GENOMIC SELECTION IN PLANT AND ANIMAL BREEDING WITH POTENTIAL APPLICATIONS TO TOXICOLOGICAL RESEARCH

Russ Wolfinger
SAS Institute, Inc.
NIEHS Seminar
December 15, 2014

Copyright © 2013, SAS Institute Inc. All rights reserved.

General Mills Products

Cereal

Snacks

Baked Goods

Meals

Dairy

Crop Biosciences

Grain

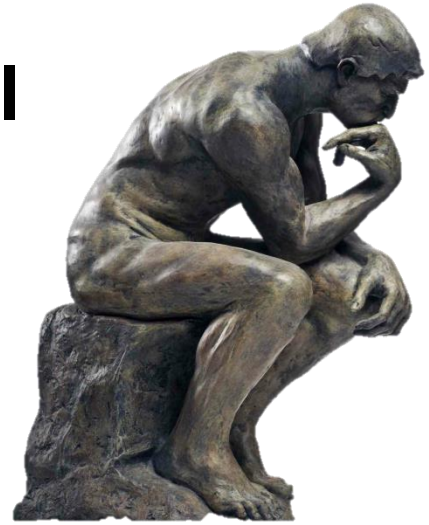
Fruit
Vegetables

Agronomy
Sustainability

Changing Landscape

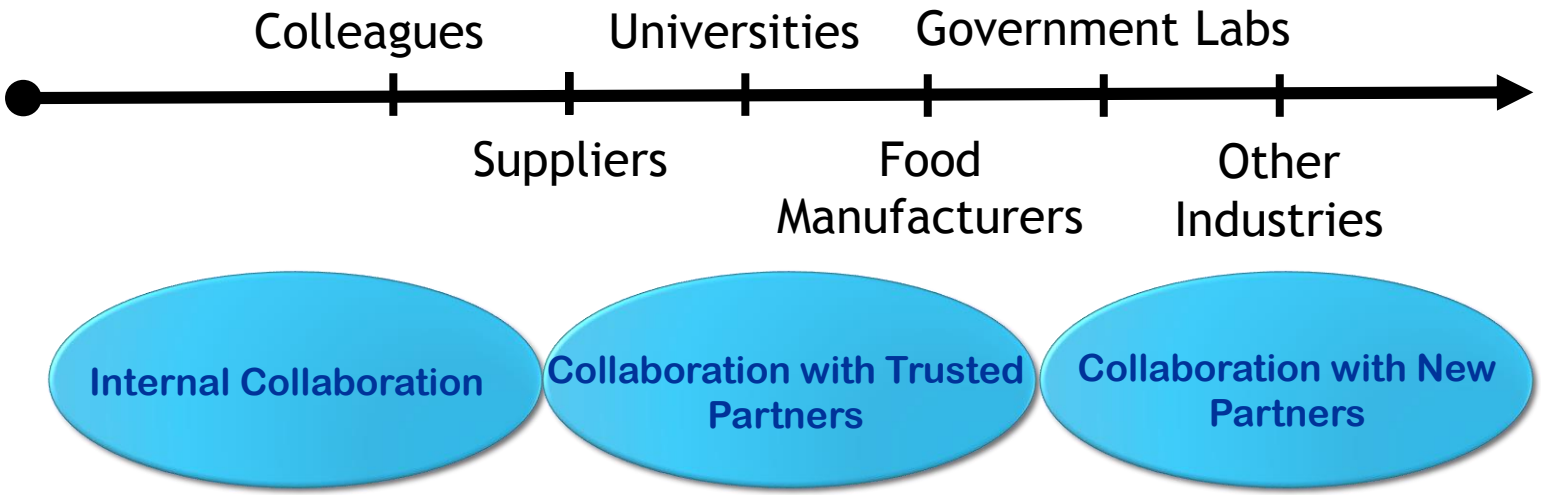
Invention Model

The lab is my world



Connection Model

The world is my lab



Copyright © 2013, SAS Institute Inc. All rights reserved.

