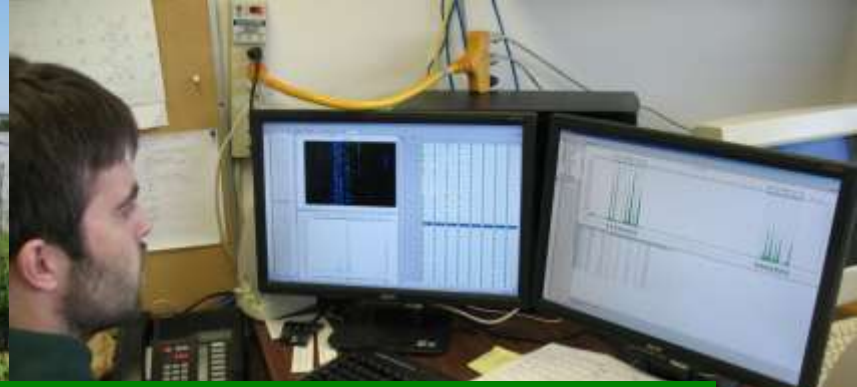


Marker Assisted Selection in Forage Breeding: A Broadened Vision

Heathcliffe Riday

U.S. Dairy Forage Research Center
Greener Horizons for Crops, Cows, and Communities
USDA-Agricultural Research Service

Animal Breeding

Animal Breeding
$$$$ to Phenotype

Smallest Selection Unit – One Animal

Inbred Line Plant Breeding

Self-Pollinated Plants
$$ to Phenotype

Smallest Selection Unit – Inbred Line

What Do You See?

1 Selection Unit

Thousands upon Thousands of Selection Units

Mass Selection

Cheapest and Easiest Breeding Program

Florex: ... **<u>One third acre</u>** of Dollard was **<u>planted</u>** ... At the end of six years, some **<u>2,000 plants</u>** were selected. These plants **<u>were lifted by digging</u>** ... Healthy, non-diseased plants were planted in a spaced, polycross nursery ... **<u>Syn 1 seed was produced on the surviving plants</u>** …

CW0401: is an ... synthetic ... with 87 parent plants. Parent plants were selected for persistence from replicated grazing tests following two and three years of ... grazing ... vegetative cuttings from … parent plants were established to produce the breeder seed …

RC9603: ... a ... trial with 18 entries was seeded ... 25 plants each of the following entries in this trial were dug ... These plants were ... transplanted into a crossing block ... and seed bulk harvested ... The resulting population was screened in the greenhouse ... for resistance ... Approximately 300 screened plants were sent to ... for seed production ...

Source: PVP Office and AOSCA – National Alfalfa and Misc. Legumes Variety Review Board

