




# What's in a Mutt? An Intro to Dog DNA Analysis

EN 601.147.13  
Intersession 2019



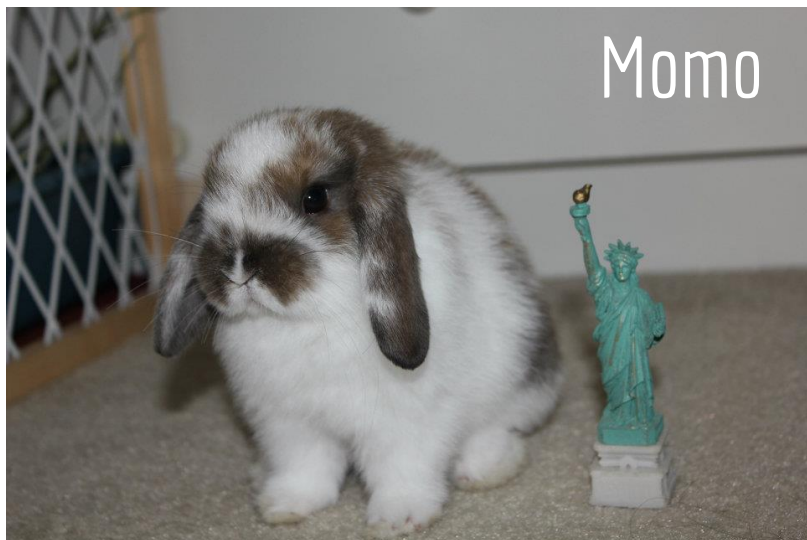
# Class Logistics

- Syllabus, assignments, and lectures in Piazza resources
- Attendance via sign in sheet; make sure you sign in each class!
- Class will be part lectures and part a data analysis project
- For the project you'll need:
  - A laptop with a Linux or Unix terminal/shell (Windows users can use a Virtual Machine)
  - A partner! Find a partner by Friday, or let me know if you'd prefer I assign you one

# A bit about me

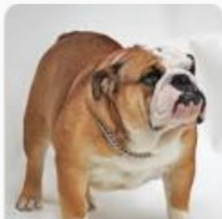
4th year graduate student in CS working on human genomics

Volunteer at the MD SPCA



Earliest documented genomic selection, 30k years ago!





Bulldog



German Shepherd



Labrador Retriever



Beagle



Poodle



Chihuahua



Golden Retriever



Pug



Border Collie



Newfoundland dog



American Staffordshire ...



Greyhound



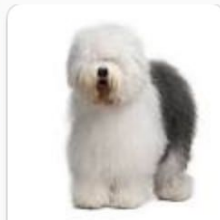
Australian Shepherd



Pembroke Welsh Corgi



Bull Terrier



Old English Sheepdog

Dog breeds, via Google Images

# What even is a breed?

Over 300 dog breeds around the world

192 breeds are recognized by the American Kennel Club (AKC)

From the AKC:

“The simplest way to define a breed is to say it always “breeds true.” That is, breeding a purebred Irish Setter to another purebred Irish Setter will always produce dogs instantly recognizable as Irish Setters.”

AKC provides breed “standards” of recognizable, defining features for breeds



AKC registered Doberman Pinschers

“Ears normally cropped and carried erect. The upper attachment of the ear, when held erect, is on a level with the top of the skull.”



“Tail docked at approximately second joint, appears to be a continuation of the spine.”

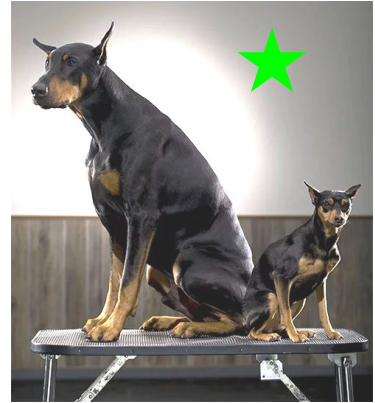
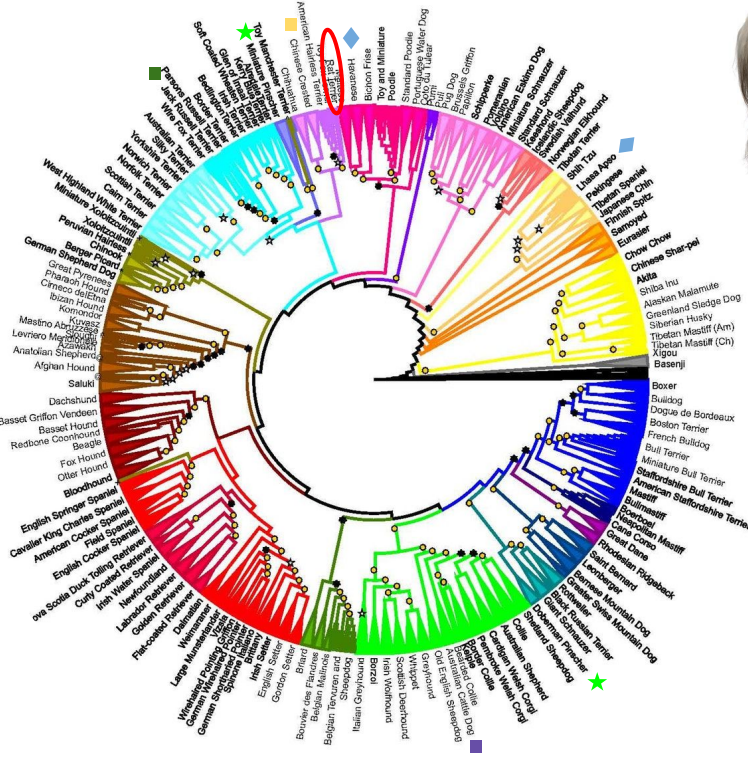
AKC breed standard for a Doberman Pinscher





Do these dogs meet the AKC Doberman breed standard?  
Do dogs who do meet the standard (cropped ears, docked tail) breed true?

# Defining breeds genetically instead of visually



Why care about breed (or why not)?

# Why care about breed (or why not)?

- Curiosity!
- Breed-specific medical concerns
  - **Greyhounds** have an extreme **sensitivity to propofol**, an anesthesia drug
  - **Herding breeds** more likely to have a mutation causing a Heartworm preventative drug, **ivermectin** to not metabolize and **build up to toxic levels**
  - **Boxers** more likely to have an **adverse reaction to the sedative acepromazine**
  - Likelihood of tumors being benign vs malignant correlates with breed
- Breed-specific legislation
- Shelter adoptions

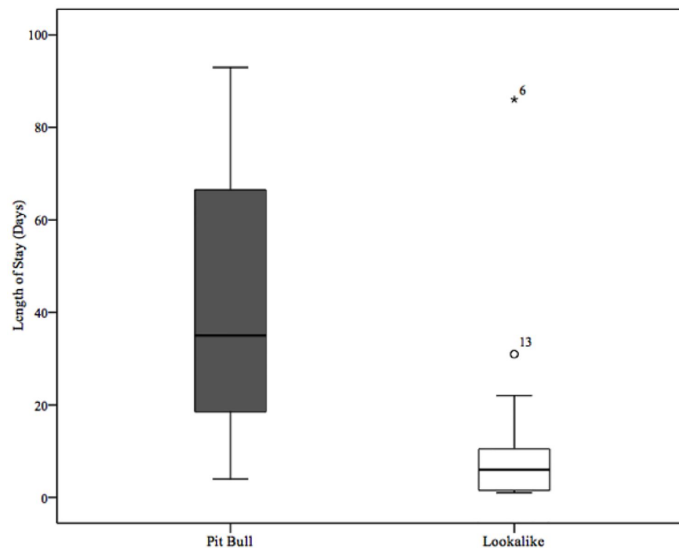
# Shelter Adoptions and Breed

Study of 120 shelter dogs (Olson *et al*, 2015):

“Whereas DNA breed signatures identified only 25 dogs (21%) as pit bull-type, shelter staff collectively identified 62 (52%) dogs as pit bull-type.”

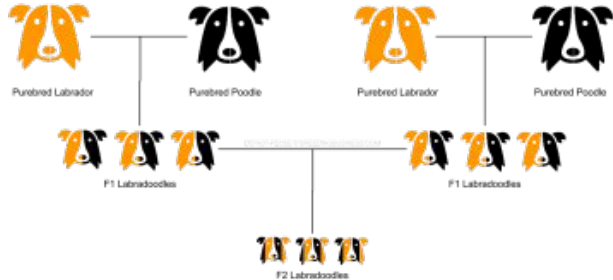
If shelters typically over identify dogs as pit bulls, might DNA testing increase adoption rates?

Study labeling pairs of similar looking dogs as pit bull vs other breed labels (Gunter, Barber, and Wynne, 2016):

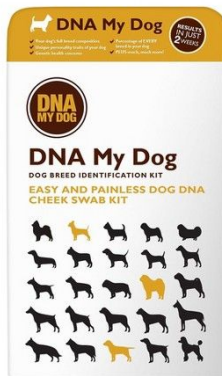
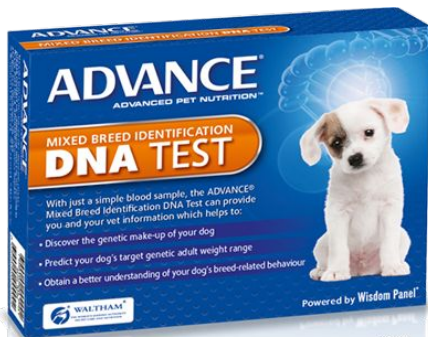


# But what about mixed breeds?

- How easily can we look at a mutt and guess the breeds?
- When does a mix become its own breed?
  - What are Labradoodles? Yorkipoos? "Cross-breeds" as per the AKC
  - If two mutts of the same makeup breed true, is it a new breed?
  - How might you characterize whether a mix is or isn't a new breed genetically?
- Is knowing the breed makeup medically informative?



