Identification of trait-associated genes related to meat quality using an RNA-seq analysis approach

Joel D Leal-Gutiérrez, Mauricio Elzo, and Raluca Mateescu
Department of Animal Sciences
University of Florida
joelleal@ufl.edu
Introduction
There is **high variability** in tenderness within quality grade classification (marbling).

It has been reported that at least 40% of this variation is due to **additive genetics**.
There is **high variability** in tenderness within quality grade classification (marbling).

It has been reported that at least 40% of this variation is due to **additive genetics**.

Riley et al, 2009
**Objectives**

1) To perform a **trait associated (TA) gene expression analysis** using a constructed meat quality index as response

2) To assess **differential gene expression (DE)** using WBSF based grouping as predictor

3) To perform an **eQTL mapping** for overall gene expression and for DE genes in skeletal muscle
Materials and methods
Population
- Multibreed Angus-Brahman (MAB) UF herd

Based on the Angus composition
1 = 100 to 80%
2 = 79-65%
3 = 62.5% (Brangus)
4 = 59 to 40%
5 = 39 to 20%
6 = 19 to 0%

Phenotypic traits
150 steers with meat quality related phenotypes in *longissimus dorsi* muscle
Recorded phenotypes:

- WBSF
- Marbling
- Tenderness
- Juiciness
- Connective tissue
- Flavor

Principal component analysis

First three PCs were used
Recorded phenotypes:

- WBSF
- Marbling
- Tenderness
- Juiciness
- Connective tissue
- Flavor

First three PCs were used

Principal component analysis

Meat quality index $i = \sum (PCS_{ij} \times PCW_j)$

$PCS_{ij} = \text{Principal component score of animal } i \text{ for the PC}_j$

$PCW_j = \text{Weight of the PC}_j$
Recorded phenotypes:

- WBSF
- Marbling
- Tenderness
- Juiciness
- Connective tissue
- Flavor

First three PCs were used

Meat quality index $i = \sum (PCS_{ij} \times PCW_j)$

PCS$_{ij}$ = Principal component score of animal $i$ for the PC$_j$
PCW$_j$ = Weight of the PC$_j$

Selected animals: 80 individuals (extremes)
mRNAseq and read mapping

Raw Data

Fragment

Paired read

@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCACACTCACAGTTT+
!'*(((**+)%%%++)(%%%%%)%1***-+*''))**55CCF>>>>>>CCCCCCC65
mRNAseq and read mapping

Raw Data → Quality control and processing
Base filtering for low sequencing quality

Read filtering for low sequencing quality

Single reads screening

On average, 34.9 million paired reads were uniquely mapped
mRNAseq and read mapping

- Raw Data
  - Quality control and processing
    - Align reads to the genome
      - Genome guided transcriptome assembly
        - Tophat and bowtie:
mRNAseq and read mapping

Raw Data → Quality control and processing

Align reads to the genome → Tophat and bowtie:

Genome guided transcriptome assembly → Alignment file

Counts
# Gene counts

<table>
<thead>
<tr>
<th>Gene</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAMLG</td>
<td>113</td>
</tr>
<tr>
<td>CAMP</td>
<td>0</td>
</tr>
<tr>
<td>CAMSAP1</td>
<td>429</td>
</tr>
<tr>
<td>CAMSAP3</td>
<td>3</td>
</tr>
<tr>
<td>CAMTA1</td>
<td>36</td>
</tr>
<tr>
<td>CAMTA2</td>
<td>1414</td>
</tr>
<tr>
<td>CAND1</td>
<td>102</td>
</tr>
<tr>
<td>CAND2</td>
<td>2886</td>
</tr>
<tr>
<td>CANT1</td>
<td>118</td>
</tr>
<tr>
<td>CAP1</td>
<td>160</td>
</tr>
<tr>
<td>CAP2</td>
<td>4326</td>
</tr>
<tr>
<td>CAPG</td>
<td>8</td>
</tr>
<tr>
<td>CAPN1</td>
<td>669</td>
</tr>
<tr>
<td>CAPN10</td>
<td>49</td>
</tr>
<tr>
<td>CAPN11</td>
<td>9</td>
</tr>
<tr>
<td>CAPN12</td>
<td>0</td>
</tr>
<tr>
<td>CAPN13</td>
<td>0</td>
</tr>
<tr>
<td>CAPN14</td>
<td>0</td>
</tr>
<tr>
<td>CAPN15</td>
<td>197</td>
</tr>
<tr>
<td>CAPN2</td>
<td>678</td>
</tr>
<tr>
<td>CAPN3</td>
<td>9615</td>
</tr>
<tr>
<td>CAPN5</td>
<td>79</td>
</tr>
<tr>
<td>CAPN6</td>
<td>99</td>
</tr>
<tr>
<td>CAPN7</td>
<td>373</td>
</tr>
<tr>
<td>CAPN9</td>
<td>1</td>
</tr>
<tr>
<td>CAPNS1</td>
<td>1059</td>
</tr>
<tr>
<td>CAPNS2</td>
<td>0</td>
</tr>
<tr>
<td>CAPRN1</td>
<td>1130</td>
</tr>
</tbody>
</table>
Gene counts