A Revolution in Breeding Dairy Cattle: Genomic Selection

$$\Delta G = khr\sigma_A$$

Genetic Gain

Selection Intensity

Heritability

Additive Genetic Variance

$$h^2 = \frac{\sigma_A^2}{\sigma_e^2 + \sigma_A^2}$$

Requires Phenotyping – Goal: Reduce σ_e

$$\Delta G = khr\sigma_A$$

$$\Delta G = kr\sigma_A$$

Correlated or Indirect Selection

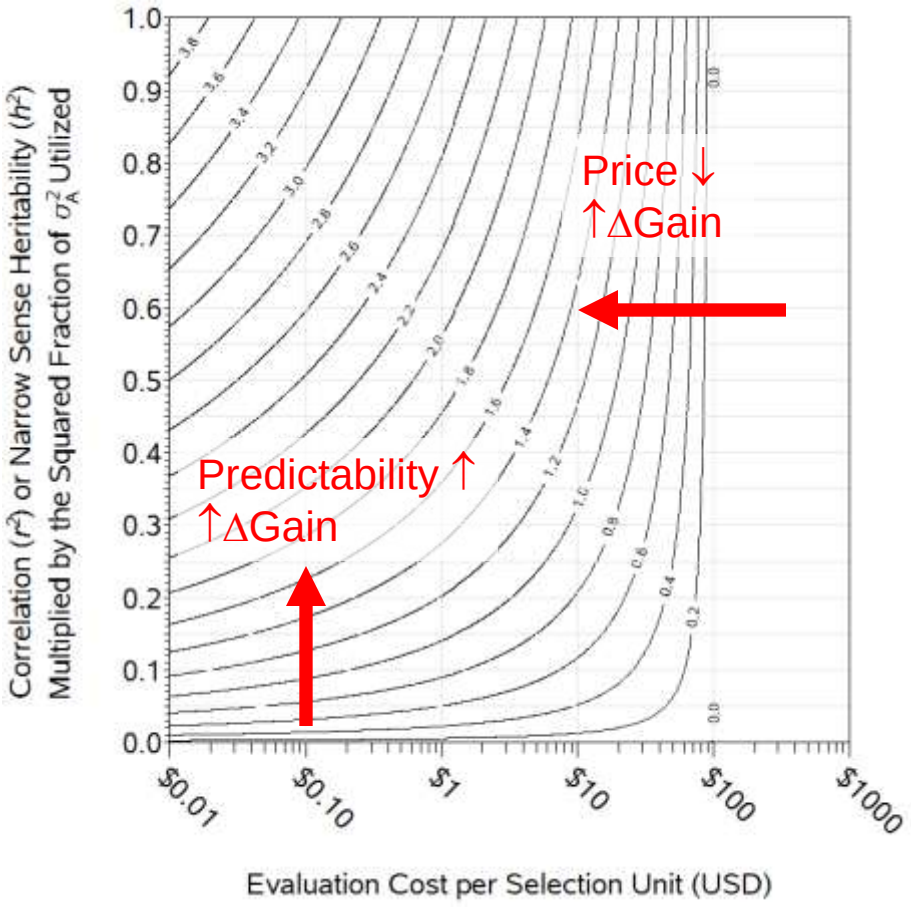
Requires no phenotyping!
However . . .

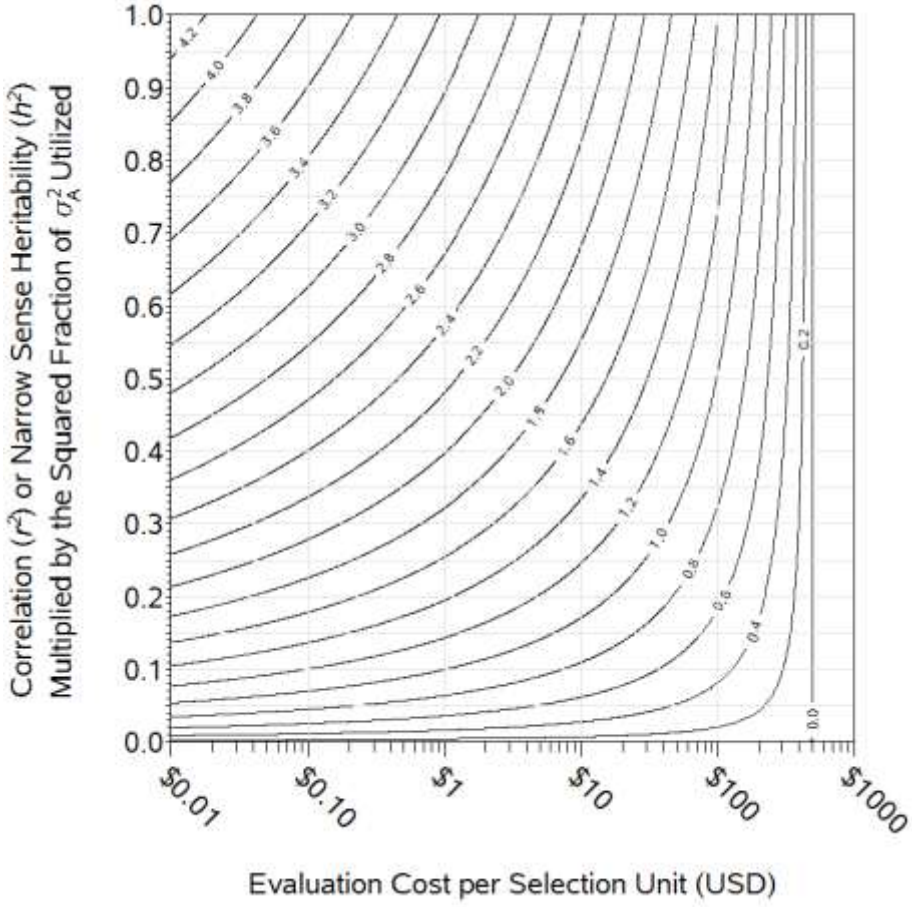
1. Requires genotyping
2. Subject to linkage
3. Single marker not usually conducive to prediction models were $r^2 = 1$