# Commercial dog DNA sequencing



# Commercial dog DNA sequencing

Breed makeup

Medical markers

Breed-specific kits

All commercial testing is  
currently unregulated

**Read Atlantic article for  
discussion Wednesday**

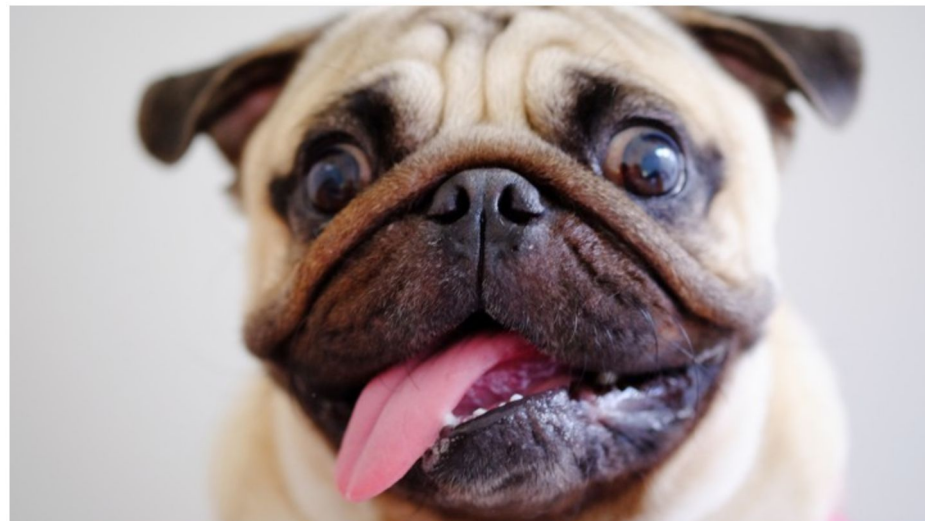


SCIENCE

## What Vets Think of '23andMe for Dogs'

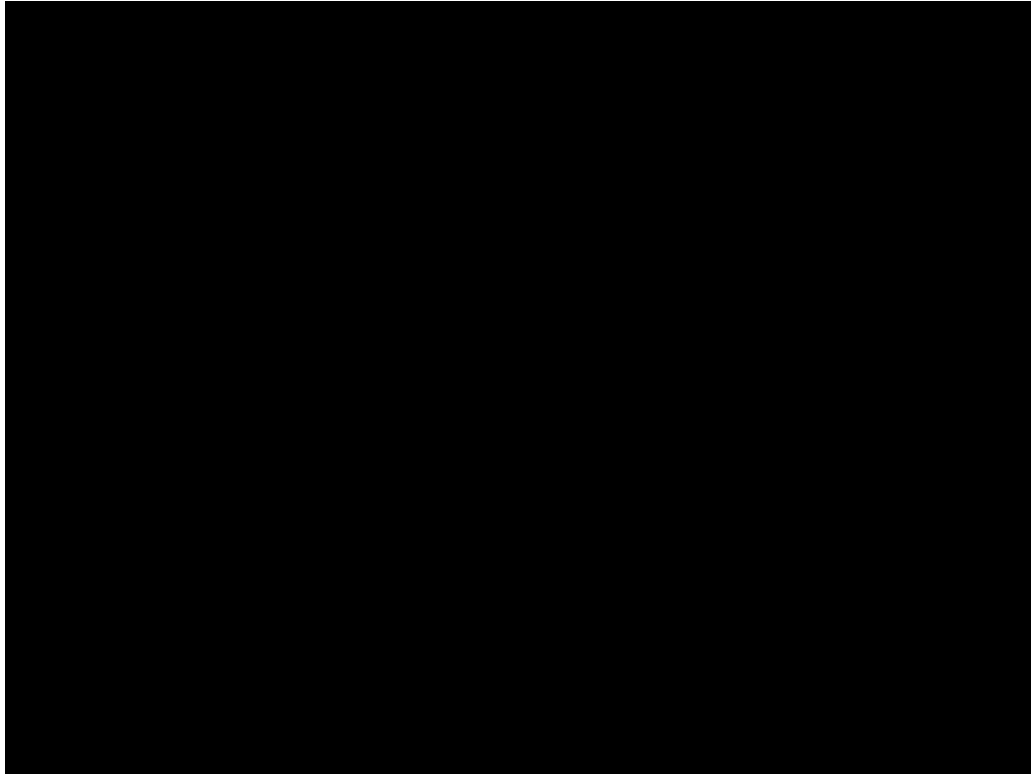
More and more companies are selling DNA-test kits for pets.


SARAH ZHANG NOV 12, 2018






# Sequencing for Research: Mutt Mix Project





Pup quiz!  
[PollEv.com/rachelsherma442](https://PollEv.com/rachelsherma442)



# Bella



Bella's shoulders are **knee high** when standing next to a person of average height.



**Other physical traits or features:**

Double coat.

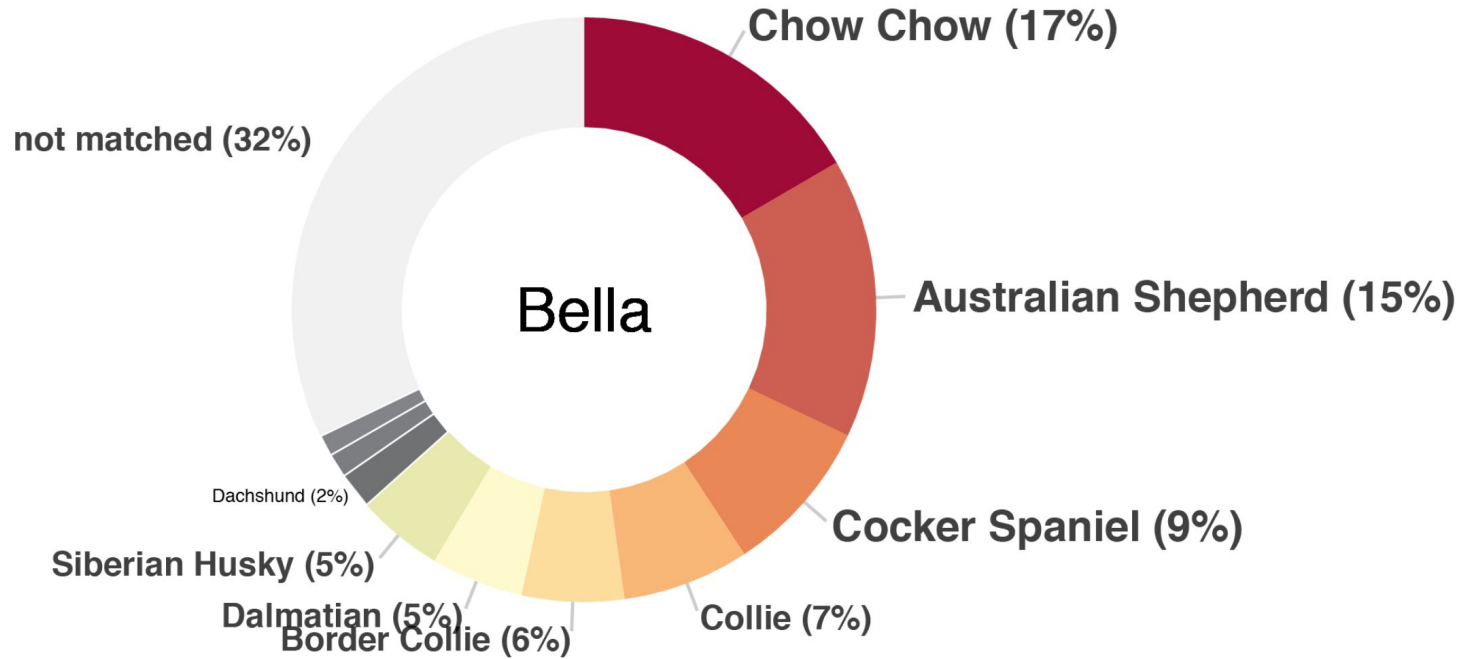
# What is Bella's top breed?

Australian Cattle Dog	<b>A</b>
Australian Shepherd	<b>B</b>
Basenji	<b>C</b>
Basset Hound	<b>D</b>
Beagle	<b>E</b>
Bernese Mountain Dog	<b>F</b>
Bichon Frise	<b>G</b>
Bloodhound	<b>H</b>
Boston Terrier	<b>I</b>
Boxer	<b>J</b>
Bull Terrier	<b>K</b>
Bulldog	<b>L</b>
Cavalier King Charles Spaniel	<b>M</b>
Chihuahua	<b>N</b>
Chow Chow	<b>O</b>
Cocker Spaniel	<b>P</b>
Collie/Border Collie	<b>Q</b>
Corgi	<b>R</b>
Dachshund	<b>S</b>
Dalmation	<b>T</b>
Doberman Pinscher	<b>U</b>
English Springer Spaniel	<b>V</b>
French Bulldog	<b>W</b>
German Shepherd	<b>X</b>
Golden Retriever	<b>Y</b>
Great Dane	<b>Z</b>
Great Pyrenees	<b>AA</b>
Greyhound	<b>AB</b>
Havanese	<b>AC</b>
Irish Setter	<b>AD</b>
Irish Wolfhound	<b>AE</b>
Italian Greyhound	<b>AF</b>
Jack Russell Terrier	<b>AG</b>
Labrador Retriever	<b>AH</b>
Lhaso Apso	<b>AI</b>
Maltese	<b>AJ</b>
Mastiff	<b>AK</b>
Newfoundland	<b>AL</b>
Old English Sheepdog	<b>AM</b>
Pomeranian	<b>AN</b>
Poodle	<b>AO</b>
Pug	<b>AP</b>
Rottweiler	<b>AQ</b>
Saint Bernard	<b>AR</b>
Shiba Inu	<b>AS</b>
Shih Tzu	<b>AT</b>
Siberian Husky	<b>AU</b>
Staffordshire Bull Terrier (Pit bull)	<b>AV</b>
West Highland Terrier	<b>AW</b>
Yorkshire Terrier	<b>AX</b>

# What is Bella's second best breed?

Australian Cattle Dog  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier

# Bella



32% of participants guessed at least one breed correctly, 2% guessed 2 breeds correctly, and 1 participant could guess all three breeds.

# Rosie

[PollEv.com/rachelsherma442](https://PollEv.com/rachelsherma442)



Rosie's shoulders are **knee high** when standing next to a person of average height.



## Other physical traits or features:

Stripe of wavy fur down center of back. Black spots on tongue.

# What is Rosie's top breed?

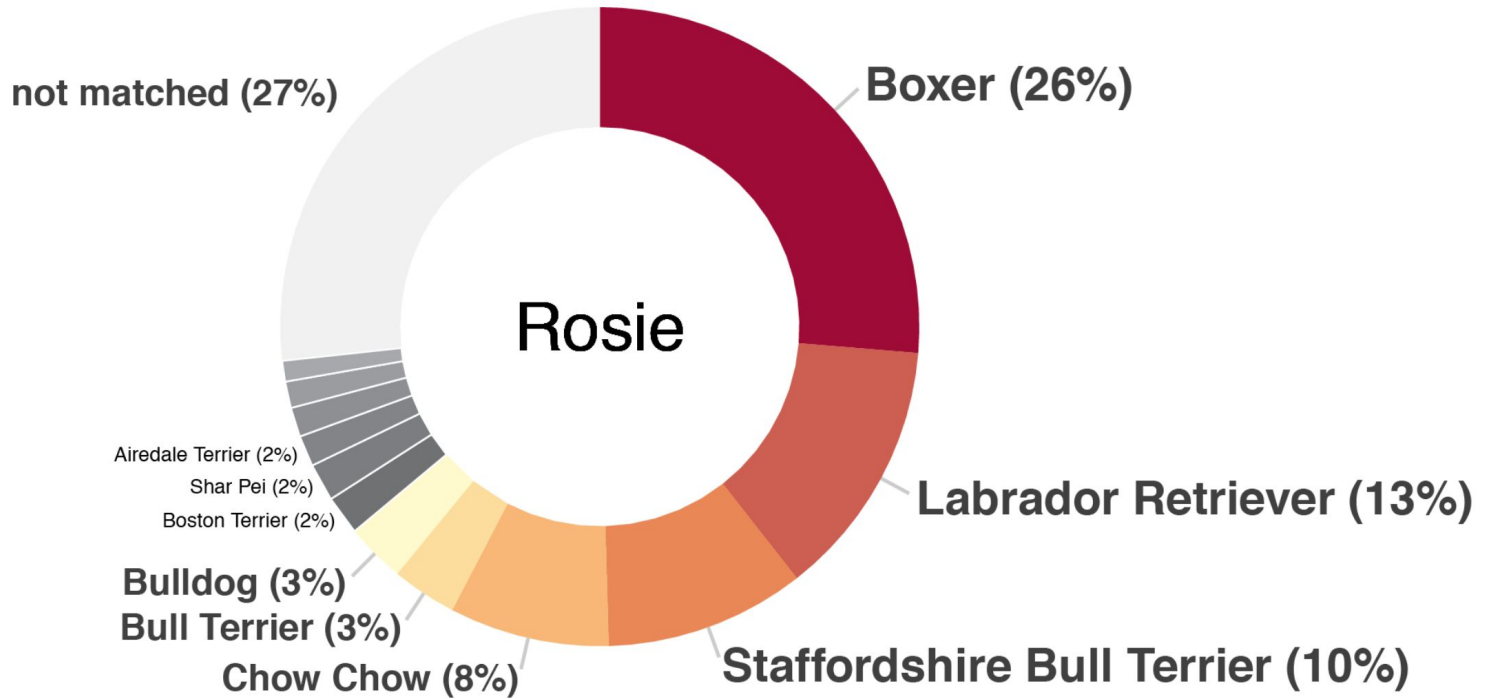
Australian Cattle Dog  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier



# What is Rosie's second best breed?

Australian Cattle Dog  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier

# Rosie



45% of participants guessed at least one breed correctly, 6% guessed 2 breeds correctly, and 2 participants could guess all three breeds.

