Use of machine learning and REIMS technology to improve beef quality

D. Andy King,
Roman L. Hruska US Meat Animal Research Center,
Clay Center, NE 68933
“Old” approach—Yield grade equation
(Murphey et al., 1960)

\[
\% \text{BCTRC} = 51.34 - (5.784 \times \text{AFT}) - (0.462 \times \text{KPH}) + (0.74 \times \text{REA}) - (0.0093 \times \text{HCW})
\]

\[
\text{YG} = 2.50 - (2.50 \times \text{AFT}) - (0.20 \times \text{KPH}) + (0.32 \times \text{REA}) - (0.0038 \times \text{HCW})
\]

- Linear regression
- Small number of variables
- **Known** relationships to trait of interest
“New approach”—yield grading

• Collect image of ribeye at 12\textsuperscript{th}-
  13\textsuperscript{th} rib
• ~20,000 variables
• ???Relationship to yield
• Find previously unknown relationships
• AI to better segment the ribeye
Machine learning in meat industry

• Process optimization
• Automation
• Value determination
• Research and Development—Understanding biology
Understanding regulation of the phenotype

- Genetics
- Nutrition
- Time on feed
- Growth promotants
- Metabolic modifiers
- Animal age
- Electrical stimulation
- Chilling rate
- State of contraction
- Postmortem aging
- Blade tenderization
- Marination

We have studied these factors and more to understand their effects on tenderness, color stability, healthfulness, ...
Component traits

Majority of variation left unexplained

WHAT ARE WE MISSING?
‘...the best way to predict an individual’s disease risk is likely to come from understanding the biology behind each disease – and that will come from a combination of genomics, proteomics, and metabolomics.’

Principal Component Analysis

- One of the most common techniques for seeing complex relationships in data
- Derives a series of variables (components) that explain variation in the original data set
- First one explains the most, etc.
- By plotting loadings of variables for each component can visualize relationships
Using clustering to identify groups of samples that behave similarly in in vitro system

- Differences in enzyme system among cluster groups
- Studying differences in metabolic capacity
- Very different across temp decline rates
- Correlated to pH decline with fast temp decline
  - IMF, OC, carbonyls in mitochondrial fraction
- Correlated at slow temp decline
  - Glucose, MHCI
Mass spectrometry

• Can be used for metabolomics or proteomics
• Requires extensive sample prep
• Time consuming
• Molecular fingerprint
• Compounds identified by libraries
• Annotation is limiting
Beef tenderness metabolomic experiment

• Carcasses selection
  • U.S. Select
  • Screened for extremes for tenderness (n = 10/class)

• Tenderness assessed using slice shear force (Shackelford et al. 1999)

• Postmortem proteolysis of desmin (Wheeler et al. 2002)

• Non-targeted proteomic profiling
  • UPLC-MS and GC-MS

• Relationships to tenderness assessed
  • ANOVA
  • Pearson correlation coefficients
  • Partial least squares regression
Slice shear force

Days postmortem

Slice Shear Force, kg

Tender

Tough
Metabolomics results

- 2,562 metabolites profiled
- 651 met Bonferroni corrected significance level for aging effect
  - 254 met nominal significance for tenderness class effect

<table>
<thead>
<tr>
<th>Significance level</th>
<th>Tenderness class</th>
<th>Aging effect</th>
<th>Tenderness class X aging effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>P &lt; 2 × 10^{-5}</td>
<td>3</td>
<td>651</td>
<td>32</td>
</tr>
<tr>
<td>P &lt; 0.05</td>
<td>254</td>
<td>651</td>
<td>173</td>
</tr>
</tbody>
</table>

- Highly significant metabolites related to days postmortem
- Evaluated correlations with and without days postmortem effect

<table>
<thead>
<tr>
<th>Significance level</th>
<th>Pearson correlation</th>
<th>Partial correlation (dpm removed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slice shear force</td>
<td>Desmin degraded</td>
</tr>
<tr>
<td>P &lt; 2 × 10^{-5}</td>
<td>206</td>
<td>294</td>
</tr>
<tr>
<td>P &lt; 0.05</td>
<td>437</td>
<td>534</td>
</tr>
</tbody>
</table>

- 102 selected for PLS regression based on partial correlation
Metabolites results

- 28 were annotated
- Free amino acids/small peptides (n = 12)
  - Greater in tender, highly associated with proteolysis
- Glycolysis/Kreb’s intermediates (n = 5)
  - More glycolytic muscles produced more tender steaks
- Glycosides (n = 4)
  - Increased postmortem, associated with proteolysis, more so in steaks classified as tender
- Fatty acids (n = 2)
  - Greater increase in abundance during aging in steaks classified as tender.
Partial least squares regression
Rapid Evaporative Ionization Mass Spectrometry (REIMS)