B'Zillion Markers! Problem Solved

Selection Predictability vs. Price

Goal: Invest $1,000 and evaluate as many selection units as possible in a population and end up with 10 genotypes





Contours = ΔG (assuming $\sigma_A = 1$)

Investing $5,000?

Marker Assisted Selection

"Structure"	"Correlation"
$\Delta G = khr\sigma_A$	$\Delta G = kr\sigma_A$
What about Here?	Current Use

"Structure" $\Delta G = khc\sigma_A$

Halfsib selection

$$\Delta G = \frac{kc¼\sigma_A^2}{\sqrt{\frac{\sigma_e^2}{r} + ¼\sigma_A^2}}$$

↑ Replication
↓ Error
↑ Heritability
↑ Selection Gain

Parental Control
(½, 1, or 2)

Typical Parental Control Values

Dig out of nursery, intermate (parental control = 1)

Intermate individuals from remnant seed of selected families (parental control = 1)

Intermate parents of selected families (parental control = 2)

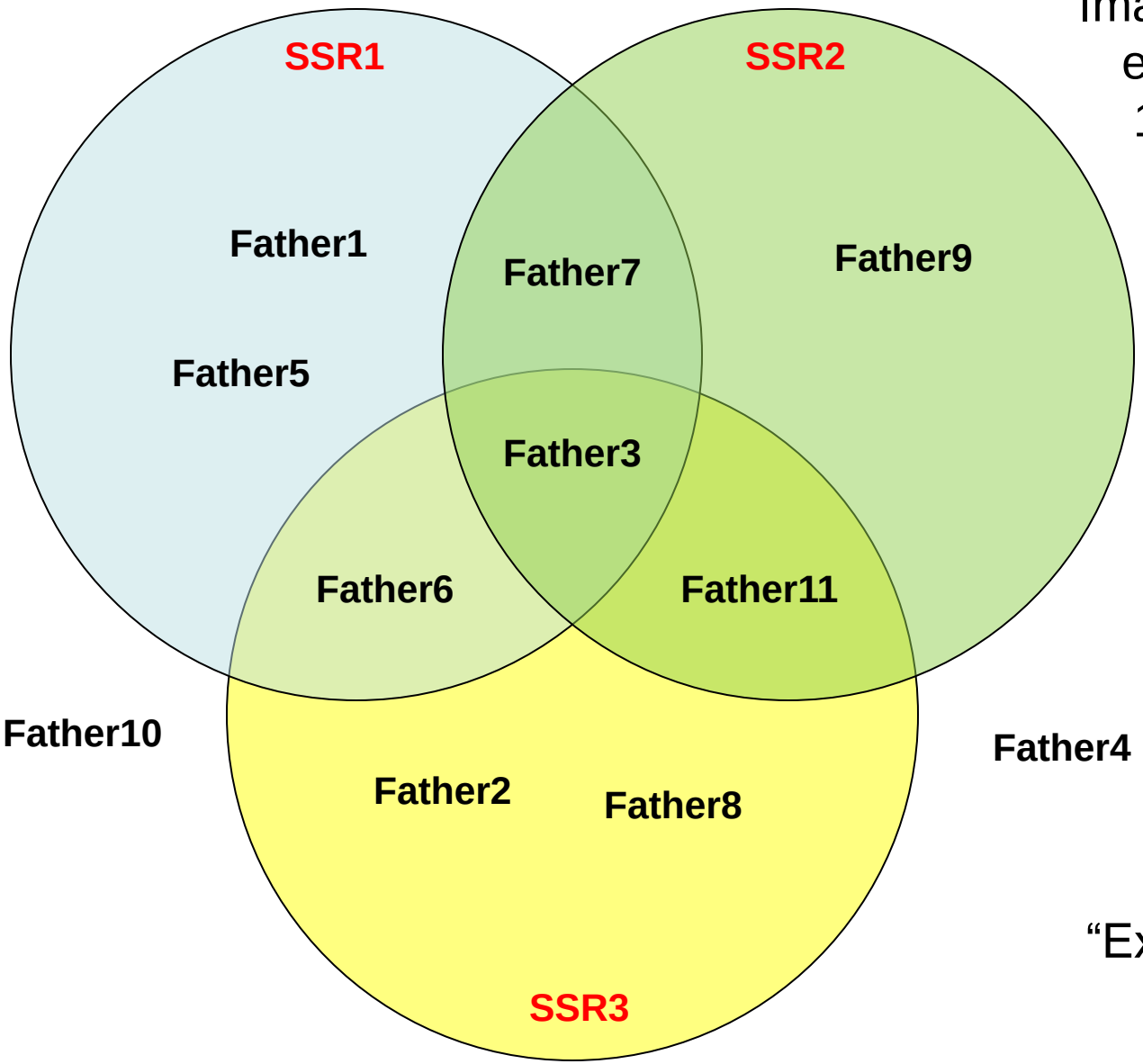
"Structure" $\Delta G = khi\sigma_A$

Halfsib selection

$$\Delta G = \frac{k\frac{1}{4}\sigma_A^2}{\sqrt{\frac{\sigma_e^2}{r_M} + \frac{1}{4}\sigma_A^2}} + \frac{k\frac{1}{4}\sigma_A^2}{\sqrt{\frac{\sigma_e^2}{r_F} + \frac{1}{4}\sigma_A^2}}$$

What if we know the father?

Paternity Testing is Easy



Imagine a Venn Diagram
each circle is an SSR
10 parent polycross

“Exclusion Analysis”

Paternity Testing Software

