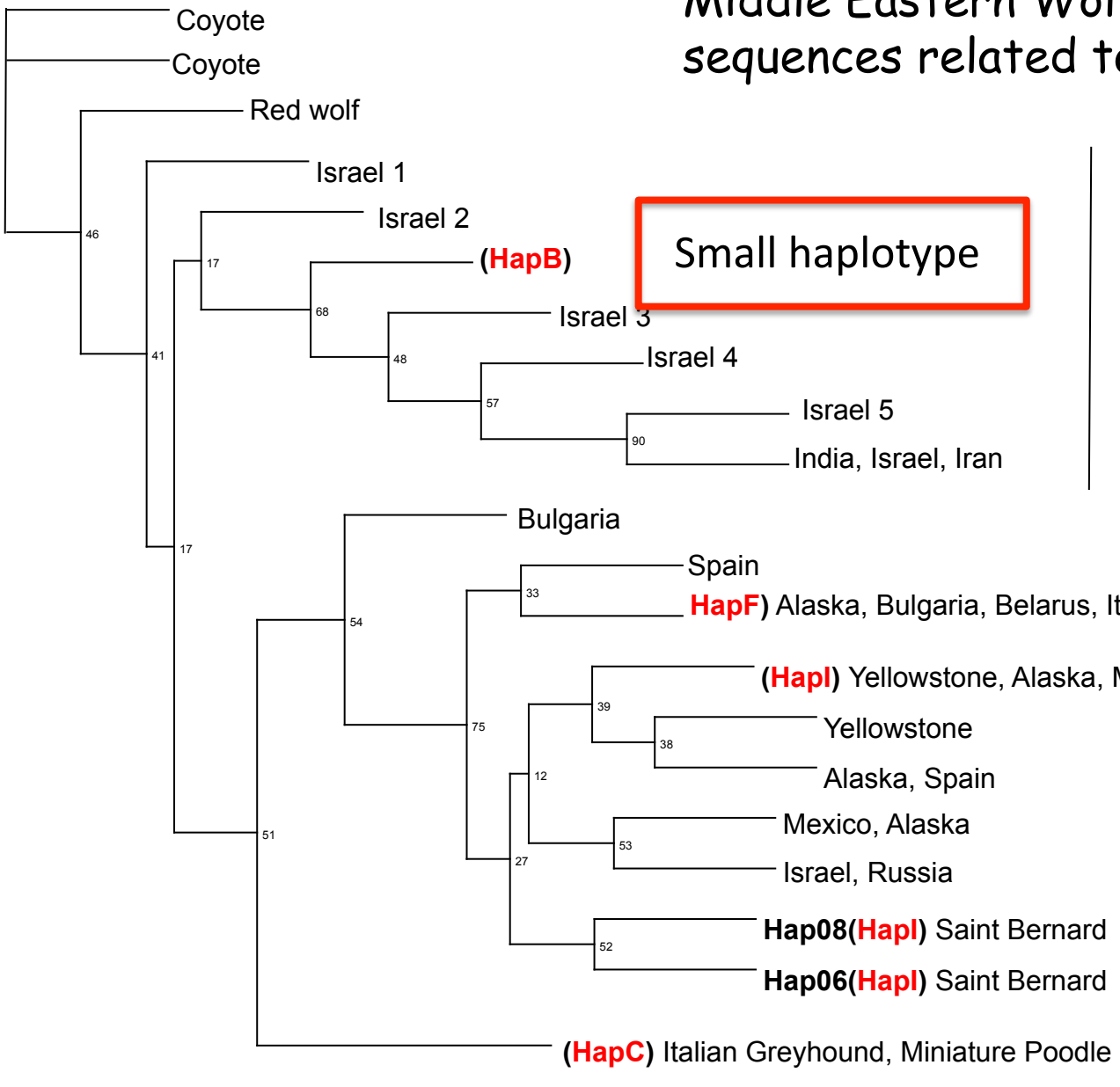
1. SINE element
2. CA_n mutation in *IGF1* promoter



IGF-1 explains more than 50% of the variation across 143 breeds

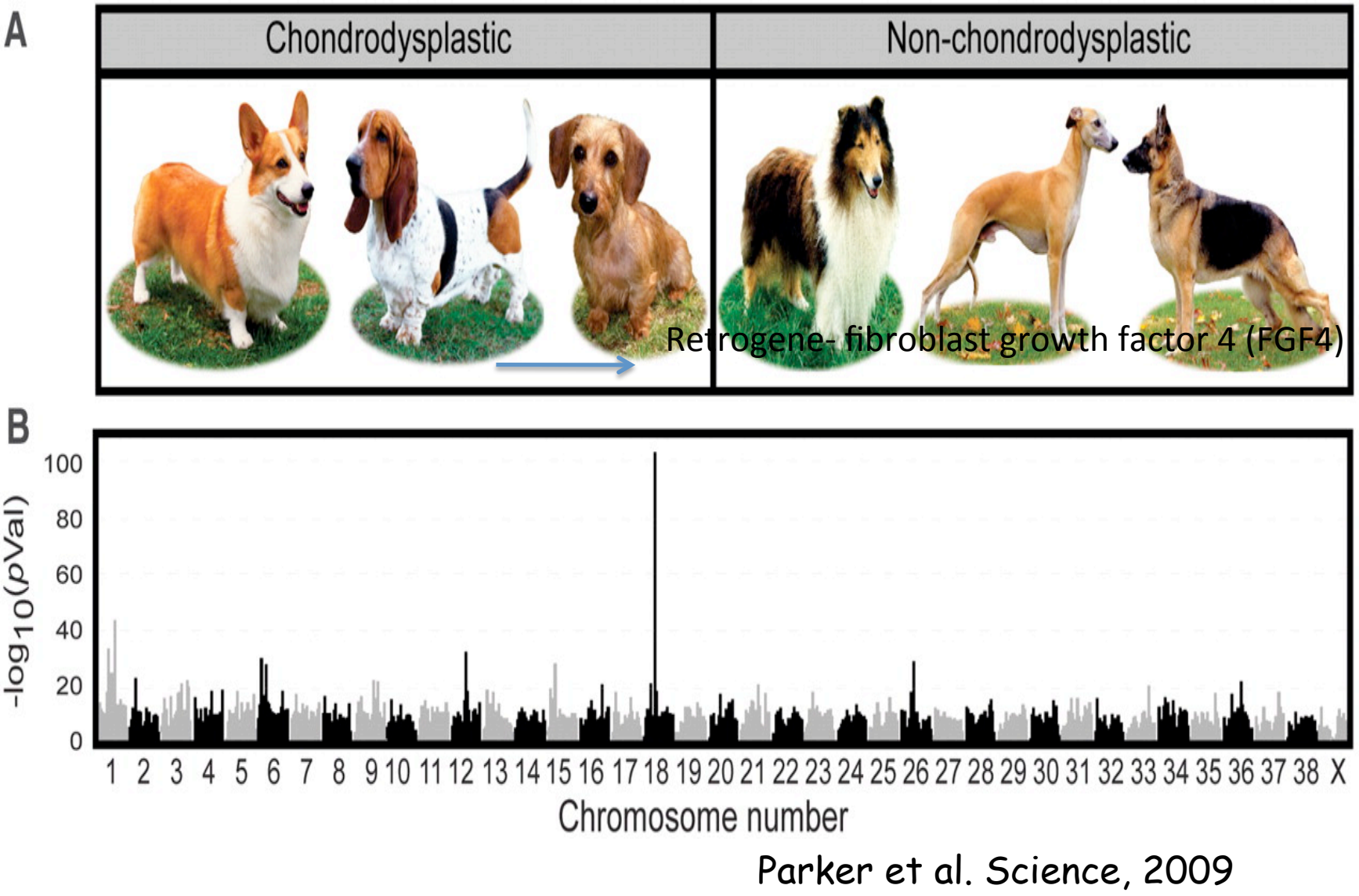


Middle Eastern Wolves have sequences related to the small allele



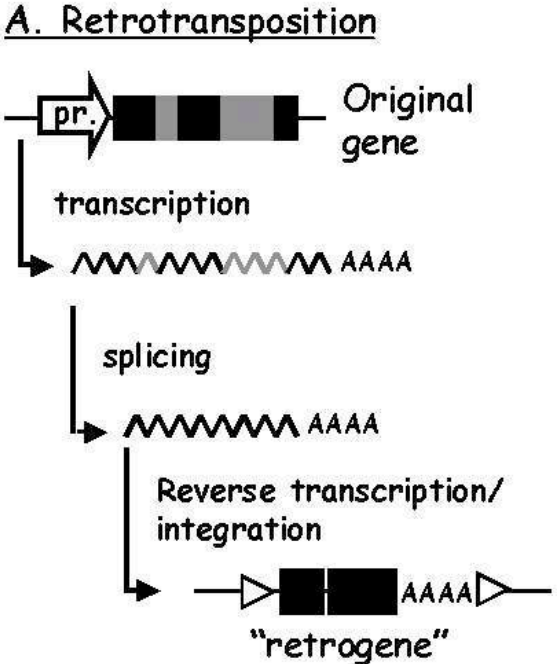
Gray et al., 2010

c. The short legs gene: association mapping



Parker et al. Science, 2009

Dogs with foreshortened limbs are genetically engineering by nature!



Pseudogene features

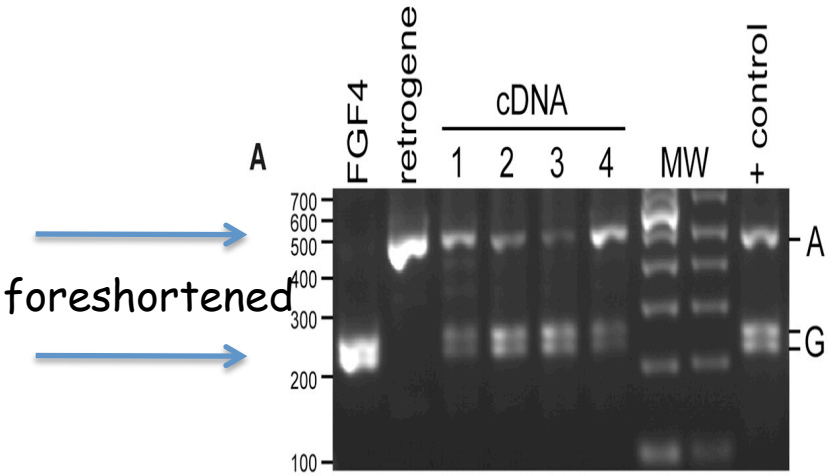
-Absence of introns and promoter sequences