- Originally developed for cancer research
- “Burns” and analyzes volatiles
- Gives Mass Spec fingerprint
- Compound ID??
- Real-time chemical fingerprint.
- Develop prediction equations
## Sampling for beef REIMS experiment

<table>
<thead>
<tr>
<th>Classifications (# of observations)</th>
<th>Model Sets</th>
<th>Breed</th>
<th>Tenderness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main</td>
<td>Specialized</td>
<td></td>
</tr>
<tr>
<td>Dark Cutter (41)</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Top Choice/Prime* (81)</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Choice/Select* (84)</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wagyu (42)</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Grass fed (42)</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Tender (215)</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Tough (74)</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Angus (159)</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Not Angus (46)</td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

1,700 bins (compounds) detected on each sample
Main data set
Accuracy of Main prediction

- LDA: 76.3 (PCA-FS)
- Svm Poly: 75.9 (FS)
- PLSDA: 75.5 (FS)
- PDA: 74.8 (PCA-FS)
- XGBoost: 72.7 (PCA-FS)
- Svm Linear: 71.8 (PCA-FS)
- Svm Radial: 70.0 (PCA-FS)
- LogbitBoost: 67.8 (PCA-FS)
- RF: 65.2 (FS)
- Knn: 53.6 (FS)
Specialized data set
Breed type data set
Breed prediction accuracy

<table>
<thead>
<tr>
<th>Breed</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLSDA</td>
<td>82.5 (PCA-FS)</td>
</tr>
<tr>
<td>XGBoost</td>
<td>82.4 (FS)</td>
</tr>
<tr>
<td>Svm Radial</td>
<td>81.3 (PCA-FS)</td>
</tr>
<tr>
<td>LDA</td>
<td>80.9 (PCA-FS)</td>
</tr>
<tr>
<td>PDA</td>
<td>80.9 (PCA-FS)</td>
</tr>
<tr>
<td>logitBoost</td>
<td>79.0 (PCA-FS)</td>
</tr>
<tr>
<td>Svm Poly</td>
<td>78.9 (PCA-FS)</td>
</tr>
<tr>
<td>Svm Linear</td>
<td>78.4 (PCA-FS)</td>
</tr>
<tr>
<td>RF</td>
<td>78.4 (FS)</td>
</tr>
<tr>
<td>Knn</td>
<td>78.0 (PCA-FS)</td>
</tr>
</tbody>
</table>
Tenderness dataset
Tenderness prediction accuracy

- Svm Poly: 90.8 (FS)
- XGBoost: 88.3 (FS)
- PLSDA: 87.2 (FS)
- RF: 86.9 (FS)
- Svm Linear: 86.9 (FS)
- PDA: 86.5 (FS)
- Svm Radial: 86.4 (FS)
- LDA: 84.5 (FS)
- Knn: 82.7 (PCA)
- LogbitBoost: 81.4 (FS)
## Final model prediction accuracies

<table>
<thead>
<tr>
<th>MODEL SET</th>
<th>DIMENSION REDUCTION APPROACH</th>
<th>NUMBER OF PREDICTORS</th>
<th>MACHINE LEARNING ALGORITHM</th>
<th>FINAL ACCURACY RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>PCA-FS</td>
<td>24 PCs</td>
<td>LDA</td>
<td>81.5%</td>
</tr>
<tr>
<td>Specialized</td>
<td>FS</td>
<td>60 mass-bins</td>
<td>SVM - Linear</td>
<td>99%</td>
</tr>
<tr>
<td>Breed</td>
<td>PCA-FS</td>
<td>38 PCs</td>
<td>SVM - Radial</td>
<td>85%</td>
</tr>
<tr>
<td>Tenderness</td>
<td>FS</td>
<td>67 mass-bins</td>
<td>XGBoost</td>
<td>90.5%</td>
</tr>
</tbody>
</table>
Conclusions

• Machine learning is a growing aspect of meat science research
• Will continue to impact all parts of the meats industry
• Meat science research will increasingly be dependent on finding relationships in complex data sets.
  • Combining of multiple levels of “omics” data
• REIMS specifically has great potential in sorting beef carcasses
• With annotation could be very useful for research purposes.
• REIMS will have to be adapted to a plant environment
• Further development of models for robustness
Acknowledgements

Steven Shackelford
Tommy Wheeler
Dan Nonneman
Gary Rohrer
Larry Kuehn
Mark Thallman
Amanda Lindholm
Brittney Keel
Warren Snelling
Aaron Dickey
J.R. Tait

Kristen Ostdiek
Casey Trambly
Patty Beska
Peg Ekeren
Megan Landes-Murphy
Jody Gallagher

Meat operations
Cattle operations,
Feedlot crew
Core lab

Jade Cooper
Carol Lorenzen
Melvin Hunt
Ranjith Ramanathan
Surendranath Suman
Devin Gredell
Jessica Prenni
Dale Woerner
Keith Belk
Steven Lonergan
Elisabeth Huff-Lonergan