Complex Breeding Focus

Grower Needs → Miller Needs → Product Needs → Consumer Needs

Advantaged Finished Product

- Yield Components
- Processing components
- Hedonic components

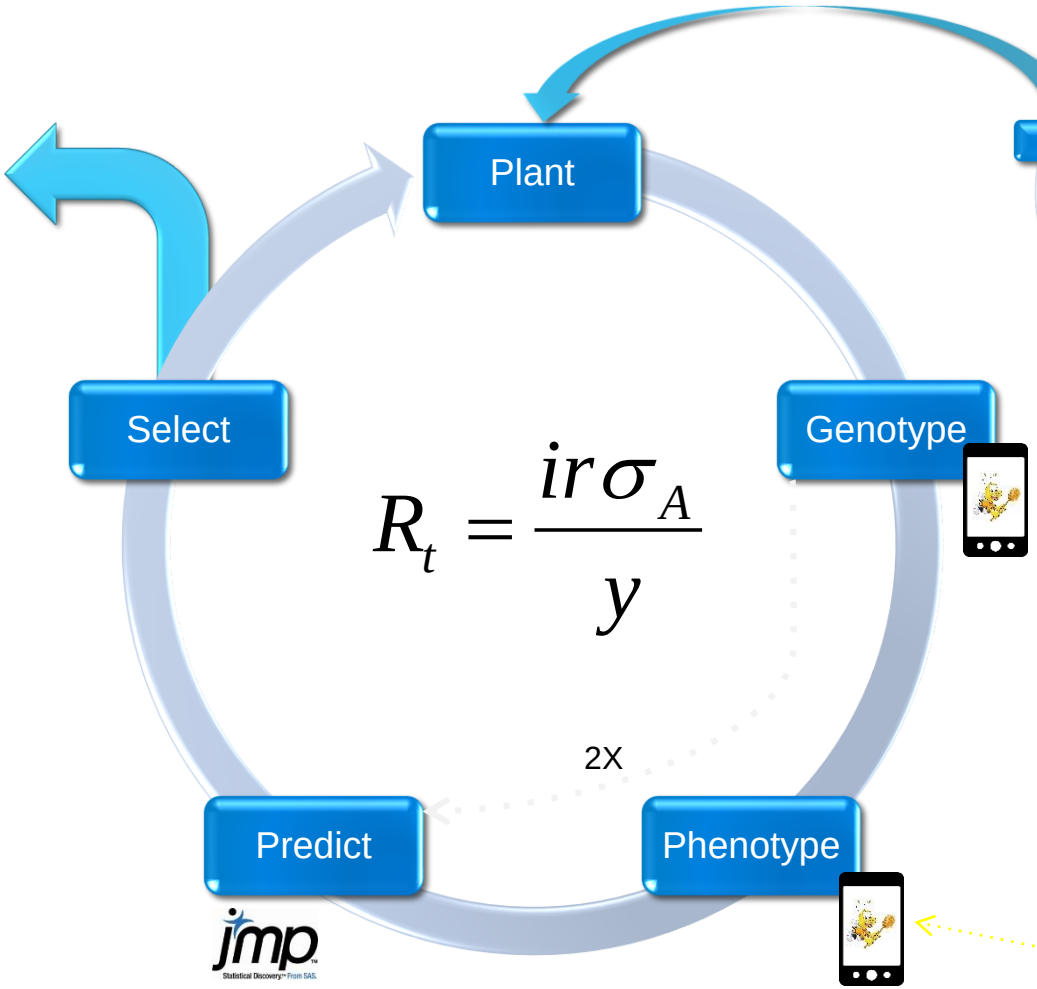
Johnson L. 2002. A dozen years of marker assisted sweet corn breeding. Presented at the International Plant and Animal Genome Meeting, San Diego, CA USA, 12 January 2002.

Copyright © 2013, SAS Institute Inc. All rights reserved.

High Paced Breeding Cycle

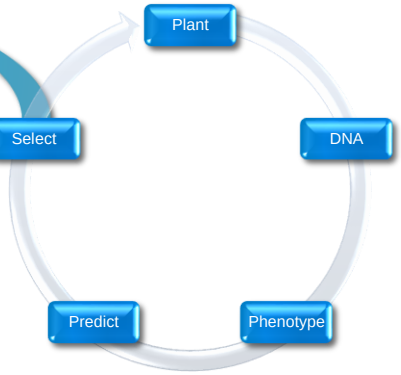
Advanced Trials

Next Generation Lines



$$R_t = \frac{ir\sigma_A}{y}$$

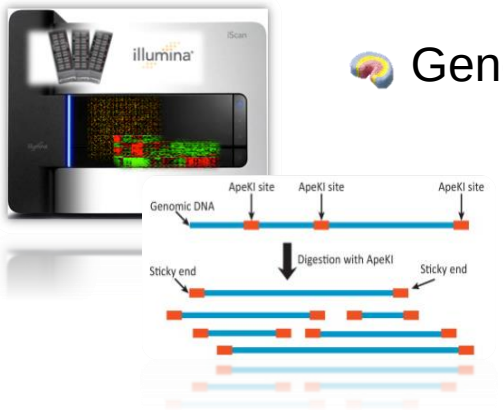
Breeding/Training Population



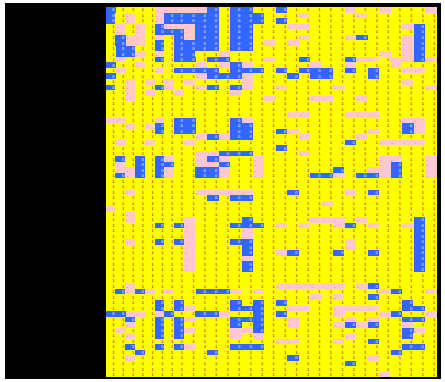
Developmental Populations

Copyright © 2013, SAS Institute Inc. All rights reserved.