- flanked by direct repeats

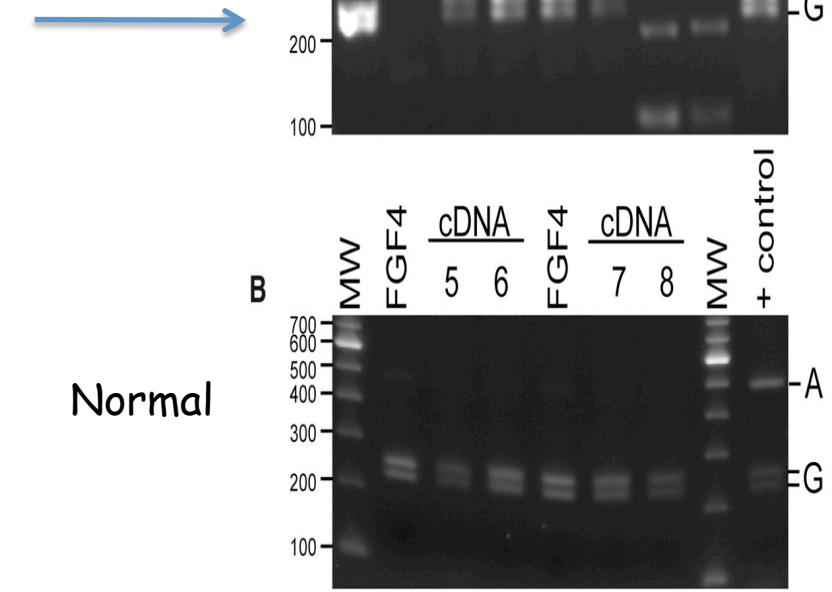
- presence of poly-A tract

- randomly integrated anywhere in the genome

Dogs with foreshortened limbs expressive retrogene and normal genes

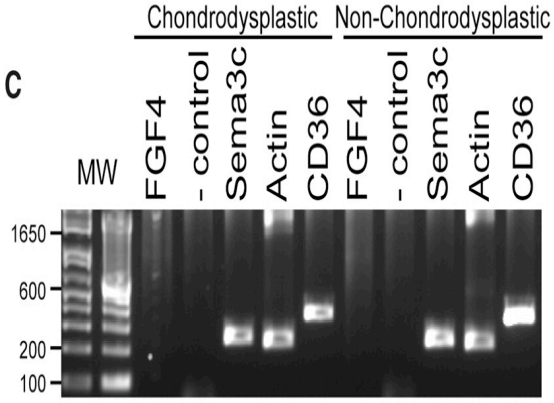


foreshortened



Normal

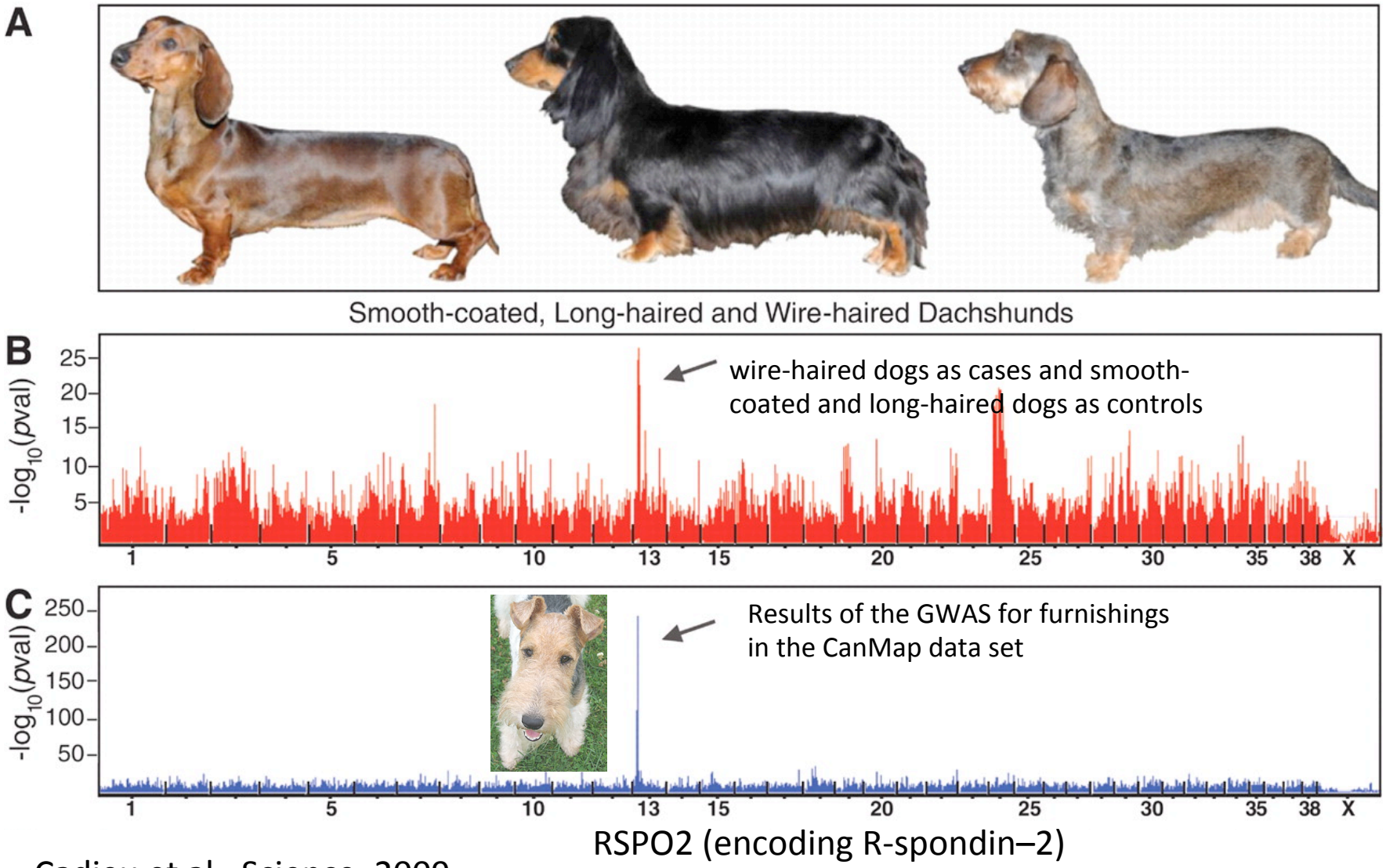
Juveniles



Adults

Parker et al. Science, 2009

d. Fur Types



RSPO2 (encoding R-spondin–2)

Cadieu et al., Science, 2009

Three mutations in various combinations explain the observed pelage phenotype of 95% of dogs sampled

	PHENOTYPE	*FGF5*	*RSPO2*	*KRT71*
A	Short	-	-	-
B	Wire	-	+	-
C	Wire and Curly	-	+	+
D	Long	+	-	-
E	Long with Furnishings	+	+	-
F	Curly	+	-	+
G	Curly with Furnishings	+	+	+

(encoding R-spondin–2, fibroblast growth factor–5, and keratin-71)

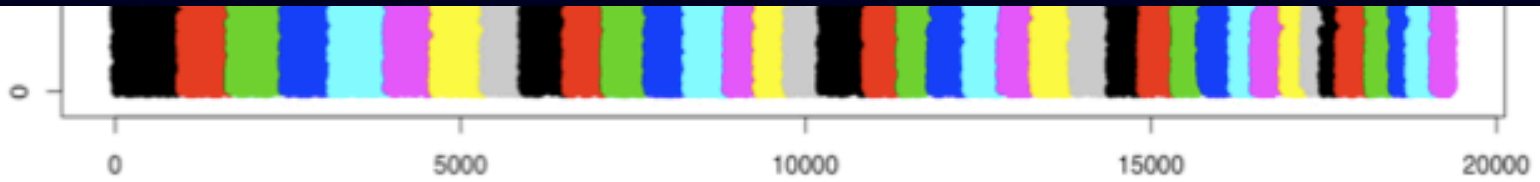
e. The number of genes

Table 1 | Estimates of heritability and number of loci for several complex traits

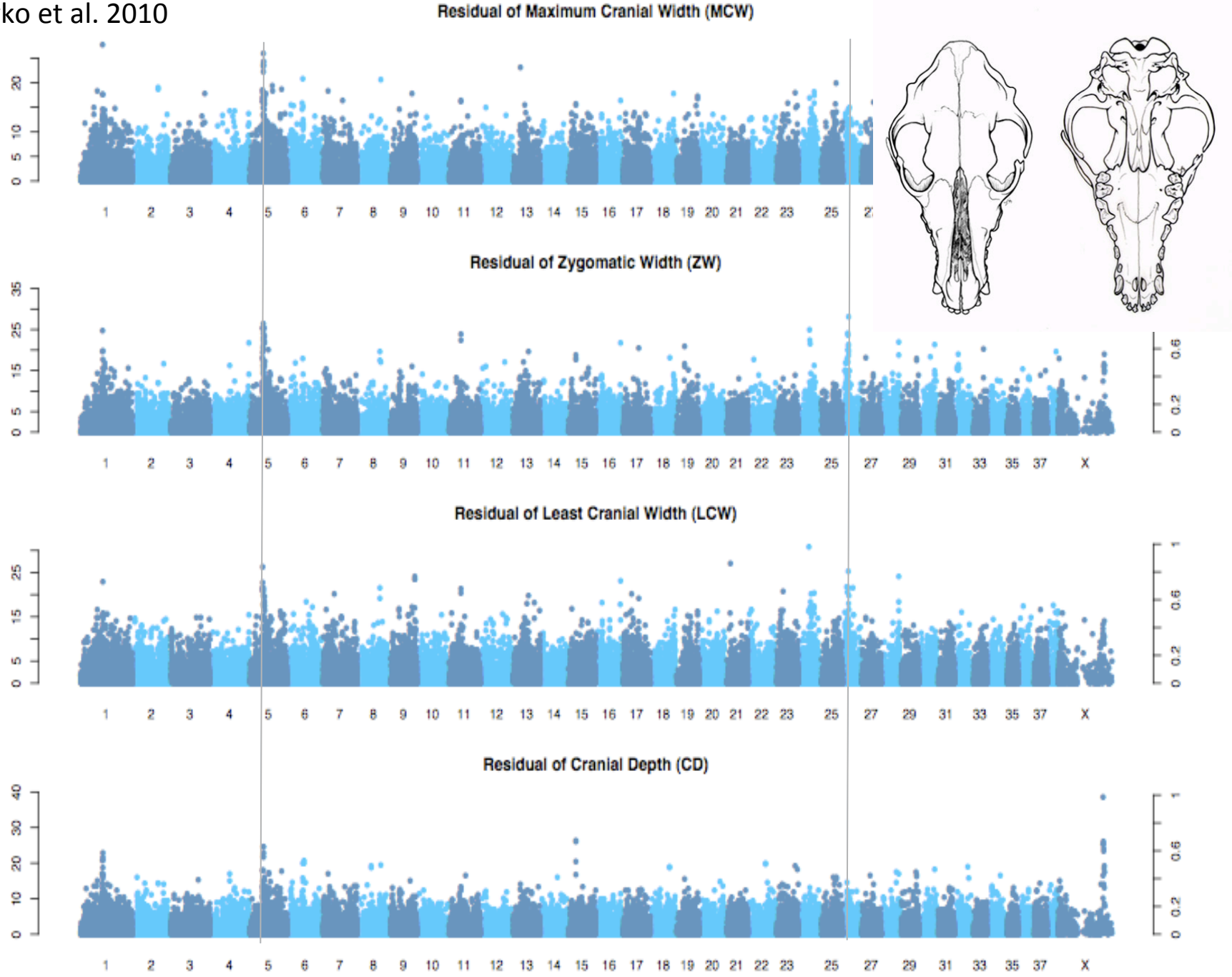
Disease	Number of loci	Proportion of heritability explained
Age-related macular degeneration[72]	5	50%
Crohn's disease[21]	32	20%
Systemic lupus erythematosus[73]	6	15%
Type 2 diabetes[74]	18	6%
HDL cholesterol[75]	7	5.2%
Height[15]	40	5%
Early onset myocardial infarction[76]	9	2.8%
Fasting glucose[77]	4	1.5%

* Residual is after adjustment for age, gender, diabetes.

