

Are Breed-specific Differences in Beef Fatty-acid Profiles a "Needle in the Haystack"?

Yuvaraj Ranganathan^{1,*}, Irena Hulova¹, Matthias Schreiner², Jan Macak¹

¹Agriresearch Rapotin Ltd., Vykrovice, Czech Republic

²Department of Food Science and Technology, University of Natural Resources and Applied Life Sciences (BOKU), Vienna, Austria

*Corresponding author: yuvaraj.ranganathan@vuchs.cz

Received August 18, 2014; Revised November 10, 2014; Accepted November 19, 2014

Abstract A healthy human nutrition stems from macronutrients that provide energy and also from micronutrients that are indispensable for basic physiological processes. Essential fatty-acids have been proved and commonly agreed in the recent years as important elements of a healthy diet. This gives impetus to look for sources of essential fatty-acids in the diet. Beef is not only considered a complete protein, but also a source of essential fatty-acids such as linoleic and linolenic acids. The levels of these fatty-acids are however variable depending on the diet and breed of cattle. In this work, we analyzed intra-muscular fatty acid profiles of two breeds (Aberdeen Angus and Blonde d'Aquitaine) as a part of long-term complex and comprehensive comparison of breed differences. Using a novel statistical algorithm we found interesting breed-specific fatty-acid profiles. The advantages of this approach over conventional approaches are discussed as well as the specific differences in fatty acid profiles between the breeds.

Keywords: Aberdeen Angus, Blonde d'Aquitaine, beef fatty acids, multivariate analysis, random forests

Cite This Article: Yuvaraj Ranganathan, Irena Hulova, Matthias Schreiner, and Jan Macak, "Are Breed-specific Differences in Beef Fatty-acid Profiles a "Needle in the Haystack"?" *American Journal of Food Science and Technology*, vol. 2, no. 6 (2014): 192-195. doi: 10.12691/ajfst-2-6-4.

1. Introduction

Although beef is considered a complete protein, health effects of beef consumption can vary based on its fat content [1]. These fats are usually present in the form of long-chain fatty acid with varying levels of saturation. From the nutritional perspective, saturated fatty acids (SFA) and polyunsaturated fatty-acids (PUFA) are the most relevant to human health [2]. The levels of these fatty-acids are however variable depending on the diet and breed of cattle [3,4]. Certain lean cuts of beef have about 50% SFAs but also contain essential n-3 and n-6 PUFAs which are necessary for human nutrition [5]. SFAs have been linked to increase in blood cholesterol levels, coronary heart disease and certain forms of cancer [6,7,8]. Beneficial effects of diet rich in PUFAs include reduction of chronic and inflammatory diseases such as cardiovascular disease and dementia [9]. There is growing interest to enrich beef with such PUFAs by manipulating factors which affect differential intra-muscular fatty acid composition [10,11], but it seems to be exceedingly difficult as the gut bacteria in ruminants tend to modify the fatty acid profile [12].

While techniques to measure and characterize fatty acids are well established [3,4,13,14], they produce large amounts of accurate data, which have to be interpreted. Most methods use pair-wise comparisons; comparing one fatty-acid at a time [2]. Rigorous statistical treatment by advanced techniques is an approach that provides promising options for improvement in this field.

In this paper, we demonstrate the use of Random Forests (RF), a machine learning algorithm belonging to the class of data-mining techniques, for the analysis of beef intramuscular fatty acid composition data. This new technique is being increasingly used in data-rich fields such as bioinformatics, chemoinformatics, medical diagnostics, food technology and astronomy to select the most appropriate candidate variables from the surrounding data [15-20]. However, to the best of our knowledge, such data mining technique has not been used for comparison of fatty acid profiles between various cattle breeds. Thus, we compare in this work differences in intramuscular fatty acid composition between Aberdeen Angus and Blonde d'Aquitaine breeds of cattle using Random Forests to find breed-specific differences.