Workflow



- Genotype
  - Illumina
  - Genotype by Sequencing
- Clean the genotypic matrix
  - Segregation
  - Minor allele frequency
  - Missing data/imputation



- Evaluate phenotypes
  - ANOVA
  - BLUP
  - Marker-trait associations

- Optimize predictive models
  - Must handle n << p
  - Cross validation

- Virtual breeding cycles
  - Cross evaluation and selection
  - Progeny simulation

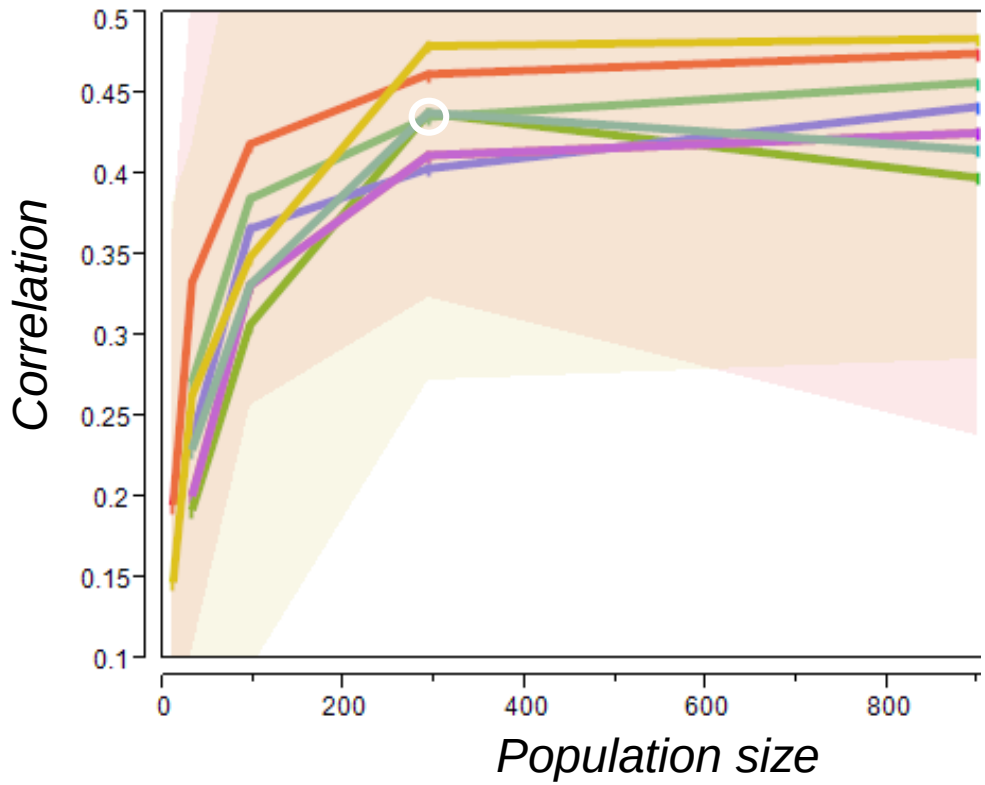
Statistical Discovery.™ From SAS.

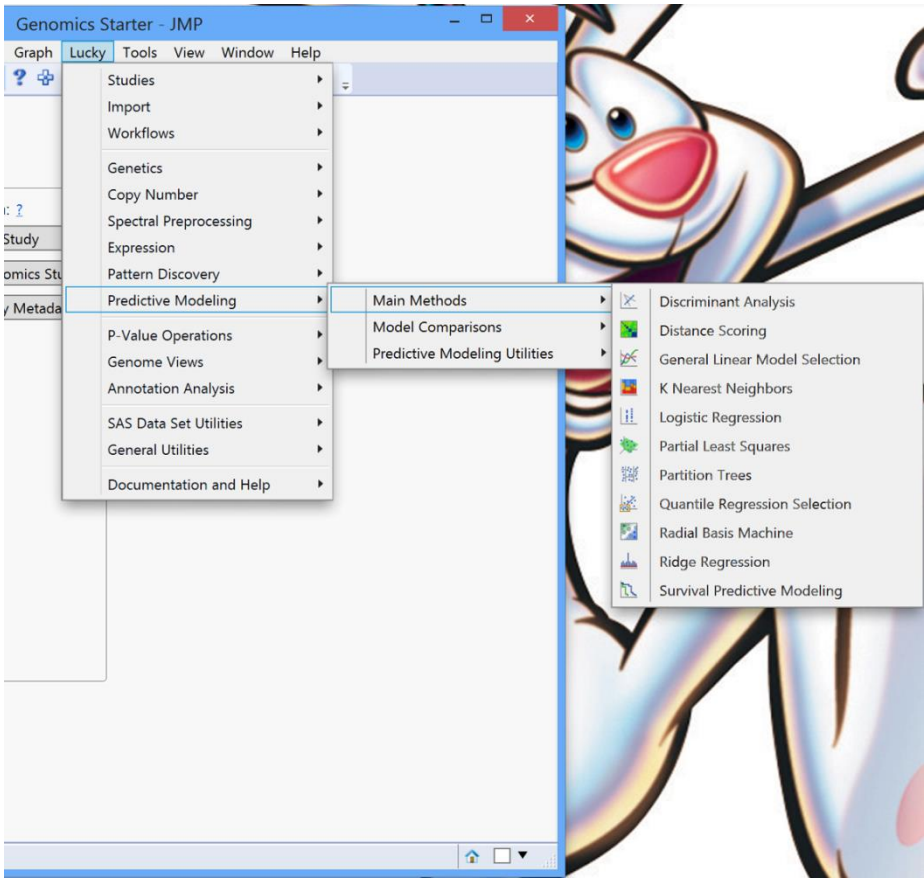
jmp
Copyright © 2013, SAS Institute Inc. All rights reserved.
sas THE POWER TO KNOW.