Manolio (2009)
Nature 46: 747

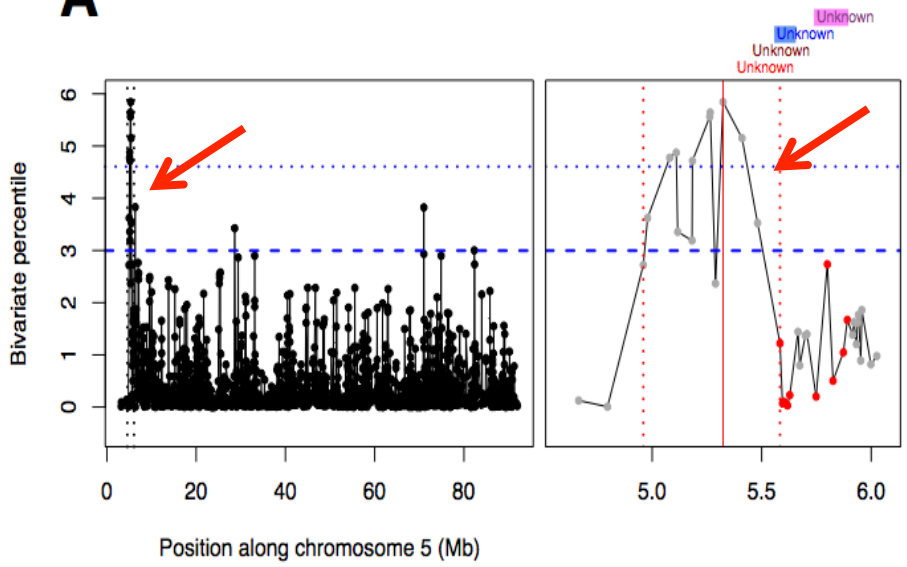


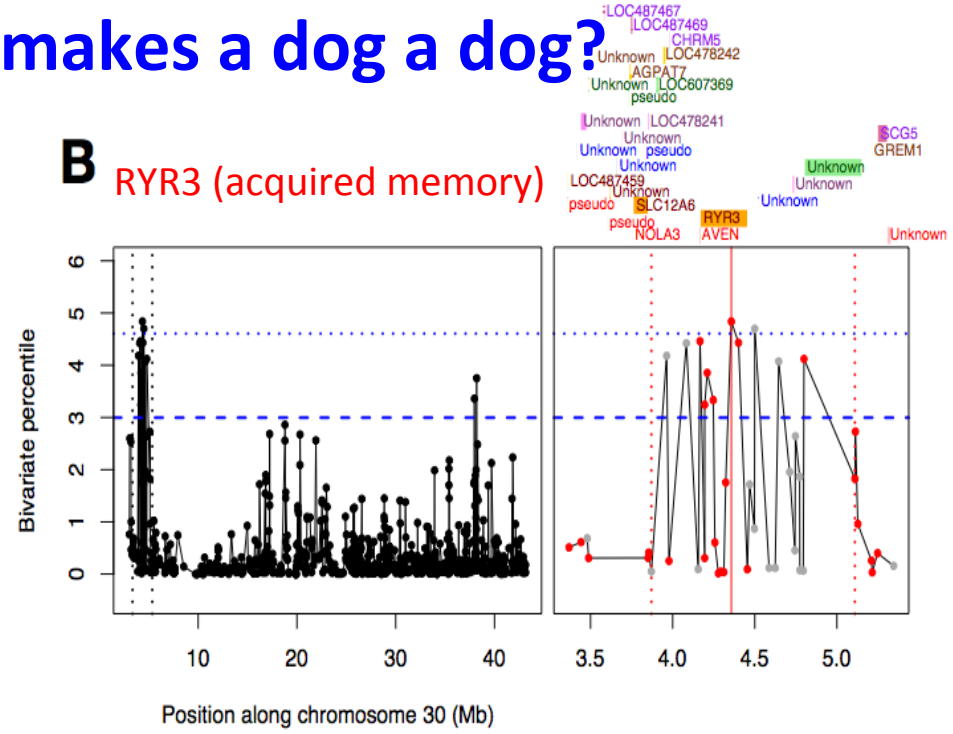
(Boyko et al. 2010

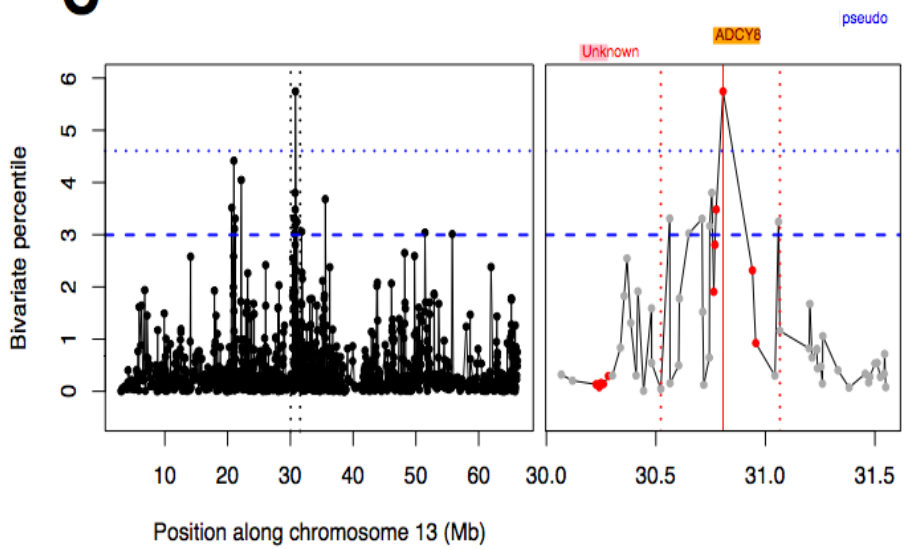


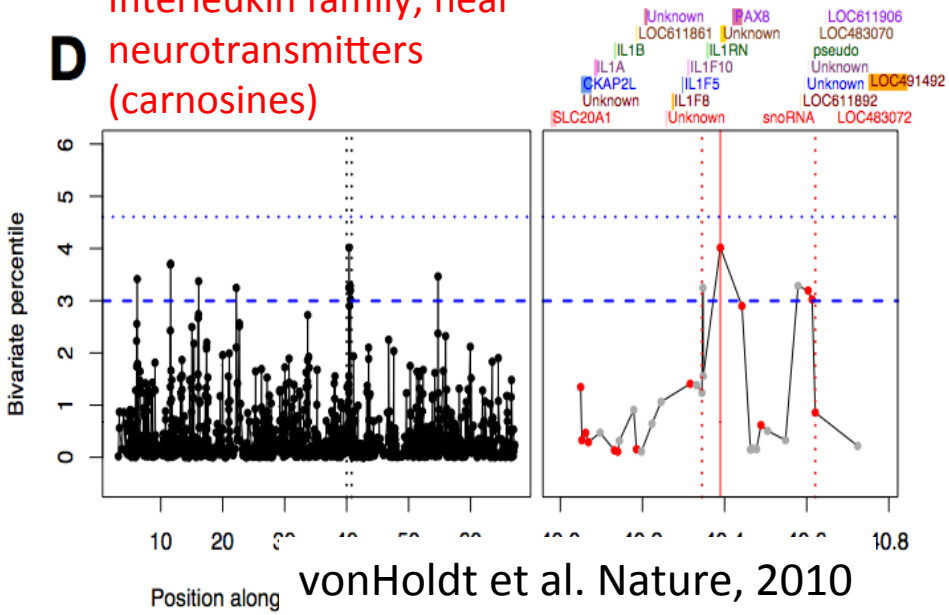
What makes a dog a dog?

AA similarity to neurotrimin an an opioid receptor (involved in the stress response process)









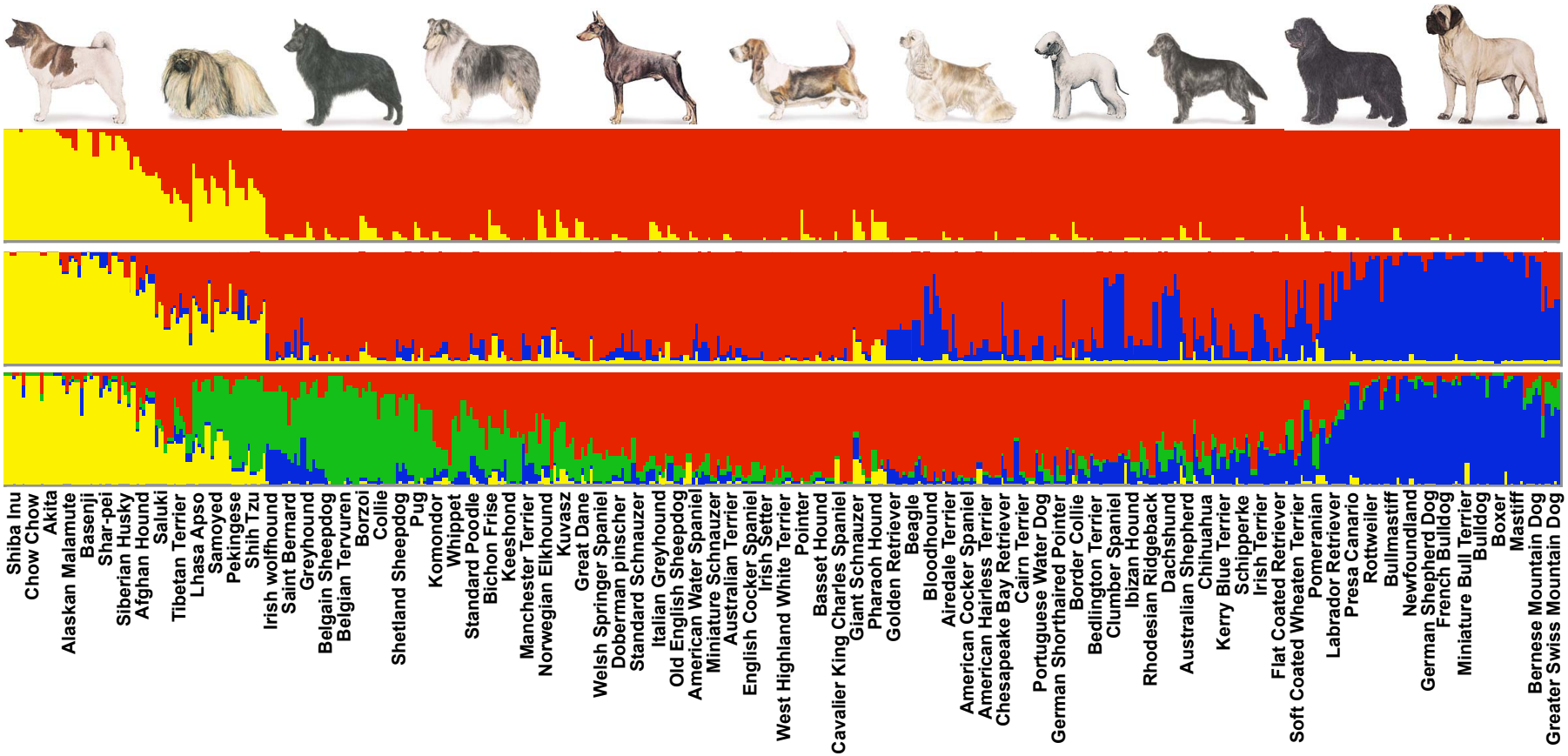
vonHoldt et al. Nature, 2010

Alluring story-telling II...

- WBSCR17: Williams syndrome (WS; also Williams-Beuren syndrome or WBS) is a rare neurodevelopmental disorder caused by a deletion of about 26 genes from the long arm of chromosome 7. It is characterized by a distinctive, "elfin" facial appearance, along with a low nasal bridge; an unusually cheerful demeanor and ease with strangers, coupled with unpredictably occurring negative outbursts; a predisposition to violent outbursts; mental retardation coupled with unusual (for persons who are diagnosed as mentally retarded) language skills; a love for music; and cardiovascular problems, such as supravalvular aortic stenosis and transient hypercalcaemia.

