

Breeding for Healthier Livestock

Hint: No silver bullet

Christian Maltecca

North Carolina State University

EAAP Meeting: Aug-31-2016, Section 36

Overview



Overview

Selection for health traits by on-farm computer systems

Selection for health traits

- Increase efficiency by decreasing input costs as opposed to increasing output of products
- Application of genomic information to health traits:
 - Genetic and genomic predictions
 - Identification of genes related to disease resistance/susceptibility
 - Cow Risk Predictions

Genomic Selection Health Traits

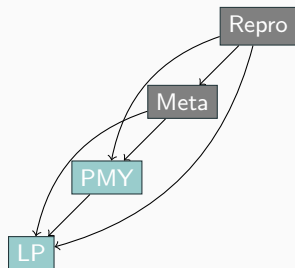
Health Event	Pedigree			Blended pedigree and genomic			
	mean	Unproven	Proven	mean	Unproven	Proven	Gain
DA	0.44	0.22	0.65	0.55	0.38	0.71	0.11
KETO	0.35	0.18	0.52	0.48	0.35	0.61	0.13
LAME	0.24	0.15	0.32	0.39	0.31	0.47	0.15
MAST	0.39	0.26	0.52	0.51	0.4	0.61	0.12
METR	0.35	0.24	0.46	0.48	0.38	0.57	0.13
RTP	0.55	0.42	0.67	0.64	0.54	0.73	0.09

Corr	DPR	PL	NM
DA	-0.35	-0.349	-0.26
Keto	-0.314	-0.318	-0.266
Lame	-0.101	-0.173	-0.237
Mast	-0.129	-0.191	-0.149
Metr	-0.226	-0.119	-0.241
Retp	-0.395	-0.307	-0.27

Causal relationships between health and production

Trait	λ
RP \rightarrow PeakD	0.095
RP \rightarrow PMY	-0.011
RP \rightarrow LP	-0.001
METR \rightarrow PeakD	0.026
METR \rightarrow PMY	-0.001
METR \rightarrow LP	0.001
KETO \rightarrow PeakD	0.02
KETO \rightarrow PMY	-0.012
KETO \rightarrow LP	-0.002
DA \rightarrow PeakD	0.018

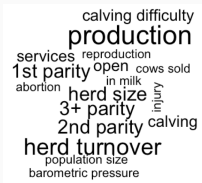
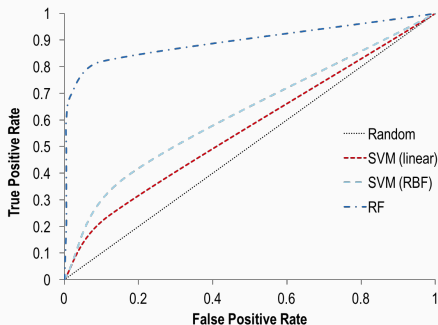
Trait	Mean	HPD
RP \rightarrow Cull	1.226	[1.091; 1.385]
METR \rightarrow Cull	0.929	[0.846; 1.018]
KETO \rightarrow Cull	1.004	[0.889; 1.145]
DA \rightarrow Cull	1.59	[1.379; 1.729]



Dhakal et al. 2014 2015 JDS, 2015
Liv.Sci.

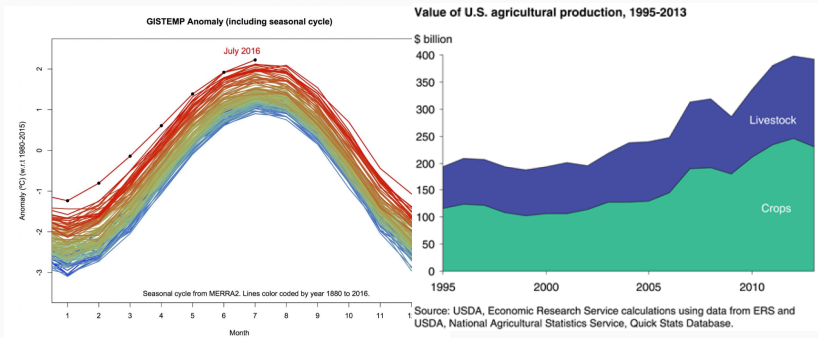
Individual Cow Risk Prediction

Mastitis	Accuracy	Sensitivity	Specificity
SVM (linear)	0.70 (0.003)	0.24 (0.002)	0.88 (0.003)
SVM (RBF)	0.70 (0.01)	0.39 (0.03)	0.83 (0.02)
RF	0.93 (0.001)	0.82 (0.003)	0.97 (0.001)