Predictive Model Selection

Try various models across all traits
- Grid of tuning parameters

Automated workflows
- Cross validation
- Learning curves

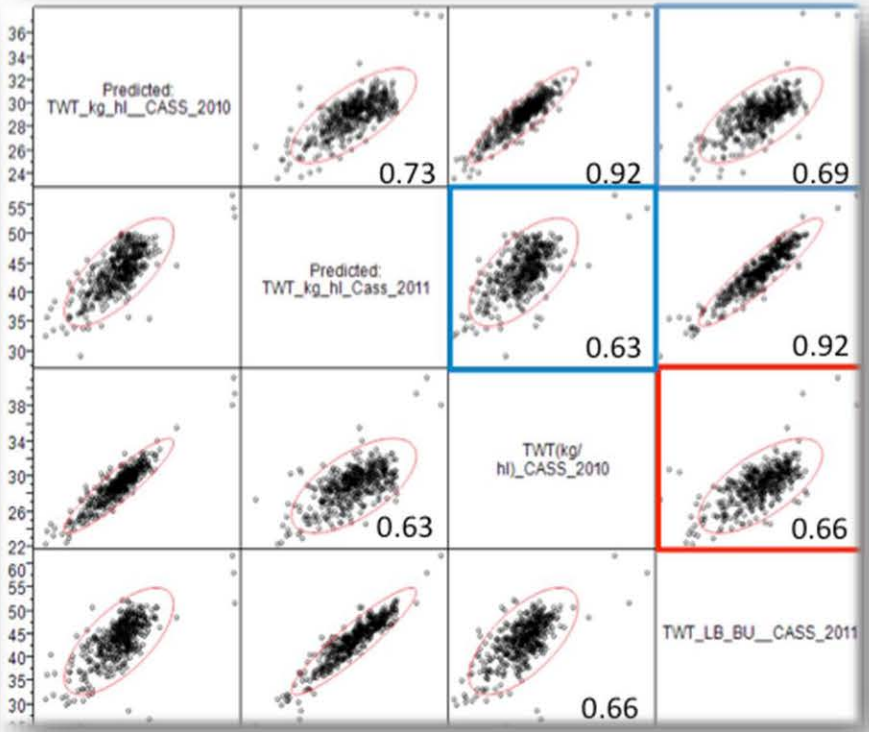


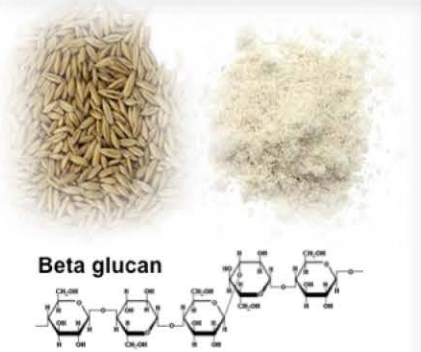


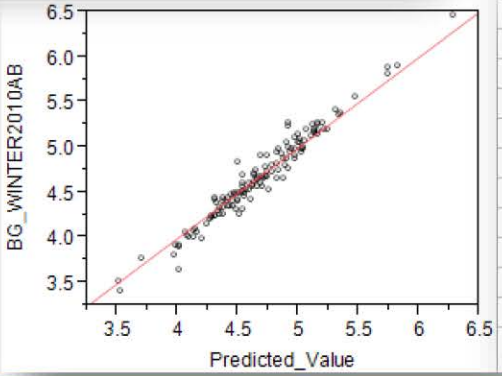
Y
CORR_M1_PLS_PLS
CORR_M2_PLS_PLS_20
CORR_M3_PT_Boosted
CORR_M4_PT_Boosted_20
CORR_M5_PT_Forest
CORR_M6_PT_Forest_20
CORR_M7_RR_RR
CORR_M8_RR_RR_20

reserved.

