Genetic and Environmental Info in goat milk FTIR spectra

B. Dagnachew and T. Ådnøy
Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences
Outline

- Background
  - Milk composition
  - Infrared (IR) spectroscopy
  - Objectives
- Methods
- Results
- Conclusion
Background

- Milk is a complex mixture of several thousand kinds of bio-molecules
  - In the current animal breeding programs, only few major milk components are included
  - Fat %, protein %, lactose %, SCC, FFA, etc

- Infrared (IR) spectrometry is used
  - Rapid method (chemical analysis are time and labor intensive)
Infrared (IR) spectra

- Fourier Transform Infrared (FTIR)
  - Improve signal to noise ratio
  - Multiplexing capabilities
  - Higher energy throughput

- Genetic variability of cow milk based on MIR spectra ([Soyeurt et al., 2010](#))

- Genetic variability of goat milk FTIR spectra have not been studied
  - Some regions are indirectly known

Objective:

Direct estimation of genetic and environmental variability of goat milk FTIR spectra
Materials and Methods

- Raw FTIR Spectra data
  - TINE (Norwegian dairies) has four D-labs performing routine FTIR analysis on milk samples
  - Raw FTIR spectra are stored since 2007 in a database

- For this study:
  - Two years (2007 and 2008)
  - Total of 73,858 observations
    - 28,269 goats
    - 271 farms
Cont...

- **Selection of FTIR spectra**
  - O-H bending region (between 1600 cm\(^{-1}\) and 1700 cm\(^{-1}\))
  - O-H stretching region (between 3025 cm\(^{-1}\) and 3500 cm\(^{-1}\))

Low signal to noise ratio

Variables

536 removed
524 selected
Limitations ...

- Simultaneous genetic analysis of selected spectra data points is not possible with the current methods available for genetic analysis e.g. DMU 31, WOMBAT 42, ASReml 21 traits, ...

- Two step analysis performed
  - Dimension (trait) reduction
    - Principal component analysis (PCA)
  - Genetic analysis
    - Mixed model
Cont ...

- **Principal component analysis (PCA)**
  
  **Goal:** to extract a set of fewer components that explain as much variation as possible of the original variation
  
  \[ Y = TP' + F \]
  
  - Y are the FTIR spectral data
  - T is score matrix
  - P is loading matrix
  - F is error term

  - **PROC PRINCOMP** in SAS (SAS Institute, 1994)
    - On correlation matrix of Y
    - Scores, T, are considered as new traits \(\rightarrow\) **PC traits**

  - **Pre-correction**
    - PC traits T corrected for:
      - farm region (12 levels)
      - kidding season (3 levels)
      - lactation stage (4 co-variables)
      - age at kidding (5 levels)
      - lactation number (3 levels)
Methods Cont...

Model

- Multi-trait linear mixed model on pre-corrected PC traits

\[ t_c = Xb + Za + Qp + Wh + e \]

- \( t_c \) is the vector of pre-corrected PC traits
- \( b \) is the vector of fixed effect (HTY and HTM)
- \( a \) is the vector of additive genetic effect
- \( p \) is the vector of permanent environment effect
- \( h \) is the vector of random herd-test-day (HTDr) effect
- \( e \) is the vector of random residual effect

- \( X \), \( Z \), \( Q \), and \( W \) are the corresponding design matrices

- AI-REML of WOMBAT  \((Meyer, 2007)\)
Results and Discussion

- **PCA**
  - Resulted in **8 PCs** which explain ~99% of the total spectral variation
  - The dramatic reduction of the number of components suggests there is high correlation among spectral variables

<table>
<thead>
<tr>
<th>PCs</th>
<th>Relative eigenvalues</th>
<th>Genetic (heritability)</th>
<th>Permanent environment</th>
<th>Herd test-day (HTDr)</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58.96</td>
<td>0.154</td>
<td>0.020</td>
<td>0.105</td>
<td>0.721</td>
</tr>
<tr>
<td>2</td>
<td>25.62</td>
<td>0.264</td>
<td>0.091</td>
<td>0.298</td>
<td>0.347</td>
</tr>
<tr>
<td>3</td>
<td>7.74</td>
<td>0.146</td>
<td>0.015</td>
<td>0.388</td>
<td>0.451</td>
</tr>
<tr>
<td>4</td>
<td>3.24</td>
<td>0.057</td>
<td>0.270</td>
<td>0.020</td>
<td>0.653</td>
</tr>
<tr>
<td>5</td>
<td>1.66</td>
<td>0.011</td>
<td>0.013</td>
<td>0.941</td>
<td>0.035</td>
</tr>
<tr>
<td>6</td>
<td>1.07</td>
<td>0.229</td>
<td>0.135</td>
<td>0.277</td>
<td>0.359</td>
</tr>
<tr>
<td>7</td>
<td>0.41</td>
<td>0.211</td>
<td>0.058</td>
<td>0.298</td>
<td>0.432</td>
</tr>
<tr>
<td>8</td>
<td>0.28</td>
<td>0.285</td>
<td>0.033</td>
<td>0.140</td>
<td>0.542</td>
</tr>
</tbody>
</table>

Table 1: Variance ratios for genetic, permanent environment, HTDr and residual random effects
Amide III (protein)
OPO asym. Stretch (mixture)
CO stretch (mixture)

Lactose
OH group (C-OH bond)

Fat A
Carbonyl group (C=O bond)

Fat B
Fatty acids chains (C–H bond)

Protein
amide II group (N-H bond)
FTIR fingerprint region

C-O and C-C stretching mode

Atmospheric CO2

No specific chemical information
Cont ...

- Average spectral heritabilities

<table>
<thead>
<tr>
<th>Traits</th>
<th>Estimated heritability (2009)</th>
<th>Spectral heritability (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactose</td>
<td>0.38</td>
<td>0.25 (0.054 – 0.38)</td>
</tr>
<tr>
<td>Fat</td>
<td>0.30</td>
<td>0.22 (0.041 – 0.385)</td>
</tr>
<tr>
<td>Protein</td>
<td>0.35</td>
<td>0.25 (0.017 – 0.39)</td>
</tr>
</tbody>
</table>

Contents are predicted from FTIR spectra

- 1000 – 1200 cm\(^{-1}\)
- 1740 – 1800 cm\(^{-1}\) (Fat A)
- 2800 – 2900 cm\(^{-1}\) (Fat B)
- 1240 – 1260 cm\(^{-1}\) (amide III)
- 1520 – 1560 cm\(^{-1}\) (amide II)
Conclusion

- There is a substantial amount of genetic variation in goat milk FTIR spectra
  - Not all FTIR regions are of breeding interest
  - But could be used in herd management
Thank you for your attention!