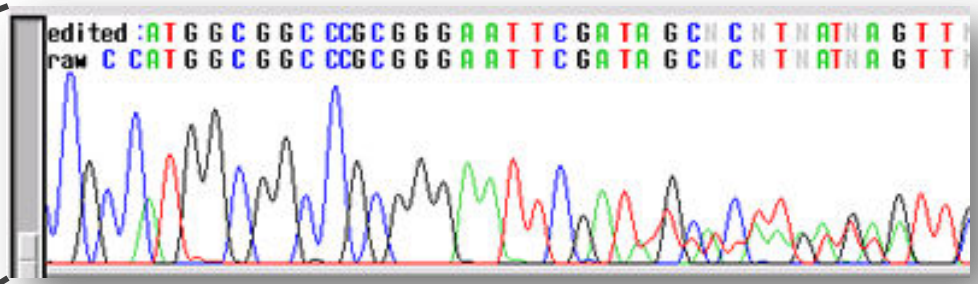
Molecular Markers



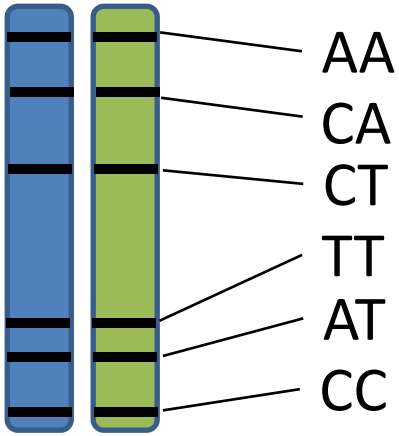
CRITFC Genetics Workshop
December 9, 2014

Molecular Markers

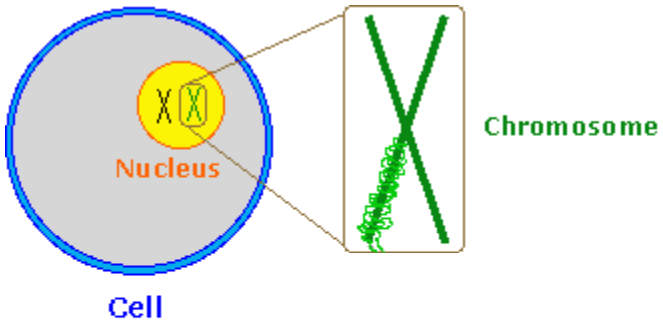
- Tools that allow us to collect information about an individual, a population, or a species
- Application in fisheries
  - mating behaviors/effective population size
  - population structure
  - parentage, genetic tagging
  - genetic basis of traits/gene mapping

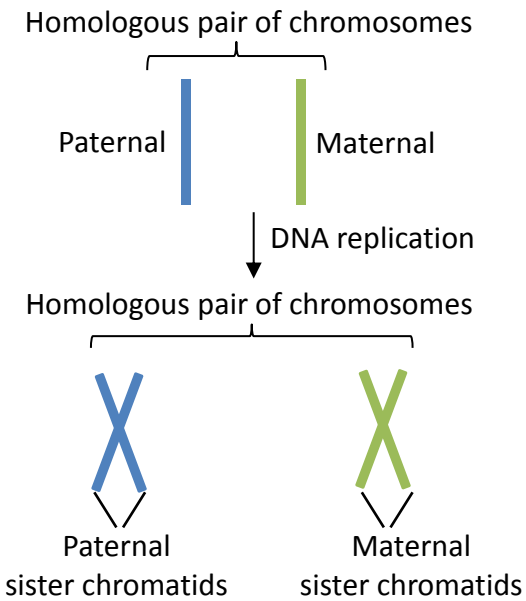
Molecular Markers

- Designed to interrogate heritable differences in DNA sequence called “polymorphisms”
- Individuals may inherit unique combinations of polymorphisms/alleles across several loci providing a distinctive DNA ID or fingerprint

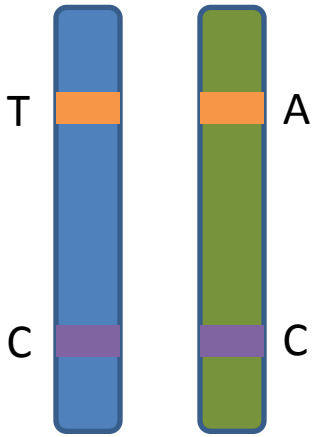


Molecular Markers





Homologous pair of chromosomes



Locus – physical location within the DNA sequence

Allele – a single variant of a locus

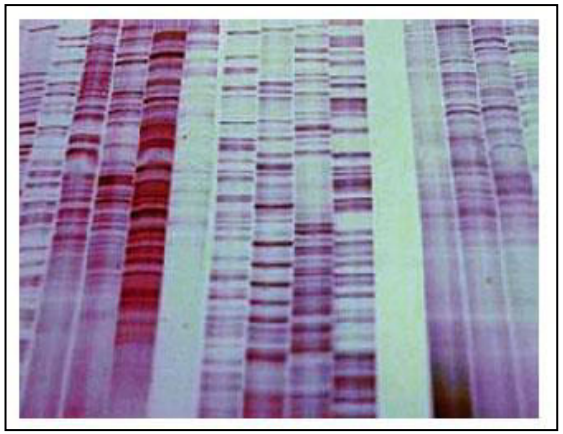
Genotype – a set of 2 alleles at a locus

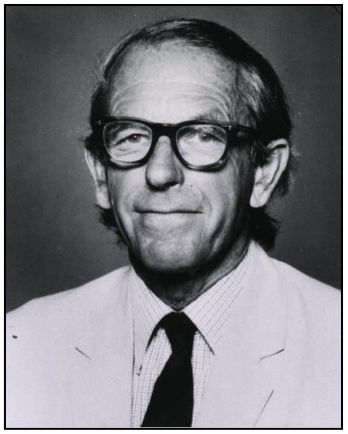
2 loci:

- TA = heterozygous
- CC = homozygous

DNA Sequence and Genetic Markers

- 1977 Sanger and colleagues describe laboratory methods of DNA sequencing





- Allowed for the direct interrogation of variation at the DNA sequence level which lead to the discovery of informative genetic markers

http://www.nature.com/nmeth/journal/v5/n1/full/nmeth1154.html

DNA Sequence Variation

Mistakes made during DNA replication leads to sequence variation (i.e., mutations)