vonHoldt et al., Nature, 2010

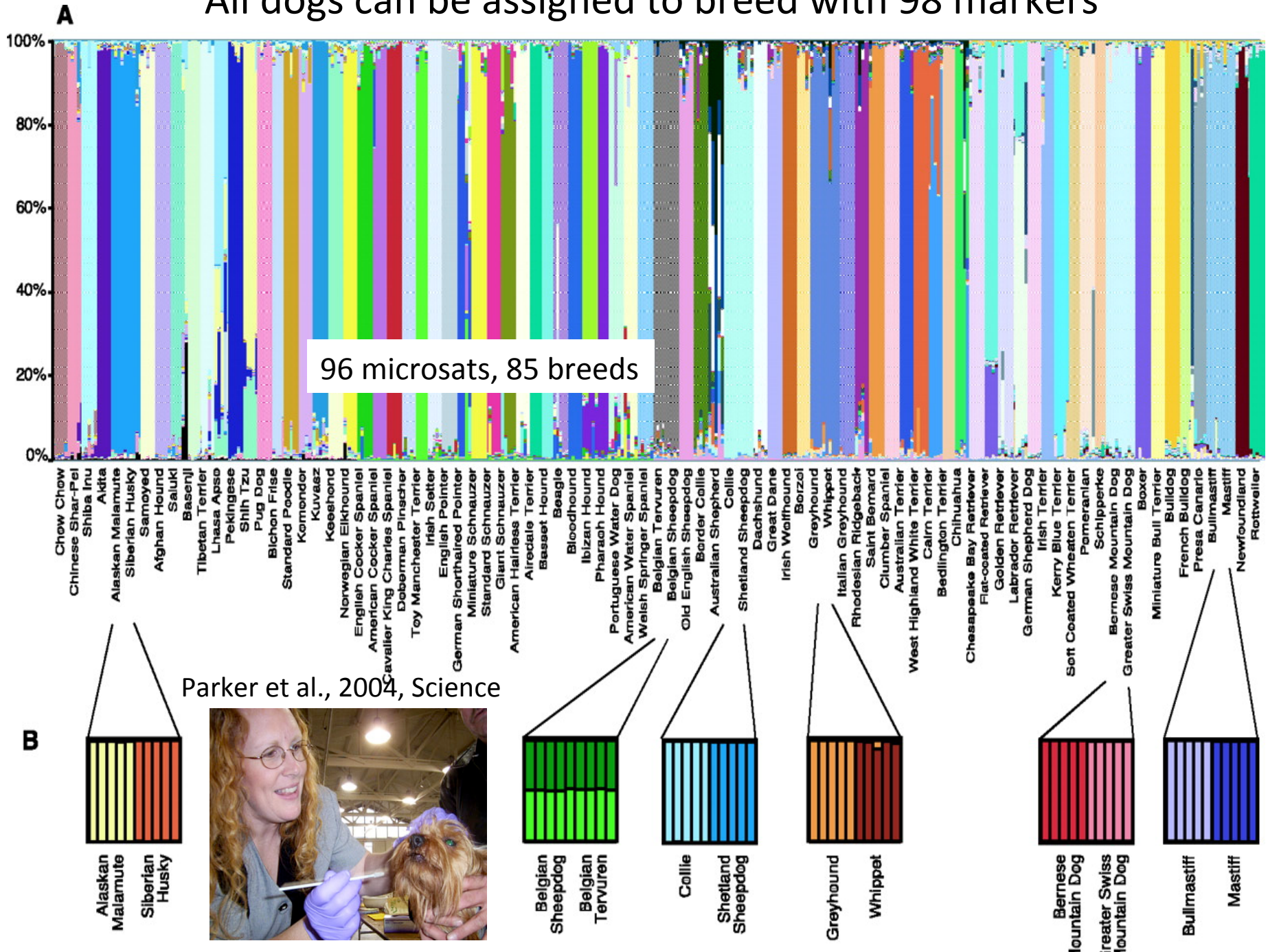
f. Evolving the Dog

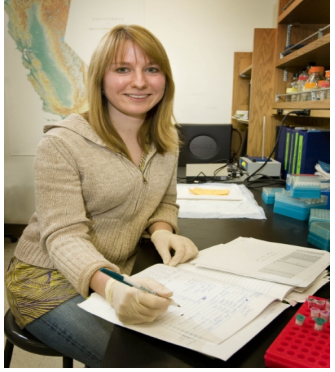


Population Structure of 85 Domestic Dog Breeds

Parker et al. Science, 2004

All dogs can be assigned to breed with 98 markers





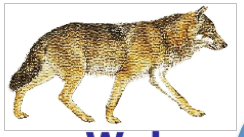
d) The Evolutionary framework

Ancient & Spitz breeds

Toy dogs

Spaniels

C



Wolves

Scent hounds

Working dogs

Sight hounds

Herding dogs

Retrievers

Mastiff-like dogs

Small terriers

Ancient & Spitz breeds

vonHoldt et al., Nature, 2010

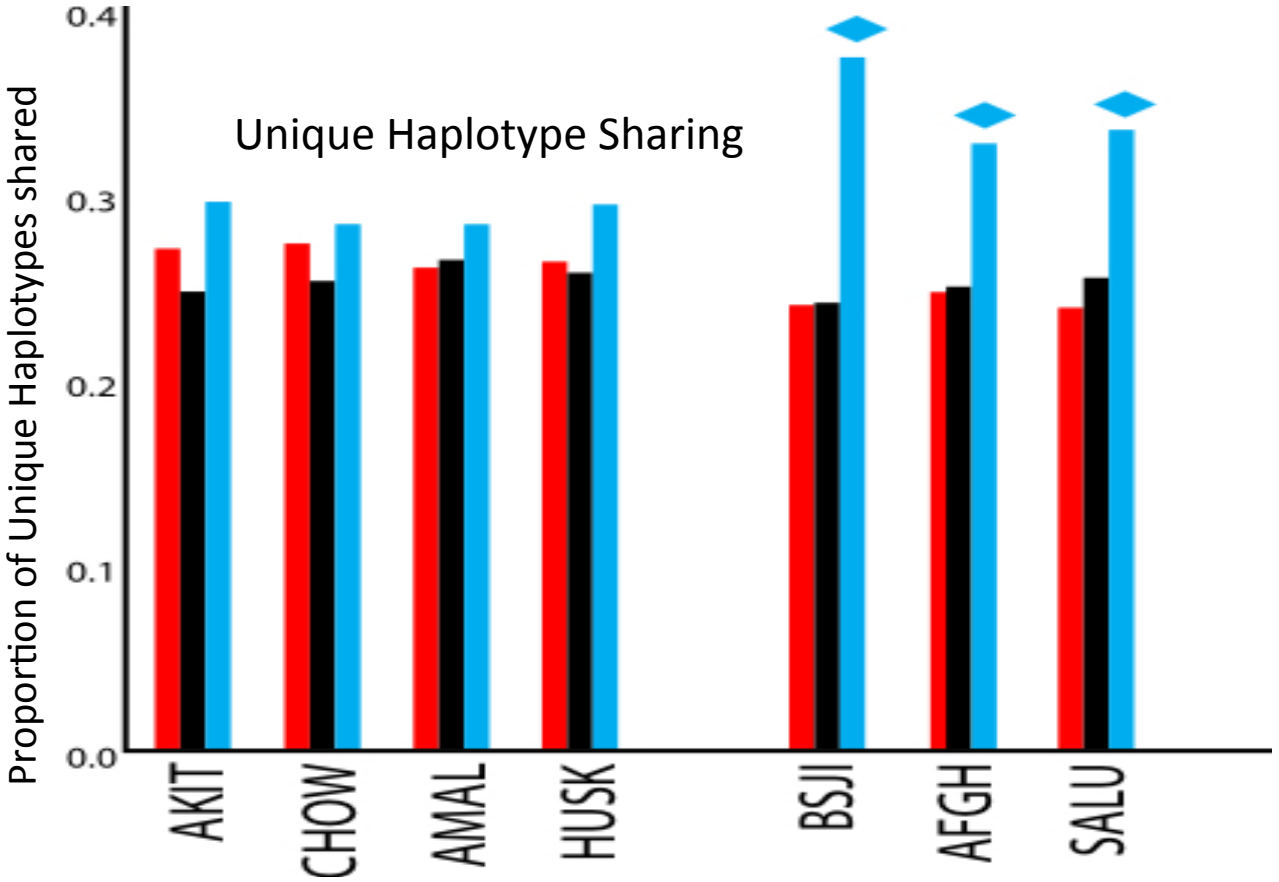
Supplemental Table 4. History of breeds with discordant phenotypic/functional and genetic group assignments[2,3].

Breed Name	Genetic Cluster (see Fig. 1)	Phenotypic/ Functional Group	Breed Information[2,3]	Concordance with Historical Evidence
Briard	Small Terrier	Herding	Possible East Asian origin from crosses with local dogs to create a new breed used for flock guarding	No historical evidence for breed admixture between small terriers and herding dogs
Brussels	Toy	Terriers	European origins from crosses with Affenpinscher (terrier) and Toy breeds (i.e. English Toy Spaniels, Yorkshire Terriers, Pekingese, or Pug) to miniaturize the breed	Evidence for breed admixture between toy and terrier breeds
Chihuahua	Toy	Ancient	Probable Chinese origins with introduction to Mexico from Spanish traders returning from East Asia	Evidence for breed admixture between East Asian Ancient and toy breeds
German Shep. Dog	Gun	Herding	European breed with recent origins	Inconclusive
Gt. Schnauzer	Gun	Herding	European origins likely from crosses with smooth-haired dogs and possibly Great Danes	Inconclusive
Glen of Imaal	Mastiff-like	Terriers	European origins from crosses of Bullterriers, Staffordshire terriers (Mastiff-like breeds) and other fighting dogs; Glen of Imaal is an aggressive hunter (e.g. badgers, rats)	Evidence for admixture between Mastiff-like and terrier breeds
Mini. Pin.	Toy	Terriers	European origins from crosses of German Pinscher (terrier) and Dachshunds or Italian greyhounds	Evidence for admixture between toy and terrier breeds
Newfoundland	Retrievers	Mastiff-like	North American origins with possible crosses to Mastiff or Portuguese Water dog; considered an ancestor of the modern Labrador Retriever	Evidence for Retriever and Mastiff-like breed admixture
Papillon	Toy	Spaniels	European origins from crosses of Spaniels and Bichon-type (toy) breeds	Evidence for admixture of toy and spaniel breeds
Pekingnese	Toy	Herding	Chinese origins; considered a dwarfed Tibetan terrier or Pug (toy)	Evidence of admixture of toy and other breeds
Pomeranian	Toy	Spitz	European origins from crossing European herding and spitz-type breeds	Inconclusive
Portuguese Water Dog	Gun	Spaniels	European origins; bred to be a water dog	Inconclusive
Pug	Toy	Mastiff-like	Chinese origins; considered a “mini-mastiff”, likely from miniaturizing the Affenpinscher (Terrier) or the English Bulldog and crossing with the Tibetan Mastiff (Mastiff-like breeds)	Evidence for breed admixture of Mastiff-like and toy breeds
Shih Tzu	Toy	Herding	Tibet/Chinese origins; considered a dwarf of Tibetan terriers or Lhasa Apsos (herding breeds)	Evidence for admixture of Toy and herding breeds
Std. Schnauzer	Gun	Herding	European origins from crossing the Standard Pinscher, Poodles, “Wolfspitzs", or Shepherds	Inconclusive

vonHoldt et al., Nature, 2010

Dog Origins Revisited
Haplotype tree (15 SNP windows)
Middle East
Ancient and Spitz dogs
Toy breeds
Spaniels
Scent hounds
Gun dogs
Mastiff-like breeds
Small terriers
Mastiff-like breeds
Retrievers
Herding dogs
Sight hounds
Func
BSJI AKIT CHOW DINGO SHAR AMAL HUSK AFGH SALU AESK SAMO POM CHIH PEKE SHIH BRUS PUG PAPI MPIN ACKR ECSP BRIT GSHP BASS BEAG BLDH PBGV DACH HAVA SPOO TPOO DOBP GSNZ SSNZ GSD PTWD BOST BOX BULD FBUL MBLT STAF GLEN BULM MAST JACK BRIA AUST YORK CAIR WHWT SCOT NOWT BMD STBD DANE ROTT FCRT GOLD LAB NEWF AUSS COLL SSHP BORD CARD PEMB OES BORZ DEER IWOF GREY WHIP ITGY IBIZ KUVZ

Dogs share more unique haplotypes with European and Middle Eastern Wolves

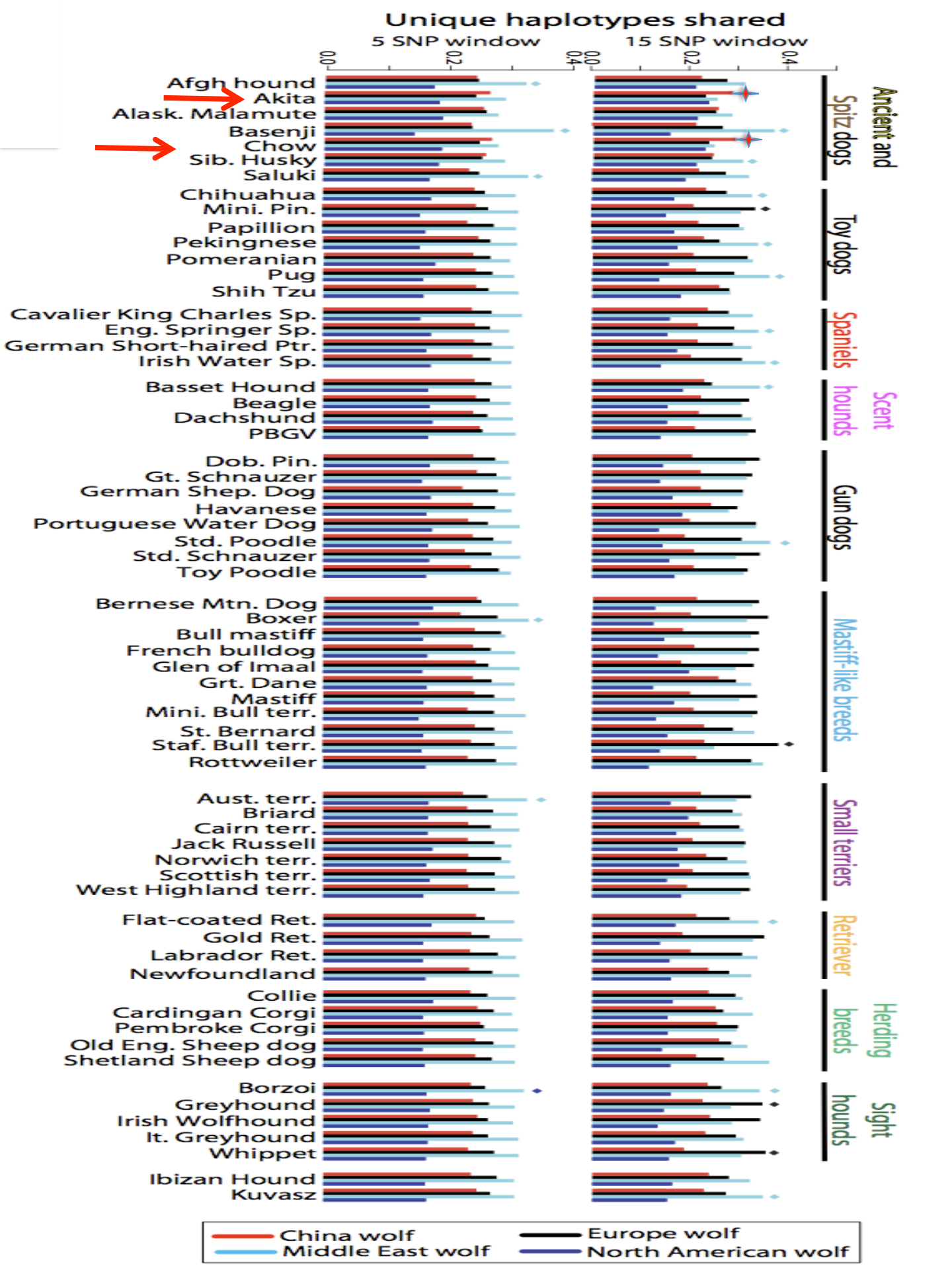


Kirk Loemueller
(Cornell/Stanford)

Shared with Chinese wolves

Shared with European wolves

Shared with Middle Eastern wolves


Unique haplotypes shared
5 SNP window
15 SNP window
0.0
0.2
0.4
Afgh hound
Akita
Alask. Malamute
Basenji
Chow
Sib. Husky
Saluki
Chihuahua
Mini. Pin.
Papillion
Pekingnese
Pomeranian
Pug
Shih Tzu
Cavalier King Charles Sp.
Eng. Springer Sp.
German Short-haired Ptr.
Irish Water Sp.
Basset Hound
Beagle
Dachshund
PBGV
Dob. Pin.
Gt. Schnauzer
German Shep. Dog
Havanese
Portuguese Water Dog
Std. Poodle
Std. Schnauzer
Toy Poodle
Bernese Mtn. Dog
Boxer
Bull mastiff
French bulldog
Glen of Imaal
Grt. Dane
Mastiff
Mini. Bull terr.
St. Bernard
Staf. Bull terr.
Rottweiler
Aust. terr.
Briard
Cairn terr.
Jack Russell
Norwich terr.
Scottish terr.
West Highland terr.
Flat-coated Ret.
Gold Ret.
Labrador Ret.
Newfoundland
Collie
Cardingan Corgi
Pembroke Corgi
Old Eng. Sheep dog
Shetland Sheep dog
Borzoi
Greyhound
Irish Wolfhound
It. Greyhound
Whippet
Ibizan Hound
Kuvasz
Ancient and
Spitz dogs
Toy dogs
Spaniels
Scent
hounds
Gun dogs
Mastiff-like breeds
Small terriers
Retriever
Herding
breeds
Sight
hounds
China wolf
Middle East wolf
Europe wolf
North American wolf