Predictive Model Validation







	NAME	COOPERATOR	Predicted: BG_WINTER2010A	BG_WINTER 2010AB
103	SECTARIAT LA495	LSU	4.92413328	4.95000
104	UFRGS077026.2	LSU	4.93758656	4.99000
105	TX07CS1268	TX A & M	4.96337543	4.92000
106	TX07CS2235	TX A & M	4.96337543	4.98000
107	TAMO 606	TX A & M	4.96587521	4.87000
108	TX05CS 542	TX A & M	4.96790619	5.11000
109	FL03001BSB.S7	LSU	4.99364088	5.05000
110	UPF95H900.10	LSU	4.99453399	5.14000
111	LA04014SBSB.39	LSU	5.01249361	5.09000
112	LA04004SBSB.121	LSU	5.01463286	5.06000
113	TROPHY	LSU	5.01745808	4.95000
114	HARRISON	LSU	5.03024168	4.98000
115	Rogers		5.03361809	5.00000
116	TX05CS 556	TX A & M	5.03476041	4.98000
117	UPFA22	LSU	5.05722227	5.07000
118	TX05CS 347.1	TX A & M	5.06391996	5.20000
119	Clintland 64		5.08907712	•
120	FL0238BSB.22	LSU	5.11027114	5.12000
121	Sure Grain	NC State	5.11245196	5.25000
122	TX07CS1039	TX A & M	5.12585795	5.16000
123	TX07CS2140	TX A & M	5.12585795	5.18000
124	UPF98H2000.4	LSU	5.12598583	5.20000
125	UFRGS087129.1	LSU	5.15376563	5.26000
126	TX07CS1402	TX A & M	5.15872817	5.15000
127	TX07CS2350	TX A & M	5.15872817	5.27000
128	LA04018SBSB.181	LSU	5.1658908	5.19000
129	Floriland	NC State	5.19391745	5.28000
130	HORIZON 270	LSU	5.21327203	5.20000
131	TX07CS1965	TX A & M	5.23528629	5.20000
132	FL03167BSBS.3	LSU	5.30234613	5.41000
133	Florida 500	NC State	5.34196285	5.36000
134	Santa Fe	NC State	5.34902826	5.38000
135	FL03053.S06.15.B.S1B	LSU	5.46843863	5.57000
136	TX07CS1948	TX A & M	5.74128564	5.88000
137	LA03046SBS7.B.S1	LSU	5.74654424	5.82000
138	LA97006GSB.59.2.4.SBS1	LSU	5.82432728	5.90000
139	FL04155.S06.31.B.S1	LSU	6.28286917	6.47000

Copyright © 2013, SAS Institute Inc. All rights reserved.