DNA sequence of three individuals:

```
GCTCTCTCTCTCTCTCTCCCTCTCTGCTTACAGACGTTAACACAGCCAAGAAGAGCCAGGATACCAGTGATACTGTAGCCTAGTCCCCG
CTCTCTCTCTCTCTCTCCCTCTCTGCTTACAGACGTTAACACAGCCAAGAAGAGCCAGGATACCAGTGATACTGTAGCCTAGTCCCCG
CTCTCTCTCTCTCTCTCCCTCTCTGCTTACAGACGTTAACACAGCCAAGAAGAG::::::::CCAGTGATACTGTAGCCTAGTCCCAG
```

microsatellite

- di-nucleotide "CT" repeat
- di- and tetra-nucleotide markers common

insertion/deletion (indel)

- Ex. 8 bp deletion
- 1 or more nucleotides

SNP

- Any variation of the 4 nucleotides

Microsatellite Markers

Repeating sequences of 2-6 base pairs of DNA
and can be hyper-variable compared to other markers

microsatellite

Flanking DNA Flanking DNA

1 

TGCCGTGCATATATCGAGCTATT (3X)=23bp
TGCCGTGCATATATCGAGCTATT (3X)=23bp

2 

TGCCGTGCATATATATATATCGAGCTATT (6X)=29bp
TGCCGTGCATATATATATATCGAGCTATT (6X)=29bp

3 

TGCCGTGCATATATATATCGAGCTATT (5X)=27bp
TGCCGTGCATATATATATCGAGCTATT (5X)=27bp

4 

TGCCGTGCATATATCGAGCTATT (3X)=23bp
TGCCGTGCATATATATATATATATCGAGCTATT (8X)=33bp

5 

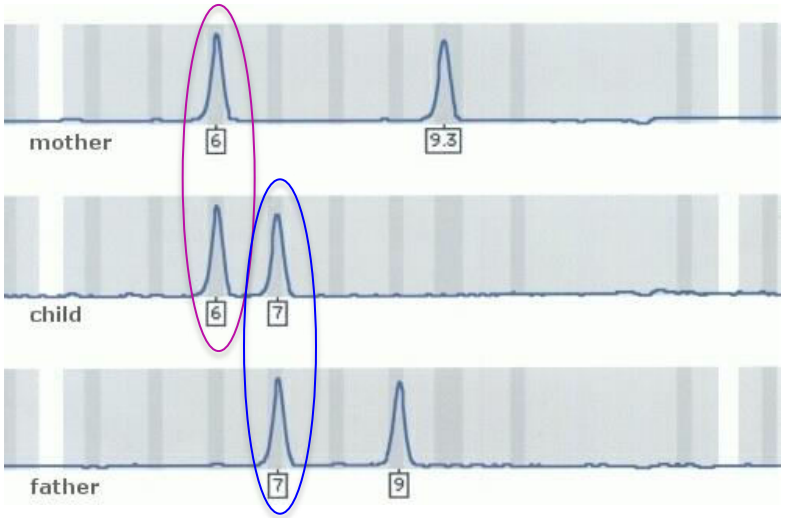
TGCCGTGCATATATATATATATATCGAGCTATT (8X)=33bp
TGCCGTGCATATATATATATATATATCGAGCTATT (9X)=35bp

Flanking DNA is identical, but <u>length</u> of microsatellite is different between individuals

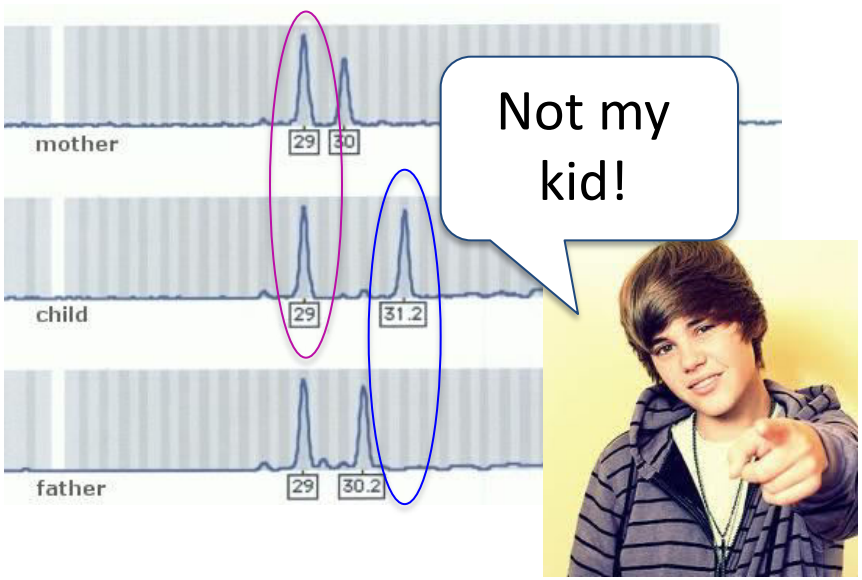
Microsatellite Markers

- Example: Paternity testing

Case 1



Case 2



http://www.paternity.be/information_EN.html

Microsatellite Markers

- Pros
  - Extremely variable
    - > 20 alleles for one locus
  - Moderately abundant in genome
  - Predominantly neutral loci
- Cons
  - Difficult to discover
    - Generally requires DNA sequence information
  - Difficult to standardize and exchange data across labs
    - PCR variation
    - Scoring variation

SNP Markers

Single nucleotide polymorphism
https://www.youtube.com/watch?v=tJjXpiWKMyA

- Variation in DNA, when a single nucleotide (A, T, C, or G) within a given sequence differs between homologous chromosomes or between individuals at homologous loci

```
ATG GCT TCG ATC GAT CTA
ATG GCC TCG ATC GAT CTA
```

```
ATG GCT ACG ATC GAC CTA
ATG GCT ACG ATC GAC CTA
```

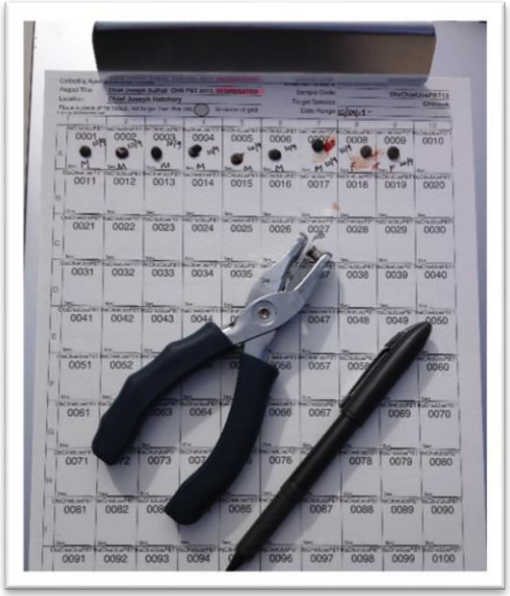
SNP Markers

- Pros
  - Codominant markers
  - Most abundant markers in the genome
  - Easy to interrogate with current high-throughput technology
  - requires little tissue
  - Highly reproducible between labs, easy to standardize, easy exchange of data
  - Can be adaptive and neutral loci

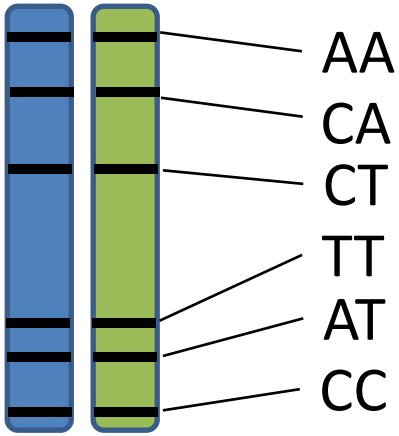
- Cons
  - Not as variable as microsatellites
  - High up front discovery/operating cost

Questions?