- Theory and software extensively developed and already in use in forensics, ecology, and other fields

- Programs Available:

  - Cervus (Kalinowski et al., 2007)
    - Advantage: user-friendly interface and data handling
    - Disadvantage: all potential fathers need to be genotyped

  - PATRI (Nielsen et al., 2007)
    - Advantage: Bayesian approach; all potential father need not be genotyped
    - Disadvantage: more difficult interface; data file input format not user-friendly

  - FAMOZ (Gerber et al., 2003)
    - Advantage: can handle dominant marker data
    - Disadvantage: can't get it to work; complex program set up and interface

Paternity Testing Software

- Other considerations for Cervus and PATRI
  - Molecular markers need to be co-dominant
  - Software can only be used in diploids
    - Although diploid sub-genome specific molecular markers in allopolyploids could be used
  - Molecular marker with null alleles need to be avoided since genotypes with single allele bands are assumed to be homozygotes (i.e. 2 alleles)

Paternity Testing in Polyploids

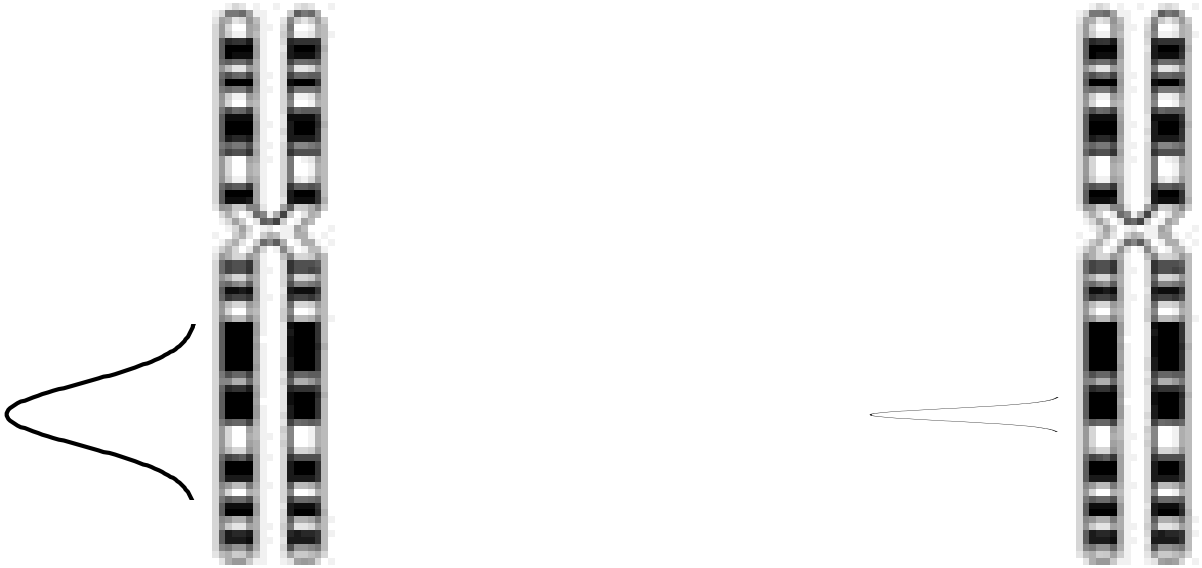
- Most perennial forage species are polyploids
- FAMOZ only polyploid paternity testing software
- Goal
  - Develop polyploid exclusion analysis SAS code
  - Develop single DNA reaction alfalfa paternity test

