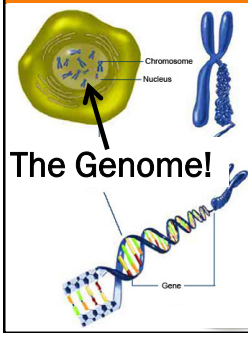
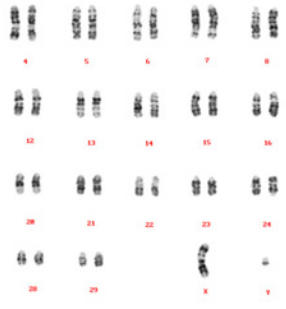
Genomics: Old, Now, and New

Megan Rolf Tonya Amen Mark McCully

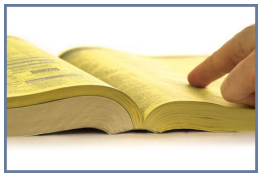
What's in a Genome?





Genomes Are Like Phone Books

- The genome is like a stack of phone books
- We represent DNA like a long string of letters (ACTG)
  - A stack of phone books is like a genome
    - Chromosome
  - Each book consists of multiple sections
    - Genes



Chromosomes

A	C
C	C
T	C
G	A
A	C
G	T
G	G
A	A
G	G
G	G
A	A
G	G
G	G
A	A
G	G
G	G

Genomes are like phone books

- Your phone book is different from my phone book
  - The chromosomes in different families differ from each other

- Within your town's phone book, from one year to another there are a few additions, but most of the book is the same
  - Each time we go through a generation, there are mutations introduced into the genome

DNA Testing

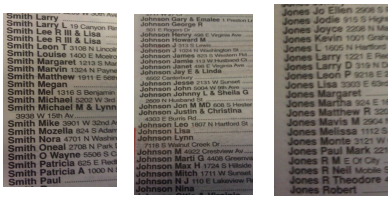
- Causal Mutations
  - The mutation that causes the phenotype
  - Genotype = Phenotype
  - Genetic defect testing
  - Coat Color
- DNA Markers
  - A mutation NEAR the causal mutation
  - Genotype = Phenotype ?
  - Typically for performance traits
  - SNP chips
  - Marker tests

Types of Markers

- We represent DNA like a long string of letters (ACTG)
- Microsatellite
  - Long strings of repeats in the DNA
  - Like Smith or Jones
  - Common on older parentage panels

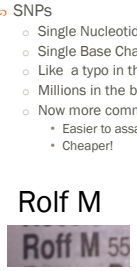
Chromosomes

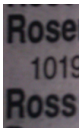
A	C
C	C
T	C
G	A
A	C
G	T
G	G
A	A
G	G
G	G
A	A
G	G
G	G
A	A
G	G
G	G

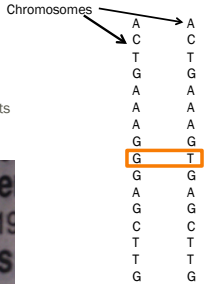


Types of Markers

- We represent DNA like a long string of letters (ACTG)
- SNPs
  - Single Nucleotide Polymorphism
  - Single Base Change in DNA
  - Like a typo in the phone book
  - Millions in the bovine genome
  - Now more commonly used for DNA tests
    - Easier to assay
    - Cheaper!

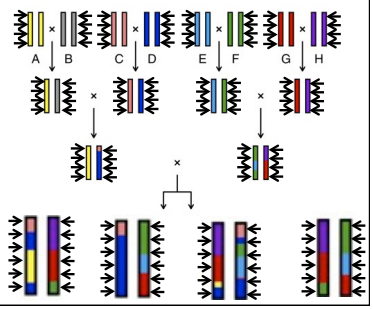






What is so Important about DNA Markers?

- Not the Causal Mutation (unfortunately)
  - Can't test directly
- Independent assortment
  - Many possible combinations
  - Recombination
  - Track inheritance



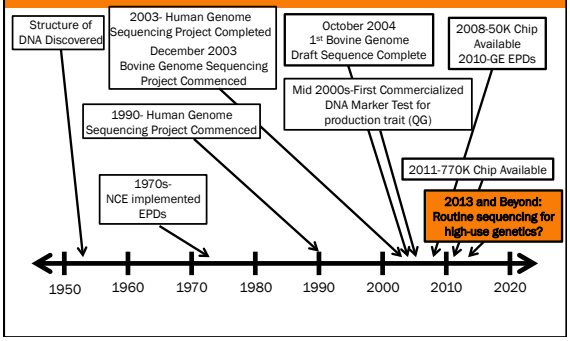
Independent Assortment

- Genes assort independently into gametes
  - Every mating can produce 1,152,921,504,606,846,976 genetically distinct full siblings!!!
  - DNA markers can help us figure out what was inherited by each animal

Independent Assortment

- Genes assort independently into gametes
  - Every mating can produce 1,152,921,504,606,846,976 genetically distinct full siblings!!!
  - DNA markers can help us figure out what was inherited by each animal

Timeline



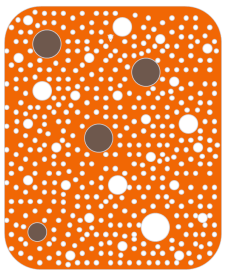
Evolution of DNA Technology

- Mid 2000s
  - First DNA marker test for a production trait
  - GeneStar Quality Grade
  - 4 markers by 2007
- 2008
  - 50K SNP Chip
- 2011
  - 770K SNP Chip
- 2013
  - Genome sequencing-includes all possible DNA markers!
  - Exome Sequencing
  - Gene Expression (RNA)
  - Rumen Microbes (Microbiome)
- Prices are Dropping as amount of data and adoption (volume) increases!!!