Laboratory techniques



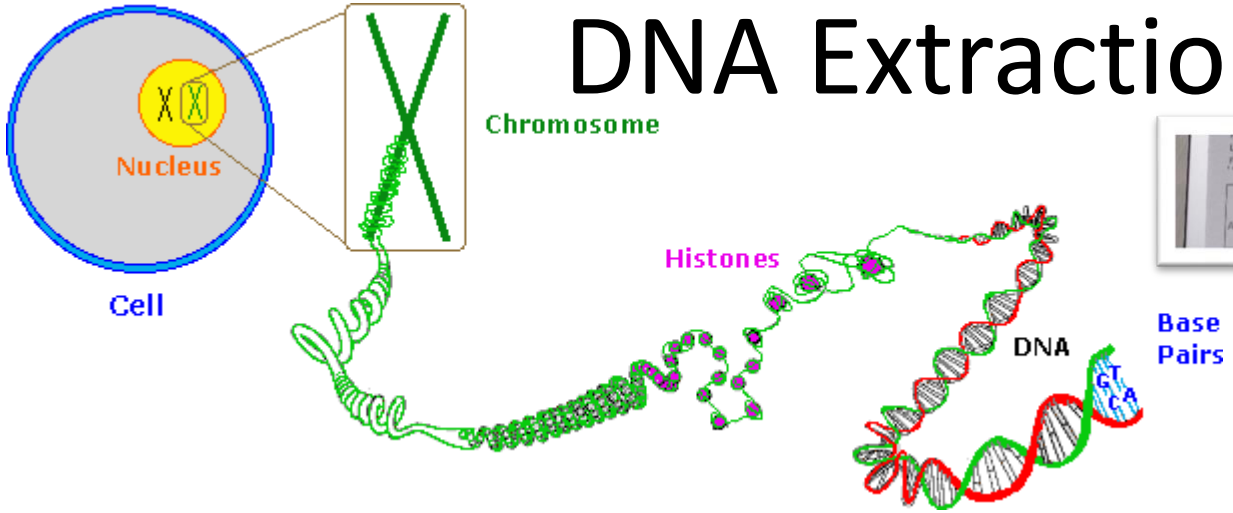
How do we get from tissue sample to multi-locus genotype?