- Results:
  - A single 16 alfalfa SSR reaction developed
  - Exclusion analysis parentage testing SAS code developed
    - No need to calculate population allele frequencies
    - Can utilize molecular marker loci containing null alleles
    - Identifies self-pollination events

"Correlated" $\Delta G = k r \sigma_A$

- With plants individually phenotyped and both parents known (i.e. “structure” defined”), the breeding nursery becomes a complex mapping population

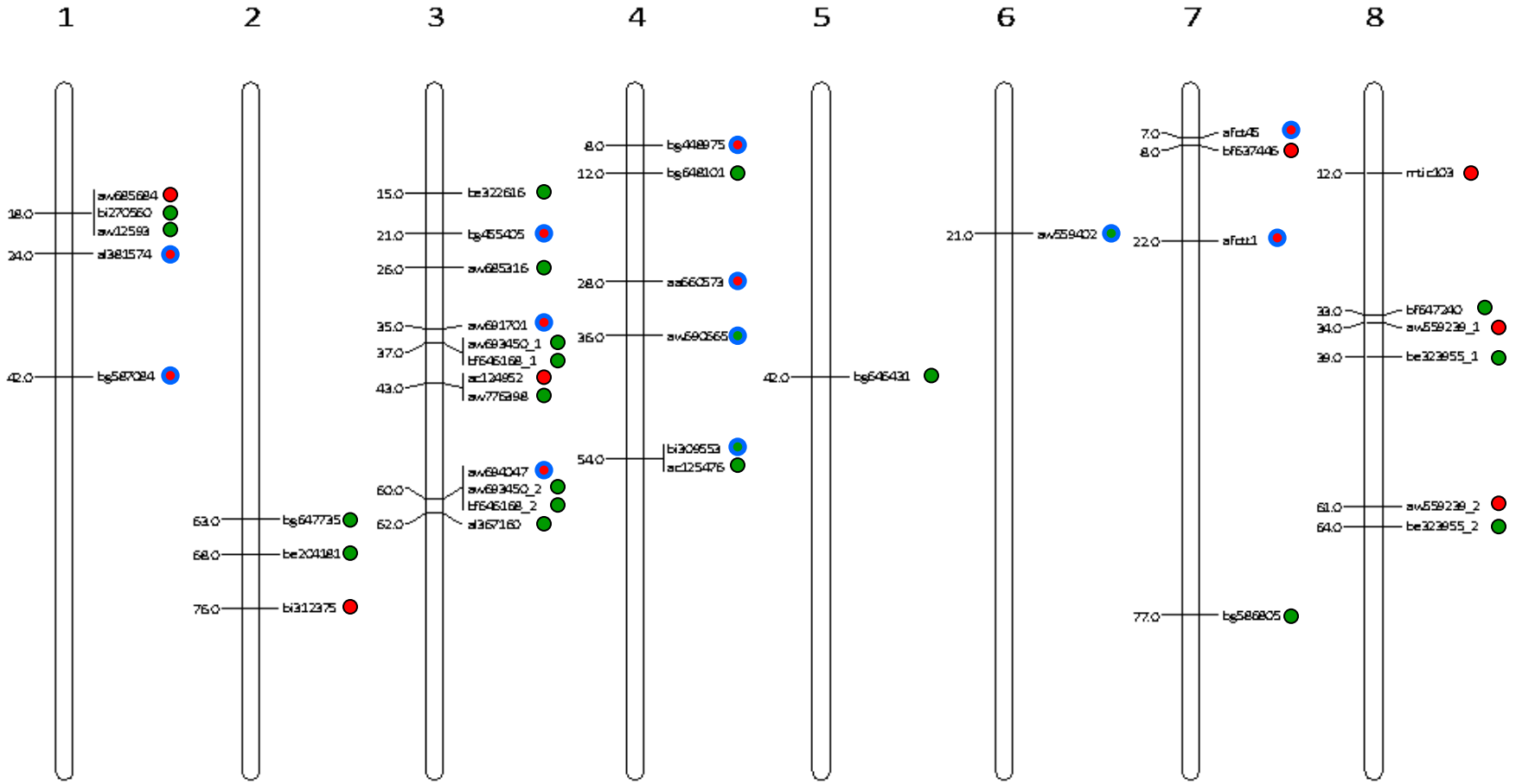
- Maximum linkage disequilibrium vs. Residual linkage disequilibrium