Cross Evaluation

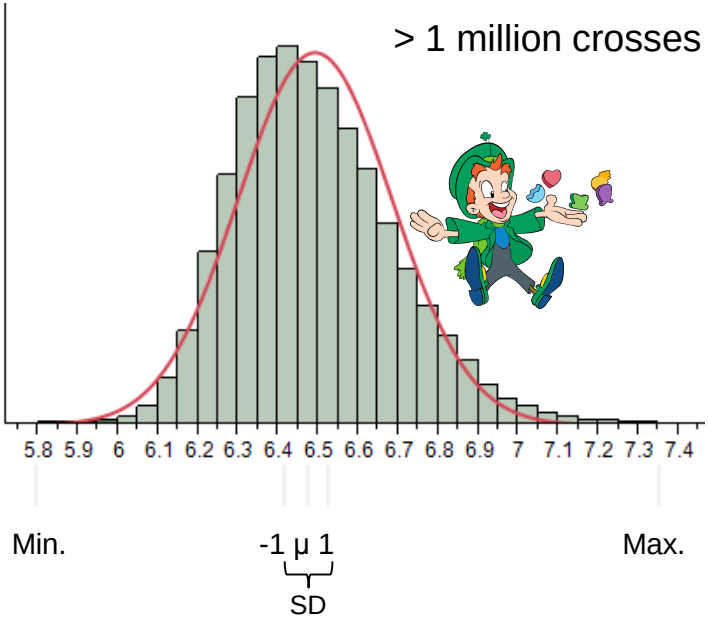
Training population

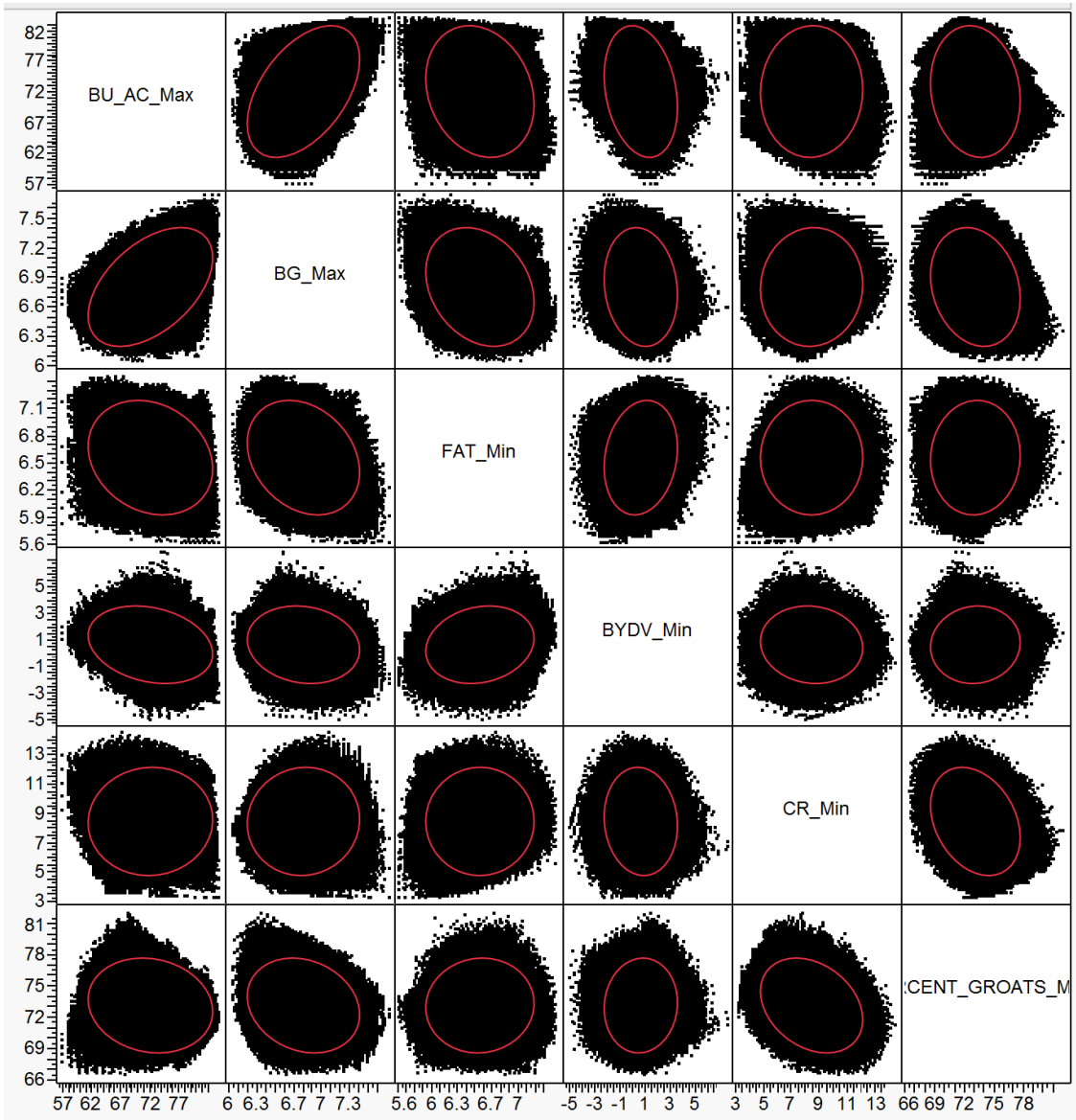
- 25 crosses
- 20 lines/cross

All pair-wise crosses