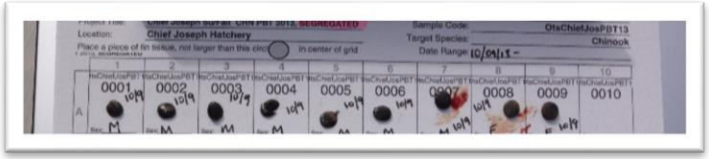


1.) DNA extraction

2.) Amplify specific regions of DNA (i.e., molecular markers) via PCR

DNA Extraction

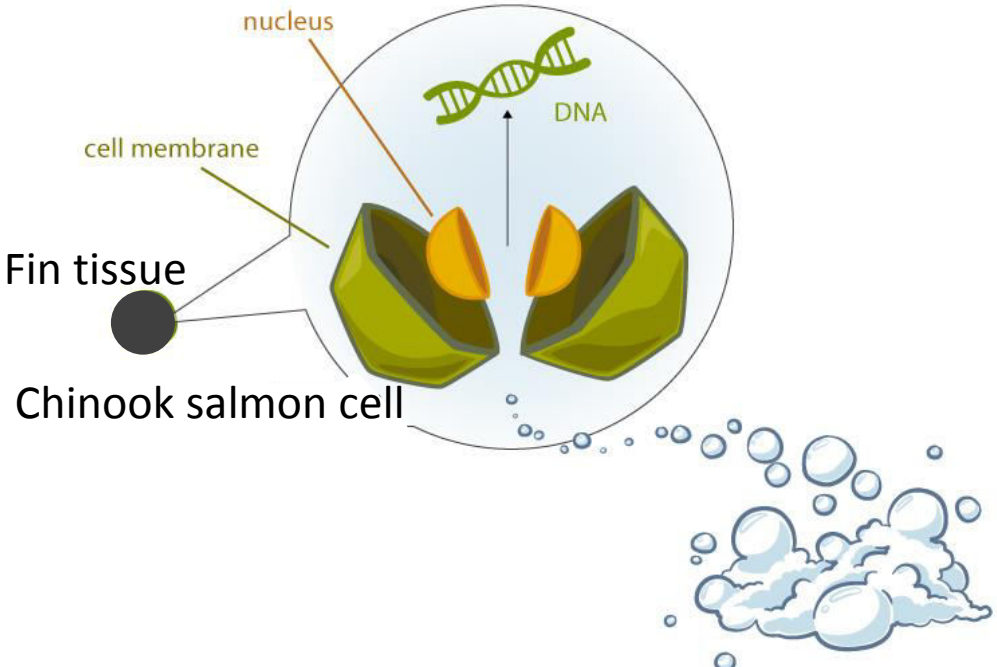




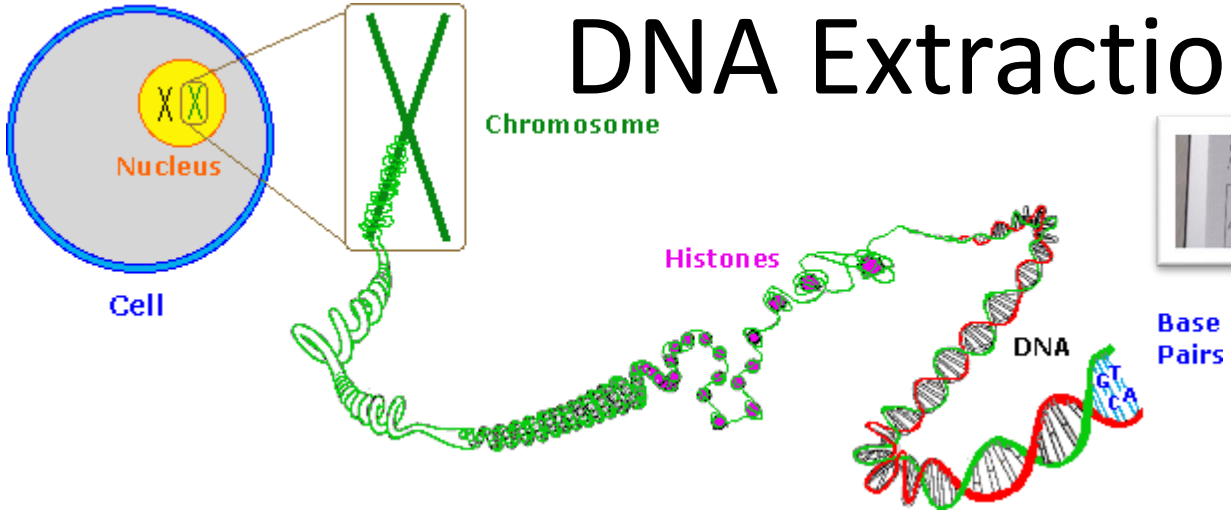
Ingredients

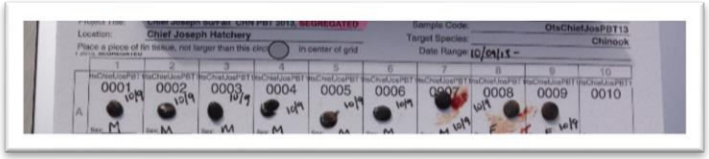
- Tissue
- Detergent & Salt
- Protease
- Alcohol

Break up cell walls, disrupt proteins



DNA Extraction

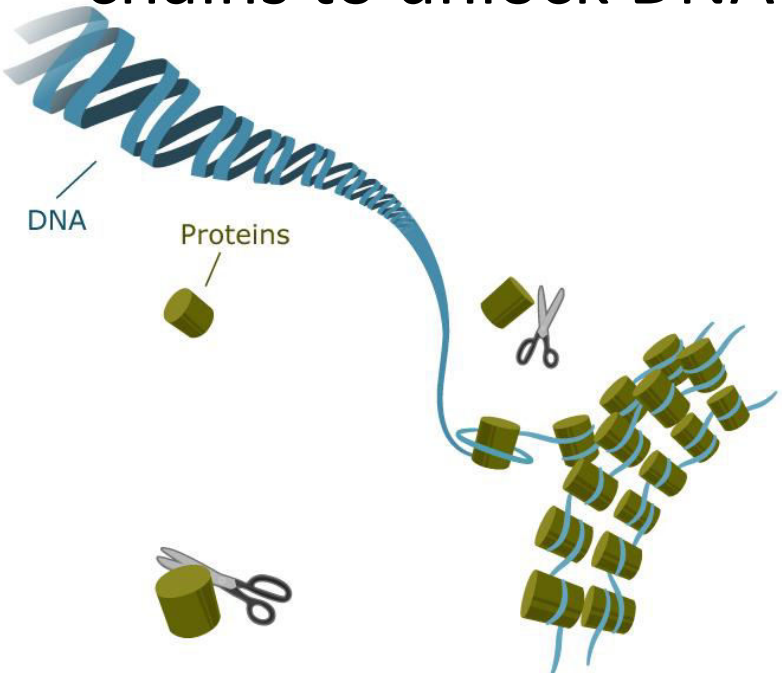




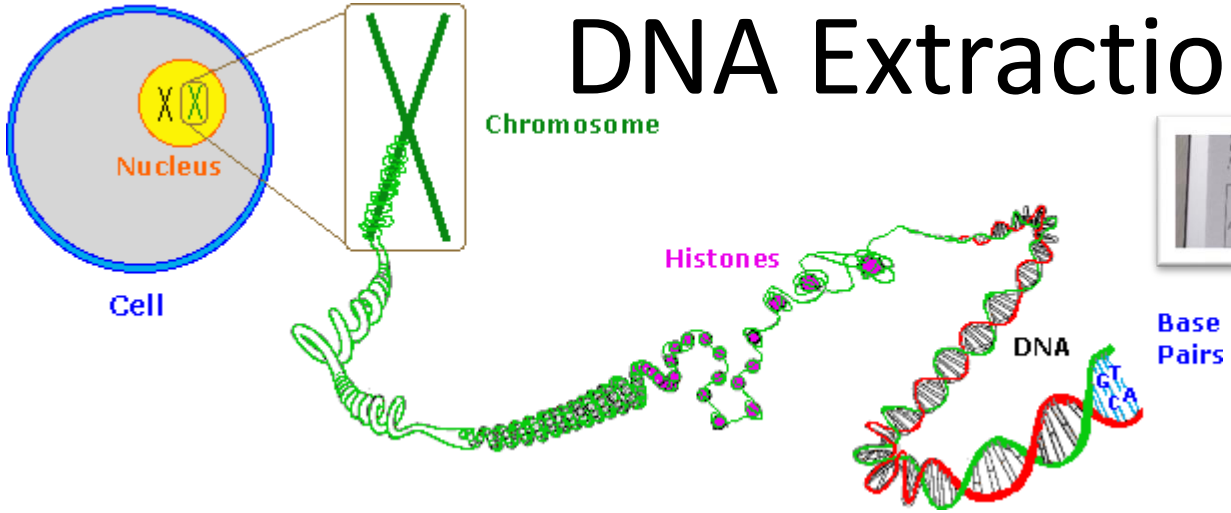
Ingredients

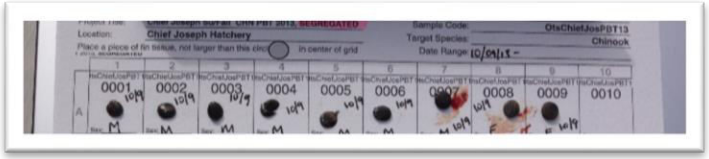
- Tissue
- Detergent & Salt
- Protease
- Alcohol