# Boone

[PollEv.com/rachelsherma442](https://www.PollEv.com/rachelsherma442)



Boone's shoulders are **knee high** when standing next to a person of average height.



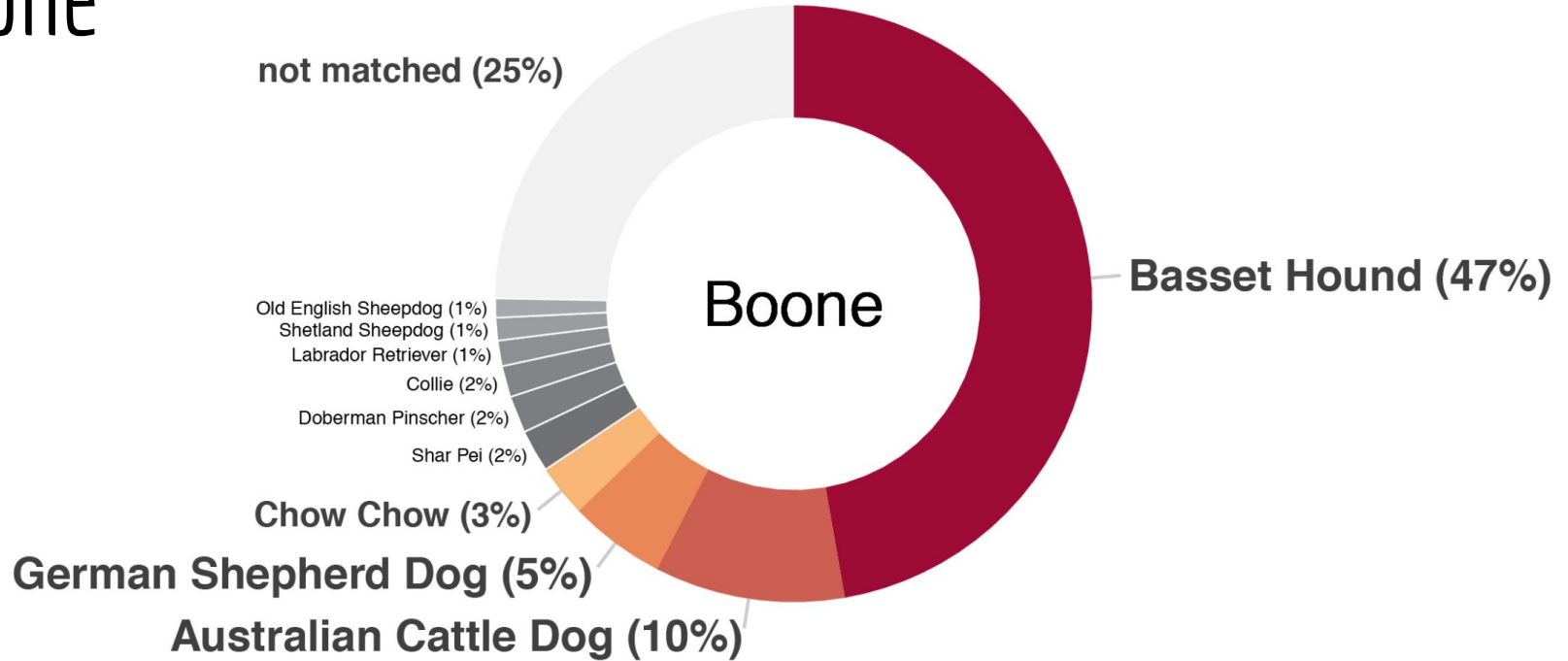
# What is Boone's top breed?

Australian Cattle Dog  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane

# What is Boone's second best breed?

Australian Cattle Dog  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier

# Boone



78% of participants guessed at least one breed correctly, 11% guessed 2 breeds correctly, and 11 participants could guess all three breeds.

# Maxine



Maxine's shoulders are **thigh high** when standing next to a person of average height.



**Other physical traits or features:**

Hair not fur

# What is Maxine's top breed?

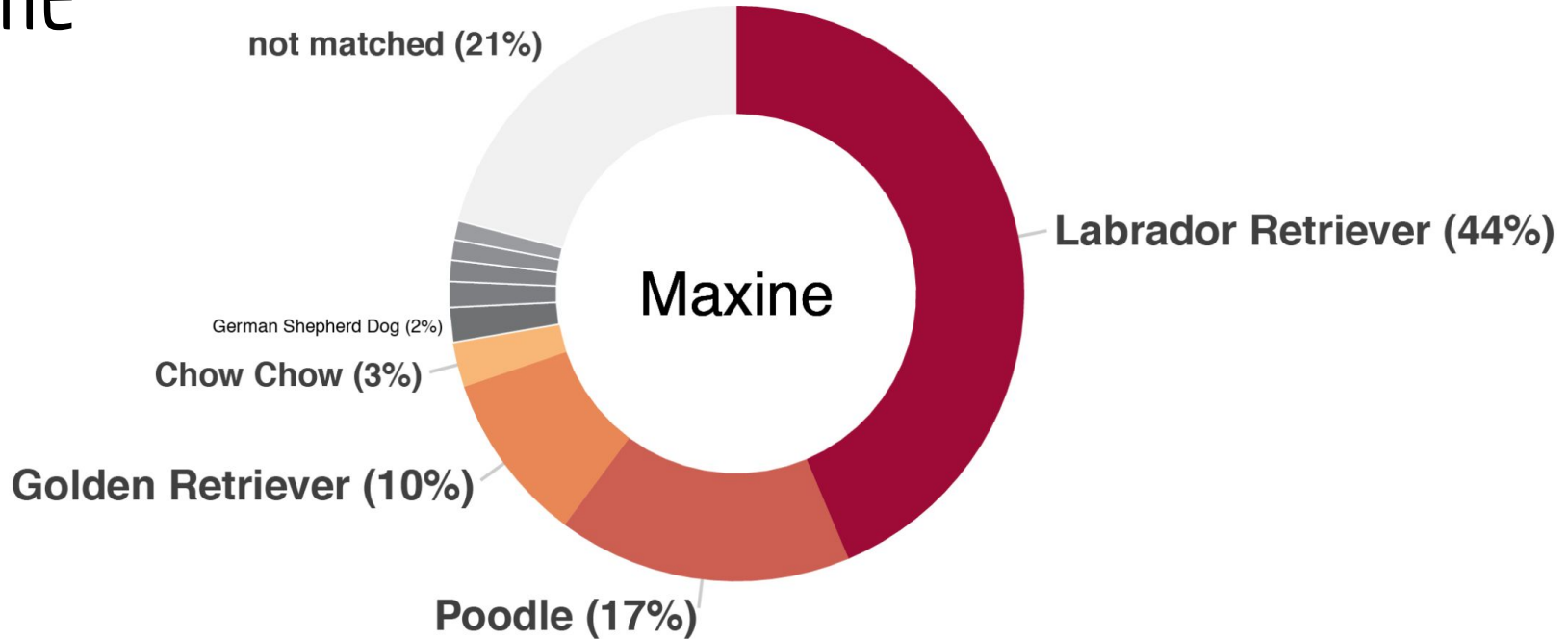
Australian Cattle Dog  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane



# What is Maxine's second best breed?

Australian Cattle Dog  
Australian Shepherd  
Basenji  
Basset Hound  
Beagle  
Bernese Mountain Dog  
Bichon Frise  
Bloodhound  
Boston Terrier  
Boxer  
Bull Terrier  
Bulldog  
Cavalier King Charles Spaniel  
Chihuahua  
Chow Chow  
Cocker Spaniel  
Collie/Border Collie  
Corgi  
Dachshund  
Dalmation  
Doberman Pinscher  
English Springer Spaniel  
French Bulldog  
German Shepherd  
Golden Retriever  
Great Dane  
Great Pyrenees  
Greyhound  
Havanese  
Irish Setter  
Irish Wolfhound  
Italian Greyhound  
Jack Russell Terrier  
Labrador Retriever  
Lhaso Apso  
Maltese  
Mastiff  
Newfoundland  
Old English Sheepdog  
Pomeranian  
Poodle  
Pug  
Rottweiler  
Saint Bernard  
Shiba Inu  
Shih Tzu  
Siberian Husky  
Staffordshire Bull Terrier (Pit bull)  
West Highland Terrier  
Yorkshire Terrier

# Maxine



29% of participants guessed at least one breed correctly, 6% guessed 2 breeds correctly, and 39 participants could guess all three breeds.

# Mutt Mix Project Results

- 26,639 participants and 432,743 guesses for 31 dogs
- Number of breeds detected doesn't correlate with guessing accuracy
  - Four of the mutts had over 19 breeds detected but two of those weren't as tough to guess as some of the dogs with far fewer breeds
- Average of **25% of guesses correct**, 28% for the 4,804 participants in dog professional fields
- Most guessed breeds: Labrador Retriever, German Shepherd Dog, Golden Retriever, Beagle, and Border Collie
- Most commonly detected breeds: Chow Chow, German Shepherd Dog, Staffordshire Bull Terrier, Golden Retriever, and Labrador Retriever

# Meet Our Mutts




Reilly




Clarence



Finch



Take a break and then we'll talk about  
how we sequence DNA!

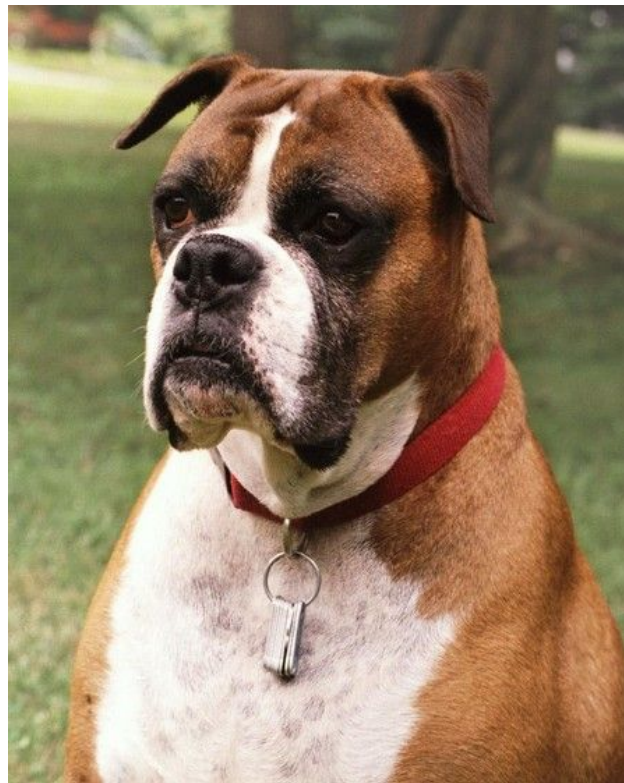


# Meet Tasha!

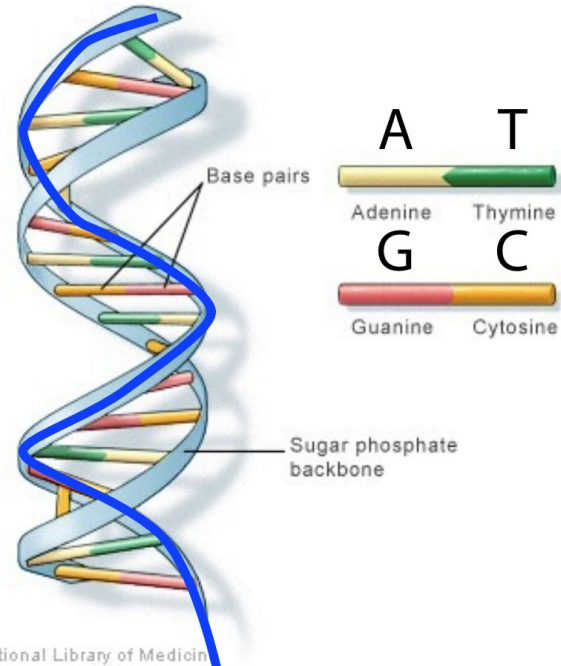
Tasha is a boxer and she's the basis of the dog reference genome, released in 2005

- 39 chromosomes
- 2.8 billion base pairs (A, C, G, T)
- 19,000 genes (approximately)

*A reference genome* is a genome we have put together to be representative of a species. We'll talk more about their creation later.