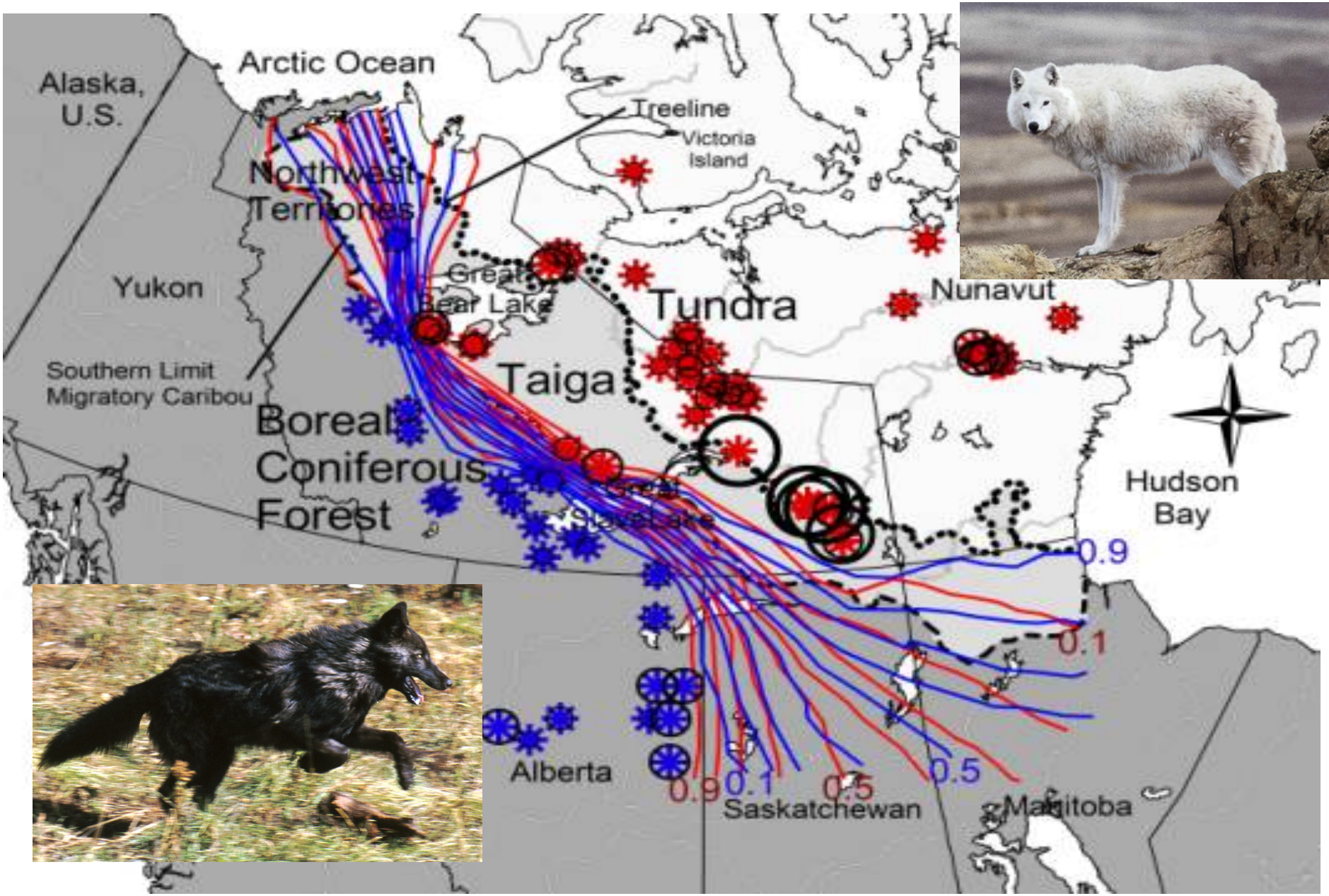
Questions

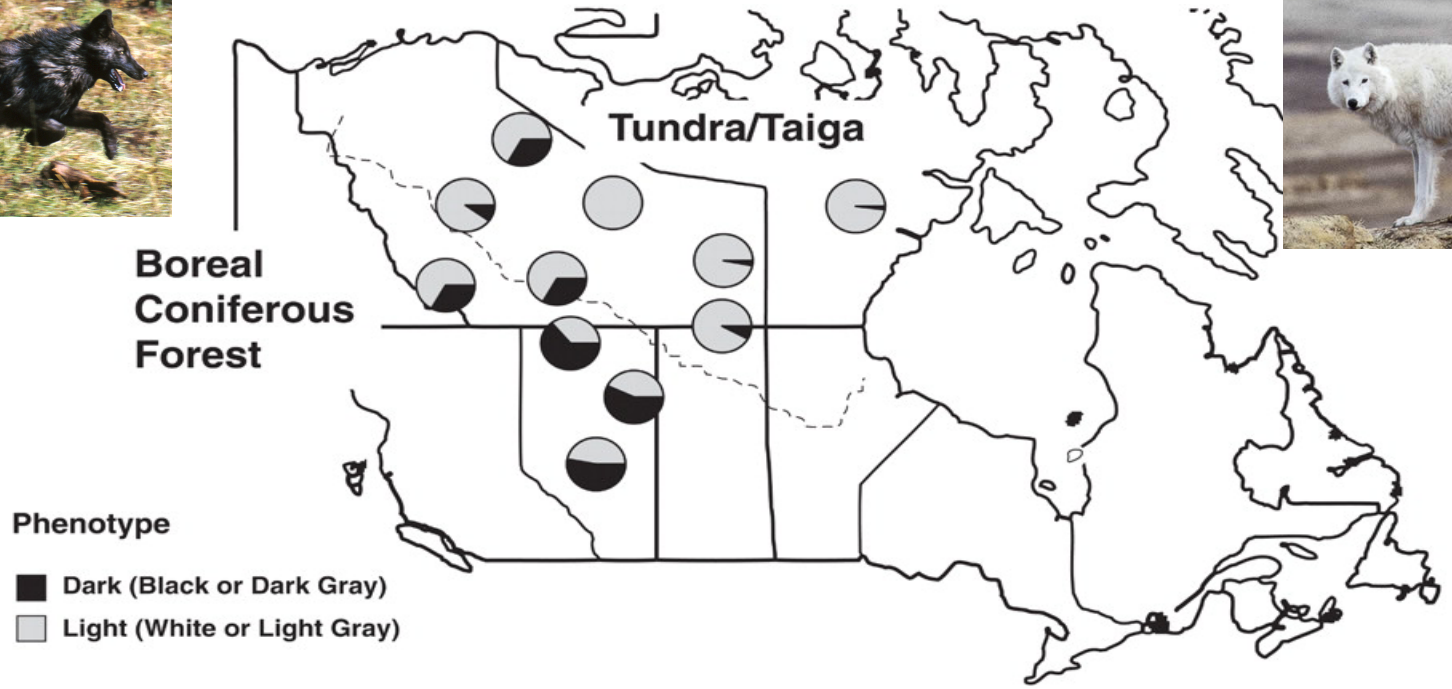
1) The primary reason body size differs among dogs breeds is:
   a. A single gene of large effect
   b. Multiple genes of small effect
   c. Nutritional differences
   d. Exercise differences
   e. None of the above

2) Unlike other domesticated species, dogs phenotypes are determined by
   a. Multiple genes of small effect
   b. A few genes of large effect
   c. One master gene
   d. Transposable elements

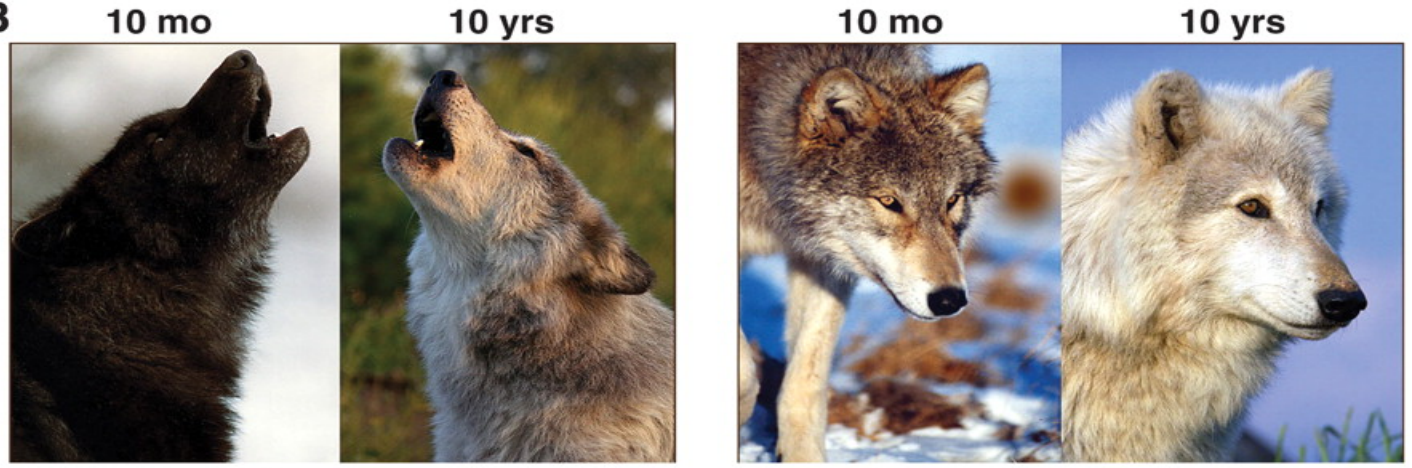
3) Diversity in dogs is limited in part by:
   a. Human creativity
   b. Expense of breeding
   c. Longevity
   d. The nature of mutations

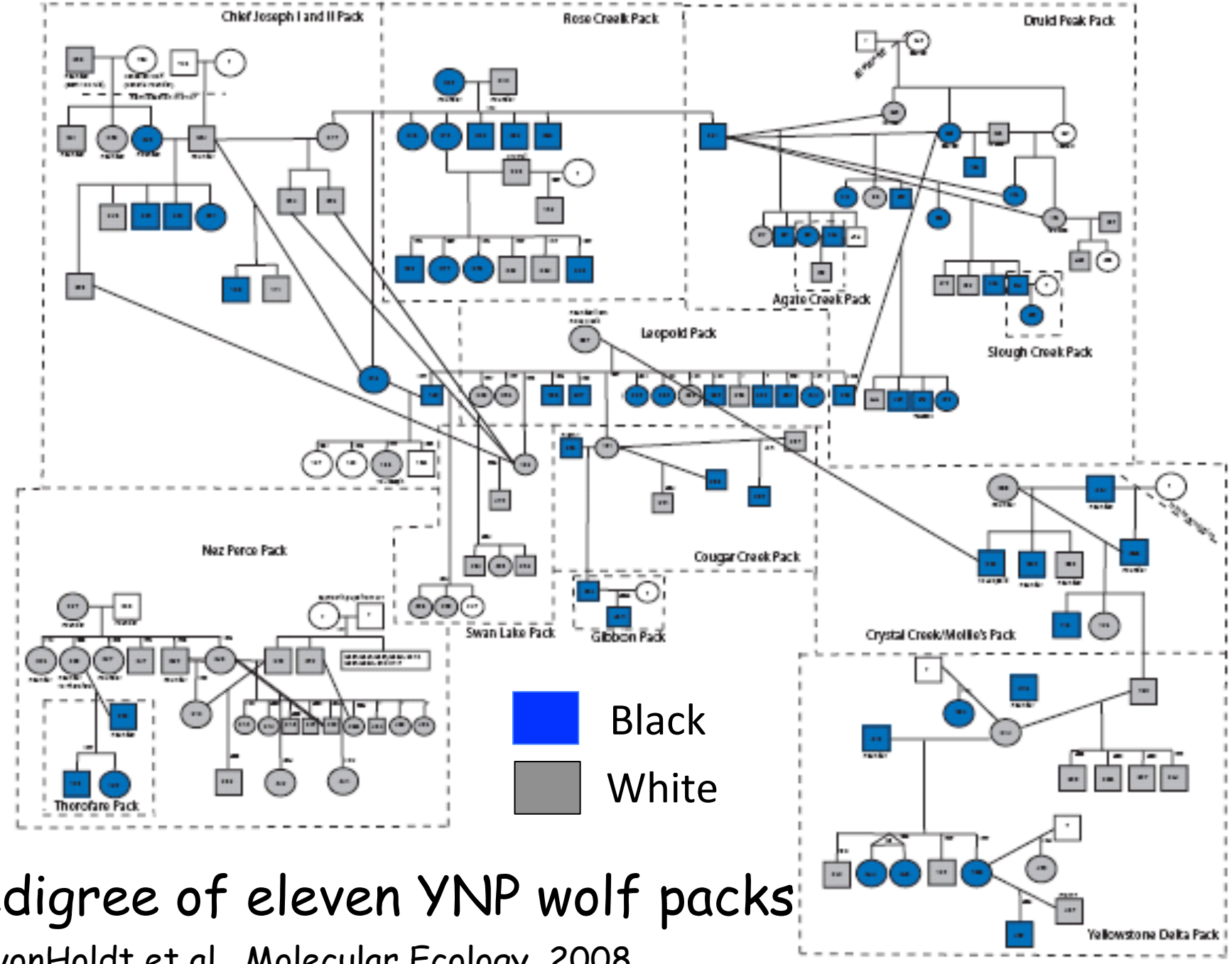
3. Engineering in the wild: black and gray coat colors of wolves




Tundra/Taiga
Boreal
Coniferous
Forest
Phenotype
Dark (Black or Dark Gray)
Light (White or Light Gray)