2. Materials and Methods

2.1. Intramuscular Fatty Acid Extraction

The fatty acids were extracted from longissimus muscle from two breeds – Aberdeen Angus (n = 13) and Blonde d'Aquitaine (n = 14). Fatty acids were extracted as follows: to 2 g of homogenized meat, 4 ml of methanol and 2 ml of dichloromethane were added and shaken well. This mixture was filtered through a funnel containing anhydrous sodium sulfate. To the filtrate, 2 ml of methanol and 2 ml distilled water was added, shaken well and the lower phase was recovered after centrifugation at 1100 rpm for 5 min at room temperature. This clear

solution was concentrated under a gentle stream of N₂. Conversion of extracted fatty acids to methyl esters was done using the methanolic KOH method: To the vial containing the extracted fat, 2 ml heptane was added, followed by 2 ml of 2M methanolic KOH, after which the vial was vortexed for 1 min. One gram of anhydrous sodium sulphate was added, shaken well and allowed to settle. The clear solution was transferred a GC vial for analysis. The acids were analyzed using a gas chromatograph (Thermo DSQ) fitted with a Supelco SP2560 column (100 m length and 0.25 mm internal diameter). The temperature program was as follows: 60°C was held for 1 min; followed by a ramp of 15°/min to 150°C and then 2.5°/min to 230°C for 5 min. The carrier gas was Hydrogen at a flow rate of 1.5 ml/min. The injector and detector (FID) was set to 250°C and was injected with a split ratio of 1:100.

2.2. Statistical Analysis

Fatty acid composition data consists of samples of beef collected over several individuals. Whatever the type of sample, or its method of collection, the key feature of such a data set is that it is analogous to a microarray gene expression dataset in the sense that there are many more variables than samples. This nature of a dataset limits the usage of classical multivariate analysis methods such as multivariate analysis of variance (MANOVA) or linear discriminant analysis (LDA). Such conventional methods require sample sizes to be proportionately increased for each added variable and also assume normality of the dataset, as well as the absence of auto-correlation between variables, besides other limitations [21].

Random Forests [22] is a classification algorithm which has the following features which makes it best suited for fatty acid analyses: 1) it allows for more variables than samples, 2) it has a good classification efficiency, even with a lot of background noise, 3) it is capable of arriving at a minimal set of variables, which can be used as predictors of that particular group, 4) it is robust to interactions and correlations among variables, and 5) it gives measures of relative variable importance. RF builds sets of decision trees using bootstrapping from the set of samples and also selects a variable set of attributes (different fatty acids in this case) at each node of the many decision trees so generated. In this way RF is also different from other tree-building methods such as PAUP: Phylogenetic Analysis Using Parsimony [23]. A major advantage of RF is that it does not overfit the data [18,22] so that even if minor fluctuations in variable strength (fatty acid proportions in this case) lead to the building of thousands of classification trees, these fluctuations are not given undue importance in the final model; thus only the minimum set of important predictor variables is obtained.

2.3. Classification Using RF

In this case, we retained the identity of the groups (e.g. Aberdeen Angus and Blonde d'Aquitaine) and attempted to find the prediction error of group membership and the variables (fatty acids). The package randomForest was used with R software version 3.0.2 for this purpose [24]. The selected variables which show differences between the groups were identified using the *importance* function. This function provides a Mean Decrease in Accuracy

(MDA) index of variables sorted according to their importance for classifying the groups. Visualization of breed-specific differences was obtained by the *MDSplot* function. For all analyses presented in this paper, only the proportions of the different fatty acids present in the samples were used.

3. Results and Discussion

Overall we were able to identify 40 fatty acids from both breeds. The multivariate plot shown in Figure 1 shows that fatty acid profiles from Aberdeen Angus and Blonde d'Aquitaine cattle were clearly separated. The MDA values give the list of fatty acids in decreasing order of importance which would be useful in differentiating the beef from the two breeds. The higher the value of MDA, the important that particular fatty acid is in differentiating the two breeds. The MDA values in the negative range indicates that these are the fatty acids present in similar levels in both breeds which means that they are not good for classifying the two breeds. Beef from Aberdeen Angus were clearly classifiable from Blonde d'Aquitaine by having higher levels of oleic acid (C18:1 cis-9), where as beef from Blonde d'Aquitaine had higher levels of linoleic acid (C18:2 n6c). These results, summarized in Table 1, demonstrate the relative importance of certain fatty acids in defining the fatty acid signature of the group.