Emzyme cuts amino acid chains to unlock DNA



DNA Extraction

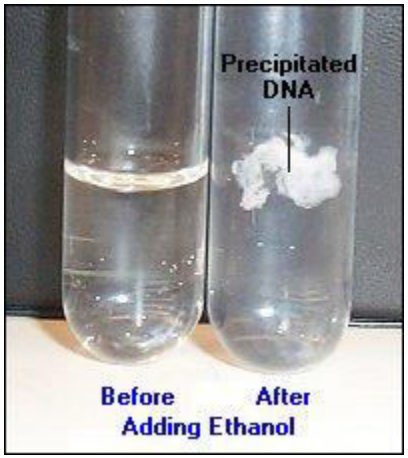




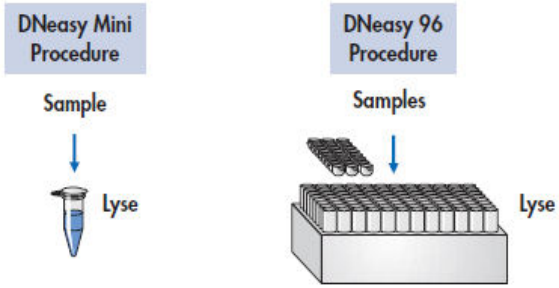
Ingredients

- Tissue
- Detergent & Salt
- Protease
- Alcohol

DNA is not soluble in ethanol, and will precipitate



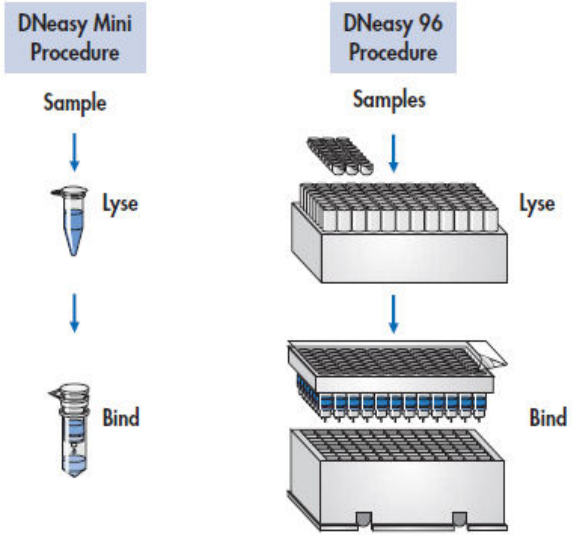
DNA Extraction Qiagen Dneasy kits



- Small piece of fin clip is added to lysis buffer
- Tissue is digested

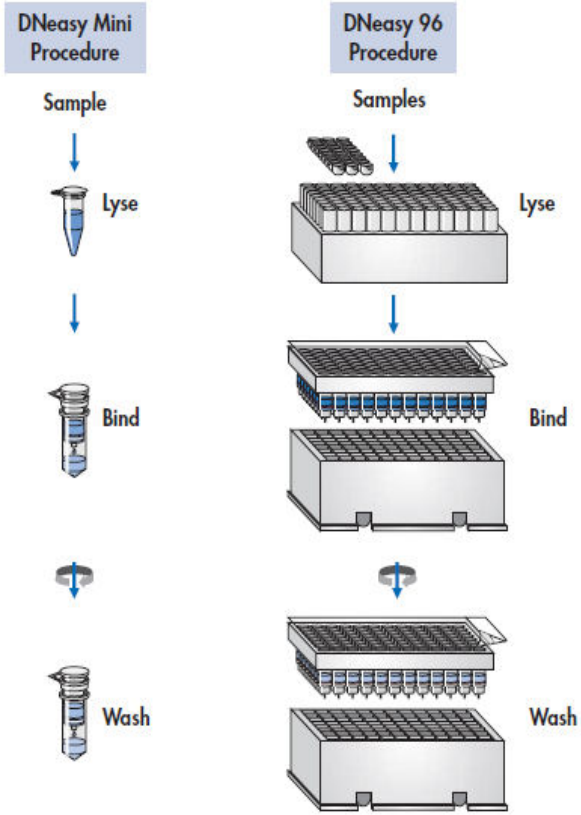
DNA Extraction Qiagen Dneasy kits

- Small piece of fin clip is added to lysis buffer
- Tissue is digested
- Add solution containing alcohol
- Precipitated DNA is caught in filter columns



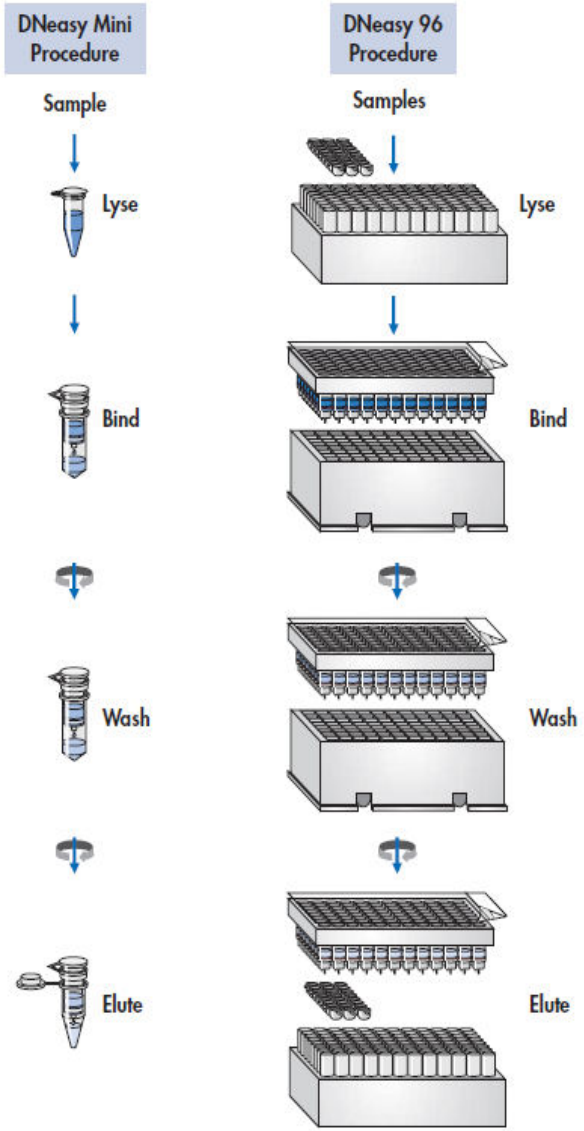
DNA Extraction Qiagen Dneasy kits

- Small piece of fin clip is added to lysis buffer
- Tissue is digested
- Add solution containing alcohol
- Precipitated DNA is caught in filter columns
- 2 wash steps to remove salts



DNA Extraction Qiagen Dneasy kits

- Small piece of fin clip is added to lysis buffer
- Tissue is digested
- Add solution containing alcohol
- Precipitated DNA is caught in filter columns
- 2 wash steps to remove salts
- DNA is released from filter by dissolving in weak buffer

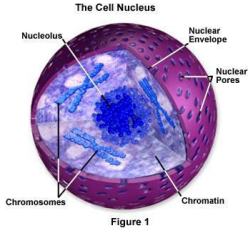


PCR song

PCR: Polymerase Chain Reaction

Review what happens in the cell (DNA replication): http://www.dnalc.org/resources/3d/03-mechanism-of-replication-basic.html