Accounting for GxE in Genomic Predictions

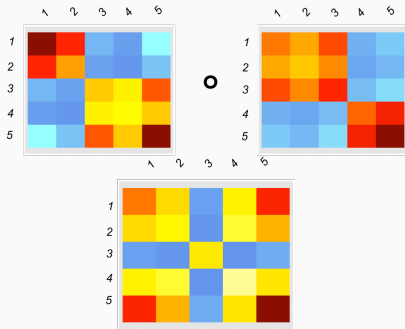
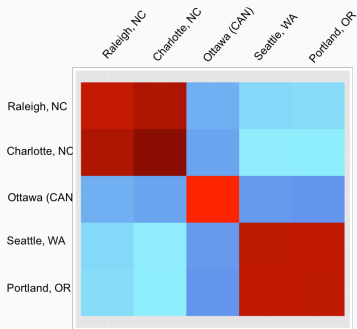
New environmental conditions



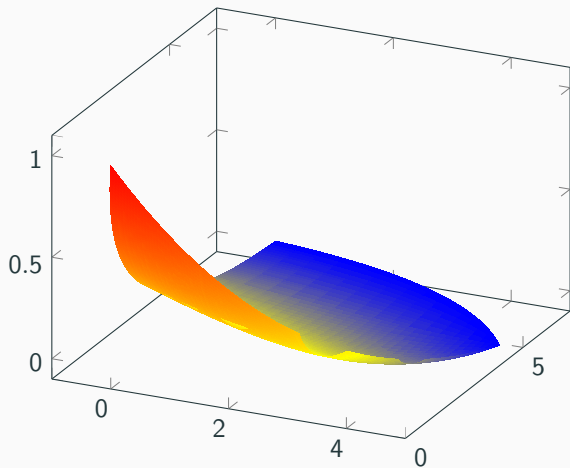
GxE in dairy cattle

- Ravagnolo and Misztal, 2000
- Hayes et al., 2009
- Hammami et al., 2009
- Norberg et al., 2014
- Streit et al., 2013
- Windig et al., 2011
- Bryant et al., 2010
- Haile-Mariam et al., 2008
- Bohmanova et al., 2008
- Oseni et al., 2004
- Fikse et al., 2003
- Calus and Veerkamp, 2003
- Mulder and Bijma, 2005
- ...

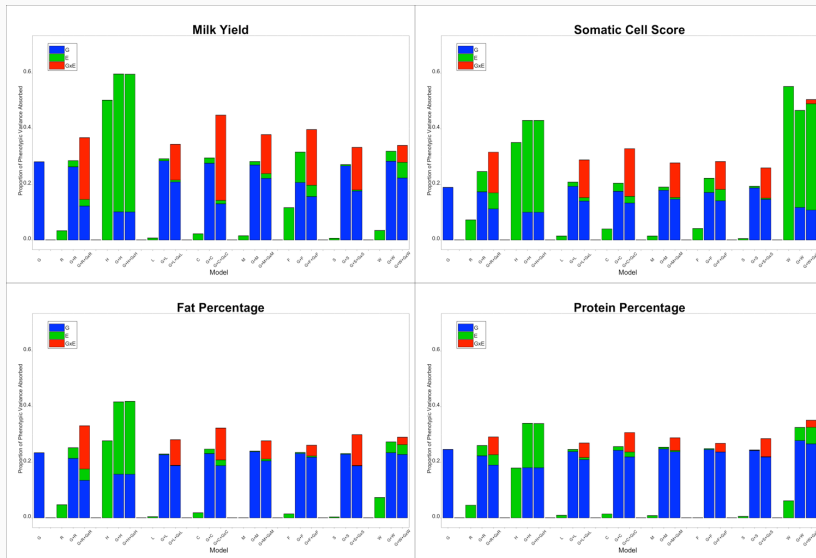
Reaction norm for high dimension genomic and environmental data



