Impact of Genomics on Dairy Cattle Breeding

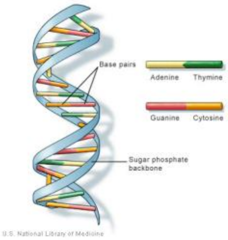
Bennet Cassell
Virginia Tech
2012 VSFA/VA Tech Nutrition Cow College

Genomic data revolutionize dairy cattle breeding

- Accuracy of selection prior to progeny data increases from 35 to over 70% for Net Merit in Holsteins
- Genome evaluations cut generation intervals in half
- Bulls entering progeny testing programs are screened much more carefully.
- Many young PT bulls are no longer "in-waiting"
- Accuracy of genomic predictions for young bulls are sufficient to use them much more heavily than in the past. Time for new habits!

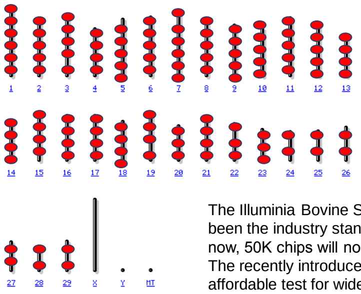
Basics of the DNA molecule

- It is long and complex
- Four nucleic acids:
  A, C, G, T
- A pairs with T and G with C
- 3 billion base pairs
- 30 chromosome pairs
- DNA translates into proteins
- Protein performs cell function



U.S. National Library of Medicine

Dense SNP maps



The Illuminia Bovine SNP50 Bead Chip has been the industry standard. Years from now, 50K chips will no longer be “dense”. The recently introduced ~~3~~ 7K chip is a more affordable test for widespread use.

Genotyping options

From Dr. John Cole, AIPL, ARS, Beltsville MD

- Illumina: a major player in chip technology
  - Bovine SNP50
    - Version 1 54,001 SNP
    - Version 2 54,609 SNP
  - HD 777,962 SNP (only 50K used currently)
  - LD 6,909 SNP (Replaced 3K)
- Standard practice: impute from lower to higher densities with high accuracy

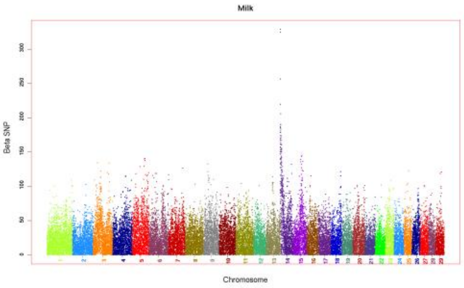
50K SNP results:
864 of 2000 marked alleles on Chromosome 1 for Elevation

```
10001112200200121110111121111011110011211000201220022220111
1202101200211122110021112001111001011011010220011002201101
1200201101020222121122102010011100011220221222112021120120
20100202202000021100011202011221112111022011110000021220200
0221012020002211220111012100111211102112110020102100022000
2201000201100002202211022112101121110122220012112122200200
0200202020122211002222222002212111121002111120011011101120
02022200011120110102111212111020221002112012110011111102111
2110211122000101101110202200221110102011121111011202102102
1211011022122001211011211012022011002220021002110001110021
1021101110002220020221212110002220102002222121221121112002
0110202001222222112212021211210110012110110200220002001002
0001111011001211021212111201010121202210101011111021102112
211111212111210110120011111021111011111220121012121101022
202021211222120222002121210121210201100111222121101
```

The first genomic predictions

- SNP sequences for old bulls with many daughters, a discovery population, produced prediction equations with 43,000 “x” variables
- Y variable was Daughter Yield Deviation, a “super phenotype” based on bull progeny
- Only 43,000 of the 50,000 SNP sites tested were considered “informative”
- Prediction equations must be updated often to reflect recent genotypes and phenotypes
- Predictions rely on genetic relationships as well as linkage to causative genes

Prediction of milk from 38,000 SNPs across 29 chromosomes in Holsteins



Original test of genomic evaluations in Holsteins

- Discovery population: 4422 older Holstein bulls born 1952-1999
- Discovery used progeny data through 2004
- Traditional methods created DYDs on old bulls for genomic predictions
- Accuracy tested using genomic and progeny data through January 2009 on 2035 young bulls
- Study included 5 yield traits, 5 health traits, 16 conformation traits, and Net Merit

Increase in Reliability compared to Parent Average from adding genomic data

Trait	HO	JE*	BS*
Net merit	24	8	9
Milk	26	6	17
Fat	32	11	10
Protein	24	2	14
Fat %	50	36	8
Protein %	38	29	10