B


10 mo
10 yrs
10 mo
10 yrs



Pedigree of eleven YNP wolf packs

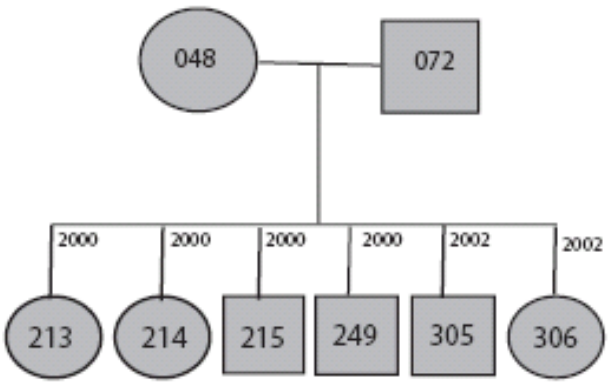
vonHoldt et al., Molecular Ecology, 2008

Coat Color and Dominant Inheritance

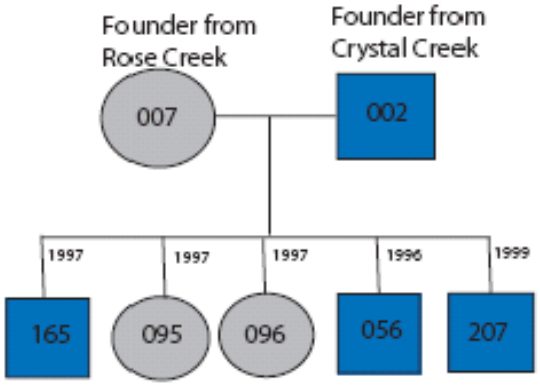
Model	Grand mu	1-1	1-2	2-2	Scale	Freq1	Freq2	LOD	Free Par.	SIC
Sporadic	1.336	----	----	----	0.6301	0.5	0.5	-233.789	5	471.1795
Genotypic	-0.7631	1.658	2.731	-4.389	0.2924	0.8038	0.1962	-199.246	5	402.0933
Dominant	1.426	-0.5353	0.5353		0.2921	0.7871	0.2129	-193.736	4	390.9771
Recessive	2.047	0.5353		-0.5353	0.2921	0.2129	0.7871	-193.736	4	390.9771
Allelic	1.336	0.5	----	0.5	0.6031	0.5	0.5	-411.234	4	825.9739

- Examples to support dominant inheritance

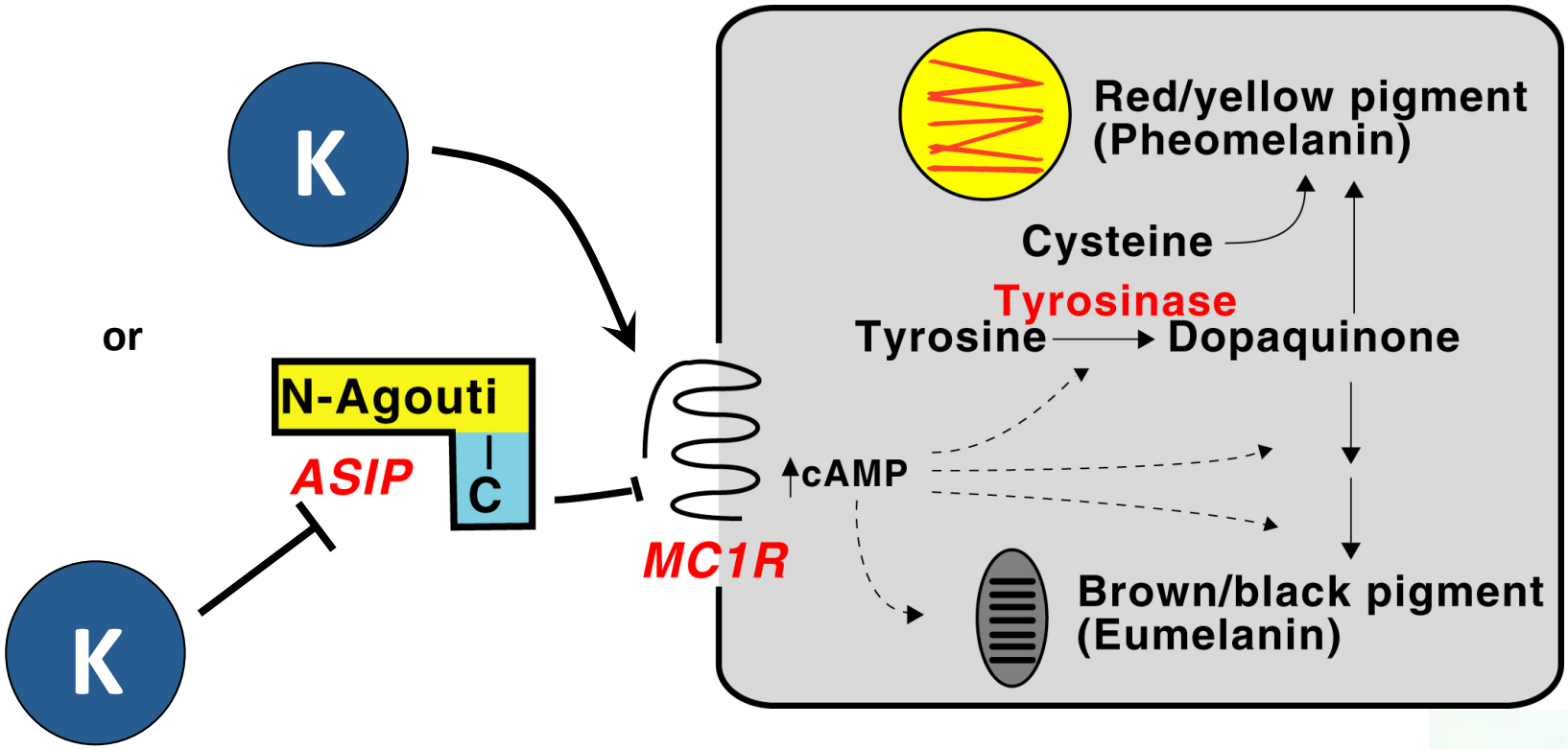
Nez Perce Pack



Leopold Pack



The K locus encodes a secreted protein that affects the Agouti-Mc1r system

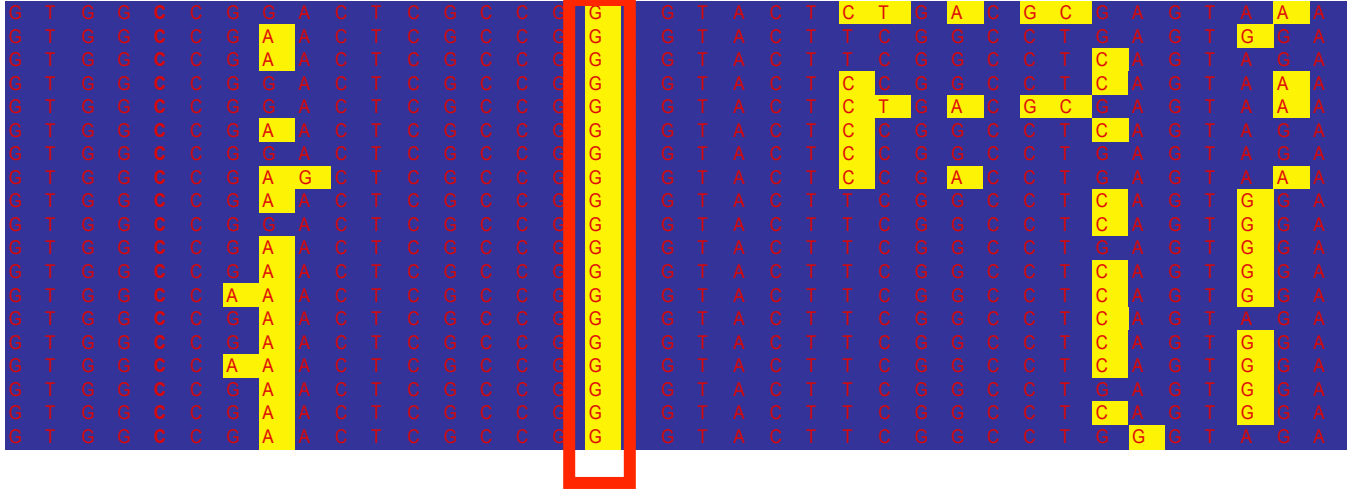


In a sample of 24 gray and 28 black wolves, there was perfect association of the K allele with black color.