Generates millions of copies of a particular sequence (i.e., microsatellite, SNP) of DNA



PCR ingredients

- Genomic DNA
- Primers
- Taq polymerase
- Nucleotides (A, C, T, G)

Cell ingredients

- Genomic DNA
- Primase enzyme
- DNA polymerase
- Nucleotides (A, C, T, G)

PCR: Polymerase Chain Reaction

Amplifying a microsatellite

pcr.exe

Ingredients

- Genomic DNA
- Primers
- Enzyme (Taq)
- Nucleotides (A, C, T, G)

Tell enzyme where to initiate DNA synthesis

DNA from which you will amplify a molecular marker

Synthesizes DNA by adding nucleotides

ATGCAGGCTGAGAGAGACTAGTCGATG

CAGCTAC

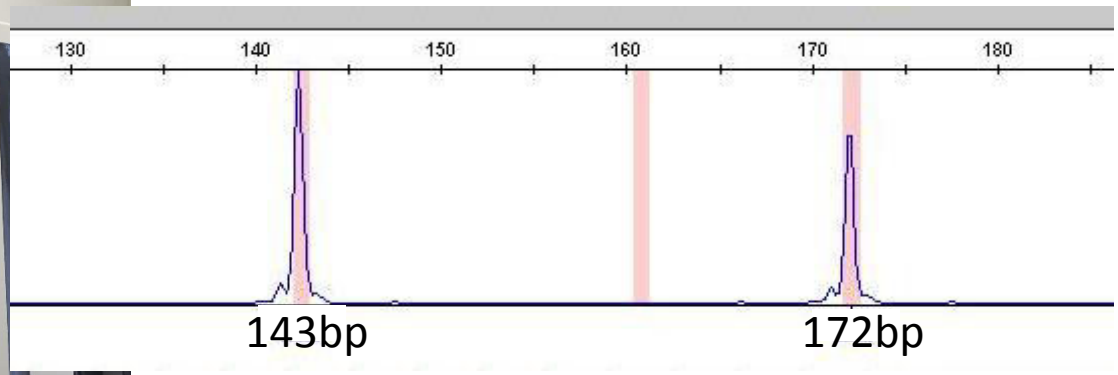
ATGCAGGCTGA

TACGTCCGACTCTCTCTGATCAGCTAC

PCR: Polymerase Chain Reaction

“Reading” a microsatellite

Sizes of PCR products determined by the 3730 DNA analyzer



	A	B	C	D	E	F	G	H	I	J	K
1	**Individual Name**	**usat1a**	**usat1b**	**usat2a**	**usat2b**	**usat3a**	**usat3b**	**usat4a**	**usat4b**	**usat5a**	**usat5b**
2	Sample 1	142	154	241	243	132	144	142	142	241	243
3	Sample 2	164	164	251	251	144	144	142	142	243	251
4	Sample 3	142	164	243	243	132	132	142	142	243	243
5	Sample 4	154	164	241	243	132	144	142	152	241	243
6	Sample 5	154	154	253	243	154	154	142	142	253	243
7	Sample 6	154	176	241	243	132	154	142	142	243	243

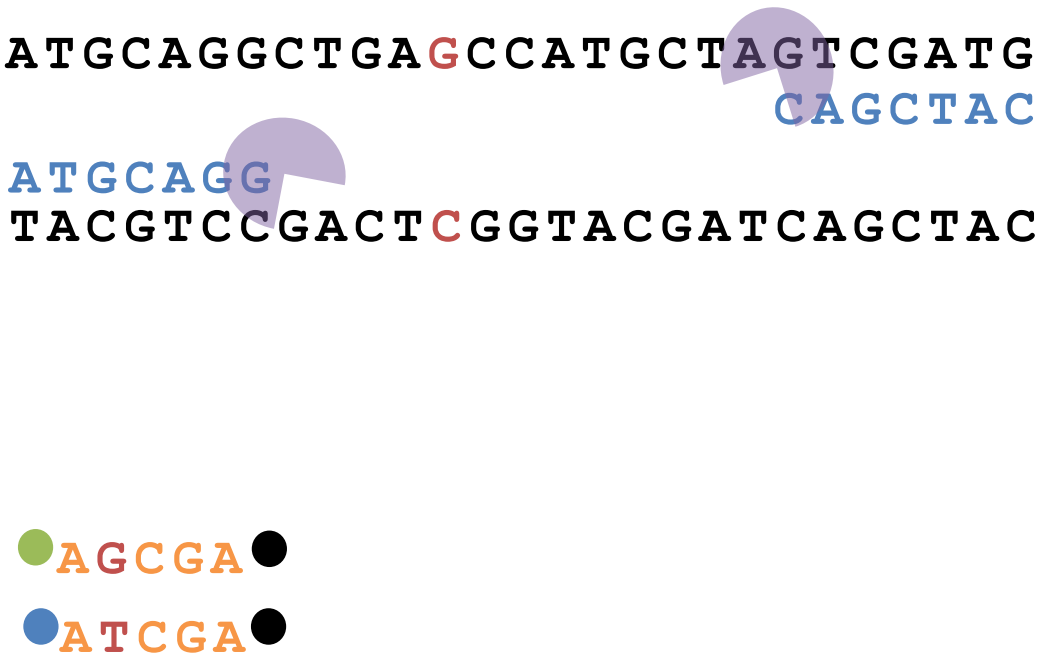
PCR: Polymerase Chain Reaction

Amplifying a SNP

http://www.lifetechnologies.com/us/en/home/life-science/pcr/real-time-pcr/qpcr-education/ask-taqman/ask-taqman-video-gallery.html

Ingredients

- Genomic DNA
- Primers
- Enzyme (Taq)
- Nucleotides (A, C, T, G)
- Fluorescent probes