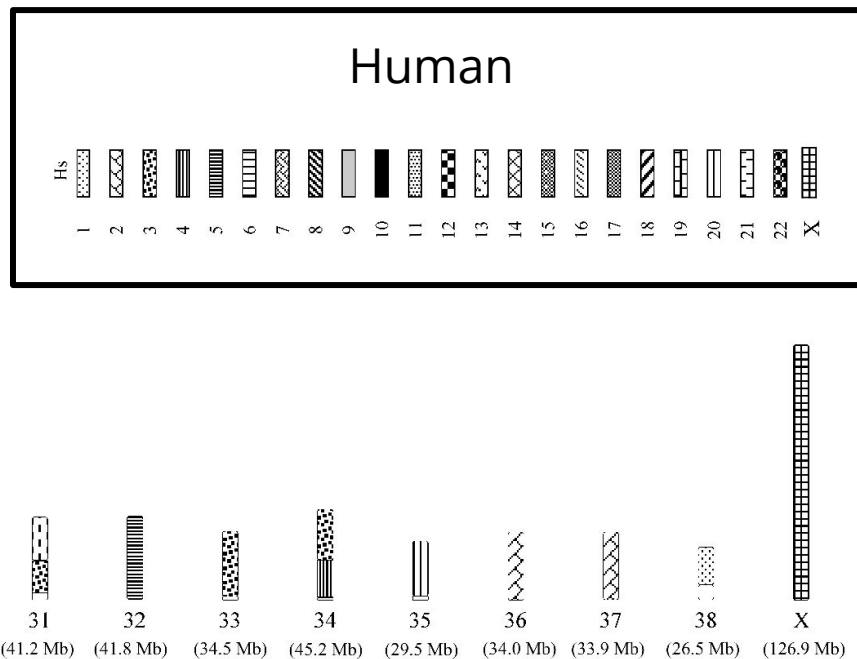
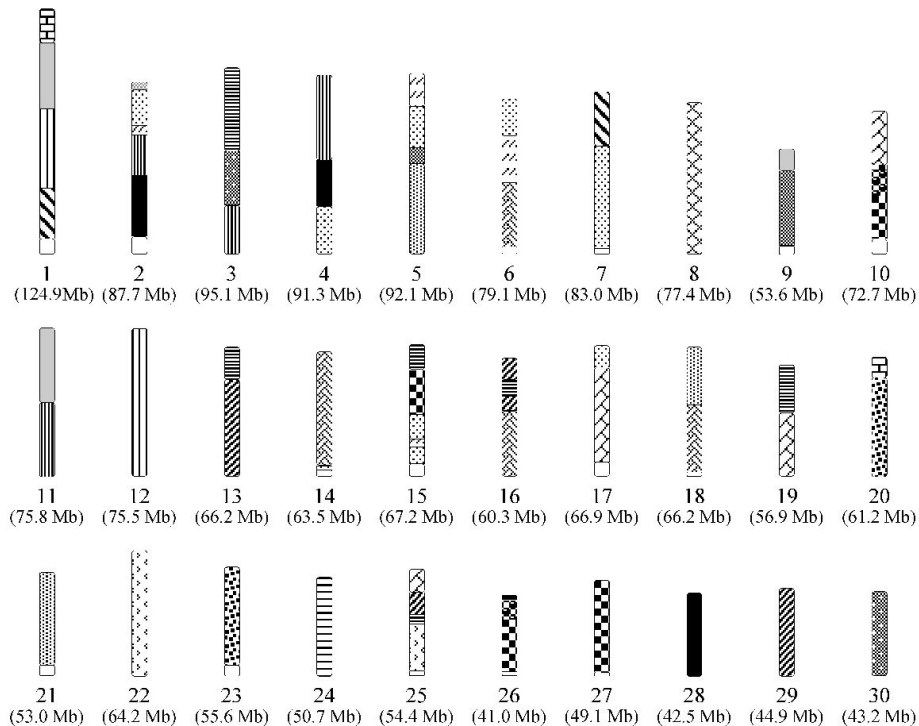


# DNA



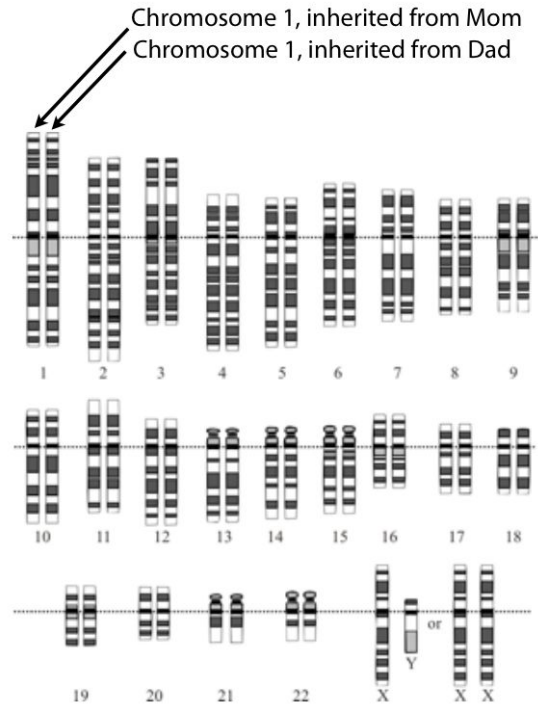
TCACACTGAGCGTGCTG

# Dog reference genome





# Genotypes and Genomes



The “genome” is all DNA on all chromosomes for an individual

Actually pairs of chromosomes, so dogs have 78 total; 76 autosomes and 2 sex chromosomes

At each base pair (2.8 billion of them in dogs), we have a *genotype* of the two bases present

# Genotypes and Genomes

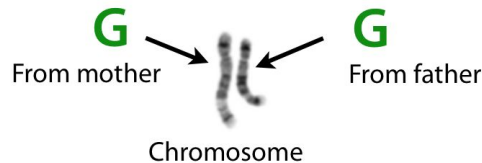
*Genotype* is a description of both copies, the possibilities for each copy are *alleles*

Interpretations are likelihoods; often the *phenotype* is affected by other factors

Who	Genotype	What It Means
	AA	In Europeans, 85% chance of brown eyes; 14% chance of green eyes; 1% chance of blue eyes.
	AG	In Europeans, 56% chance of brown eyes; 37% chance of green eyes; 7% chance of blue eyes.
Benjamin Langmead	GG	In Europeans, 72% chance of blue eyes; 27% chance of green eyes; 1% chance of brown eyes.

Sources: 23andme.com, Ben's genome

A and G are *alleles*



# Single Nucleotide Polymorphisms (SNPs)

Tasha

. . . ATCGGAATAGCGAGTA . . .

dog of interest  
(one copy)

AT**G**GGGAATAG**C**TAGTA

But SNPs don't occur this often - on average, we see 1 SNP every ~1,000 bases between dogs.

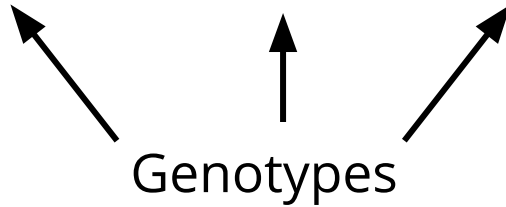
# Single Nucleotide Polymorphisms (SNPs)

Tasha

. . . ATCGGAATAGCGAGTA . . .

dog of interest  
(two copies)

AT**G**GGGAATAGCT**T**AGTA  
AT**G**GGGAATAGC**G**AGTA



# Single Nucleotide Polymorphisms (SNPs)

Tasha

. . . ATCGGAATAGCGAGTA . . .

dog of interest  
(two copies)

G  
G

T  
G

When we look at SNPs, especially commercially, we use SNP chips aka genotyping arrays. These **only** look at predetermined sites.

# Single Nucleotide Polymorphisms (SNPs)

Tasha

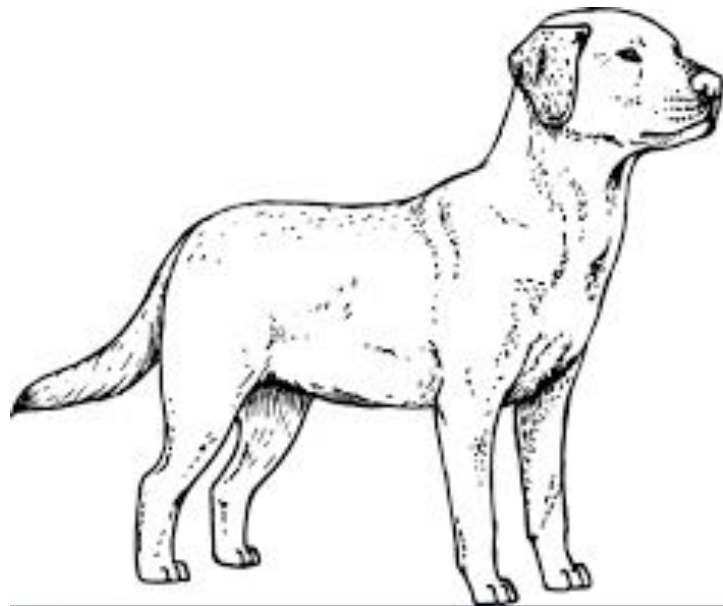
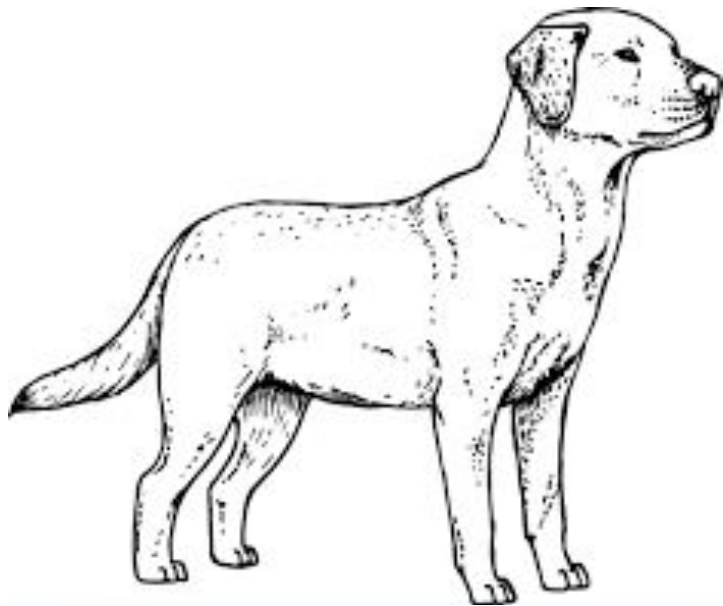
. . . ATCGGAATAGCGAGTA . . .

dog of interest  
(two copies)