This multivariate approach is more suitable to data sets similar to what is illustrated here. This is because when data is presented as percentages, an increase in one will automatically mean a decrease in the other. This non-independence of variables makes conventional univariate statistics unsuitable for comparing one variable at a time. Another advantage this approach provides is that variable importance is ranked and this gives relative importance which provides a starting point for further investigation. In one analysis, such a wealth of information is not given by conventional univariate tests such as T-test or F-test.

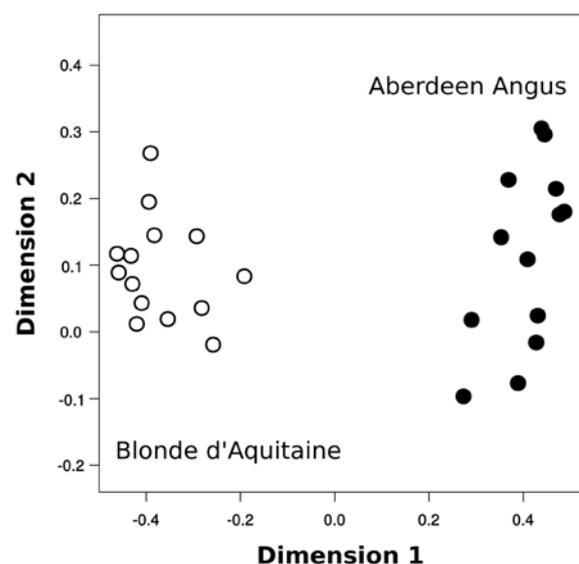


Figure 1. Multivariate plot of breed-specific differences in beef fatty acid profiles. Open circles are from Blonde d'Aquitaine (n = 14) and filled circles are from Aberdeen Angus (n = 13)

Investigating the reasons for the underlying causes of such breed-specific differences is beyond the scope of this

study as it is discussed elsewhere [13,14]. What we have attempted to do here is to provide the scientific community with possible statistical tools to improve their analysis and interpretation of datasets similar to ours.

Table 1. Fatty acid profiles of Blonde d'Aquitaine and Aberdeen Angus cattle with mean decrease in accuracy (MDA)