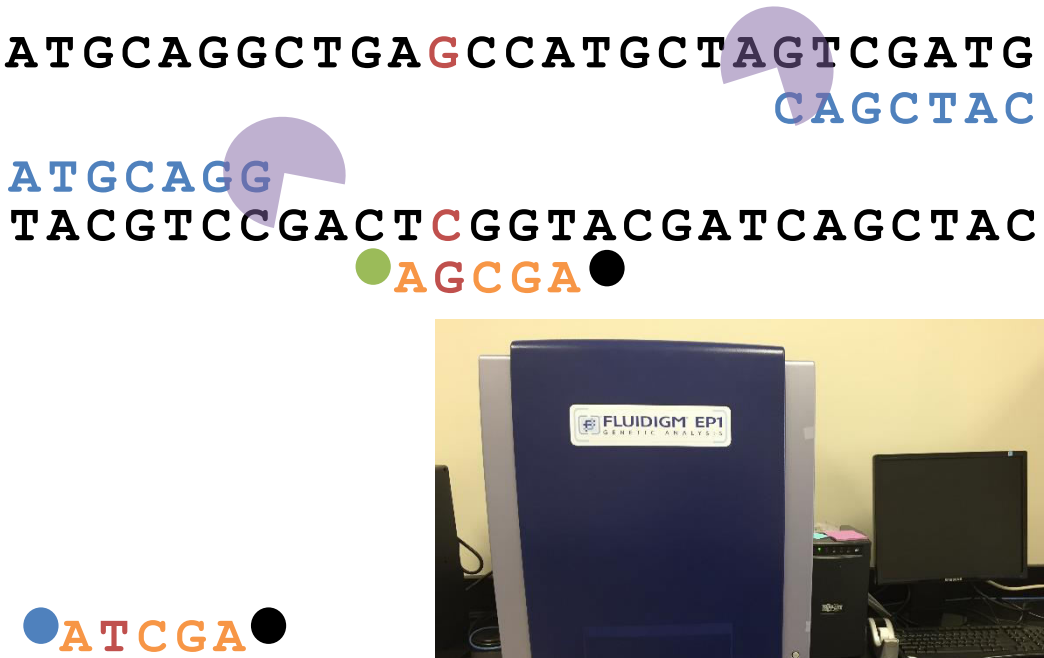


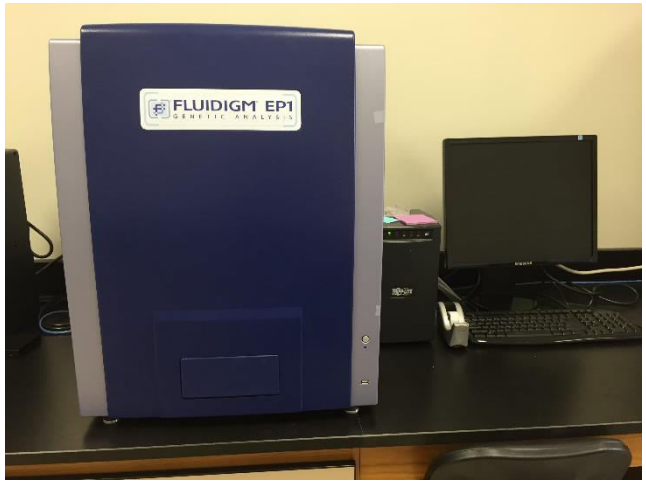
PCR: Polymerase Chain Reaction

Amplifying a SNP

Ingredients

- Genomic DNA
- Primers
- Enzyme (Taq)
- Nucleotides (A, C, T, G)
- Fluorescent probes



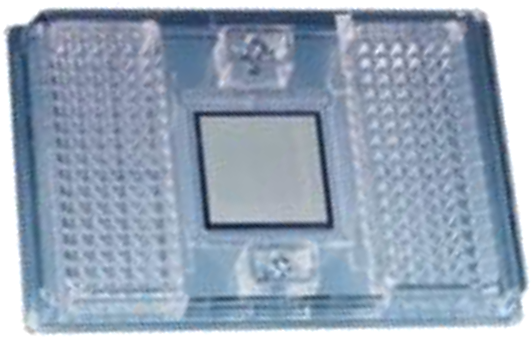


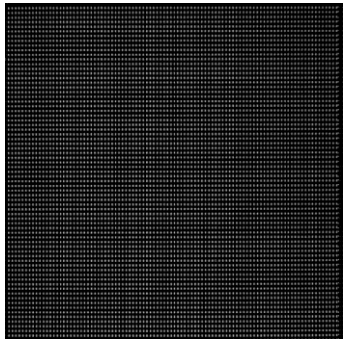
PCR: Polymerase Chain Reaction

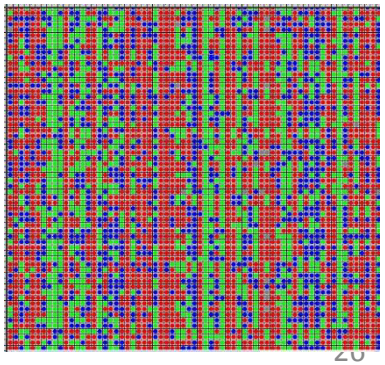
Amplifying a SNP

General Steps

- Pre-amplify DNA (boost starting copy; 14 cycles)
- 1 chip = 96 SNP assays, 96 samples
- Array is loaded with PCR reagents and the pre-amplified samples
- PCR resumes inside the chip for 50 cycles in the FC-1 thermal cycler
- Chip is then read using the camera inside the EP-1 instrument.

