ANIMAL SCIENCE | CIAG | IOWA STATE UNIVERSITY | Animal Breeding & Genetics

How Whole Genome Selection may affect NCE

Dorian Garrick

What is Whole Genome Selection?

- WGS (or more correctly in this context, Whole Genome Evaluation) involves, in its simplest form, the calculation of an EPD for chromosome fragments (rather than animals)
- The EPD of an animal is the sum of the EPDs for all the chromosome fragments that are carried by the animal
- If it worked simply and perfectly, it would not require any pedigree or performance data

What is National Cattle Evaluation?

PEDIGREE & PERFORMANCE DATA

BREED ASSOCIATION SERVICING ENTITIES

EPD

R & D KNOW HOW

RELIABLE TIMELY COST-EFFECTIVE ESTIMATES OF TOTAL ECONOMIC MERIT

What is Whole Genome Selection ?

- WGS uses knowledge of the collective *inheritance* of chromosome *fragments* from the *entire genome* to predict genetic merit(s) (eg EPD, Comb. Abil.)
  - The size of the fragments is influenced by available marker technologies
    - Currently 50k SNP chip
  - Knowledge comprises recent (LA) and ancestral (LD) fragment inheritance/variation & analytical experience
    - Immature discipline

What is needed for WGS ?

- **Molecular technology** to reliably (and cost effectively) identify chromosome fragments
- High levels of linkage between identifiable chromosome fragments and genes contributing to variation in the trait
  - improved by denser marker panels (eg 50k / 500k / 1m / individual sequence)
  - influenced by recent & ancestral population structure

What is needed for WGS ?

- **Phenotypes** on suitable *resource* populations to enable calculation of the *effects* of chromosome fragments. Resource populations must consider
  - at least 4x animals as there are causal genes
  - sex, stage of life, envt, breed considerations
    - Dominance, Epistasis, GxE
- **Analytical technology** to robustly estimate effects with more explanatory variables (SNPs) than data

Operational WGS

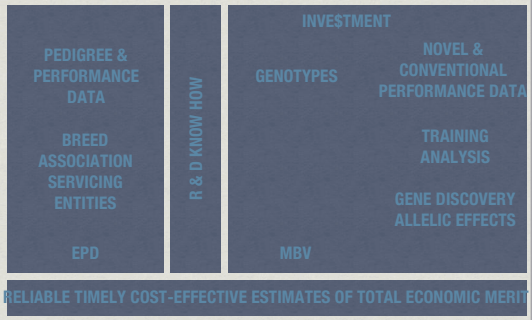
- Two-step Process
  - Step One: Bayesian Training
    - Molecular, phenotypic and analytical prerequisites
  - Step Two: MBV prediction
    - "Know how" from step one, and genotypes on new animals

Current Status of Bovine WGS

- Extensive research analysis being undertaken
- Correlations of MBV and EPD in range 0.5-0.7
  - 25% to 50% genetic variation can be accounted for using 50k markers
    - 50% to 75% variation not explained
      - can be predicted from PA &/or PT
- Nature and scope of routine delivery in the beef industry is uncertain

Industry Infrastructure

Industry Infrastructure



Future Status of WGS

- Compared to today:
  - Denser SNP panels (if not genomic sequence)
  - Heterogeneous panels according to animal "status" (i.e. more dense information on legacy sires and less dense panels on non-parents)
  - Greater account of available variation
    - Less value from adding NCE
  - Extension to epigenetic prediction

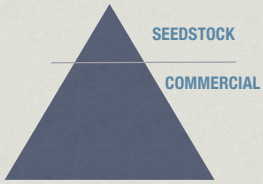
Role of Phenotypes in WGS

- Training (retraining)
  - Ongoing for new traits/environments
  - May need repeating (or continuous training) for epistasis
- Accounting for variation not captured by WGS
  - This will diminish if/as the predictive capability of WGS increases