<table>
<thead>
<tr>
<th>Gene</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAMRA1</td>
<td>0</td>
</tr>
<tr>
<td>CAMLG</td>
<td>113</td>
</tr>
<tr>
<td>CAMP</td>
<td>0</td>
</tr>
<tr>
<td>CAMSAP1</td>
<td>429</td>
</tr>
<tr>
<td>CAMSAP3</td>
<td>3</td>
</tr>
<tr>
<td>CAMTA1</td>
<td>36</td>
</tr>
<tr>
<td>CAMTA2</td>
<td>1414</td>
</tr>
<tr>
<td>CAND1</td>
<td>102</td>
</tr>
<tr>
<td>CAND2</td>
<td>2886</td>
</tr>
<tr>
<td>CANT1</td>
<td>118</td>
</tr>
<tr>
<td>CAP1</td>
<td>160</td>
</tr>
<tr>
<td>CAP2</td>
<td>4326</td>
</tr>
<tr>
<td>CAPG</td>
<td>0</td>
</tr>
<tr>
<td>CAPN1</td>
<td>669</td>
</tr>
<tr>
<td>CAPN10</td>
<td>49</td>
</tr>
<tr>
<td>CAPN11</td>
<td>9</td>
</tr>
<tr>
<td>CAPN12</td>
<td>0</td>
</tr>
<tr>
<td>CAPN13</td>
<td>0</td>
</tr>
<tr>
<td>CAPN14</td>
<td>0</td>
</tr>
<tr>
<td>CAPN15</td>
<td>197</td>
</tr>
<tr>
<td>CAPN2</td>
<td>678</td>
</tr>
<tr>
<td>CAPN3</td>
<td>9615</td>
</tr>
<tr>
<td>CAPN4</td>
<td>79</td>
</tr>
<tr>
<td>CAPN5</td>
<td>99</td>
</tr>
<tr>
<td>CAPN6</td>
<td>373</td>
</tr>
<tr>
<td>CAPN7</td>
<td>1</td>
</tr>
<tr>
<td>CAPN8</td>
<td>1</td>
</tr>
<tr>
<td>CAPNS1</td>
<td>1059</td>
</tr>
<tr>
<td>CAPNS2</td>
<td>0</td>
</tr>
<tr>
<td>CAPRN1</td>
<td>1122</td>
</tr>
</tbody>
</table>

Lowly expressed genes were excluded.

Total number of included genes: 8,671
Association assay

TA gene expression analysis:

1. Continuous response variable (Meat quality index)
2. Gene counts as a predictor
3. 80 animals
4. Robust regression: less affected by outliers

DE genes

1. Gene counts as response variable
2. WBSF based grouping (tender and tough) as a predictor
3. 40 animals were used, 20 tough and 20 tender
4. Regression using the negative binomial distribution
Methods: TA gene expression

Y = Meat quality index

Low index

High index
Methods: TA gene expression

Y = Meat quality index

**Low index**

- 1.15
- Tougher
- Dryer
- More connective tissue
- Less marbling

**High index**

- 3.35
- More tender
- More juicy
- Less connective tissue
- More marbling
Methods: TA gene expression

\[ Y = \text{Meat quality index} \]

**Low index**
- 1.15
- Tougher
- Dryer
- More connective tissue
- Less marbling

**High index**
- 3.35
- More tender
- More juicy
- Less connective tissue
- More marbling

\[
\text{MeatQualityIndex}_i = \mu + \beta_0 + \beta_1 \times \text{year}_{1i} + \beta_2 \times \text{PC}_{2i} + \beta_3 \times \text{geneCounts}_{3i} + e_i
\]
Results: TA gene expression

Gene expression $-\log_{10}(p)$ vs. Transcription unit location in the genome.
Results: TA gene expression

Gene expression $-\log_{10}(p)$

Transcription unit location in the genome

Gene names:
- ARHGAP10
- TMEM120B
- ARRDC4
- LOC100848872
This gene controls the Arp2/3 complex and F-actin dynamics at the Golgi complex by regulating the activity of the small GTPase Cdc42.
TMEM120 proteins are expressed preferentially during adipocyte differentiation.

