Increasing Genetic Gains for Food Security in the Developing World

The roles of genomic tools for genetic improvement (in animals) and their potential application in developing countries

Dorian Garrick
dorian@iastate.edu
J.L. Lush Chair in Animal Breeding & Genetics

Above-average Individuals

- Performance in a system context (life-cycle of a plant or animal population) always depends upon multiple traits
  - What traits should comprise the portfolio that distinguishes an above-average individual? (Known as the Breeding Objective)
  - What information can be obtained to rank the candidates? (Known as the Selection Criteria)

Genetic Improvement

Is the simple result of using above-average individuals as parents of the next generation

Two kinds of production systems

- Those where animals are typically
  - in their thermo-neutral zone, have adequate quality and quantity of feed, and are treated for (or maintained) generally free of disease (High octane environments)
  - subjected to climatic stress (heat or cold), nutritional deprivation in some seasons or stages of life, one or often more diseases (endo- or ecto-parasites) virus, bacterial and protozoal challenges, or often all of these at once (Stressful environments)
This creates two challenges

- First, the definition of the breeding objective may differ between the two kinds of production systems.
- Second, and more importantly, the criteria needed to identify above-average individuals would likely be very different in these two kinds of production systems.

Practical Improvement

- The (time and monetary) cost of measuring every potential selection candidate greatly exceeds the benefit for anything other than “natural selection.”
- So we concentrate our measurement efforts in as few selection candidates as possible, then we multiply and disseminate their genes or genotypes to as much of the world-wide population as possible.
- This is known as a tiered nucleus breeding structure.

Nucleus Breeding Structure

- There are two critical aspects:
  - The germplasm nucleus, from which future generations will be descendants.
  - The information nucleus, from which selection criteria are obtained in order to identify above-average selection candidates.
- In plants, the germplasm resource (accessions) and the information resource (field studies) are physically separated and comprise different activities.
- In animals, information has typically been measured directly on those individuals that represent the nucleus germplasm.

Genotype-Environment (GxE)

- In plant breeding, it is feasible to undertake yield trials of the genotypes over the range of target environments before deployment of the genotypes.
- In animal breeding, the information for selection decisions is typically collected in either the high-octane environment or the stressful environment, and deployment is by gametes rather than genotypes.

Nucleus Breeding Programs

- For a variety of reasons, successful breeding programmes (those without market failure) have been more commonly achieved in high-octane rather than stressful environments.
- In stressful environments the options are to use the germplasm from selection in the high-octane environment (seldom works), or to develop a program specifically for the stressful environment (seldom works – even in developed countries).

Local Assessment of Germplasm

- In plant breeding, genotypes are deployed in stressful environments, typically after having been characterized as above-average under similar conditions.
- In animal breeding, sire gametes are deployed in stressful environments, and assessed in terms of F1 offspring that benefit from locally adapted dams without having been previously characterized in that environment.
  - Although the F1 may be alright, F2 and BC are not.
Issues

• Not one of “developed” vs “developing” countries, but high-octane vs stressful environments
• Market failure of improvement in stressful environments is common in developed countries
• Considering stressful environments, we must distinguish circumstances where
  • There has already been successful improvement prior to the availability of genomics tools
  • There has not been any history of genetic improvement

No history of genetic improvement

• This is the area with the greatest challenges and with the greatest opportunities

A quick fix?

If a conventional breeding program in a stressful environment has not achieved genetic improvement, it is unlikely that genomics tools alone will make any difference

Genomics tools can make a successful breeding program better

Opportunities for Genomics

• Parentage testing
• Characterizing inbreeding & diversity
• Marker-assisted Selection (MAS)
• Marker-assisted Back Crossing (MABC)
• Genomic Prediction
• Genome-wide association studies (GWAS)
• Gene-editing

Genomic Prediction

• More accurate estimates of breeding values at young ages — without progeny testing
• Opportunity for separation of the germplasm nucleus from the information nucleus
  • Obtain predictions of merit for animals in stressful environments without having to manage the germplasm nucleus itself in that environment
  • But it requires high quality phenotypic data

Theoretical Basis for Accuracy

Reliable prediction requires large training populations of genotyped and phenotyped individuals

Predictive Ability = Accuracy (r) = correlation true & predicted merit

Goddard & Hayes (Nature Reviews Genetics, 2009)
Genomics in Stressful Environments

• In the post-genomics era, phenotype is king!