Application of WGS

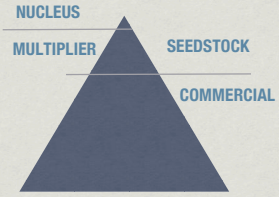
- Which animals would benefit from WGS?
  - Nucleus, multiplier, commercial parents
    - What is benefit relative to cost?
- Which people (or organizations) benefit from WGS?
  - Financial benefit is required to fund the servicing & other infrastructure costs

Everyone Must Benefit

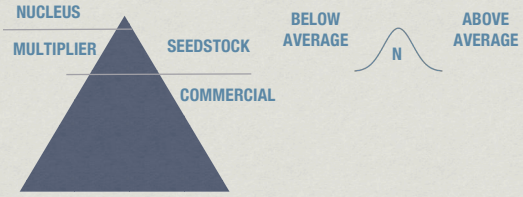


GENOTYPING IN SEEDSTOCK TIER MUST BE PAID FOR IN BULL REVENUE

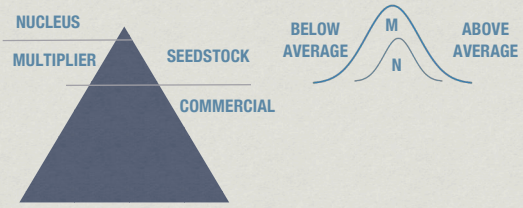
Everyone Must Benefit



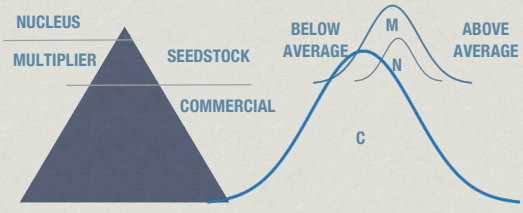
Everyone Must Benefit



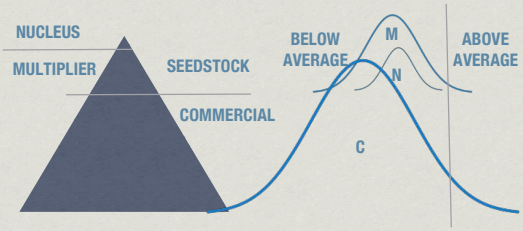
Everyone Must Benefit



Everyone Must Benefit

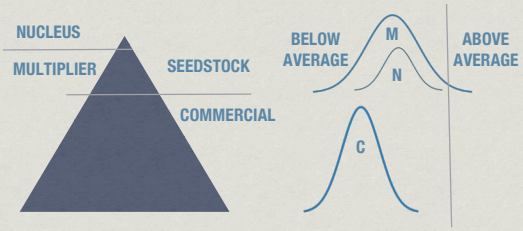


Everyone Must Benefit



THERE ARE MORE ELITE ANIMALS IN COMMERCIAL THAN NUCLEUS

Current EPDs



THERE ARE MORE ELITE ANIMALS IN COMMERCIAL THAN NUCLEUS
BUT WE DON'T KNOW WHICH ONES THEY ARE

WGS

- Will empower the evaluation of prospective sires that do not currently enjoy EPDs (eg Commercial sector)
- The portfolio of WGS MBV will be more economically relevant (and futuristic) than the current portfolio of NCE EPD
- The commercial market size is greater than the seedstock (appealing to genomics companies)
- WGS will reduce demand for "average" seedstock bulls

Stakeholders in NCE

- Breed Associations
  - Bull Breeders (nucleus & multiplier)
  - Bull Buyers
  - AI Companies
- EPD Servicing Entities
- Researchers interested in NCE
- Others (BIF, NBCEC, Cooperative Extension)

Current Business Model

- Cash
  - Bull breeders collect data at own expense
  - Bull breeders pay Breed Associations
    - Breed Associations (partly) pay Service entity
- Non cash
  - Bull breeders "freely" contribute the IP contained in their pedigree/performance data for the collective benefit of others
  - Bull breeders receive EPDs that have benefited from collective data collection efforts of others
  - Researchers improve analyses motivated by data