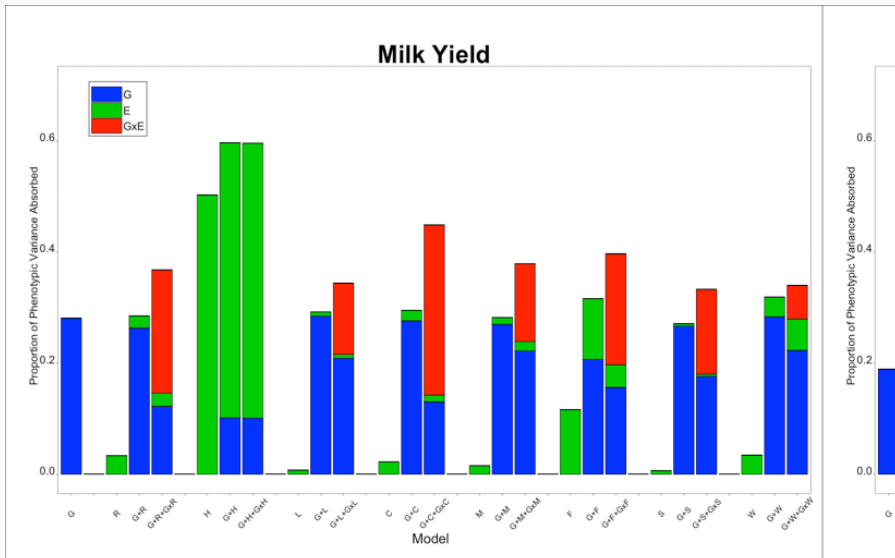
Reaction norm for high dimension genomic and environmental data



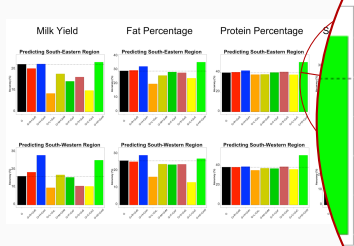
Reaction norm for high dimensional genomic and environmental data



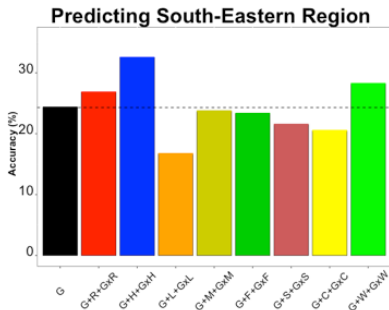
Reaction norm for high dimensional genomic and environmental data



Reaction norm for high dimensional genomic and environmental data



Somatic Cell Score



Predicting South-Western Region

**Managing genomic diversity on
pedigreed populations
undergoing selection for complex
traits.**

Managing genomic diversity

- Genomic information to constrain inbreeding and monitoring losses of genetic variance
- Genomic information to date has been primarily used as a tool to rank individuals on their genetic merit.
 - Works in principle
- Lack of effective implemented strategies
- Three pillars of genetic diversity management:
 - Understanding the basis and consequences of genetic diversity
 - Managing the population by controlling its effective size
 - Optimize genetic variability use through mating plans

Inbreeding Depression Heterogeneity

- Since inbreeding (and inbreeding depression) are function of dominance one would be tempted to just estimate marker effects
 - With genomic information that should be possible
- A few problems
 - Low freq.
 - Small effects
 - Cumulative effect (non linearity of inbreeding depression)
- Still can be attempted

Inbreeding Depression Heterogeneity

- Alternative metric that characterizes long stretches of inbreeding in the form of a run of homozygosity (ROH)
- Simulation has shown to be most associated with the recessive mutation load (Keller et al. 2011) in comparison to other metrics.

Haplotype Finder

Relies on Two Principles

- An ROH genotype is a replicate of two identical haplotypes.
 - Due to this regression can be done using ROH genotypes string instead of relying on haplotype based models.
- ROH haplotypes have a nested structure therefore methods that capitalize on this can be utilized.
 - An ROH is generated when chromosome segments are inherited that are derived from a common ancestor.
 - Due to this individuals that have the same unique ROH segment are expected to have a core segment that is consistent across individuals and can be used as a proxy for the whole ROH segments that may differ outside of the core segment.

Haplotype Finder

Stage 1

- Step 1: Tabulate Means of non-ROH and unique ROH for sliding windows of 60 SNP.
- Step 2: Combine nested windows

Before:

```
Start End Haplotype
132 182 220022200022002222000002222202202222022000200202
133 183 20022200022002222000002222022022220220002002020
134 184 0022200022002222000002222022020222202200020020200
.....
138 188 200022002222000002222202202222022000200202000022
139 189 00022002222000002222202202022220220002002020000220
140 190 00220022220000022222022020222202200020020200002202
```

After:

```
132 190 2200222000220022220000022222022022220220002002020002202
```

- Step 3: Reduce window size by 5 until 20 is reached.
- Step 4: Combine nested windows .

Before (i.e. each haplotype contains same set of animals):

```
Start End Haplotype
614 656 200020220022002222200000200022020002000020
614 646 20002022002200222220000020002202
614 651 2000202200220022222000002000220200020
614 661 2000202200220022222000002000220200020000200222
614 671 20002022002200222220000020002202000200002002220220002
614 666 200020220022002222200000200022020002000020022202202
614 641 200020220022002222200000200
```

After:

```
614 641 200020220022002222200000200
```

Stage 2

- Determine the significance of each window that passed Stage 2 using a model that allows:
 - Fixed Environmental Effects.
 - Additive effect of animal based on pedigree.
 - Permanent effect of animal.
- Contrast between each unique ROH and non-ROH.