• Phenotypes are required for genomic selection and for genome-wide association studies
  • Large numbers of phenotyped individuals are required
    • Ideally >10,000
  • Individuals need to be within sizeable cohorts that had consistent environments

• If improvement was achieved pre genomics, these phenotypes are likely already available

Priorities required for applying genomics in Developing Countries

• In stressful environments without a history of improvement, the major needs are
  • Production Phenotypes (in cohorts)
  • More Phenotypes (nature and scope)
    • endo phenotypes that more directly reflect adaptation to climatic, nutritional or disease stress than measuring yield
    • life time production phenotypes beyond the immediate yield phenotypes

Sequence Variants

Case Study 1: Offspring from a genomic selected bull

Caused by a de novo (new) mutation - not present in ancestors

>25% affected did not lactate

700k SNP mapped to BTA23

Only one of 7 novel SNP in Halcyon and Matrix WGS, and that mapped to PRL
SNP showed complete concordance in 1,045 affected and 1,160 wild type progeny
Prolactin Mutation reduces Sweating

Mutant animals produced less than half as much sweat as wild type

Exome sequence analysis
Cold tolerance variants in Yak and Bison?
• Subset variants to PRL and PRLR
• Those present in Yak, absent in all other breeds:
  • PRLR p.Cys354Phe – deleterious (SIFT)
  • PRLR p.Tyr427His – benign (SIFT)

Senepol Crosses exhibit dominant slick

 Steven Lukefahr Texas A&M Kingsville
 Tul (0.5), Senepol (0.25), and Red Angus (0.25)

Yaks
“...the function of the sweat glands is poorly developed...the absence of sweating in the yak assists cold tolerance but helps make the yak intolerant to heat...”
Prolactin Pathway involved in Heat and Cold Stress

- What other private variants exist among locally adapted breeds that are suited to stressful climatic conditions in other parts of the world?
  - Hot areas such as in Africa?
  - Cold areas such as Tibetan plateau?

- These would be obvious targets for marker-assisted introgression – not lonely from local breeds

### Case Study 2: Holstein vs Jersey Haplotypes

![Graph showing effects of hidden haplotype states](image)

Experiment with 800 F2 animals in NZ

### Birthweight in Nellore in Brazil

![Graph showing birthweight](image)

Utsunomiya et al BMC Genetics 2013:52

### PLAG1 in Holstein vs Jersey

![Graph showing PLAG1 in Holstein vs Jersey](image)

Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature


### PLAG1 effect in topically-adapted breeds (Australian data)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Substitution Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at 1st Corpus Luteum (CL)</td>
<td>38 days</td>
</tr>
<tr>
<td>Post partum anoestrous interval</td>
<td>15 days</td>
</tr>
<tr>
<td>Presence CL before weaning</td>
<td>-5%</td>
</tr>
<tr>
<td>Weight at CL</td>
<td>36 lb</td>
</tr>
<tr>
<td>Age at 26 cm Scrotal Circumference</td>
<td>19 days</td>
</tr>
</tbody>
</table>

### Whole-genome Re-sequencing (WGR)

- WGR of locally adapted breeds
  - does not rely on extensive existing phenotypes
  - will expose private variants that might be responsible for known breed level adaptation to stressors
    - Nutritional stress (including water deprivation in camels)
    - Heat stress (e.g. prolactin pathway variants)
    - Disease stress
  - New improved breeds could be produced by introgressing (or gene editing) these variants

Genome—wide Association Studies (GWAS)

• Particularly valuable among crosses between locally adapted parents outcrossed to parents developed through selection programs in less stressful environments
• Local breeds (e.g., Nellore or Brahman) can also generate useful information when combined with knowledge obtained in selection programs from improved environments (e.g., PLAG1)
• Requires reliable phenotypic data on sizeable cohorts of animals

Summary

• The limitations for using genomics to improve animal performance in stressful environments such as exist in many developing countries are quite different from the limitation in plants

Summary

• The major limitation to improving animal performance in developing countries is the lack of reliable pedigree and performance records from cohorts of animals
  • Productivity, reproduction, longevity, disease etc
• An immediate and ongoing opportunity exists to leverage knowledge of exonic variants obtained from next-generation sequencing of locally-adapted breeds
• Given sufficient numbers of reliable performance records, genome-wide association studies will provide useful QTL information for use in improvement