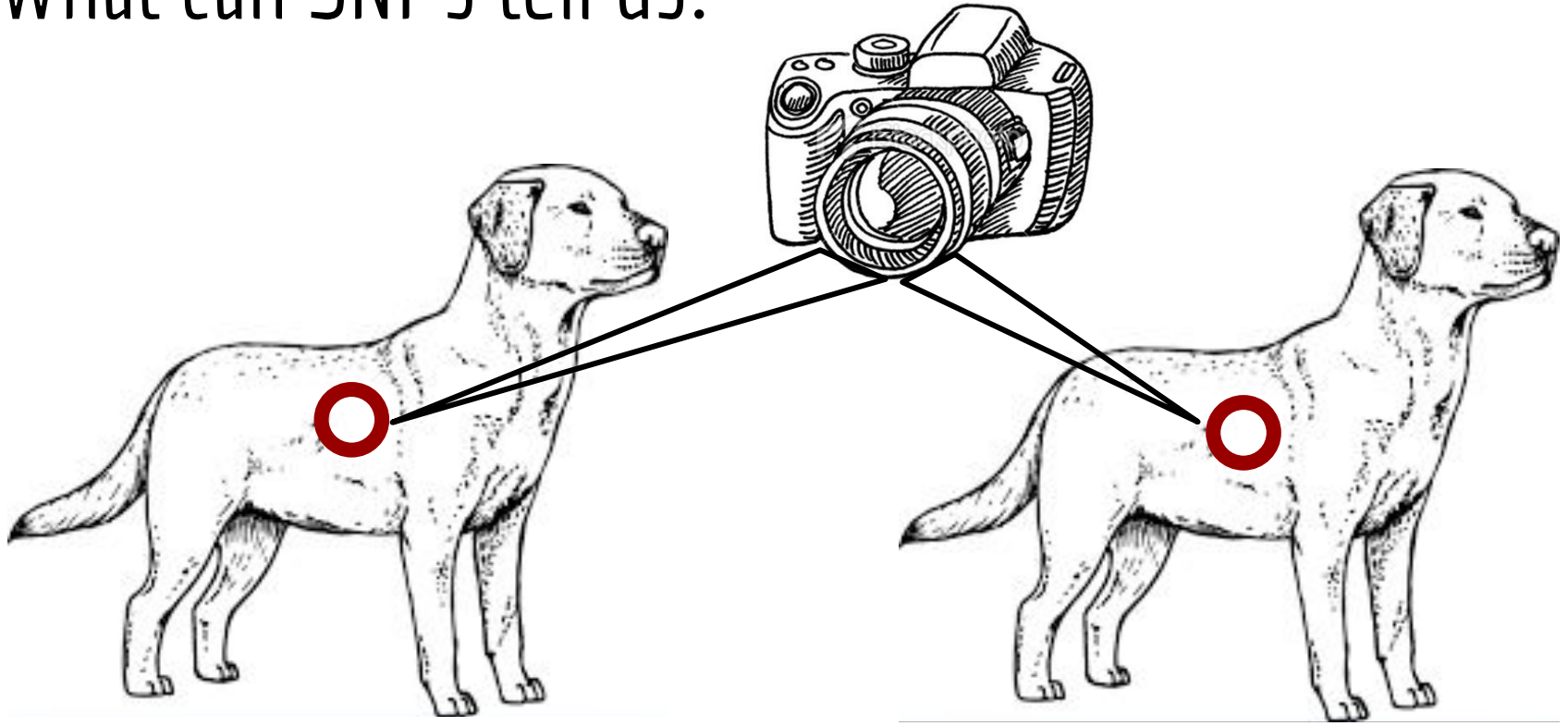
G  
G

So if we don't look at a site with a SNP, we won't know it's there.

# What can SNPs tell us?

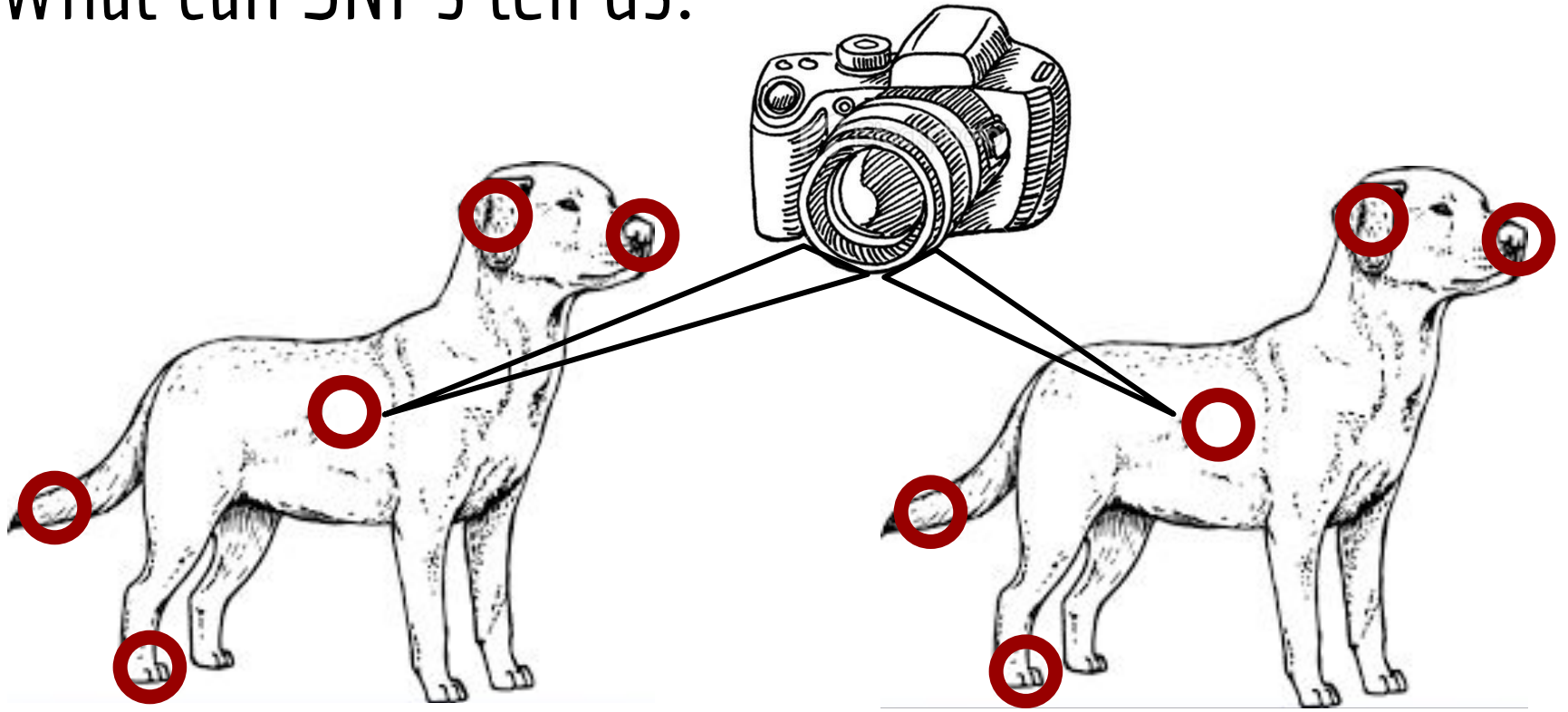


# What can SNPs tell us?

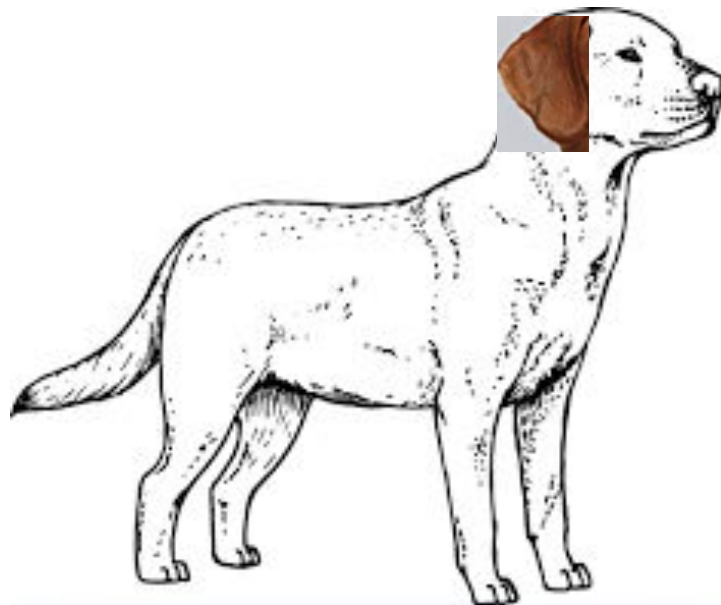
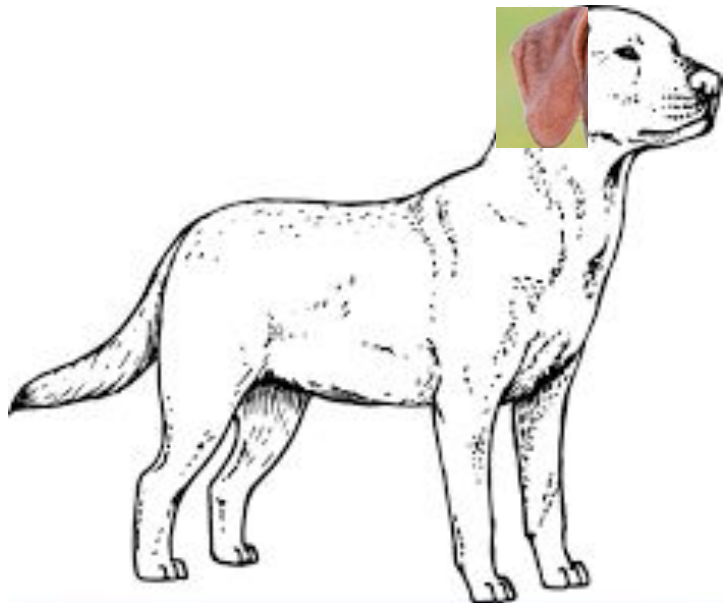




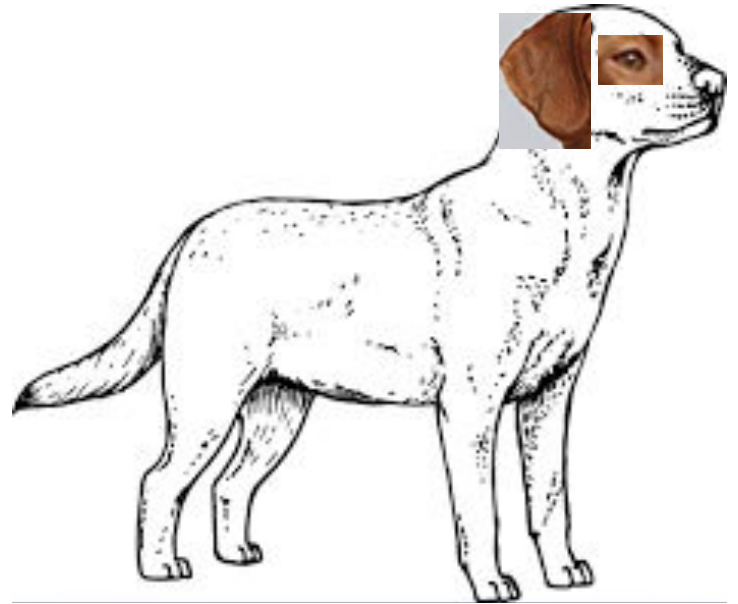
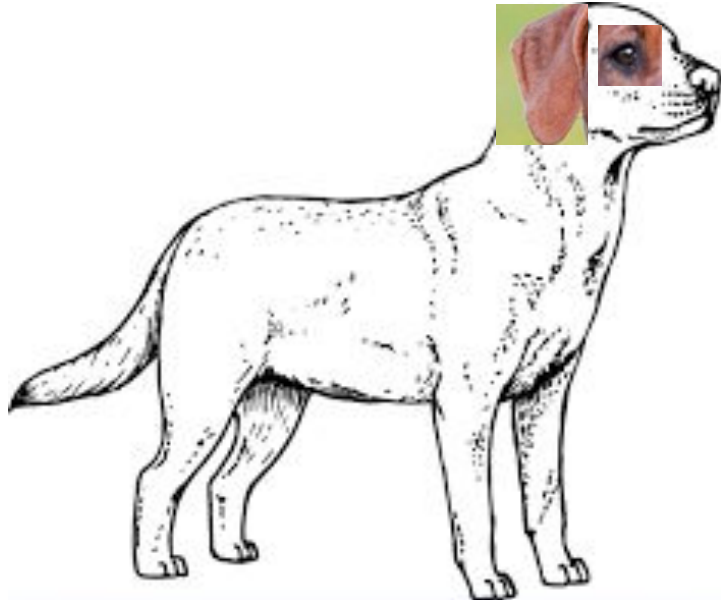
# What can SNPs tell us?