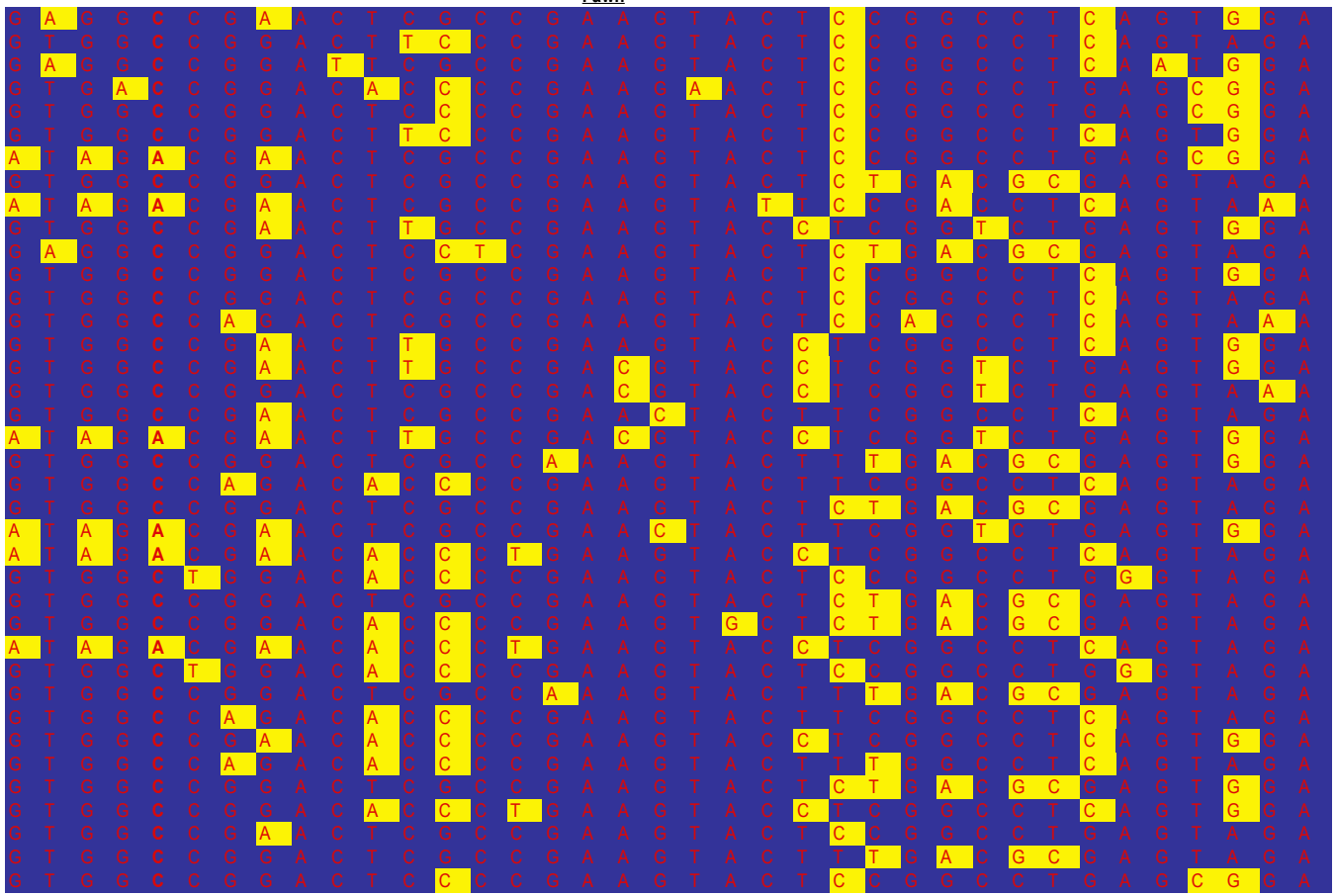
Anderson et al., Science, 2009

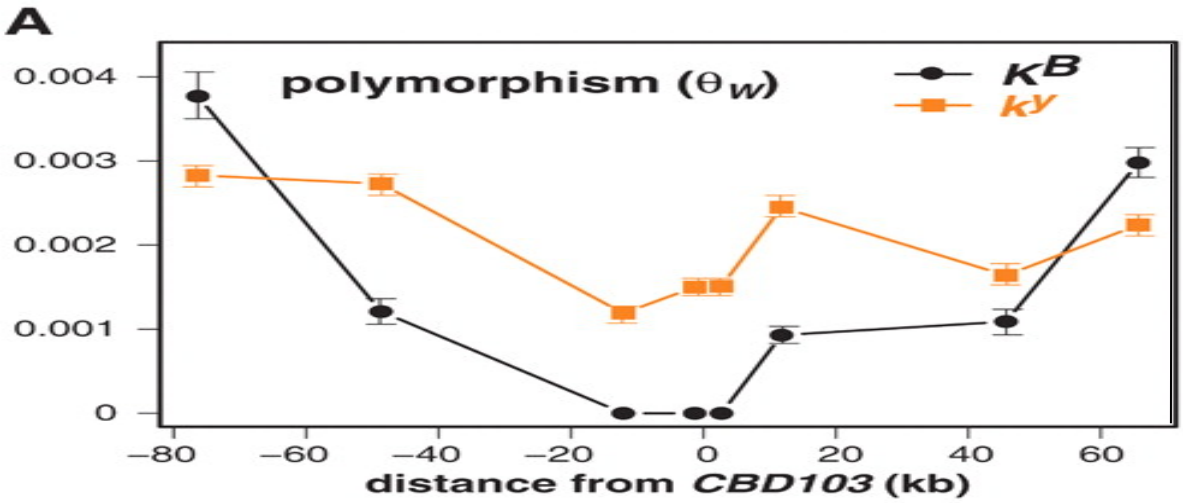
Tovi Anderson

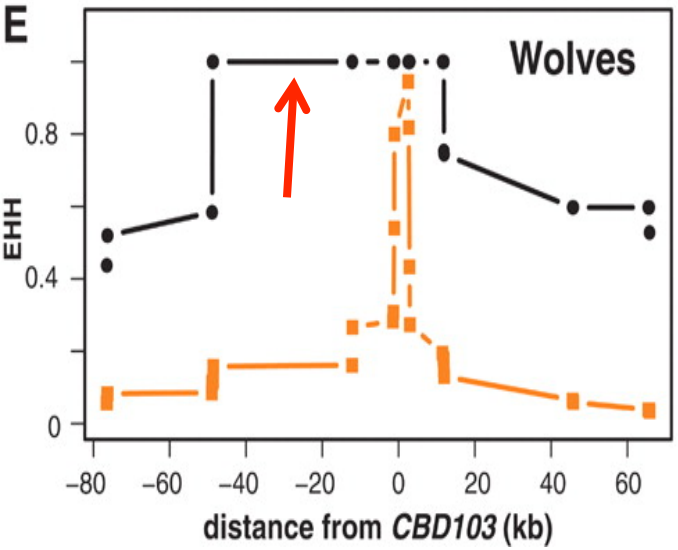
SNP variation near the K locus evidence for selective sweep

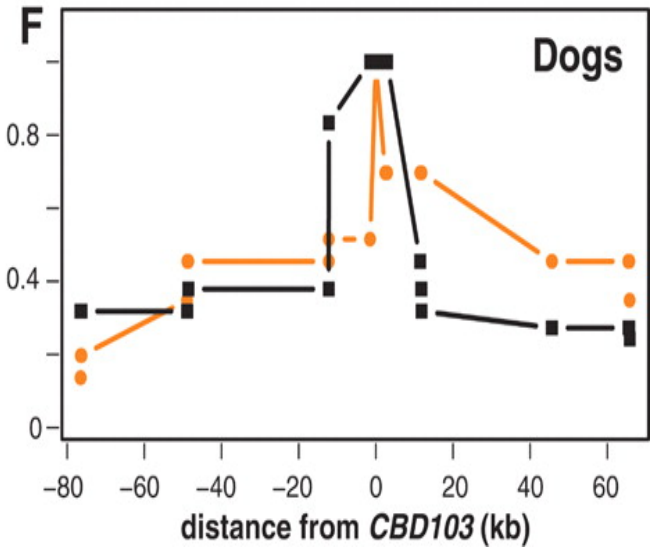


Fawn



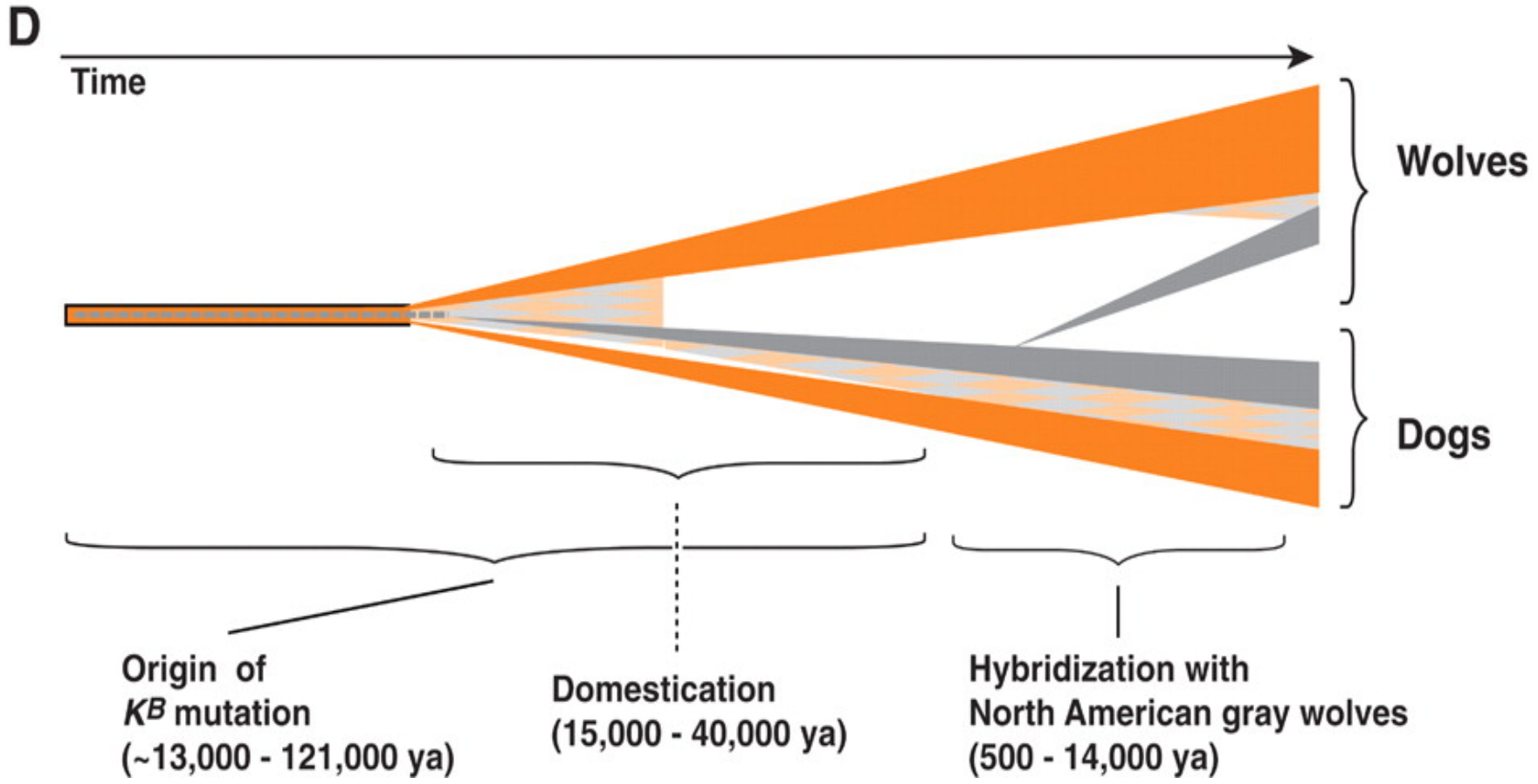






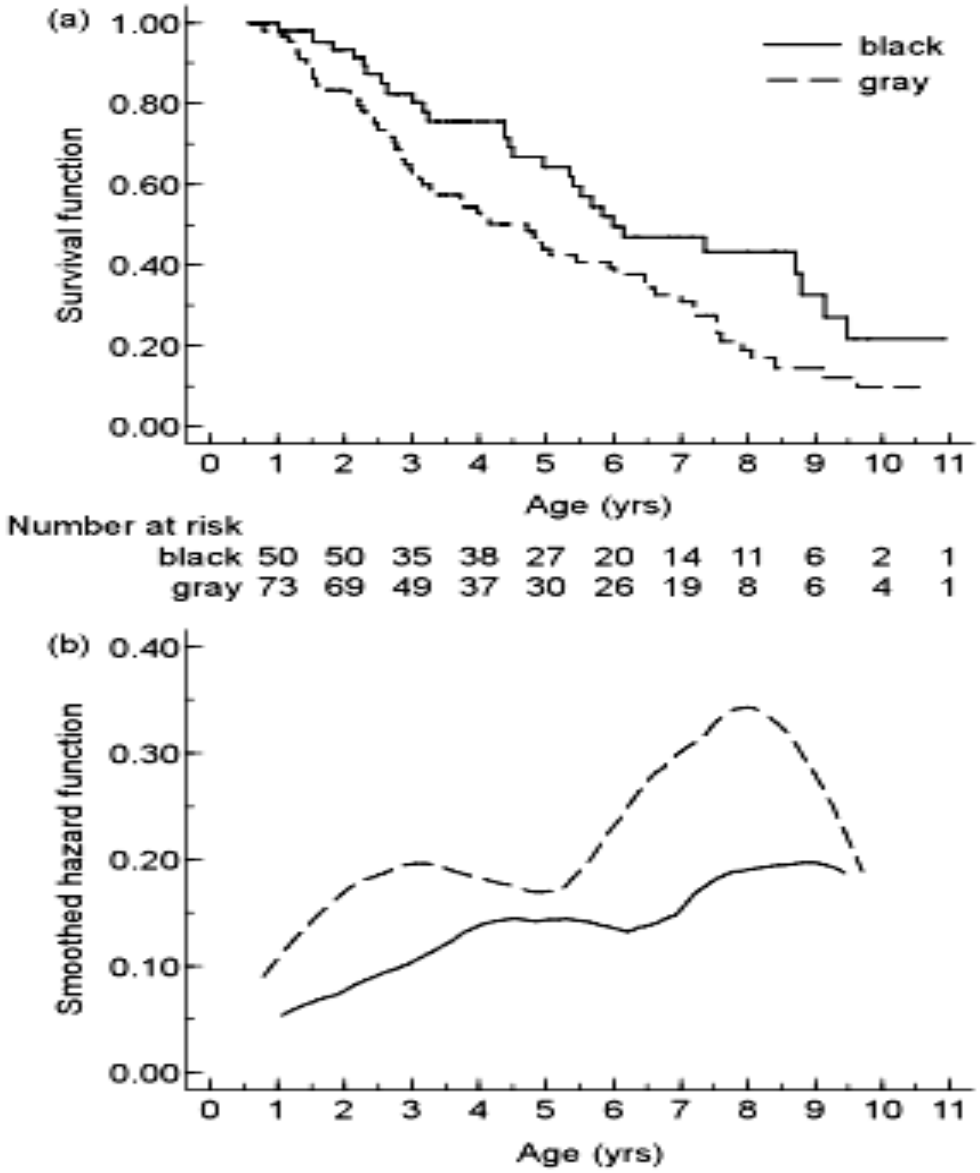
Anderson et al., 2009, Science

Dogs have enriched the genetic heritage of wolves



Genetic variants that are selected under domestication can transform wild progenitor populations

Black Wolves Live longer



Dan Stahler

Coulsin et al., 2012, Science

CONCLUSIONS

1) The engineering of the dogs has involved a relatively small number of genes of large effect.

2) These genetic variants have been transferred to established breeds by crossing. The results is the expression on different genetic backgrounds and enhanced phenotypic diversity.

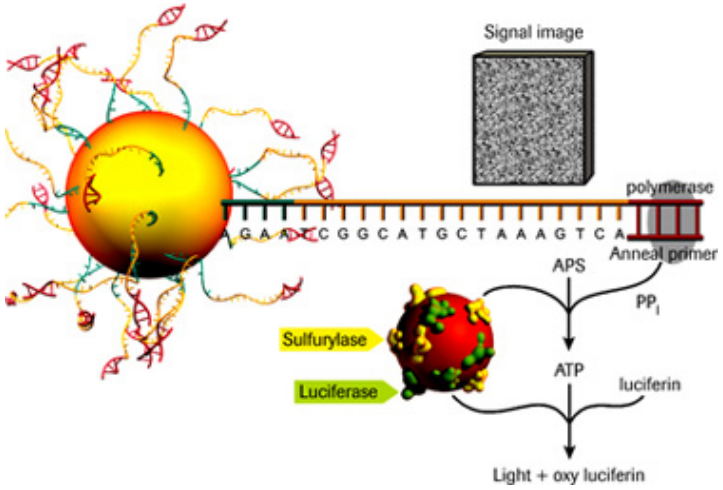
3) Mutation, as well as goal driven selection, limits diversity in dogs.

4. Variants selected under domestication can transform their wild kin.

4. The future, Genomic and expression studies

Mass sequencing technology has become a practical tool

Roche GS FLX 454 Genome Sequencer



Sequencing reaction of the Genome Sequencer System. Millions of copies of a single clonal fragment are contained on each DNA Capture Bead.

