Developing a custom genotyping chip for dairy and beef cattle in Ireland

M.P. Mullen¹*, M.C. McClure²*, F. Kearney⁵, S.M. Waters³, R. Weld⁴, P. Flynn⁴, C.J. Creevey³, A. Cromie⁵ and D.P. Berry⁶


INTERBULL BULLETIN NO. 47. Nantes, France, August 23 - 25, 2013
It is now possible to develop an inexpensive technology to include all the above!!
Objectives

Develop a custom SNP chip catering for both dairy and beef cattle that can:

1. impute microsatellite parentage data from SNPs
2. screen for lethal recessives, congenital disorders and genes with major effects
3. improve imputation accuracy to high density SNP panels
4. facilitate a research component
IDB content (V1)

- Genomic selection - imputation to HD ~12,500 SNPs
  - Standard ~7,000 Illumina LD 'base' panel
  - Additional 5,500 for imputation

- Parentage ~2,200 SNPs
  - 150 - 200 SNPs per microsatellite (n=12)
  - Heaton panel also included in 'base' panel (~120 snps)

- Lethal recessives (n=4)

- Congenital disorders (n=33)

- Major genes (n=16)

- Research component 1873 variants
Materials and Methods

• Improve imputation accuracy to HD (777K)
  • + 5,500 SNPs selected from high density (HD) panel based on:
    • High frequency across 3,122 dairy and beef cattle
      • AA (269), BB (196), CH (710), HE (234), HF (719), LM (730), SI (264)
    • Low correlation with existing Illumina ‘base’ panel snps
    • Increased coverage at chromosomal ends
Materials and Methods

• Improve imputation accuracy cont.
  • High density (777K) SNPs imputed from low density (7K), custom LD+5,500 (12K) and medium density (50K) (Berry et al. 2013)

Calibration: 2424 dairy and beef bulls
Validation: 698 bulls born since 2005
Materials and Methods

- Sire microsatellite imputation

Imputation of microsatellite alleles from dense SNP genotypes for parental verification

Matthew McClure¹, Tad Sonstegard¹, George Wiggans² and Curtis P Van Tassell¹*

¹ Bovine Functional Genomics Lab, United States Department of Agriculture, Agriculture Research Service, Beltsville, MD, USA
² Animal Improvement Program Lab, United States Department of Agriculture, Agriculture Research Service, Beltsville, MD, USA

- Unique haplotypes ≥ 4
- Linked to 1 MS 100% or ≥ 90%
- Increasing haplotype size by 10 SNP
  - ≤ 1% increase in haplotypes
- 2,176 SNPs flanking 12 ISAG MS
Materials and Methods

• Lethal recessives, congenital disorders, major genes and research
  • Lethal recessives (n=4)
    • CVM
    • BLAD
    • DUMPS
    • Brachyspina
Congenital defects – undesirable traits (n=33)

- Curly calf
- Fawn calf
- Spiderleg
- Tibial Hemimelia
- Hairlessness (Hypotrichosis)
- Polledness
- CMDI & II
- Mulefoot
- Black/Red coat colour/Red factor
- Red recessive coat colour
- Pulmonary Hypoplasia
Major genes (n=16)

- Myostatin (19 variants)
- DGAT1
- A1/A2 β-casein
- ABCG2
- Kappa casein
- Fertility haplotypes
  - (HH1, HH2, HH3, JH1)
- STAT1, 3 & 5
Materials and Methods

Research

- 1873 variants (Mullen et al. 2012a,b)
- The somatotrophic axis is a key regulator of post-natal growth and development in mammals (Etherton 2004)
- Central to nutrient partitioning, lactogenesis and fertility (Lucy et al. 2008; 2009)
Results - Stats

• 9852 samples from March to 21\textsuperscript{st} End July 2013

• 16,882 SNPs

• 154 million SNPs after poor sample call rates removed
# Breed composition

<table>
<thead>
<tr>
<th>Breed</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>5714</td>
</tr>
<tr>
<td>AA</td>
<td>49</td>
</tr>
<tr>
<td>BB</td>
<td>59</td>
</tr>
<tr>
<td>CH</td>
<td>1773</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Breed</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM</td>
<td>1478</td>
</tr>
<tr>
<td>SI</td>
<td>296</td>
</tr>
<tr>
<td>HE</td>
<td>394</td>
</tr>
<tr>
<td>Other</td>
<td>89</td>
</tr>
</tbody>
</table>
Animal call rates

Median: 0.989
<95%: 5.94%
No breed differences
SNP call rate - Illumina

The Irish Agriculture and Food Development Authority
SNP call rate - Illumina

The Irish Agriculture and Food Development Authority
SNP GT scores - Illumina

Median: 0.86
GT scores V call rate

The Irish Agriculture and Food Development Authority
SNP call rate - Non-Illumina

![SNP call rate graph]

The Irish Agriculture and Food Development Authority
SNP call rate - Non-Illumina

258 <90% call rate

Number of SNPs

Call rate

The Irish Agriculture and Food Development Authority
SNP GT scores – Non-Illumina

Median: 0.815

Average Gt score
SNPs that never made it

Illumina SNPs

832 never made it to manifest

Selection criteria

- Call rate > 95%
- GTscore > 0.55
- In HWE
Results

- Improve imputation accuracy to high density for genomic selection

![Bar chart showing proportion allele concordance for different scenarios: Neither sire or MGS, Sire, and Sire + MGS. The x-axis represents genotyped on high density, and the y-axis represents proportion allele concordance. The bars are color-coded: LD, 50k, and IDB.]
Results