# What can SNPs tell us?

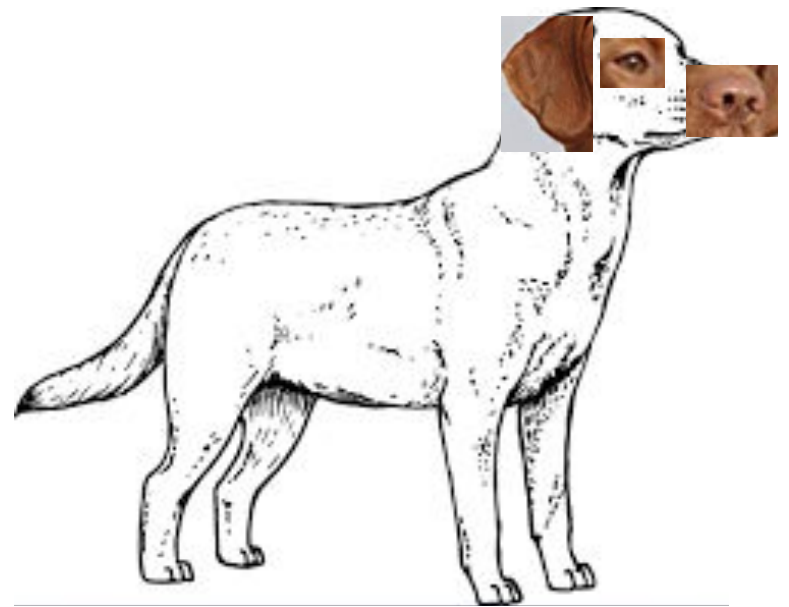
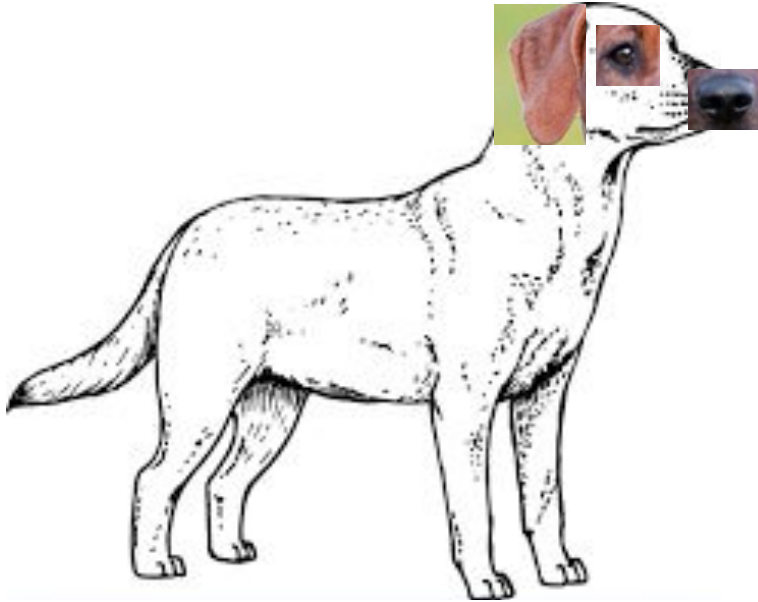


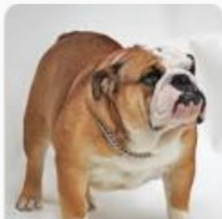
# What can SNPs tell us?



# What can SNPs tell us?

Different nose color!





Bulldog



German Shepherd



Labrador Retriever



Beagle



Poodle



Chihuahua



Golden Retriever



Pug



Border Collie



Newfoundland dog



American Staffordshire ...



Greyhound



Australian Shepherd



Pembroke Welsh Corgi



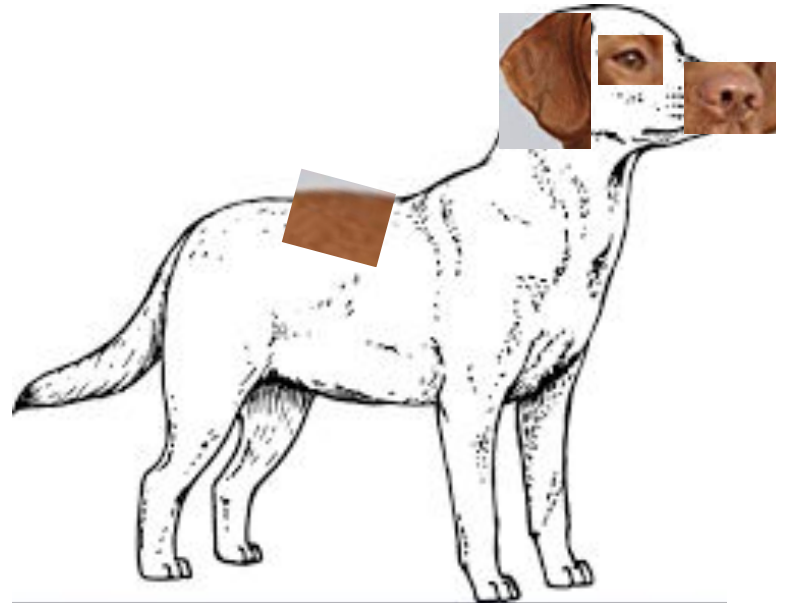
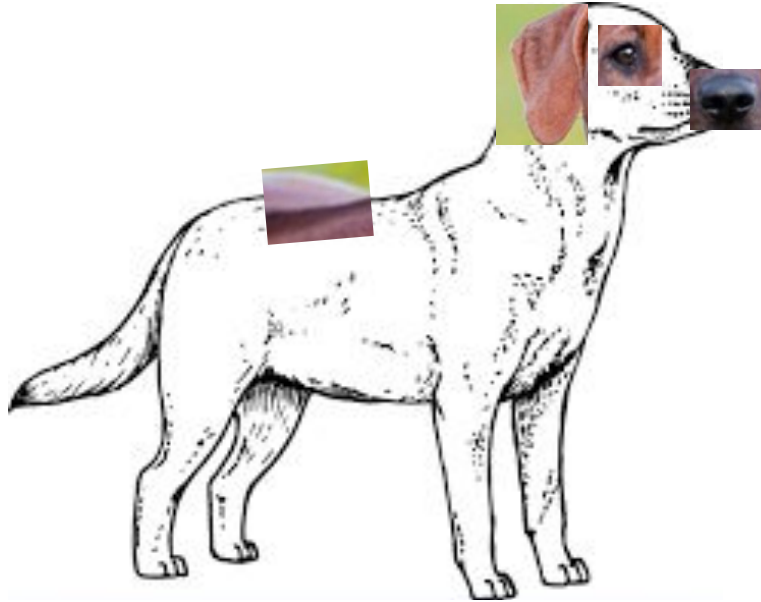
Bull Terrier



Old English Sheepdog

Dog breeds, via Google Images, all black noses!

# What can SNPs tell us?



# What can SNPs tell us?

**Rhodesian Ridgeback**



# What can SNPs tell us?

We can look for associations between SNPs or patterns of SNPs and:

Breeds

Traits

Diseases








First Two Traits Mapped With SNP Chip in Dogs. Two traits were mapped as a proof of principle test of the canine SNP chip. (a) White coat color in boxers was mapped to the gene *MITF*, a developmental gene involved in pigimentary and auditory disorders, using 10 white dogs and 9 pigmented dogs. (b) The dorsal hair ridge in the Rhodesian ridgeback breed was mapped to a duplication of 4 genes, including three fibroblast growth factor genes, using 12 ridged dogs and 9 ridgeless dogs

# SNP Arrays

<b>Dog:</b>	1	2	3	4	5	...
SNP1	AG	AA	GG	AA	AA	
SNP2	CC	CA	CC	CC	CC	
SNP3	AA	AA	TT	AA	AT	
SNP4	AT	TT	TT	TT	TT	
SNP5	GG	GG	GG	GG	GG	
SNP6	GG	CC	GG	GG	GG	
...						

We'll only consider **biallelic** SNPs; this means there are 3 possibilities per SNP

AA, Aa, aa  minor allele

Say we want to differentiate 100 dog breeds. How many SNPs do we need?

# SNP Arrays

<b>Dog:</b>	1	2	3	4	5	...
SNP1	AG	AA	GG	AA	AA	
SNP2	CC	CA	CC	CC	CC	
SNP3	AA	AA	TT	AA	AT	
SNP4	AT	TT	TT	TT	TT	
SNP5	GG	GG	GG	GG	GG	
SNP6	GG	CC	GG	GG	GG	
...						

We'll only consider biallelic SNPs; this means there are 3 possibilities per SNP

AA, Aa, aa

Say we want to differentiate 100 dog breeds. **How many SNPs do we need?**

# SNP Arrays

<b>Dog:</b>	1	2	3	4	5	...
SNP1	AG	AA	GG	AA	AA	
SNP2	CC	CA	CC	CC	CC	
SNP3	AA	AA	TT	AA	AT	
SNP4	AT	TT	TT	TT	TT	
SNP5	GG	GG	GG	GG	GG	
SNP6	GG	CC	GG	GG	GG	
...						

genotype possibilities

$$3^x \geq 100$$
$$x = 5$$

What's the problem here?

Say we want to differentiate 100 dog breeds. How many SNPs do we need?

# SNP Arrays

<b>Dog:</b>	1	2	3	4	5	...
SNP1	AG	AA	GG	AA	AA	
SNP2	CC	CA	CC	CC	CC	
SNP3	AA	AA	TT	AA	AT	
SNP4	AT	TT	TT	TT	TT	
SNP5	GG	GG	GG	GG	GG	
SNP6	GG	CC	GG	GG	GG	
...						

This would require us to see every  
combo of possible SNP patterns!