Stage 3

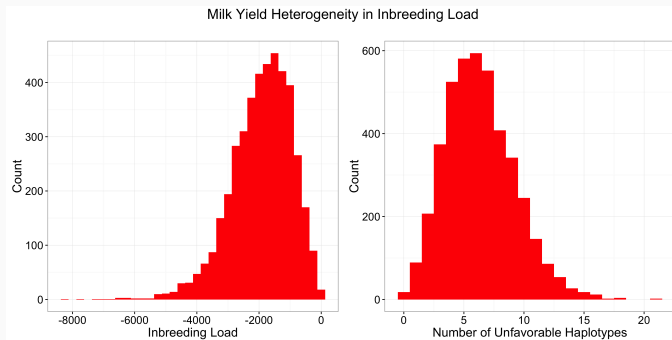
- Remove nested windows

Example: (only keep Window 1)

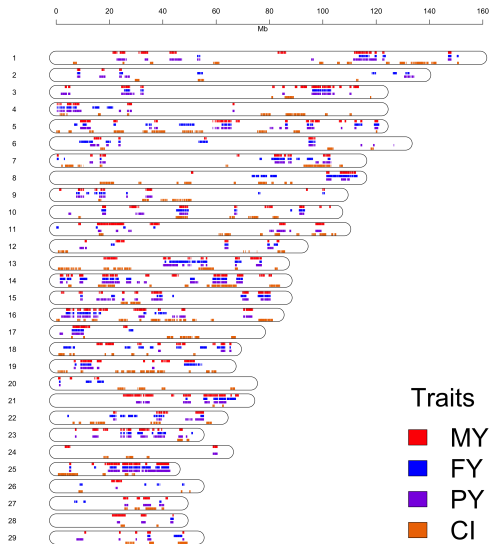
<u>Window 1</u>	<u>Window 2</u>
Animal	Animal
1	1
2	2
3	3
4	4
5	5
6	6
-	7
-	8
-	9

Example with Dairy Traits

- ≈ 200 haplotypes that when homozygous result in reduced performance across all four traits.
- Low Frequency within set of genotypes utilized:
 - Mean (Minimum - Maximum): 0.032 (0.007 - 0.13)
- Potential to further understand the variation in genetic load across individuals.



Exampe with Dairy Traits



Example with Dairy Traits

- Unfavorable haplotypes within candidate regions found previously (Howard et al. 2015).
 - Represented as deviations (lbs.) from non-ROH least square mean.

Chromosome	Location	MY	FY	PY	CI
8	109.13-110.04	-390.6	-22.1	-13.9	-
14	60.68-61.42	-331.0	-20.0	-14.3	14.8
19	9.12-9.75	-622.7	-33.4	-21.5	-
23	37.99-38.61	-489.1	-21.3	-16.5	-
25	25.16-25.63	-1026.6	-39.7	-36.7	-
25	29.94-30.42	-779.7	-32.3	-28.2	-

Understanding inbreeding

A large number of simulation programs have been developed that are suitable for testing alternative selection and/or mating strategies that are primarily based on the additive genetic effects for a quantitative trait.

- QMsim (Sargolzaei and Schenkel, 2009).
- AlphaSim Suite (Hickey et al., 2014).
- ms2gs (Perez-Enciso and Legarra, 2016).
- FREGENE (Chadeau-Hyam et al., 2008).
- XSim (Cheng et al. 2015)
- etc.....

Understanding inbreeding

In the context of animal breeding there is currently a lack of simulation programs tailored towards:

- Identifying "best practice" management decisions to manage a population at the genetic level in the form of:
 - Genetic Diversity
 - Fitness Effects
 - Additive and Dominance Effects
- The optimal use of dense marker information to manage a population at the genomic for populations that are routinely genotyped.

Understanding inbreeding

Due to this we have developed a simulation tool that:

- Generates quantitative and/or fitness traits with additive and dominance effects.
- Utilizes computationally efficient routines to generate dense marker based relationship matrices and their associated inverse.

Geno-Diver Overview

Historical Population

- Use MaCS (Chen et al. 2009) to generate founder sequences.
- Generate QTL architecture based on founder sequences.

Recent Population

- Select progeny.
- Generate gametes.
- Generate progeny.
- Cull parents.