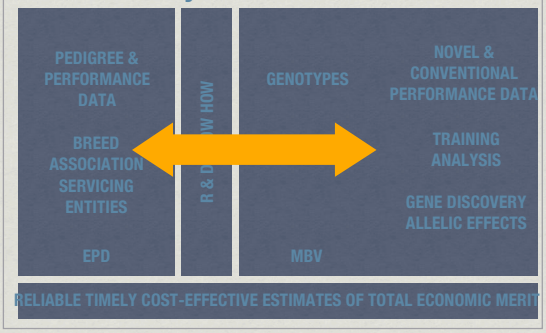
Stakeholders in WGS

- Genotyping Companies (currently Illumina)
- Genomics Companies
  - Bovigen/Catapult/Pfizer
  - Igenity/Merial
  - MMI Genomics/Metamorphix
- Breed Associations
- Researchers/Wannabe Gene Discoverers
- Owners/Controllers of Training Data

Future Business Model

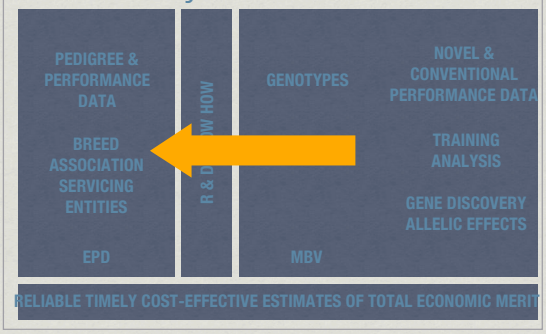
- Who/How are bull breeders rewarded for better EPDs?
  - Need to pay for genotyping/analysis costs
- Who invests in collection/analysis of training data
  - USDA ARS, companies
- Who rewards those who invest in training data?
- Who does the servicing (EPDs, MGVs, merged)

Industry Infrastructure



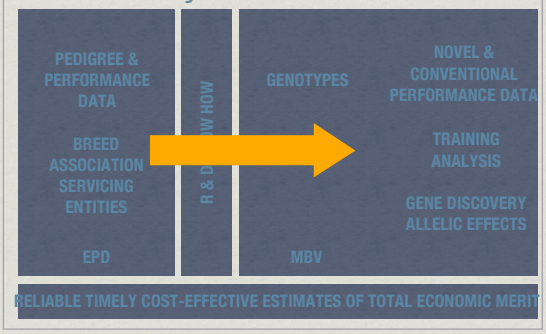
Industry Infrastructure

Industry Infrastructure

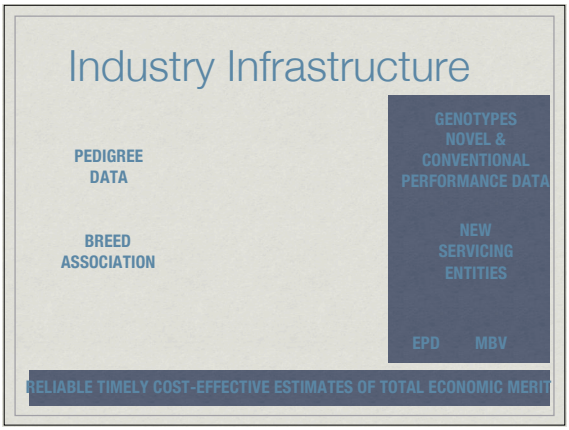


Industry Infrastructure

Industry Infrastructure



Industry Infrastructure



Summary

- Ultimately, the consumer/society will benefit
- In the short to medium-term, some early adopters will benefit, others may lose their shirt
- Late adopters will miss out on benefits
- The ultimate structure will be dictated by the collective action of a number of cultural, technological, economic and political factors operating in a period of many uncertainties