Related breeds likely to have similar  
patterns/differ in fewer SNPs

A given SNP might only be  
informative for some (or no) breeds



# SNP Arrays

2.5 million documented dog SNPs (~10% of the genome)

Data we'll be working with is 148,737 SNPs (~0.005% of the genome)

**How were the ~150k sites chosen?**

**How do we actually look at (genotype) these SNP sites?**

# SNP Arrays

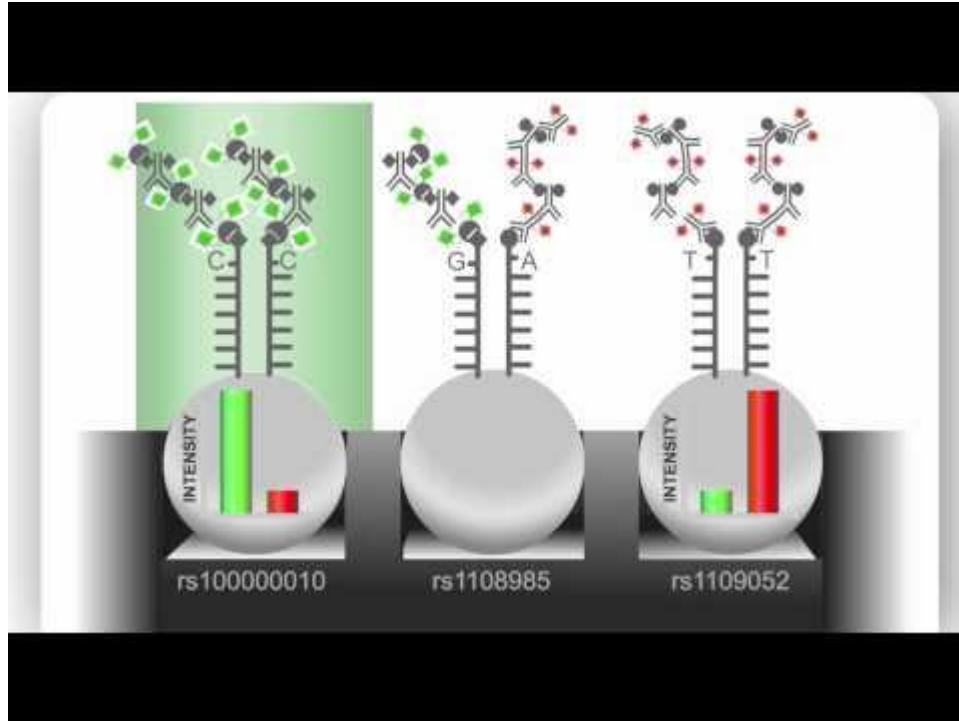
2.5 million documented dog SNPs (~10% of the genome)

Data we'll be working with is 148,737 SNPs (~0.005% of the genome)

**How were the ~150k sites chosen?**

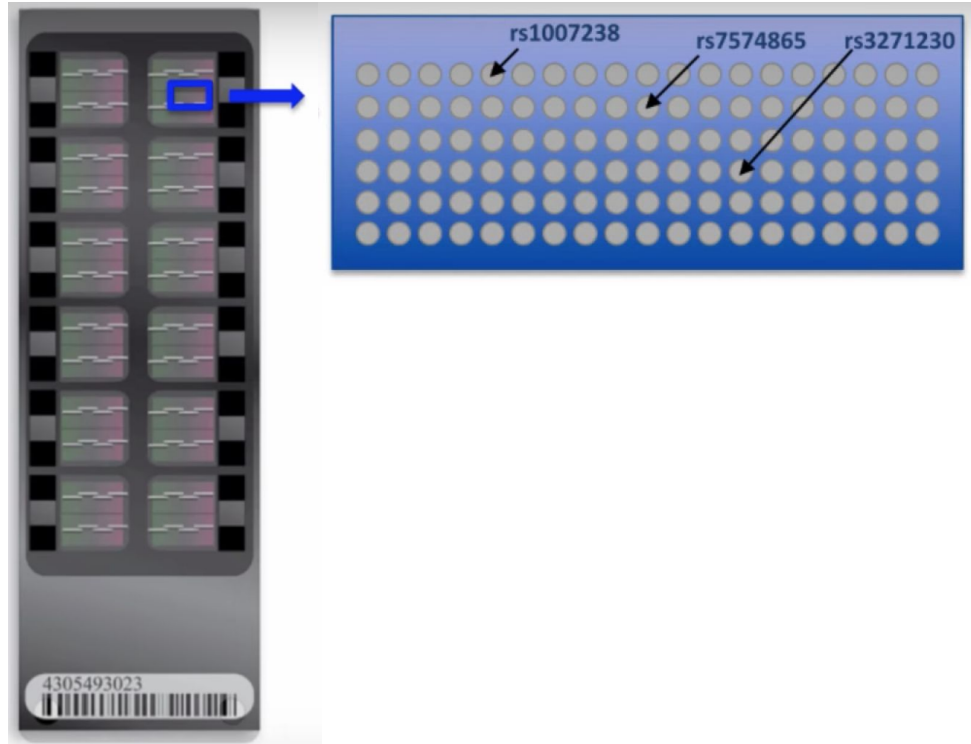
**How do we actually look at (genotype) these SNP sites?**

# Illumina "Bead" SNP Array

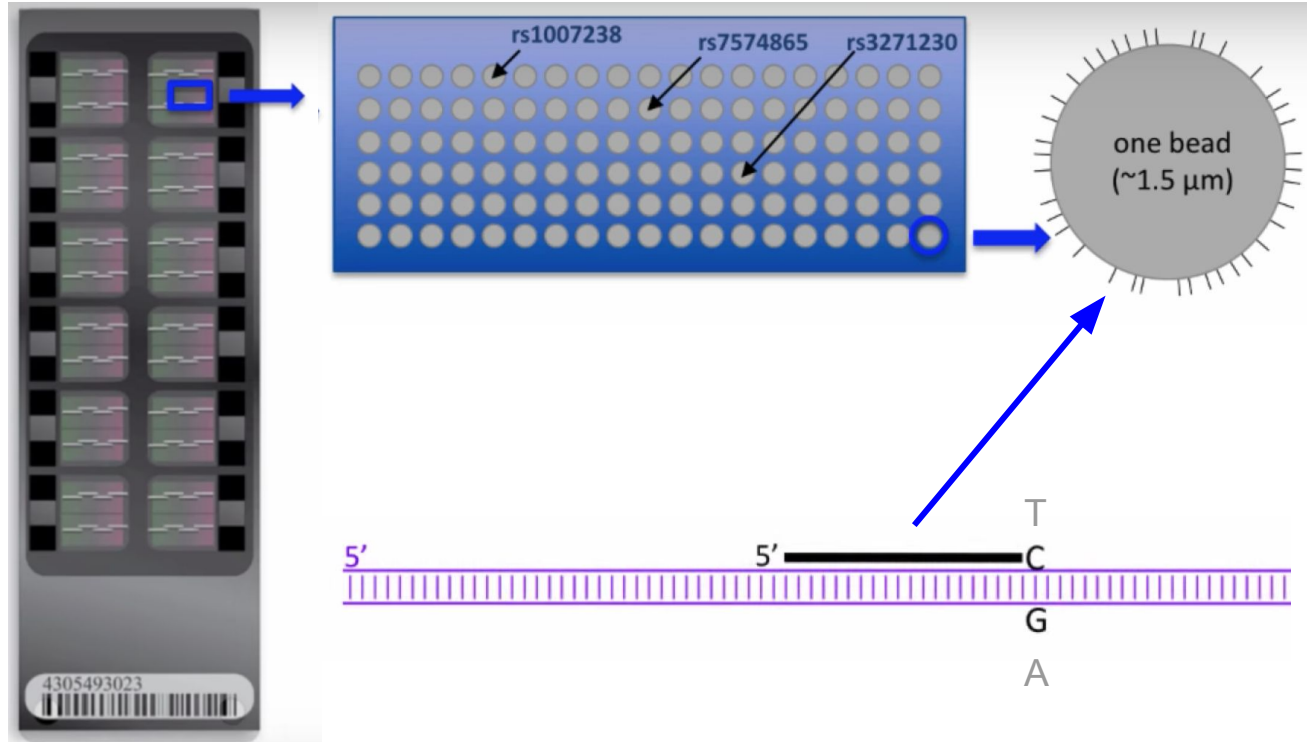




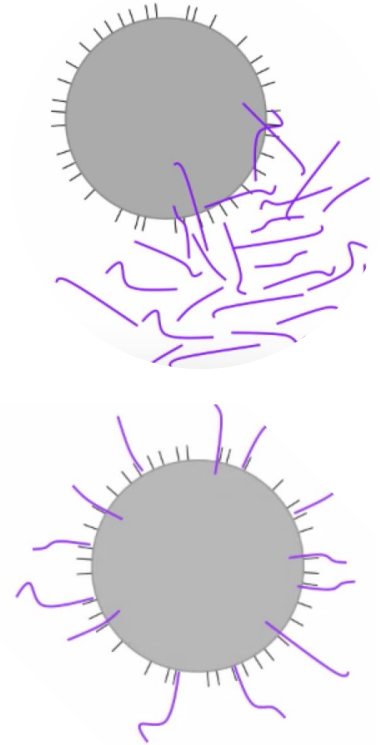
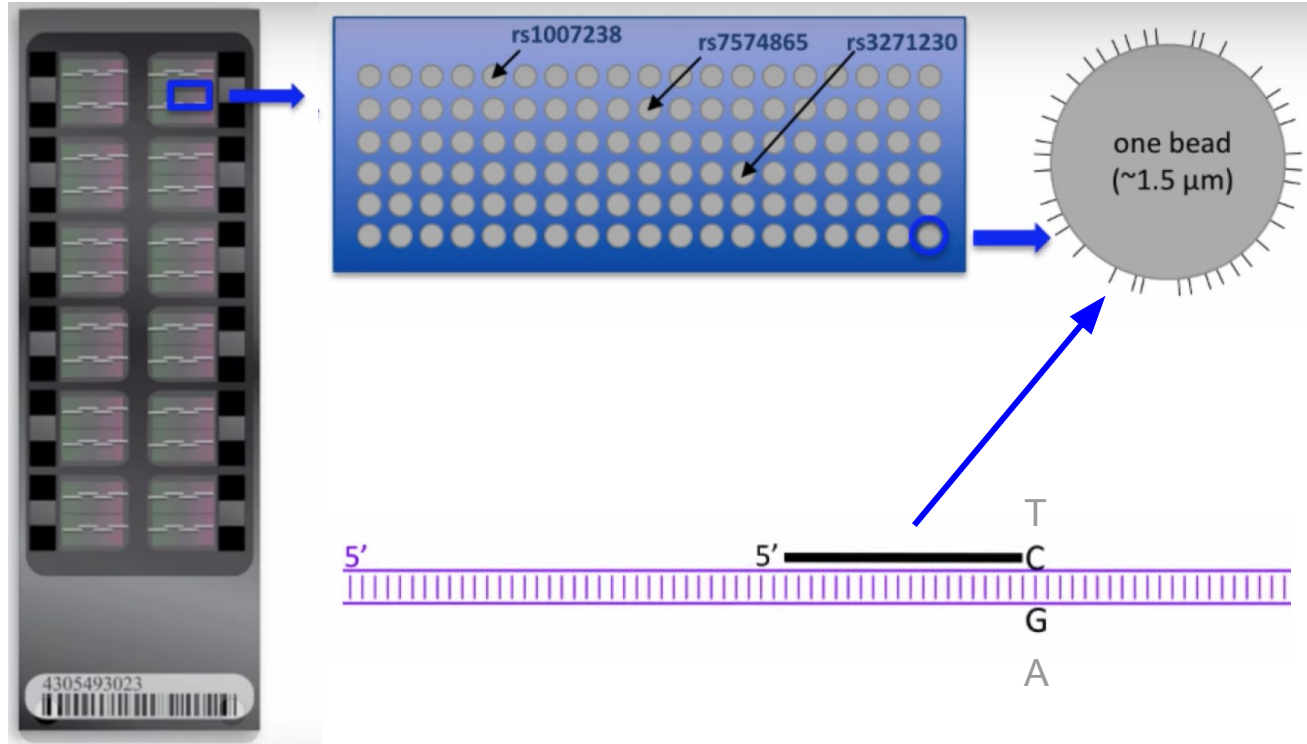
# Illumina Bead SNP Array