**Graph:**

- X-axis: Meat quality index
- Y-axis: Expression

The graph shows a negative correlation between meat quality index and expression levels of TMEM120B gene.
Methods: DE genes

GeneCounts = μ + β₀ + β₁ * year₁ + β₂ * BreedGroup₃ + β₃ * WBSFGroup₃ + eᵢ
Comparison: tough vs tender

Results: DE genes

1,286 genes
adjPvalue < 0.05
Results: DE genes

Differential gene expression for WBSF

Comparison: tough vs tender

HABP2

1,286 genes
adjPvalue < 0.05
Comparison: tough vs tender

Results: DE genes

<table>
<thead>
<tr>
<th>GO term</th>
<th>pValue</th>
<th>n.anno</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0045333</td>
<td>4.24E-07</td>
<td>42</td>
<td>Cellular respiration</td>
</tr>
<tr>
<td>GO:0015980</td>
<td>6.45E-06</td>
<td>50</td>
<td>Energy derivation by oxidation of organic compounds</td>
</tr>
<tr>
<td>GO:0044429</td>
<td>7.38E-06</td>
<td>146</td>
<td>Mitochondrial part</td>
</tr>
<tr>
<td>GO:0005739</td>
<td>1.33E-05</td>
<td>211</td>
<td>Mitochondrion</td>
</tr>
<tr>
<td>GO:0006091</td>
<td>2.33E-05</td>
<td>66</td>
<td>Generation of precursor metabolites and energy</td>
</tr>
<tr>
<td>GO:0070469</td>
<td>5.20E-05</td>
<td>30</td>
<td>Respiratory chain</td>
</tr>
<tr>
<td>GO:0017144</td>
<td>2.93E-04</td>
<td>70</td>
<td>Drug metabolic process</td>
</tr>
<tr>
<td>GO:0055114</td>
<td>3.76E-04</td>
<td>120</td>
<td>Oxidation-reduction process</td>
</tr>
<tr>
<td>GO:0005759</td>
<td>4.51E-04</td>
<td>69</td>
<td>Mitochondrial matrix</td>
</tr>
<tr>
<td>GO:1990204</td>
<td>4.73E-04</td>
<td>33</td>
<td>Oxidoreductase complex</td>
</tr>
<tr>
<td>GO:0044281</td>
<td>6.42E-04</td>
<td>152</td>
<td>Small molecule metabolic process</td>
</tr>
<tr>
<td>GO:0022904</td>
<td>7.22E-04</td>
<td>20</td>
<td>Respiratory electron transport chain</td>
</tr>
<tr>
<td>GO:0005740</td>
<td>7.67E-04</td>
<td>105</td>
<td>Mitochondrial envelope</td>
</tr>
<tr>
<td>GO:0031966</td>
<td>9.35E-04</td>
<td>101</td>
<td>Mitochondrial membrane</td>
</tr>
<tr>
<td>GO:0044455</td>
<td>1.31E-03</td>
<td>40</td>
<td>Mitochondrial membrane part</td>
</tr>
</tbody>
</table>

Gene enrichment analysis

1,286 genes

adjPvalue < 0.05

Differential gene expression for WBSF

HABP2

Comparison: tough vs tender
Methods: eQTL analysis

- Genome wide association that uses gene expression as response variable
- Goal: to identify genomic regions with control on gene expression

Input:

8,671 genes expressed in skeletal muscle, 1,286 of which were DE

115,287 polymorphic markers
Results: eQTL analysis

Cis-eQTL

Gene 1

SNP A
Results: eQTL analysis

Cis-eQTL

Trans-eQTL

Gene 1

SNP A

Gene 1

SNP B

Gene 2
Transcript genomic location

Marker genomic location

eQTL
- Trans
- Trans DE
- Cis
- Cis DE

Hot spot
Actin Binding LIM Protein 1

May act as scaffold protein.
May play a role in the development of tissues.

Associated with 42 genes = 3.2% of DE genes
Conclusions
Conclusions

❖ TA gene expression analysis:

The expression of 4 genes were determined as associated with meat quality index in the present population

❖ DE gene analysis:

A total number of 1,286 genes were DE in the tough vs tender comparison

❖ eQTL mapping:

Eight eQTL hot spots for global expression were identified and one hot spot for DE genes was found
Acknowledgments
Financial Support

Florida Beef Council
Florida Cattlemen's association
UF ANS Hatch Project

University of Florida

Tracy Scheffler  
Jason Scheffler  
Dwain Johnson  
Danny Driver  
Francisco Peñagaricano  
Mesfin Gobena

Victoria Velinsky  
Zaira Estrada  
Heather Hamblen  
Sarah Flowers  
Kaytin Sarlo  
Eduardo Rodriguez
Thanks.
Robust regression:

Association analysis

https://www.statsmodels.org
Align reads to the genome

Genome guided transcriptome assembly

Tophat and bowtie:

Reads

Alignment to transcriptome

Unaligned Reads
Raw Data → Quality control and processing

Align reads to the genome → Genome guided transcriptome assembly

Tophat and bowtie:

Reads → Alignment to transcriptome

Unaligned Reads → Alignment to genome

Unaligned Reads →
Raw Data

Quality control and processing

Align reads to the genome

Genome guided transcriptome assembly

Tophat and bowtie:

Reads

Alignment to transcriptome

Unaligned Reads

Alignment to genome

Unaligned Reads

Spitted into pieces
Raw Data → Quality control and processing → Align reads to the genome → Genome guided transcriptome assembly

Tophat and bowtie:
- Reads → Alignment to transcriptome → Unaligned Reads
- Unaligned Reads → Alignment to genome → Spitted into pieces → Unaligned Reads
Gene expression

Exon usage