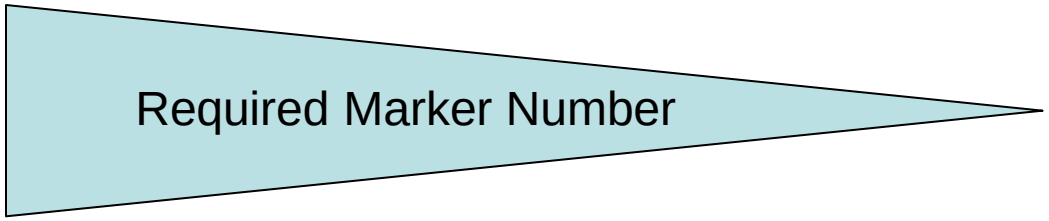
- Residual linkage disequilibrium → Whole genome selection

"Correlated" $\Delta G = k r \sigma_A$

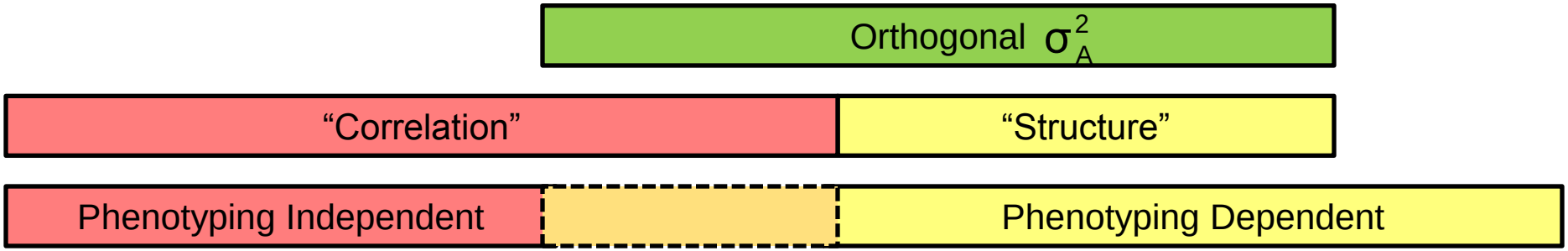
- Alfalfa 16 SSR multiplex reaction
- Alfalfa 18 SSR multiplex reaction
- Published: biomass yield, height, or regrowth QTL