- Summary stats of distribution

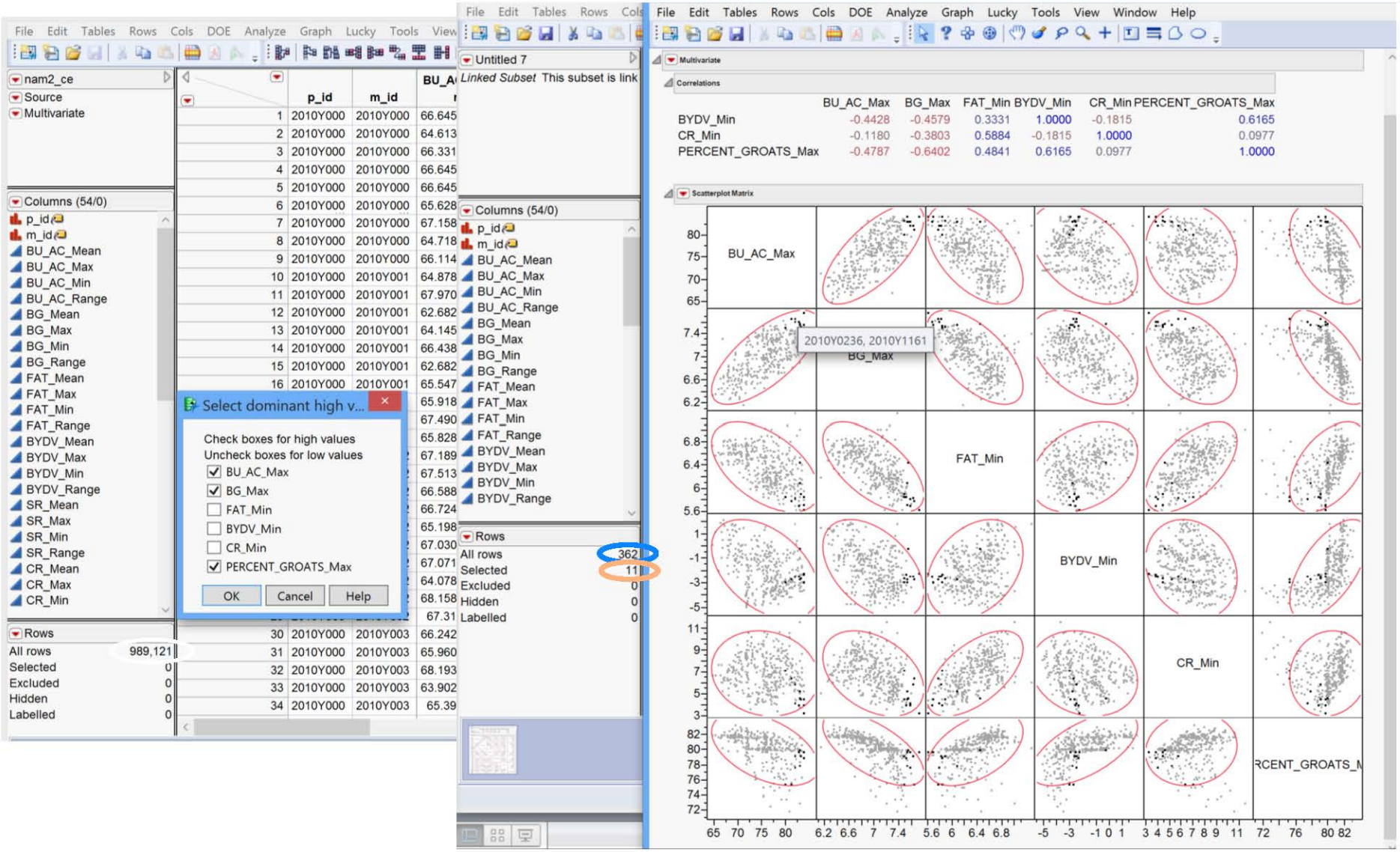
> 1 million crosses





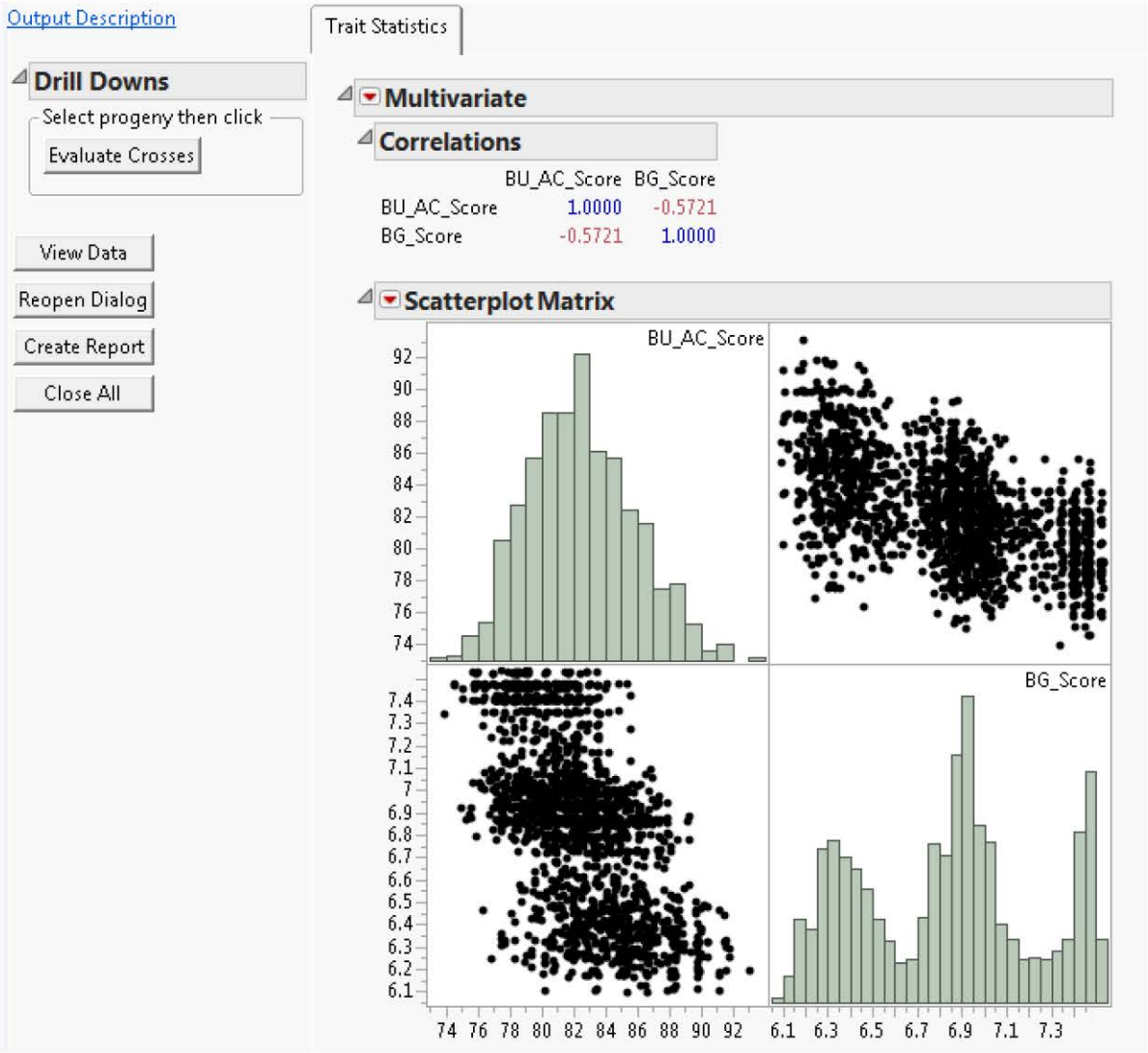
Copyright © 2013, SAS Institute Inc. All rights reserved.