Marker Tests: Then vs. Now

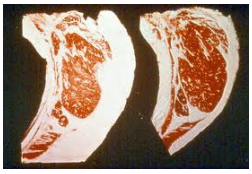
- Then: Marker Assisted Selection
- A few markers
- Large effects

Marker Assisted Selection

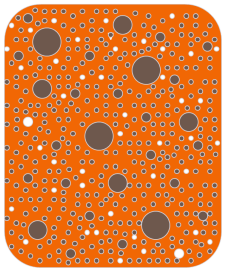


Marker Tests: Then vs. Now

- Now: Genomic Selection
- Thousands of Markers
- SNP Markers
  - Cheap(er)
- More Options!

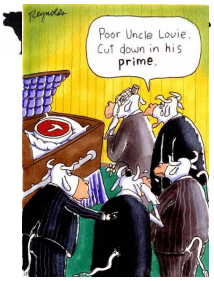


Genomic Selection



Marker Tests: The Options

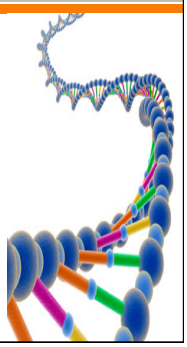
- Parentage
- Simple Traits:
  - Genetic Defect testing
- Complex Traits:
  - Genomic data into NCE (for some panels)
  - Targeted panels
    - GeneMax (not in NCE) + Others
  - Small Chips
    - 9K
  - Large panels
    - 50K



- Sequencing was key to developing all of these tools!

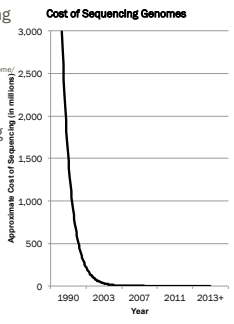
Where do we go from here?

- Maintaining current prediction accuracy will require constant phenotype collection and retraining with more genotypes
- Adding more markers on a larger chip will most likely show diminishing returns
  - 50K to 770K ~3-5% increase in prediction accuracy (Kachman)
    - Strategies for improvement?
  - Still markers, not causal mutations
  - Only a subset of all the mutations in the genome
- Whole Genome Sequencing
  - All mutations at one time (could be millions!)
  - Not just what is on the chip
  - Includes causal and rare mutations
  - Find regions (chip) vs. Find Causal Mutations (sequencing)



Sequencing Projects

- 1990-2003: Human Genome Sequencing Project
  - 13 years
  - ~$437 Million (http://www.ornl.gov/sci/techresources/Human_Genome/project/budget.shtml)
  - More than $3 Billion by some estimates!
- 2003-2005: Bovine Genome Sequencing Project
  - ~2 years (to get 6x draft assembly)
  - $53 million (http://www.genome.gov/12512874)
  - Still working to improve assembly and gene annotations
- Current Cost for 1 Genome: ~$10,000
  - Data only, no assembly
- Rumored cost for 1 Genome: ~$1,000
  - Data only, no assembly

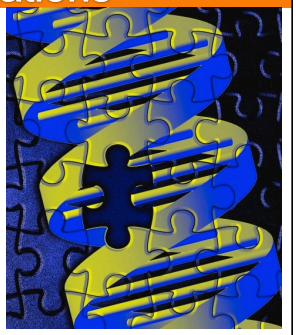


How Big is the Genome?

- How much information are we talking about?
  - 3,200 bases/sheet of paper
  - 1 copy of each chromosome
    - 3 billion base pairs
  - 4.6 pallets of paper
  - Each cell has 2 copies of each chromosome

Genome Sequencing Limitations

- Limited more by data storage requirements and computing power for analysis than by data generation
  - Like putting together a puzzle with millions of small pieces and no picture to tell you what it looks like
  - Need to use computer to predict where genes are
  - Identify mutations



Uses of Sequencing Data

- Causal mutations
  - Are all included in the genome sequence
    - Develop better marker panels in the future
  - No need for additional DNA testing
- Genetic evaluation
  - Information can be used in genetic evaluation
  - Can also be used to augment DNA testing information for relatives
    - Fill in the gaps (imputation)
- Fertility
  - Address early embryonic mortality by combo sequencing and DNA chips
- Data repository
  - Each carrying ~8 recessive lethal mutations
  - If issues do arise, already have the tools needed to solve the problem quickly

Questions?

Dr. Megan M. Rolf
201h Animal Science Building
Stillwater, OK 74078
(405)744-9292
mrolf@okstate.edu