# Illumina Bead SNP Array

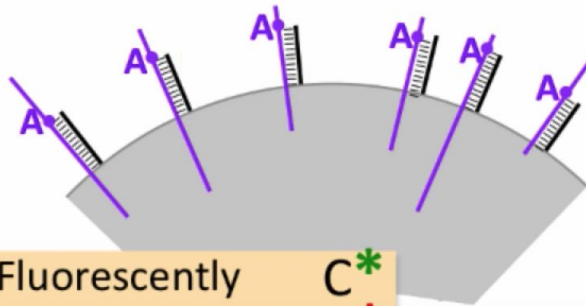


# Illumina Bead SNP Array



# Illumina Bead SNP Array

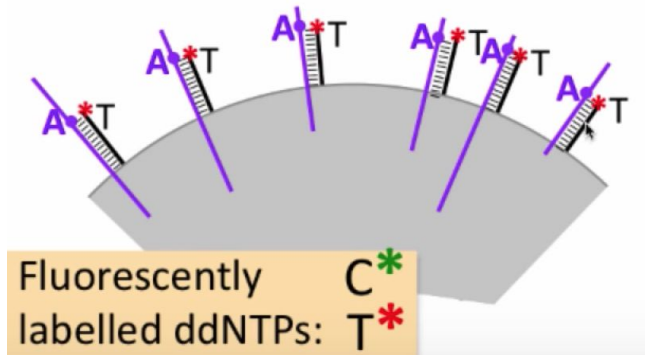
SNP  
genotype: **AA**



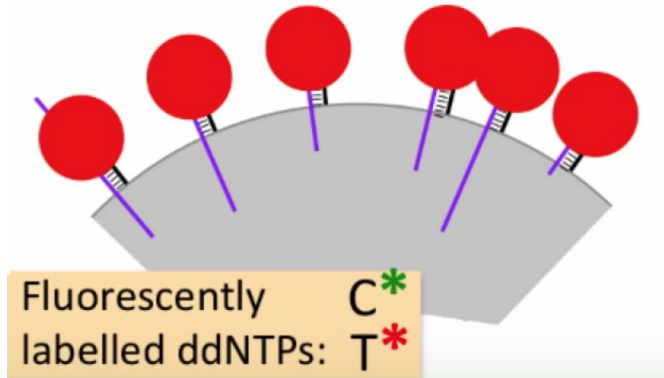
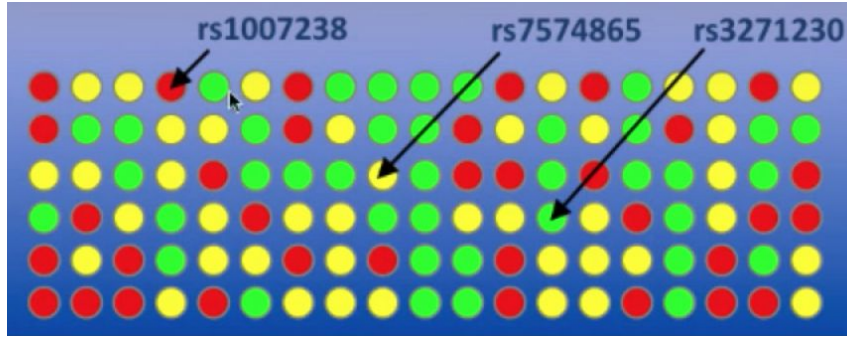
Fluorescently  
labelled ddNTPs: C\*  
T\*

# Illumina Bead SNP Array

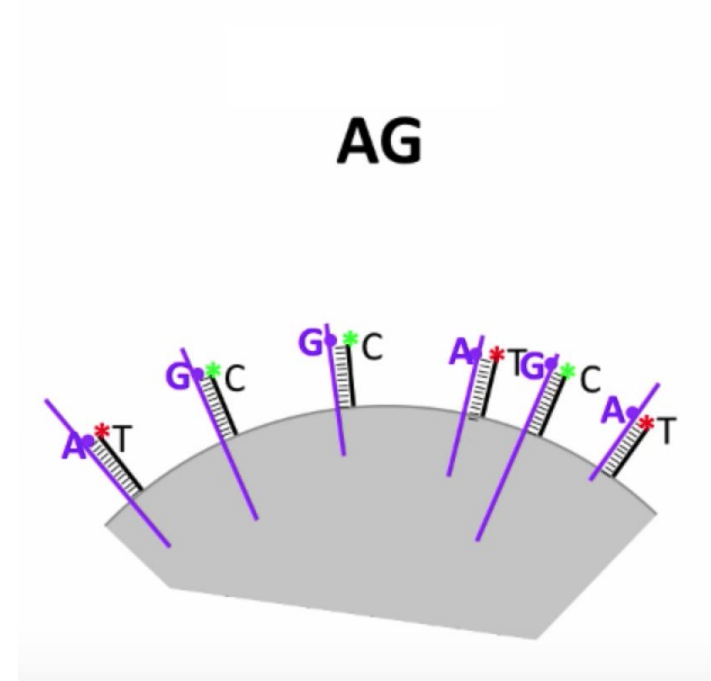
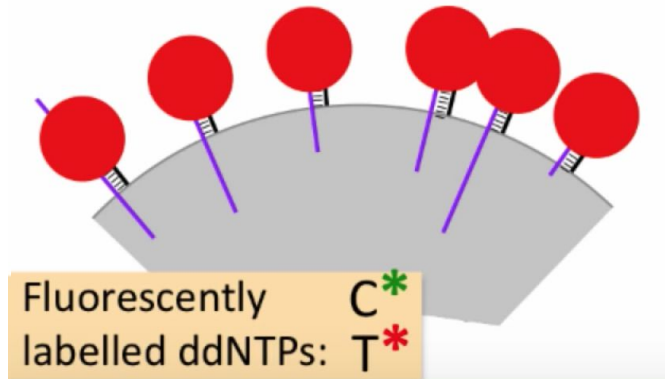
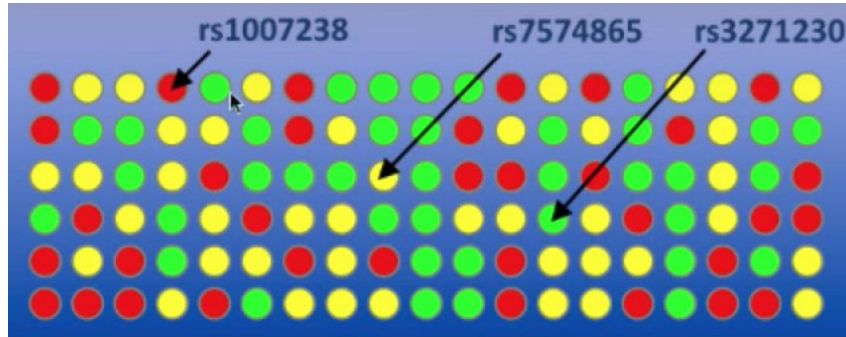
SNP  
genotype: **AA**



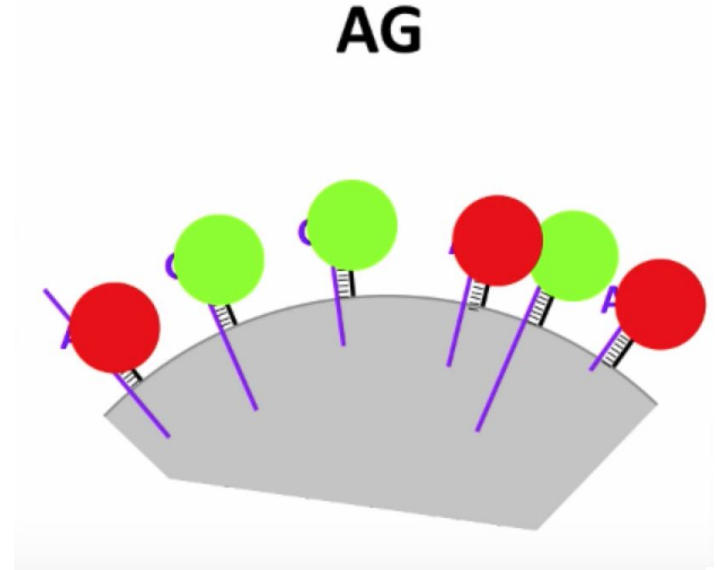
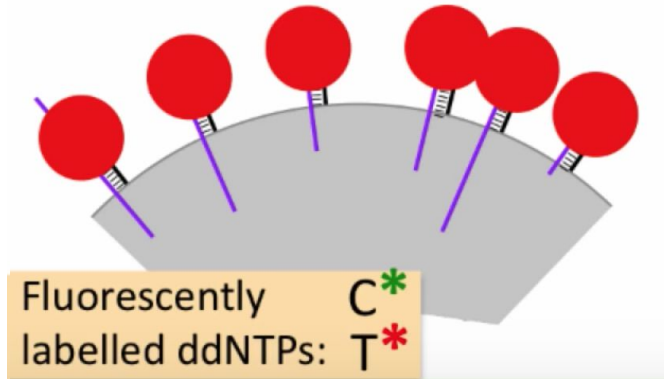
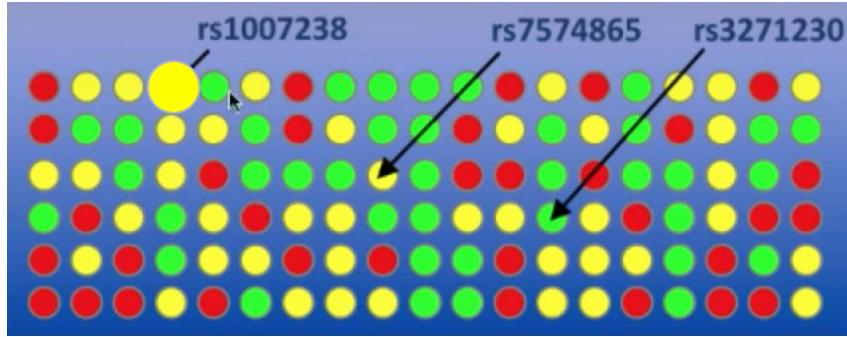
# Illumina Bead SNP Array



# Illumina Bead SNP Array

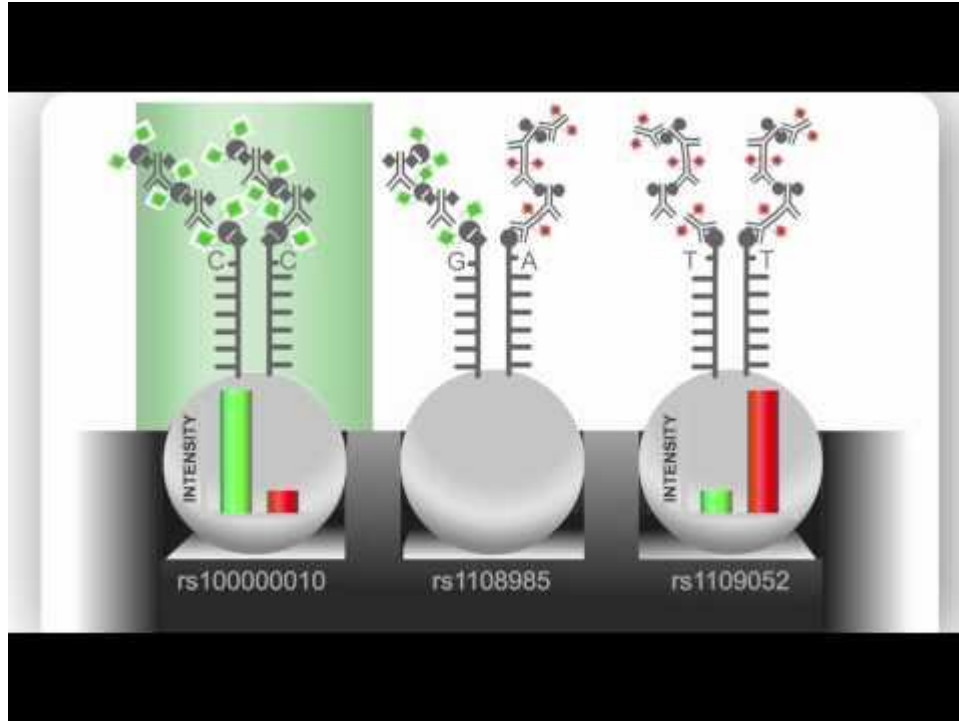


# Illumina Bead SNP Array





# Illumina Bead SNP Array Recap



# Before Wednesday

- Make sure you're on Piazza: [piazza.com/jhu/spring2019/en60114713](https://piazza.com/jhu/spring2019/en60114713)
- Fill out a (short!) survey on your background, and if this lecture was the right speed/level, posted on Piazza
- Read Atlantic article: *What Vets Think of '23andMe for Dogs'*
- Start looking for a partner!