- Pyrosequencing technology
- Clonal amplification of starting material
- **500 million bp per run (1 million reads at average 400-500bp length)**
- 10 hour run
- Ability to run 2, 4, 8 & 16 separate lanes (physical separation)
- 10bp tags built into 454 adaptors allow further sample pooling (160 unique tags available)
- Cost ~$15/1000 reads

4. THE WOLF GENOME PROJECT

Rena Schweizer

John Novembre

The Wolf Genome Project

Reference Dog

5 canid genomes being sequenced by UCLA-based team

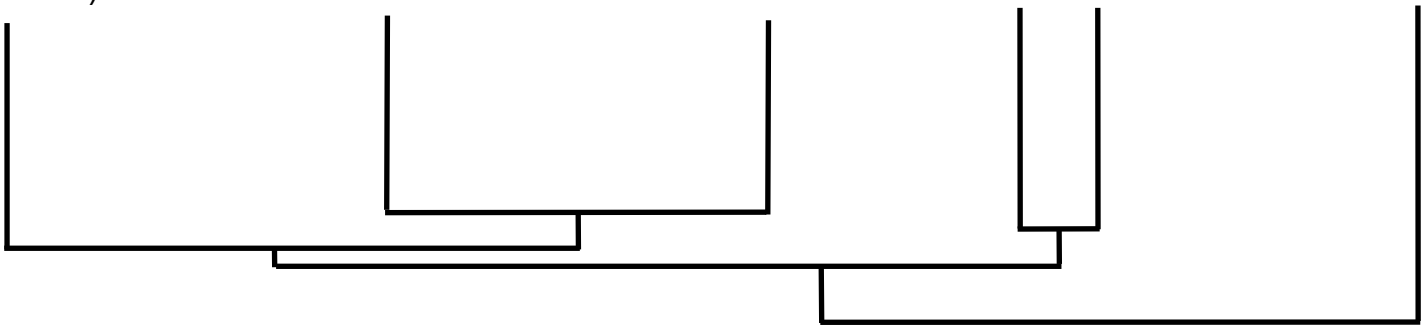
Boxer ("Tasha")

Dingo

Basenji

2 Gray wolves

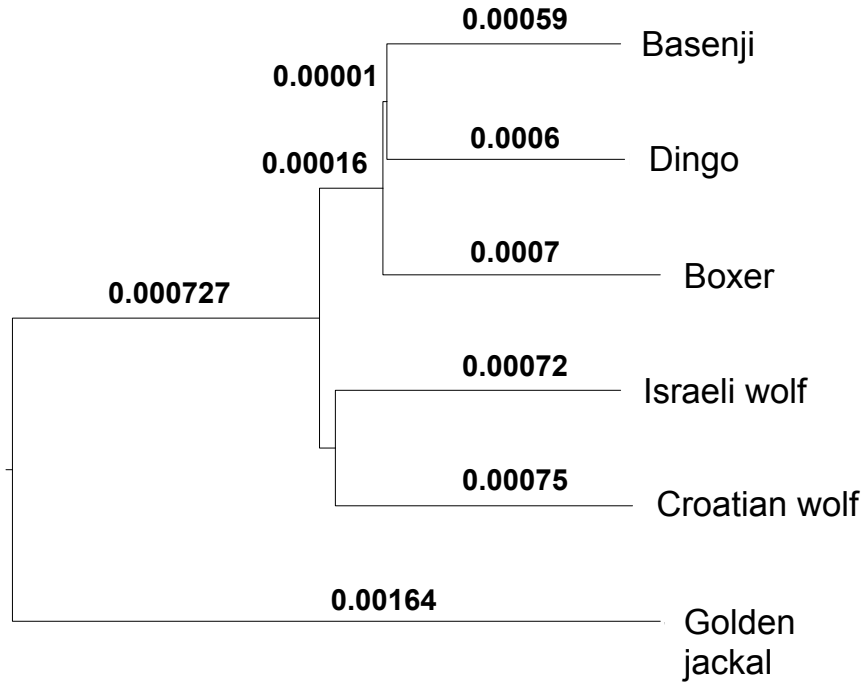
Golden jackal



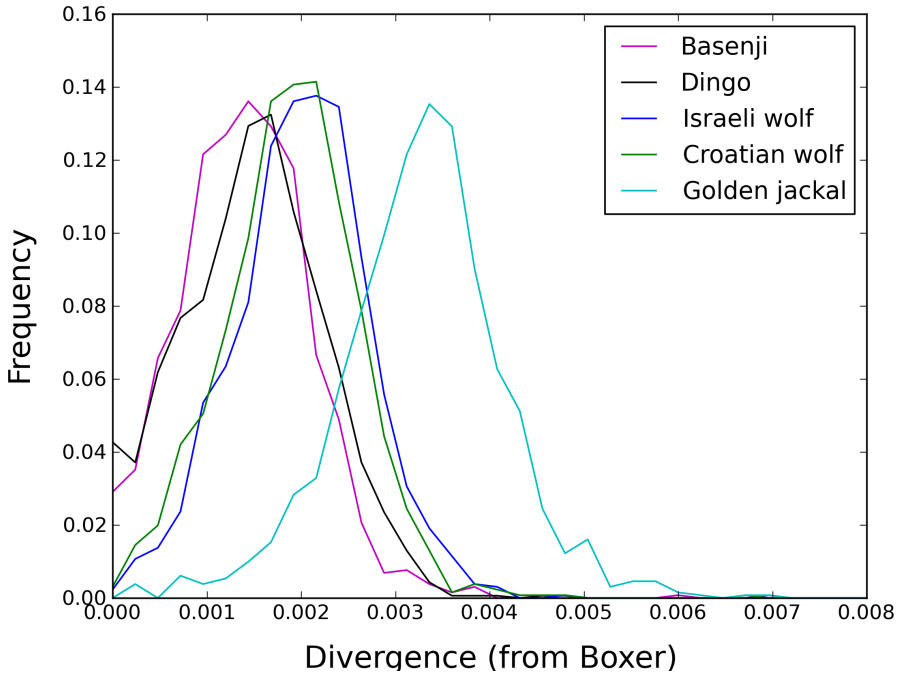
SOLiD Long mate-pair 1.5kp insert 50bp/50bp	SOLiD Fragment 50bp (75bp)	HiSeq Paired-end 400bp insert 100/100bp
1 slide	1-2 slides	1 lane

Genome-wide divergence

NJ Tree

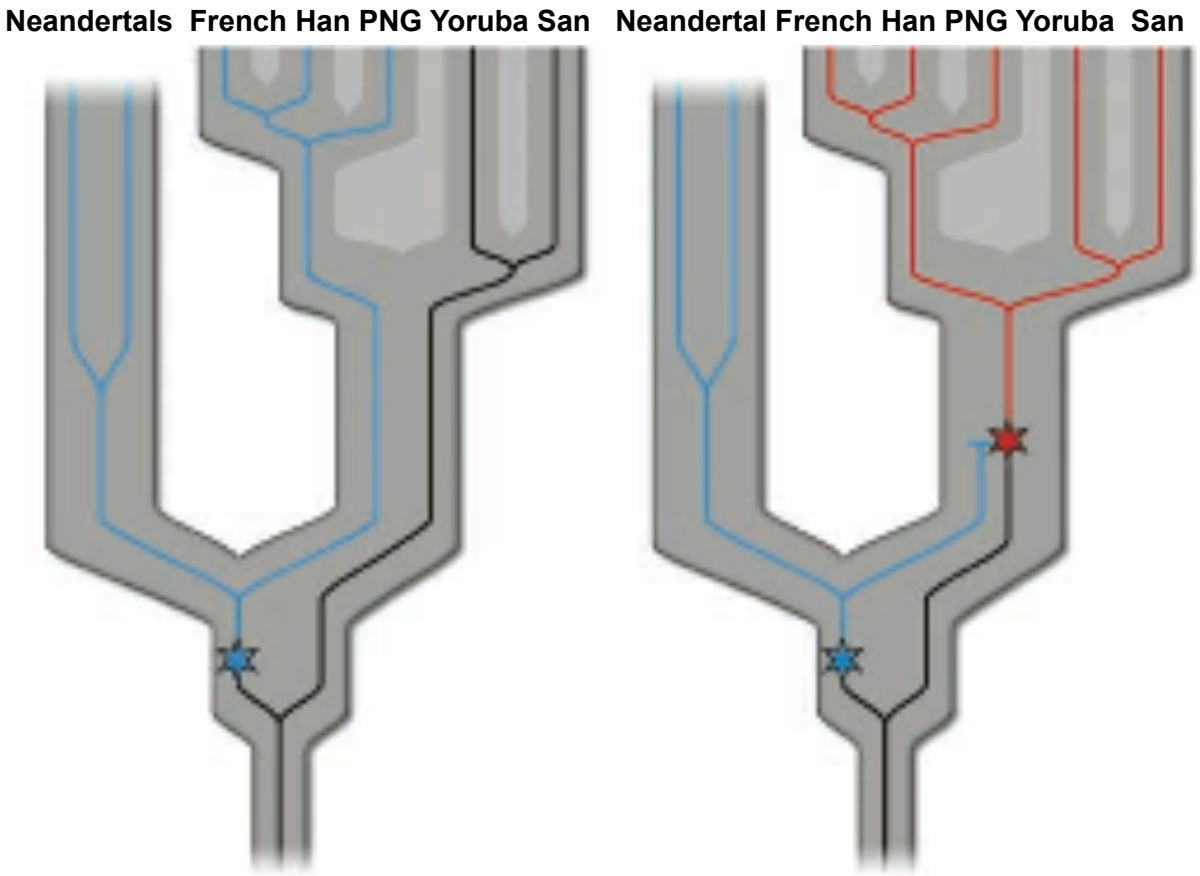


100kb windows



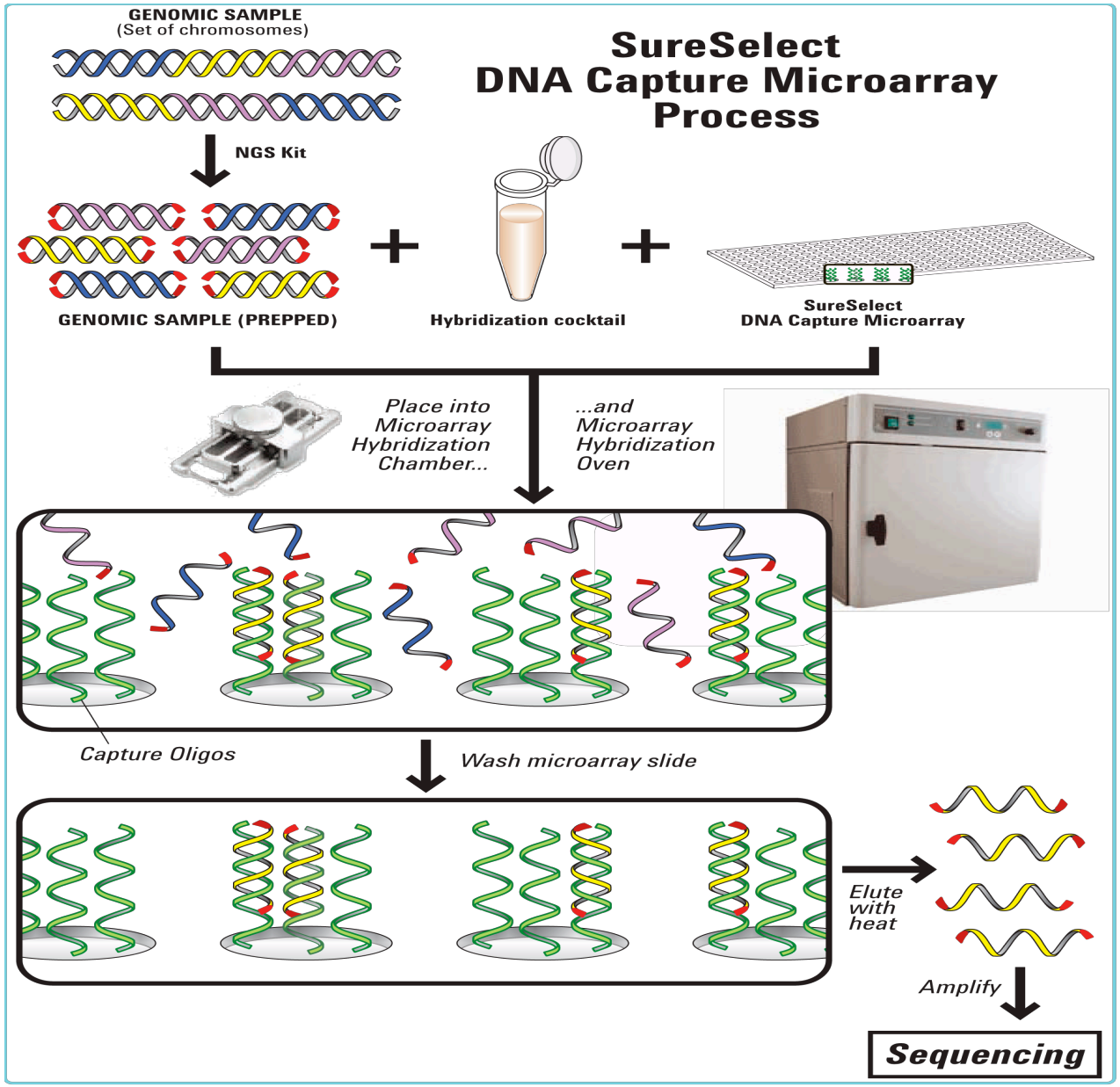
- If Gen. Time 1.5 yrs and μ =10^{-8} , wolf-dog divergence ~ 120,000 years
- Sequencing error inflates divergence, but can calibrate to SNP chip

Detecting older sweeps



Green et al. Nature, 2010

Population Genomics



Domestication of the red fox

THE FUTURE OF DOG EVOLUTION

Acknowledgements

NHGRI
Elaine A. Ostrander