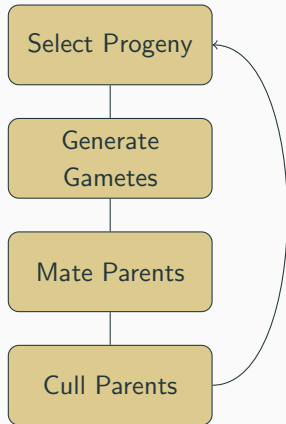
Genetic Architecture

- After the founder sequences have been created QTLs are assigned to a random set of SNP
 - Quantitative Trait (Quan):
 - Additive effect (Add).
 - Dominance effect (Dom): |Additive effect| * Degree of Dominance.
 - Fitness Trait (Fit):
 - Lethal.
 - Sub-lethal.
 - Relationship between Quantitative and Fitness effects
 - Proportion with quantitative and fitness effects.
 - Correlation.

$$\text{Quan(Add)} \leftrightarrow \text{Quan(Add + Dom)} \leftrightarrow \text{Quan(Add + Dom) + Fit} \leftrightarrow \text{Fit}$$

Recent Population

- Selection and culling within a generation based on either estimated breeding value (EBV), true breeding value, phenotype or random.
- EBV generated from pedigree or genomic relationship matrix.
- A marker array is generated from SNP that aren't QTL.
- Multiple options are available to make the simulation more realistic:
 - Maximum number of full-sibling kept within a family.
 - Differential mate allocation by age of sire.
 - Avoidance matings.



Summary Statistics

A number of summary statistics are created within each generation relating to:

- LD decay metrics.
- Mean phenotype and genetic values.
- QTL frequency.
- Number of founder or new mutations fixed or still segregating.
- Inbreeding metrics based on genomic and pedigree.
- Mean number of lethal or sub-lethal genotypes in the homozygous or heterozygous state.
- Mean fitness value of an animal and lethal equivalents.

Computing Procedures

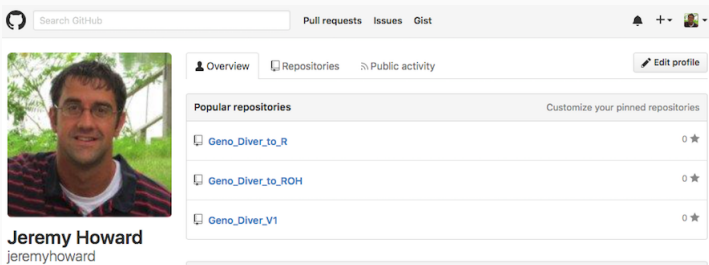
- Intel MKL libraries for matrix multiplications
- Allows for multithreading.
- Generates SNP-by-SNP relationship matrices based on strategies outlined by Aguilar et al. (2011).
- Generates inverse by updating previous generation based on either Meyer et al. (2013) or Misztal et al. (2016).
- Input is sequence information and has been tested for 1,000,000+ marker panel.

- Minimize relationships.
- Simulate sex-limited traits.
- Incorporate real genotype data.
- Incorporate external breeding value predictions.
- Incorporate the use of advanced reproductive technologies.



Source Code and Executable

- Source code for Geno-Diver can be found at:
 - <https://github.com/jeremyhoward>
- A linux executable is available.
- A comprehensive user manual with examples are also included.
- Any questions or inquiries can be directed to:
 - jthoward@ncsu.edu



The screenshot shows the GitHub profile page for Jeremy Howard. At the top, there is a search bar and navigation links for 'Pull requests', 'Issues', and 'Gist'. The profile header includes a profile picture of Jeremy Howard, his name 'Jeremy Howard', and his GitHub handle 'jeremyhoward'. Below the header, there are tabs for 'Overview', 'Repositories', and 'Public activity', with an 'Edit profile' button. The 'Popular repositories' section is visible, listing three repositories: 'Geno_Diver_to_R', 'Geno_Diver_to_ROH', and 'Geno_Diver_V1', each with a star count of 0.

Repository Name	Star Count
Geno_Diver_to_R	0
Geno_Diver_to_ROH	0
Geno_Diver_V1	0

Final Remarks

Final Remarks

- Healthy has broad definition
- Focus should include different aspects
- Several indexes do a good job in including "cow" related aspects
- Coping mechanisms and changing conditions still poorly known
- Same for the role of declining diversity and the ability to recruit new variability

Aknowledgements

- Jeremy Howard
- Francesco Tiezzi
- Jennie Pryce
- Kristen Parker Gaddis
- John Cole
- Kumud Dhakal
- DRMS
- CDCB
- US Jersey Association

Funding

- USDA NIFA
- Select Sires
- National Pork Board
- North Carolina Pork Council
- Smithfield Premium Genetics
- The Maschhoffs
- DEDJTR

Questions?