Cross Selection



Copyright © 2013, SAS Institute Inc. All rights reserved.

Progeny Simulation



Copyright © 2013, SAS Institute Inc. All rights reserved.

Applications to Tox Research

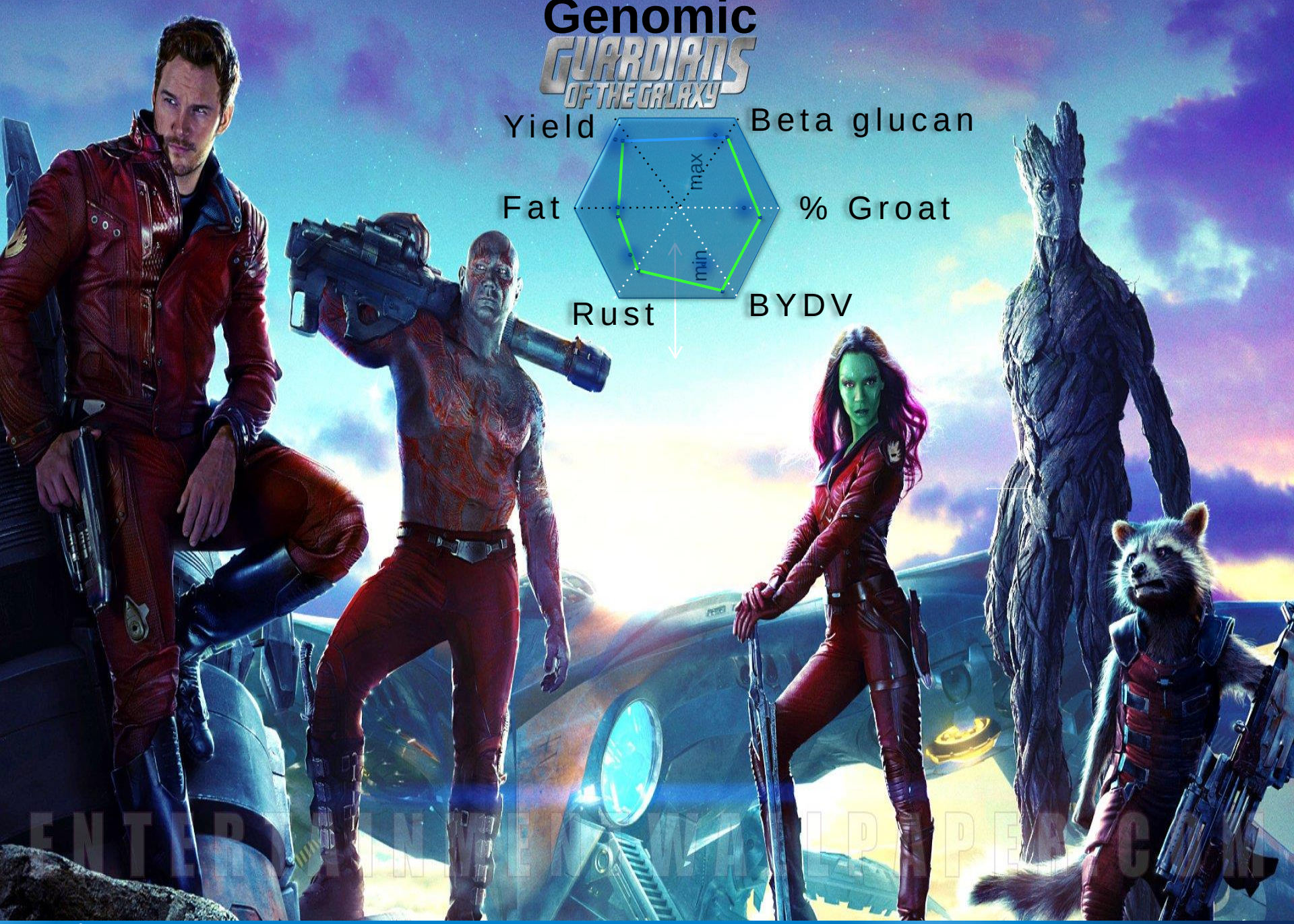
- Build “genomic selection” predictive models on mouse or rat lines
- Optimize animal line selection and potentially create new breeds for desired traits
- Dittos for microbes and viruses that interact with the environment

Copyright © 2013, SAS Institute Inc. All rights reserved.

STARK INDUSTRIES

Copyright © 2013, SAS Institute Inc. All rights reserved.

Genomic



Copyright © 2013, SAS Institute Inc. All rights reserved.

Thank You

General Mills

Eric Jackson
Joe Lutz
Lane Johnson
Ryan Brown

SAS

Kelci Miclaus
Luciano Silva
Wenjun Bao
Susan Shao

Copyright © 2013, SAS Institute Inc. All rights reserved.