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
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, ADPL, 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
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

<table>
<thead>
<tr>
<th>Trait</th>
<th>HO</th>
<th>JE*</th>
<th>BS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net merit</td>
<td>24</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Milk</td>
<td>26</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>Fat</td>
<td>32</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Protein</td>
<td>24</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Fat %</td>
<td>50</td>
<td>36</td>
<td>8</td>
</tr>
<tr>
<td>Protein %</td>
<td>38</td>
<td>29</td>
<td>10</td>
</tr>
</tbody>
</table>

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

### Increase in Reliability compared to Parent Average from adding genomic data

<table>
<thead>
<tr>
<th>Trait</th>
<th>HO</th>
<th>JE</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productive life</td>
<td>32</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Somatic cell score</td>
<td>23</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Daughter pregnancy rate</td>
<td>28</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Final score</td>
<td>20</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Udder depth</td>
<td>37</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Foot angle</td>
<td>25</td>
<td>11</td>
<td>-1</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Selection</th>
<th>Average NM$ 2009</th>
<th>Difference from NM$ 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young bulls, traditional PA</td>
<td>$395</td>
<td>-$278</td>
</tr>
<tr>
<td>Young bulls, gPTA</td>
<td>$516</td>
<td>-$130</td>
</tr>
<tr>
<td>Proven bulls, traditional PA</td>
<td>$381</td>
<td>-$96</td>
</tr>
<tr>
<td>Proven bulls, gPTA</td>
<td>$463</td>
<td>-$30</td>
</tr>
</tbody>
</table>

*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

<table>
<thead>
<tr>
<th>Pathway</th>
<th>% selected</th>
<th>Intensity</th>
<th>Accuracy</th>
<th>Generation Interval</th>
<th>I X r</th>
<th>% from pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sires of bulls</td>
<td>5</td>
<td>2.06</td>
<td>0.99</td>
<td>6.5</td>
<td>2.04</td>
<td>44%</td>
</tr>
<tr>
<td>Sires of cows</td>
<td>20</td>
<td>1.40</td>
<td>0.75</td>
<td>6.0</td>
<td>1.05</td>
<td>22%</td>
</tr>
<tr>
<td>Dams of bulls</td>
<td>2</td>
<td>2.42</td>
<td>0.60</td>
<td>5.0</td>
<td>1.45</td>
<td>31%</td>
</tr>
<tr>
<td>Dams of cows</td>
<td>85</td>
<td>0.27</td>
<td>0.50</td>
<td>4.25</td>
<td>0.14</td>
<td>3%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>21.75</td>
<td>4.68</td>
<td></td>
</tr>
</tbody>
</table>

\( \Delta G/yr = 4.68/21.75 = 0.215 \text{ genetic SD of progress per year} \)


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

<table>
<thead>
<tr>
<th>Pathway</th>
<th>% selected</th>
<th>Intensity</th>
<th>Accuracy</th>
<th>Generation Interval</th>
<th>I X r</th>
<th>% from pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sires of bulls</td>
<td>5</td>
<td>2.06</td>
<td>0.75</td>
<td>1.75</td>
<td>1.54</td>
<td>33%</td>
</tr>
<tr>
<td>Sires of cows</td>
<td>20</td>
<td>1.40</td>
<td>0.75</td>
<td>1.75</td>
<td>1.05</td>
<td>23%</td>
</tr>
<tr>
<td>Dams of bulls</td>
<td>2</td>
<td>2.42</td>
<td>0.75</td>
<td>2.00</td>
<td>1.82</td>
<td>39%</td>
</tr>
<tr>
<td>Dams of cows</td>
<td>75</td>
<td>0.42</td>
<td>0.50</td>
<td>3.25</td>
<td>0.21</td>
<td>5%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>8.75</td>
<td>4.62</td>
<td></td>
</tr>
</tbody>
</table>

\( \Delta G/yr = 4.62/8.75 = 0.53 \text{ genetic SD of progress per year} \)

This is 2.46 times the rate of genetic progress using traditional techniques


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

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: $202200200200200202020200200200202200$

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

<table>
<thead>
<tr>
<th>Source of genetic prediction</th>
<th>Genomic information available</th>
<th>Projected Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent average</td>
<td>None — traditional method</td>
<td>35%</td>
</tr>
<tr>
<td>3K chip on animal*</td>
<td>Sire/MGS with 50K test, dam not tested</td>
<td>60%</td>
</tr>
<tr>
<td>3K chip on animal*</td>
<td>Both parents tested with 50K chip</td>
<td>65%</td>
</tr>
<tr>
<td>50K chip on animal</td>
<td>Imputation not needed</td>
<td>72%</td>
</tr>
<tr>
<td>800K on animal</td>
<td>Imputation not needed</td>
<td>74%</td>
</tr>
</tbody>
</table>

* 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 Britain 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