Imputation accuracy to microsatellites

Allowing 1 mismatch
IDB 95.6%

McClure et al. 2013
90.1% imputed MS
95.9% MS

Increase reference population = increase imputation accuracy


Imputation of Microsatellite Alleles from Dense SNP Genotypes for Parentage Verification Across Multiple Bos taurus and Bos indicus breeds

Lethal recessives / Major genes

Holstein-Friesian

2% Brachyspina carriers

4% CVM carriers

<<1% Citrullinaemia; Osteopetrosis; Mulefoot

DUMPs free -

21% A1/A1, 46% A1/A2, 33% A2/A2

45% DGAT1 carriers, 15% DGAT1 (K)

56% Kappa Casein
Lethal recessives / Major genes

CH/AA/BB/HE*/LM/SM

No Brachyspina - No CVM - No DUMPs

No fawn or curly calf

Myostatin - 19 variants

<table>
<thead>
<tr>
<th>Mutation</th>
<th>CH</th>
<th>AA</th>
<th>BB</th>
<th>HE</th>
<th>LM</th>
<th>SM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q204X</td>
<td>27</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>FL94</td>
<td>27</td>
<td>2.5</td>
<td>6</td>
<td>-</td>
<td>98</td>
<td>50</td>
</tr>
<tr>
<td>nt821</td>
<td>&lt;1</td>
<td>2.5</td>
<td>33</td>
<td>-</td>
<td>6</td>
<td>-</td>
</tr>
</tbody>
</table>

*2% Hairlessness carriers

The Irish Agriculture and Food Development Authority
From just 1 hair sample!

<table>
<thead>
<tr>
<th>Calf ID</th>
<th>CVM</th>
<th>BLAD</th>
<th>DUMPS</th>
<th>Brachyspina</th>
<th>A1/A2</th>
<th>DGAT1</th>
<th>Parentage</th>
<th>Suggested sire</th>
<th>Genomic EBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>IE123 456</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>✓</td>
<td>n/a</td>
<td>€220</td>
</tr>
<tr>
<td>IE135 7911</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>x</td>
<td>BJY</td>
<td>€230</td>
</tr>
</tbody>
</table>
Conclusions

• IDB – market leader
• Eliminating need for standalone testing
  • Lethals/disorders/major genes; Parentage
  • Genomic selection & Research
• Traceability

• IDB v2, for release Jan 2014:
  • + 47 causative mutations
  • + Breed assignment
Acknowledgements

• Chip design:
  • Breed Societies and AI stations
  • Paul VanRaden, USDA
  • Jon Beever, Illinois university

• Funding:
  • Research Stimulus Fund (RSF-06-353; 11/S/112)
  • SFI (07/SRC/B1156)
  • Breed Societies and AI stations
  • ICBF & Teagasc
IDB SNP CHIP
INTERNATIONAL DAIRY & BEEF
SNP CHIP

Designed in association with the Irish Cattle Breeding Federation (ICBF), Teagasc, Weatherbys and USDA’s Agricultural Research Service.

This custom chip is the very latest design catering for both Beef and Dairy.

The chip consists of the Illumina LD (7K) base content plus a further 10,000 (10K) SNPs carefully selected to ensure very high imputation accuracy to HD & to convert to Microsatellite data for parentage verification. This extra panel of SNPs provides the very latest dual product for both Beef & Dairy breeds.

Both the core and additional ISAG recommended SNP parentage panels are present on the chip.

The IDB also contains a comprehensive selection of genetic markers to screen for genetic disorders & major genes.

CHIP CONTENTS FOR DISEASES & TRAITS

**Lethal recessives**
1. CMV-Complex Vertical malformation
2. DLNPS
3. Ehrlichia
4. ILAD

**Congenital disorders**
1. Arrhenogran (Calyx Cell)*
2. Achondroplasia (Apert Syndrome)*
3. Hypertrophic (AML-T)
4. Hypertrophic in Reticulated Calloway, BephHAPSNP
5. Hypertrophic (HRT)*
6. Spinal Cord, MDA51 gene, Simmental
7. Spinal Cord, SCOP gene, Beaufort Farms
8. Polledness
9. Mili Folt
10. Total Mutilations (TM)*
11. Black/Red Coat Color Red Factor
12. Red/Red Composite coat colour (different to red factor)
13. Silver Color Dilution
14. Dna Color
15. RKN-221 (affects growth and stature)
16. Unresponsive (Male Breeding)
17. Pink Eye (Hindossa Bronco K aragonensis)
18. Porphyrin, Fracture-like Gene (Phenocytolysis)
19. SMB-Genetic marker
20. Beta Lactoglobulin
21. Beta Melanocortin
22. Alpha Melanocortin
23. Citrullina
24. CMH-1 Congenital muscular dystrophy
25. CMH-2 Congenital muscular dystrophy II
26. Caudal tail Syndrome
27. Factor XI
28. Heterochromic (White Eye)
29. SOM- Spinulose renophtrosis STRAP Gene
30. VealPenta Dilution *
31. Pedigree Hypothesis
32. Wastey
33. Neurotrophic hypoxia (waste head syndrome)

**Major genes**
1. DGAT1
2. MTN2 EGF (Double Muscle*)
3. AA/AD beta casein +
4. Fertility (H1, H1a, H1b, H1c, H1d)
5. Kappa Casein I
6. Kappa Casein II
7. ABO2
8. GH141 and GH239 (Mating growth rate)*
9. IGF-1 947/4143
10. STAT1*
11. STAT5*
12. STAT3*
13. Calpain (Tendonitis) lox

* Royalty fees may apply

For more details Contact: Weatherbys Ireland DNA Laboratory

+353(0)45875521
jillynn@weatherbys.ie

WEATHERBYS
Ireland

The Irish Agriculture and Food Development Authority