

Genomic predictions using more markers and gene tests (Abstract #...)

G.R. Wiggans¹, P.M. VanRaden², D.J. Null², and J.B. Cole²

¹ Council on Dairy Cattle Breeding, Bowie, MD, USA

² Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, USA

paul.vanraden@ars.usda.gov

Topics

- **Increase in markers used in US genomic predictions**
- **Methods for selecting better SNPs and removing poorer SNPs**
- **Gene tests added and effects on traits**
- **Conversion of locations to new reference map**
- **Results when applied to 5 breeds (HO, JE, BS, AY, GU)**
- **Correlations and reliability gains with more SNPs**

Growth of markers used

- **Numbers of markers used have increased over the years**
 - **2008: separate lists within each breed based on MAF**
 - **Numbers were 38,416 for HO, 31,628 for JE, 34,593 for BS**
 - **2009: common list of 43,385 SNP for multi-breed research**
 - **2014: increased to 60,671 SNP with HD SNPs and gene tests**
 - **2018: increased to 79,276 SNP with added sequence, HD, and gene tests**
- **Future lists may include more HD and sequence SNPs selected from other breeds, from other traits, and more data**

Choosing the SNPs to include

- **Large effect HD and sequence SNPs from 2 previous studies**
 - **Wiggans et al., 2016 J. Dairy Sci. 99:4504-4511 [HD SNPs]**
 - **VanRaden et al., 2017 Genet. Sel. Evol. 49:32 [sequence] to be added in next update**
- **Gene tests that were added to recent chips**
- **Causal alleles within lethal haplotypes**
- **Removed previous SNPs with high parent-progeny conflict rates, missing genotypes, Hardy-Weinberg within breeds, etc.**
- **Set SNPs to missing on individual chips with poor properties**

Gene tests and sequence SNPs with large effects

- **Gene tests for DGAT1, ABCG2, β -casein, and β -lactoglobulin had large effects on yield traits and net merit.**
- **Gene tests for Holstein cholesterol deficiency, complex vertebral malformation, brachyspina, and calpain had large effects on somatic cell score, udder cleft, protein yield, and gestation length, respectively.**
- **A new sequence marker on chromosome 3 had the largest effect on Holstein final score, foot angle, feet-and-legs score, and rear legs (rear view).**
- **For Jerseys, bGHR had a large effect on productive life.**

Added HD SNPs with large effects

HD SNP	Chr	ARS-UCD1 location	Breed	Traits
BovineHD0300030795	3	106529191	HO	PL
BovineHD0500026662	5	93515983	BS	Fat
BovineHD4100004671	6	38289304	AY	Stature, Strength, Dairy Form
BovineHD1000009940	10	30246462	BS	CCR
BovineHD1200019197	12	69402495	GU	Protein
BovineHD1300003829	13	13295907	AY	CCR
BovineHD1400000401	14	1368064	BS	Fat (near DGAT1)
BovineHD1700020789	17	69337705	BS	HCR
BovineHD1900014702	19	51839023	AY	SCS
BovineHD2900011160	29	36354055	AY	Fat

New SNPs with larger effects than 50K

- **Top 5 largest SNP effects**
 - within each trait
 - within each breed
- **How many were on the original 50K chip?**
- **New SNPs were added from HD, sequence, or gene tests**
- **Percentage of 50K SNPs for each breed, across traits**

Breed	Traits	50K SNPs as % of Top 5 SNPs
Holstein	41	34%
Jersey	27	55%
Brown Swiss	30	53%
Ayrshire	27	33%
Guernsey	27	38%
Weighted average	152	42%

Correlations of 80K with 60K predictions

- **Genomic predictions improved more for breeds with larger reference populations such as HO and JE.**
- **Individual predictions changed most for animals less related to the US population, with less complete pedigrees, or genotyped with lowest density chips such as 3K.**
- **Correlations of 80K with 60K predictions for young animals were**
 - **about 0.99 for Jersey, Holstein, and Guernsey yield traits**
 - **higher for Ayrshire and Brown Swiss yield traits**
 - **lower for many other traits, about 0.98**

Reliability

- **No change in calculation of reliability**
 - Recent research suggested that reliability was over estimated
 - Wiggans et al (2016) found 1 to 3 point increase with 77 K SNP

Computation

- **Missing alleles were imputed for Holsteins by:**
 - first imputing all bulls and their ancestors
 - using those haplotypes as priors to impute the remaining 2 million females
 - required 9 d to compute with 25 processors and 270 GB of memory.
- **The 80K list increased computing times for other key programs by about 30%**


Conclusions

- **Important variants now included directly had large effects on yield traits and the net merit index.**
 - e.g., in *DGAT1*, *ABCG2*, β -casein, and β -lactoglobulin
 - For HO, 66% of top effects are from added SNPs, only 34% from 50K chip
- **Revised SNP list also improved quality control and used new reference map**
- **Using 80K instead of 60K SNPs increases computing by ~ 30% for some steps**
- **Reliability is expected to increase by 1 to 3 percentage points**
- **Future predictions could assign more prior emphasis to the known QTLs**

Acknowledgments and disclaimers

- **The Council on Dairy Cattle Breeding** (Bowie, MD) provided the evaluation data.
- Genotyping labs provided data for gene tests and additional selected SNPs.
- USDA-ARS project ARS 8042-31000-002-00, “Improving dairy animals by increasing accuracy of genomic prediction, evaluating new traits, and redefining selection goals”
- Mention of trade names or commercial products in this presentation is solely for the purpose of providing specific information and does not imply recommendation or endorsement by USDA

Questions?



AIP web site:
<http://aipl.arsusda.gov/>

**Holstein and Jersey crossbreds graze on American Farm Land Trust's
Cove Mountain Farm in south-central Pennsylvania**

Source: ARS Image Gallery, image #K8587-14; photo by Bob Nichols

Annual Meeting of the American Dairy Science Association®, Cincinnati, OH, 26 June 2019 (13)