| Fatty acid | MDA | Blonde d'Aquitaine % (mean ± SD) | Aberdeen Angus % (mean ± SD) |
|-------------|-------|-------------------------------------|---------------------------------|
| C18:2n6c | 12.49 | 21.21 ± 2.28 | 11.66 ± 3.42 |
| C18:1 cis-9 | 8.55 | 22.97 ± 3.52 | 31.12 ± 3.83 |
| C16:1 | 6.81 | 1.97 ± 0.55 | 2.70 ± 0.47 |
| C18:1n11c | 6.3 | 1.64 ± 0.15 | 1.25 ± 0.10 |
| C16:0 | 4.87 | 19.75 ± 1.87 | 22.24 ± 1.80 |
| C23:0 | 4.13 | 0.38 ± 0.13 | 0.24 ± 0.13 |
| C20:0 | 4.04 | 0.10 ± 0.02 | 0.20 ± 0.26 |
| C14:1 | 3.82 | 0.20 ± 0.18 | 0.35 ± 0.14 |
| C20:5n3 | 3.68 | 0.78 ± 0.23 | 0.54 ± 0.30 |
| C17i | 2.54 | 0.78 ± 0.35 | 0.73 ± 0.42 |
| C20:3n6 | 2.21 | 1.12 ± 0.30 | 0.77 ± 0.34 |
| C20:4n6 | 1.93 | 4.94 ± 0.94 | 3.74 ± 1.66 |
| C22:0 | 1.86 | 0.21 ± 0.03 | 0.15 ± 0.05 |
| C14:0 | 1.42 | 1.29 ± 0.39 | 1.79 ± 0.45 |
| C18:3n3 | 0.87 | 1.82 ± 1.08 | 1.20 ± 0.89 |
| C17:0 | 0.73 | 1.54 ± 0.25 | 1.76 ± 0.62 |
| C18:2n6t | 0.72 | 0.54 ± 0.08 | 0.60 ± 0.12 |
| C15i | 0.71 | 0.09 ± 0.02 | 0.10 ± 0.05 |
| C24:1 | 0.68 | 0.58 ± 0.15 | 0.63 ± 0.26 |
| C18:0 | 0.64 | 10.41 ± 1.52 | 10.02 ± 1.52 |
| C18:1n9t | 0.18 | 0.84 ± 0.65 | 1.18 ± 1.00 |
| C17ai | 0.15 | 0.25 ± 0.12 | 0.22 ± 0.08 |
| C22:6n3 | 0.14 | 0.11 ± 0.04 | 0.12 ± 0.11 |
| C18:2c9t11 | 0.07 | 0.25 ± 0.10 | 0.34 ± 0.15 |
| C20:1 | -0.07 | 0.94 ± 0.85 | 1.29 ± 1.04 |
| C20:3n3 | -0.08 | 0.06 ± 0.07 | 0.07 ± 0.07 |
| C15:1 | -0.16 | 0.01 ± 0.02 | 0.01 ± 0.02 |
| C24:0 | -0.27 | 0.04 ± 0.03 | 0.04 ± 0.05 |
| C16i | -0.78 | 0.18 ± 0.05 | 0.16 ± 0.06 |
| C18:1nXt | -0.79 | 0.44 ± 0.63 | 0.54 ± 0.38 |
| C15ai | -0.89 | 0.14 ± 0.04 | 0.13 ± 0.05 |
| C22:5n3 | -0.93 | 1.52 ± 0.43 | 1.09 ± 0.39 |
| C20:2 | -1.2 | 0.63 ± 0.29 | 0.65 ± 0.41 |
| C22:1n9 | -1.27 | 0.07 ± 0.16 | 0.10 ± 0.25 |
| C22:2 | -1.37 | 0.23 ± 0.16 | 0.32 ± 0.27 |
| C21:0 | -1.58 | 0.01 ± 0.02 | 0.01 ± 0.03 |
| C15:0 | -1.66 | 0.46 ± 0.12 | 0.47 ± 0.23 |
| C17:1 | -1.79 | 0.13 ± 0.08 | 0.14 ± 0.12 |
| C18:1nXc | -1.9 | 0.77 ± 0.16 | 0.81 ± 0.23 |
| C18:3n6 | -1.92 | 0.54 ± 0.51 | 0.45 ± 0.46 |

4. Conclusions

Here we demonstrated the utility of improved statistical approaches to rigorously distinguish various beef fatty acids. Based on this illustrative example, we believe that this approach can be applied elsewhere, where data sets are represented as composition data to overcome constraints of classical statistical approaches.

Statement of Competing Interests

The authors have no competing interests

Acknowledgement

Financial support by the project "Development of Human Resources in Food Analysis using Gas Chromatography" of the Ministry of Education, Youth

and Sport, Reg. Nr. CZ.1.07/2.3.00/30.0008., is gratefully acknowledged.