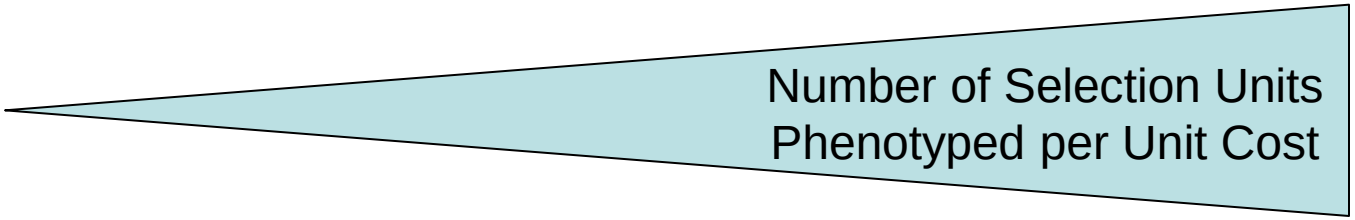


Maximizing Genetic Gain with MAS



Residual LD unknown loci + Residual LD known loci + Within maternal and paternal halfsib family max LD + Paternal halfsib family + Maternal halfsib family + Mass Selection

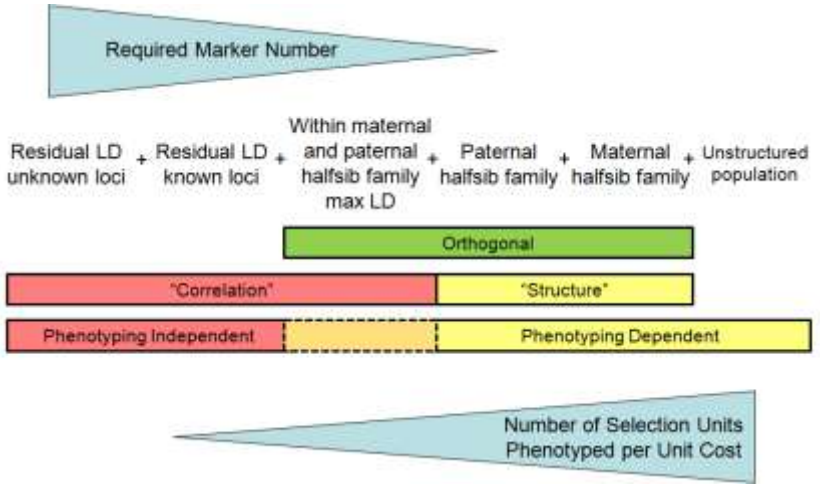




MAS Implementation

Sward

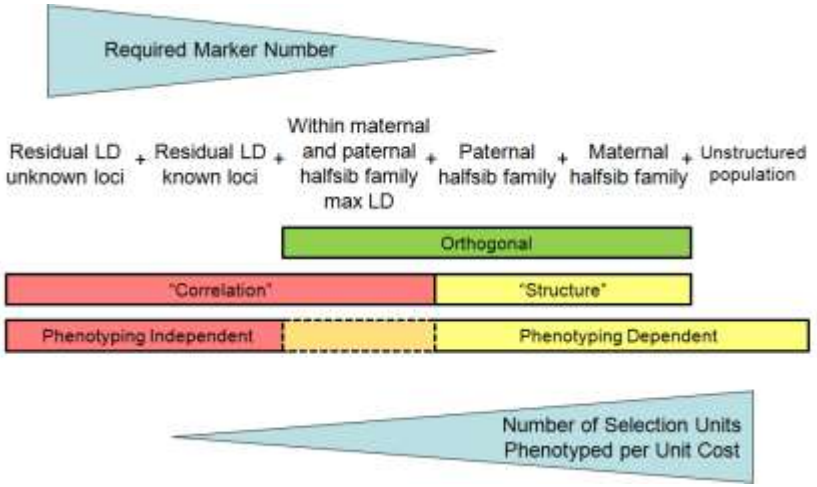
Space Plant



MAS In Swards

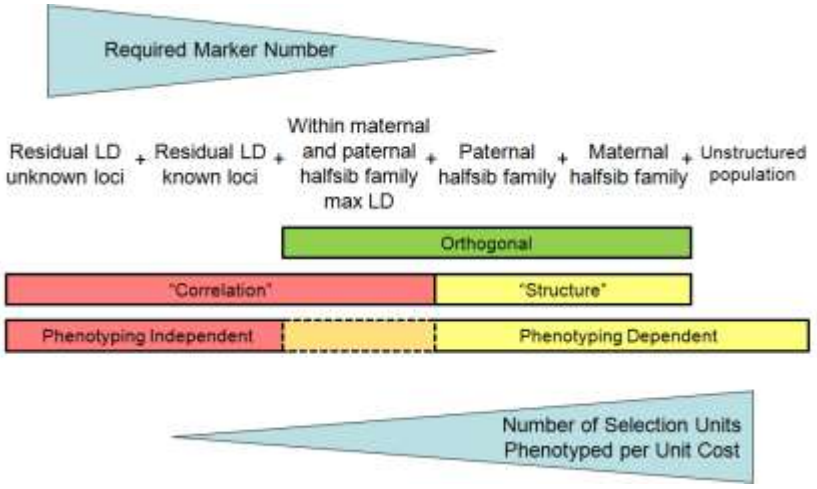
- If sward is grown from syn 1 seed and all parents are known
  - Phenotype a subset of plants
  - Phenotyped plants are also genotyped
  - MAS based selection based on “collective” stronger group phenotypic means, rather than individual “weak” phenotypes

- Many variations are possible variations possible
  - Particularly by comparing pre-planting “correlation” or “structure” frequencies to the same frequencies at the end of the sward trial



MAS Strategy and Polycross Size

- Smaller polycrosses (< 20 parents) vs. larger polycrosses (> 60 parents)
  - Lean towards “correlation” strategies in smaller polycrosses
  - Lean towards “structure” strategies in larger polycrosses



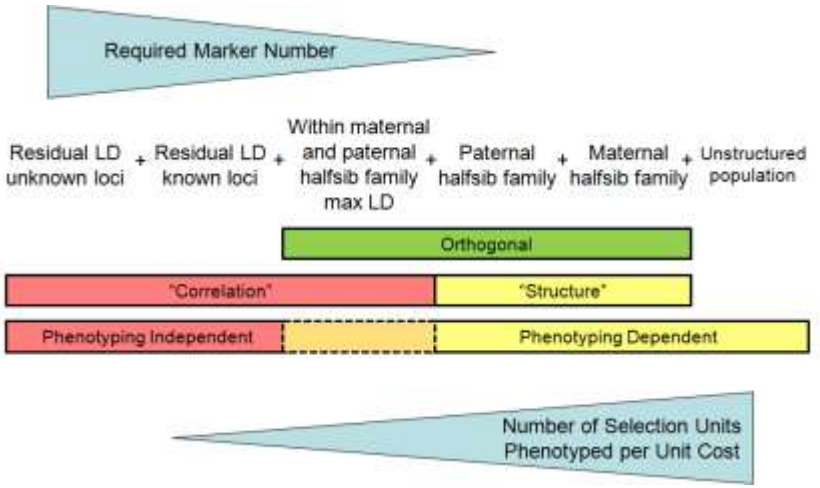
Selection on Cryptic Structure

"Structure"

$$\Delta G = khr\sigma_A$$

"Correlation"

$$\Delta G = kr\sigma_A$$



Collateral Benefits of Genotyping

- Genotyping reveals field crew sins
  - Cross-contamination between halfsib families
  - Threshing halfsib family mix-ups
  - Transplanting errors
  - (Field information helps reveal genotyping errors)
- Polycross pollination information
  - Incidence of self-pollination
  - Specific male gamete pollination distribution and frequencies
- Ability to check selected plants after 3-4 years of field evaluation for 'Volunteer' status

The Future is Now

Inventory of Freeze Dried Phenotyped Tissue

Inventory of Remnant Seed

- With paternity testing immediate meaningful marker assisted selection is possible in any forage breeding program
  - Requires genotyping infrastructure implementation
- Gateway marker assisted selection procedure
  - With genotyping infrastructure in place risks decreases and opportunities increase for implementing resource intensive MAS methodologies
- Phenotyped tissue can be warehoused like remnant seed for future exploitation when genotyping prices have declined

Acknowledgements

- Andy Krohn, John Raasch, Rebecca Heidelberger
- David Johnson (Cal/West Seeds)

- In depth discussions with numerous forage breeders about their programs, motivations, and philosophies

What's the Hurry?

What is your resource ($) environment?
Is there competition?
Efficiency vs. Speed?

Command Economy?
(Patronage System?)

Cost Reduction, Convenience and Logistics

- Central goal is to have one DNA extraction and one lab procedure per genotype