*Discovery and test populations are much smaller than for the Holstein breed

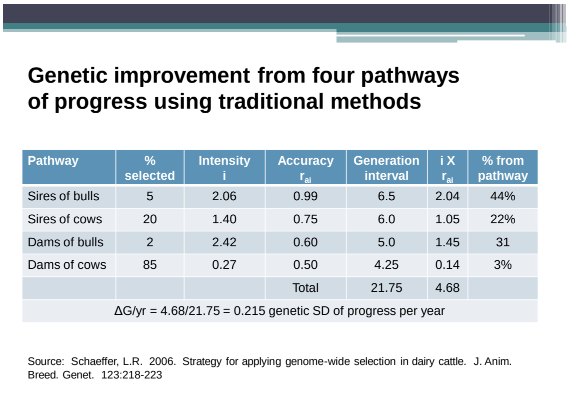
Increase in Reliability compared to Parent Average from adding genomic data

Trait	HO	JE	BS
Productive life	32	7	12
Somatic cell score	23	3	17
Daughter pregnancy rate	28	7	18
Final score	20	2	5
Udder depth	37	20	8
Foot angle	25	11	-1

Selection based on combinations of genomic and progeny data on AI young sires and proven bulls in A.I.

Select the Top 20 Holstein bulls by four methods based on 2004 data*	Average NM$ 2009	Difference from NM$ 2004
Young bulls, traditional PA	$395	-$278
Young bulls, gPTA	$516	-$130
Proven bulls, traditional PA	$381	-$96
Proven bulls, gPTA	$463	-$30

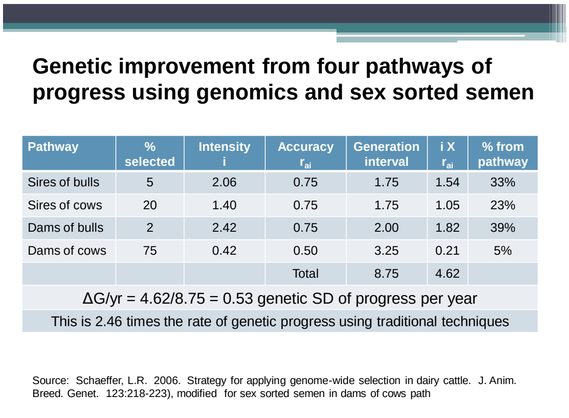
*Five years information is added between the two sets of proofs. Young bulls add progeny test daughters while older bulls add many second crop daughters.



Genetic improvement from four pathways of progress using traditional methods

Pathway	% selected	Intensity i	Accuracy r_{ai}	Generation interval	i X r_{ai}	% from pathway
Sires of bulls	5	2.06	0.99	6.5	2.04	44%
Sires of cows	20	1.40	0.75	6.0	1.05	22%
Dams of bulls	2	2.42	0.60	5.0	1.45	31
Dams of cows	85	0.27	0.50	4.25	0.14	3%
			Total	21.75	4.68	
ΔG/yr = 4.68/21.75 = 0.215 genetic SD of progress per year						

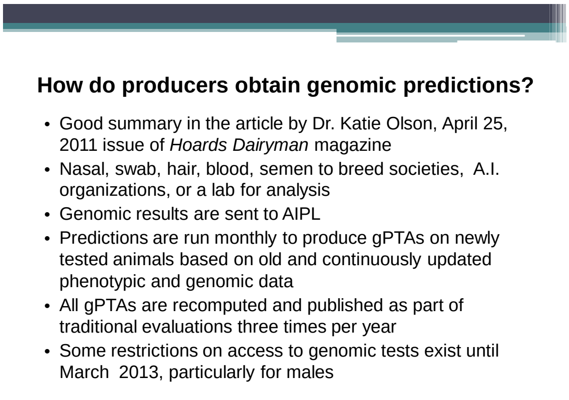
Source: Schaeffer, L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. J. Anim. Breed. Genet. 123:218-223



Genetic improvement from four pathways of progress using genomics and sex sorted semen

Pathway	% selected	Intensity i	Accuracy r_{ai}	Generation interval	i X r_{ai}	% from pathway
Sires of bulls	5	2.06	0.75	1.75	1.54	33%
Sires of cows	20	1.40	0.75	1.75	1.05	23%
Dams of bulls	2	2.42	0.75	2.00	1.82	39%
Dams of cows	75	0.42	0.50	3.25	0.21	5%
			Total	8.75	4.62	
ΔG/yr = 4.62/8.75 = 0.53 genetic SD of progress per year						
This is 2.46 times the rate of genetic progress using traditional techniques						

Source: Schaeffer, L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. J. Anim. Breed. Genet. 123:218-223), modified for sex sorted semen in dams of cows path



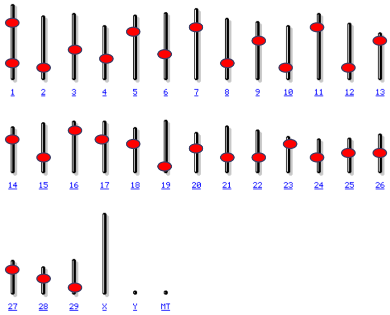
How do producers obtain genomic predictions?