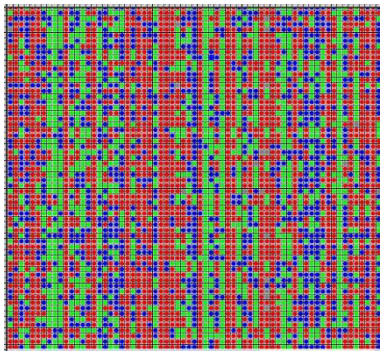




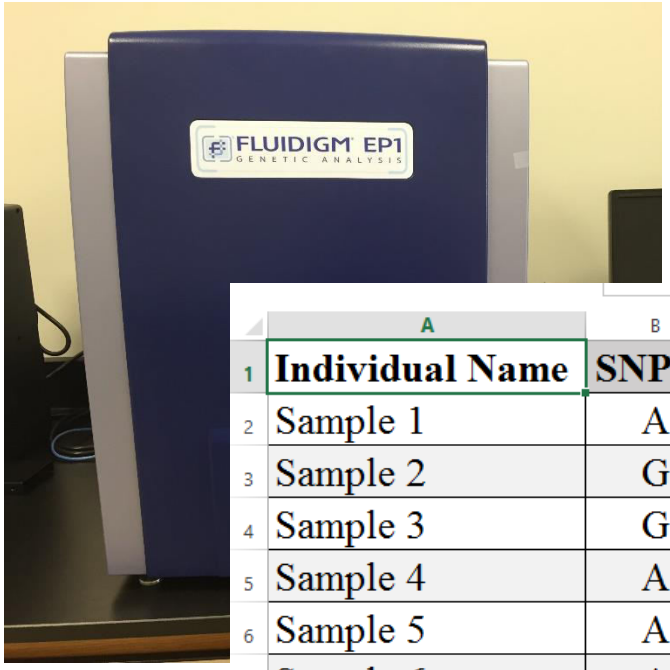


PCR: Polymerase Chain Reaction

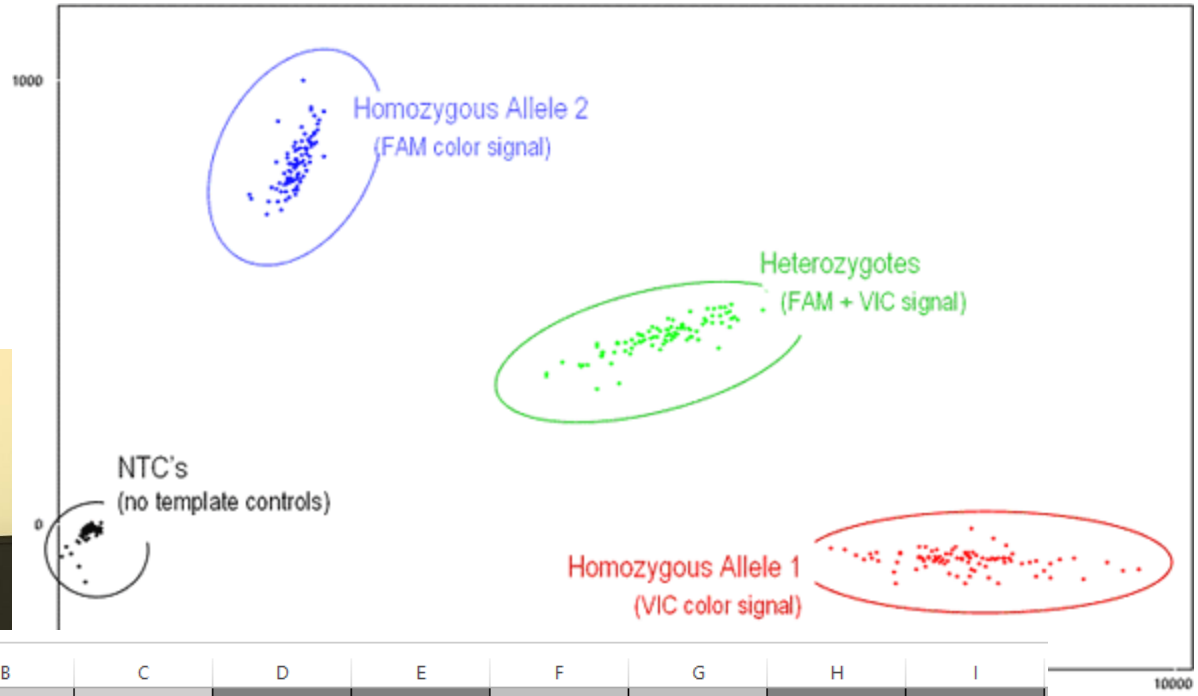
“Reading” a SNP



Fancy machine reads the chip



Example plot – 96 samples for 1 SNP

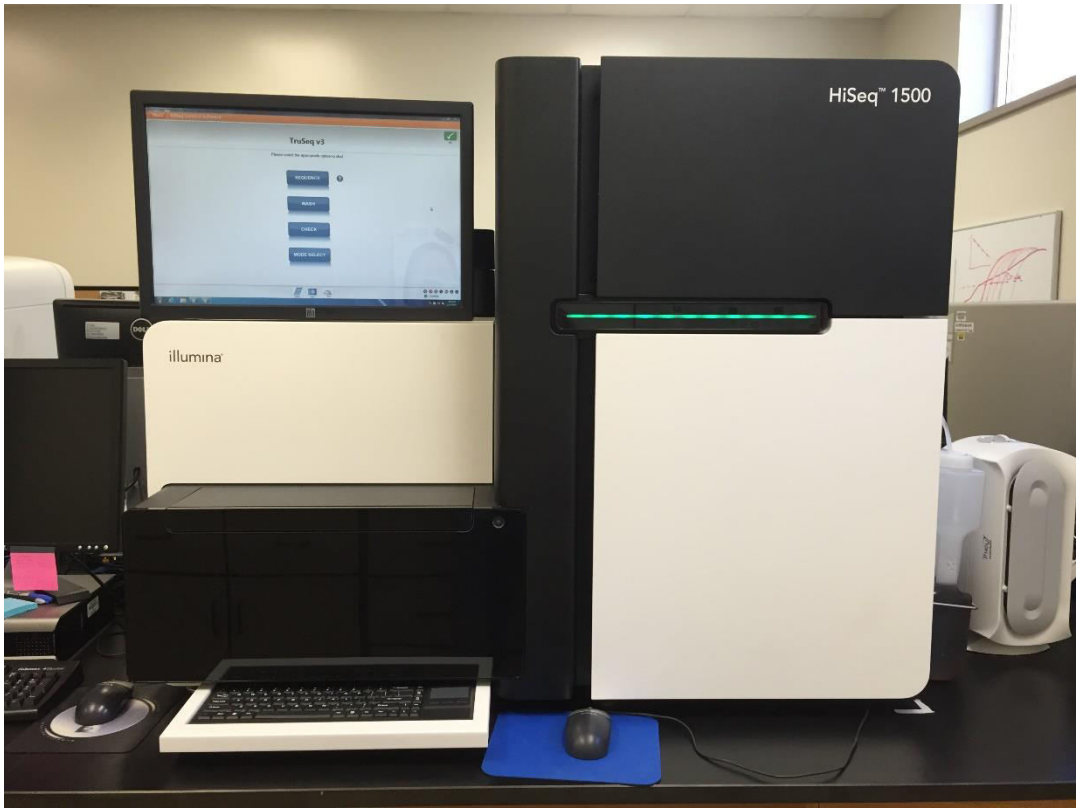


	A	B	C	D	E	F	G	H	I
1	**Individual Name**	**SNP1a**	**SNP1b**	**SNP2a**	**SNP2b**	**SNP3a**	**SNP3b**	**SNP4a**	**SNP4b**
2	Sample 1	A	A	C	G	C	C	C	C
3	Sample 2	G	G	G	G	T	T	C	C
4	Sample 3	G	G	C	C	C	T	C	C
5	Sample 4	A	A	G	G	C	T	C	C
6	Sample 5	A	G	C	C	C	T	C	C
7	Sample 6	A	G	C	G	C	C	C	C

SNPs and next generation sequencing

"Reading DNA" – determine order of the 4 nucleotides

http://www.dnalc.org/view/15479-Sanger-method-of-DNA-sequencing-3D-animation-with-narration.html



- Genome sequencing and assembly
- RNA-seq: mRNA transcript analysis (differences in gene expression)
- RAD-seq: method to discover and genotype ~10K SNP markers per 50-100 samples
- GT-seq: genotype by sequencing; ~2,000 samples in lane, ~192 SNPs