- Good summary in the article by Dr. Katie Olson, April 25, 2011 issue of *Hoards Dairyman* magazine
- Nasal, swab, hair, blood, semen to breed societies, A.I. organizations, or a lab for analysis
- Genomic results are sent to AIPL
- Predictions are run monthly to produce gPTAs on newly tested animals based on old and continuously updated phenotypic and genomic data
- All gPTAs are recomputed and published as part of traditional evaluations three times per year
- Some restrictions on access to genomic tests exist until March 2013, particularly for males

Low density SNP chips

- Prices of 50K chip tests restrict application to elite males and females
- A chip testing fewer sites might deliver acceptable accuracy at substantially reduced prices

Low density SNPs – fewer marked locations on chromosomes



Low density SNP chips: applications

- Prescreen more bull calves to locate most promising for evaluation with 50K chips
- Equally useful to screen outliers among females while they are young with potential for maximum genetic impact
- Improve culling decisions: why raise the bottom end?
- Reduce guesswork when buying replacements
- Parentage verification has enormous implications

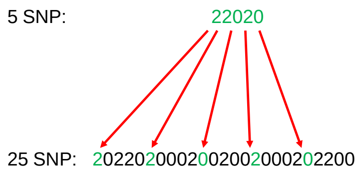
Imputation

- Genomic information from ancestors is used to expand SNP results in progeny tested with 6K chips
- Conversely, genomic information on progeny and one tested parent can be used to “impute” SNP results on the untested parent
- Basic idea: split the genotype into a paternal or maternal chromosome, segments of which are called haplotypes
- Mendelian sampling creates different haplotypes from the same animal. More progeny, more haplotypes
- Haplotypes are traced to progeny and used to fill in missing genotypes from LD genomic tests

Imputation

from Dr. John Cole, AIPL, ARS, Beltsville, MD

- Identify haplotypes in population using many markers, 50K, 800K, etc
- Track haplotypes with fewer markers
- e.g., use 5 SNP to track 25 SNP



5 SNP: 22020

25 SNP: 2022020002002002000202200

Should commercial herds use NM$ proofs with 70% REL?

- Image of 70% REL proofs was established years ago by privately proven bulls – it no longer applies
- Use groups of bulls with high genomic evaluations
- Use the best bulls more heavily later– if they deserve it. Better genomic-tested young bulls will probably be available
- Using young bulls selected for high gPTA’s will produce daughters at least two years quicker than waiting for progeny tests

Reliability of genetic predictions from 3K and other tests for milk in Holsteins

Source of genetic prediction	Genomic information available	Projected Reliability
Parent average	None – traditional method	35%
3K chip on animal*	Sire/MGS with 50K test, dam not tested	60%
3K chip on animal*	Both parents tested with 50K chip	65%
50K chip on animal	Imputation not needed	72%
800K on animal	Imputation not needed	74%
* A new 7K chip, designed to replace the 3K chip, will increase Reliability to within about 2% points of the 50K chip		

Traditional genetic evaluations are critical to genomic predictions

- Genomic predictions are based on traditional genetic evaluations of genomically tested animals
- The discovery population was originally dominated by high REL AI bulls with 50K genomic tests
- Many new genomic tests are being performed, most based on the 7K low density chip, mostly on females
- Accuracy of traditional evaluations in the discovery population is changing and results must be monitored to protect the prediction system

Traditional PTAs have two parts

- A Parent Average (PA), or the expected genetic merit of an average progeny of a single mating between animals with traditional evaluations
- An estimate of Mendelian Sampling (MS), which is the deviation of true transmitting ability from PA
- Estimates of MS for females are more extreme than for males with the same amount of information
- MS based on progeny in many herds for bulls. MS in cows is based primarily on her records.
- MS in cows is adjusted to have variance similar to bulls so genomically tested cows can be used

Using cows to develop genomic predictions

- Cows will be increasingly important in future “discovery” populations.
- Traditional evaluations over-evaluated the best cows
- Traditional evaluations under-evaluated the poor cows
- In April 2011, MS in cows was adjusted to have variance similar to bulls of equal REL
- Traditional PTAs are now less extreme for cows, up and down
- Adjustments should improve the utility of both traditional PAs and genomic predictions.

Conclusions

- The progeny test will remain a gold standard for some time but expense and delay clearly reduce its utility. Days may be numbered for the progeny test.
- Genomic predictions are independent of management decisions on individual animals
- Genomic results are NOT as accurate as progeny testing.
- Young sires with exceptional genomic proofs have a much more important role in herd improvement than in the past

Conclusions

- More data and more animals genotyped means better predictions. Is it time to ignore some national borders? Cooperation with Great Britian and Italy already exists

- Combining genomic data on cows from different breeds?
  - USDA Animal Model merged breeds in May 2007 with the first all-breed genetic evaluation in its history
  - New Zealand and US research shows very different prediction equations from Jersey than from Holstein data

- Genomic prediction is work in progress- procedures will change and results will improve