References

- Ascherio A. and Willett W.C., "Health effects of trans fatty acids," *The American Journal of Clinical Nutrition*, 66, 1006S-1010S. 1997.
- Daley C.A., Abbott A., Doyle P.S., Nader G.A. and Larson S., "A review of fatty acid profiles and antioxidant content in grass-fed and grain-fed beef," *Nutrition Journal*, 9, 10. 2010.
- Bures D., Barton L., Zahradkova R., Teslik V., and Krejcová M., "Chemical composition, sensory characteristics, and fatty acid profile of muscle from Aberdeen Angus, Charolais, Simmental, and Hereford bulls," *Czech Journal of Animal Science*, 51, 279-284. 2006.
- Barton L., Bures D. and Kudrna V., "Meat quality and fatty acid profile of the musculus longissimus lumborum in Czech Fleckvieh, Charolais and Charolais x Czech Fleckvieh bulls fed different types of silages," *Czech Journal of Animal Science*, 55, 479-487. 2010.
- Wyness L., Weichselbaum E., O'Connor A., Williams E.B., Benelam B., Riley H. and Stanner S., "Red meat in the diet: an update," *Nutrition Bulletin*, 36, 34-77. 2011.
- Brouwer I.A., Wanders A.J. and Katan M.B., "Effect of animal and industrial trans fatty acids on HDL and LDL cholesterol levels in humans – a quantitative review," *PLoS One*, 5, e9434. 2010.
- Astrup A., Dyerberg J., Elwood P., Hermansen K., Hu F.B., Jakobsen M.U., Kok F.J., Krauss R.M., Lecerf J.M., LeGrand P., Nestel P., Riserus U., Sanders T., Sinclair A., Stender S., Tholstrup T. and Willett W.C., "The role of reducing intakes of saturated fat in the prevention of cardiovascular disease: where does the evidence stand in 2010?," *The American Journal of Clinical Nutrition*, 93, 684-688. 2011.
- Bastide N.M., Pierre F.H. and Corpet D.E., "Heme iron from meat and risk of colorectal cancer: a meta-analysis and a review of the mechanisms involved," *Cancer Prevention Research*, 4, 177-184. 2011.
- Welch A.A., Shakya-Shrestha S., Lentjes M.A., Wareham N.J. and Khaw K.T., "Dietary intake and status of n-3 polyunsaturated fatty acids in a population of fish-eating and non-fish-eating meat-eaters, vegetarians, and vegans and the precursor-product ratio of α -linolenic acid to long-chain n-3 polyunsaturated fatty acids: results from the EPIC-Norfolk cohort," *The American Journal of Clinical Nutrition*, 92, 1040-1051. 2010.
- McGregor E.M., Campbell C.P., Miller S.P., Purslow P.P. and Mandell I.B., "Effect of nutritional regimen including limit feeding and breed on growth performance, carcass characteristics and meat quality in beef cattle," *Canadian Journal of Animal Science*, 92, 327-341. 2012.
- Franco D., Crecente S., Vázquez J.A., Gómez M. and Lorenzo J.M., "Effect of cross breeding and amount of finishing diet on growth parameters, carcass and meat composition of foals slaughtered at 15 months of age," *Meat Science*, 93 (3), 547-556. 2013.
- Shingfield K.J., Bonnet M. and Scollan N.D., "Recent developments in altering the fatty acid composition of ruminant-derived foods," *Animal*, 7, 132-162. 2013.
- Cifkova E., Holcapek M., Lisa M., Ovcacikova M., Lycka A., Lynen F. and Sandra P., "Nontargeted quantitation of lipid classes using hydrophilic interaction liquid chromatography-electrospray ionization mass spectrometry with single internal standard and response factor approach," *Analytical Chemistry*, 84, 10064-10070. 2012.
- Lisa M. and Holcapek M., "Characterization of triacylglycerol enantiomers using chiral HPLC/APCI-MS and synthesis of enantiomeric triacylglycerols," *Analytical Chemistry*, 85, 1852-1859. 2013.
- Svetnik V., Liaw A., Tong C., Culbertson J.C., Sheridan R.P. and Feuston B.P., "Random Forest: A classification and regression tool for compound classification and QSAR modeling," *Journal of Chemical Information and Computer Sciences*, 43, 1947-1958. 2003.
- Cannon E.O., Bender A., Palmer D.S. and Mitchell J.B.O., "Chemoinformatics-based classification of prohibited substances

- employed for doping in sport," *Journal of Chemical Information and Modeling*, 46, 2369-2380. 2006.
- [17] Díaz-Uriarte R. and de Andrés S.A., "Gene selection and classification of microarray data using random forest," *BMC Bioinformatics*, 7, 3. 2006.
- [18] Granitto P.M., Biasioli F., Aprea E., Mott D., Furlanello C., Märk T.D. and Gasperi F., "Rapid and non-destructive identification of strawberry cultivars by direct PTR-MS headspace analysis and data mining techniques," *Sensors and Actuators B*, 121, 379-385. 2007.
- [19] Zhang J., Sokal I., Peskind E.R., Quinn J.F., Jankovic J., Kenney C., Chung K.A., Millard S.P., Nutt J.G. and Montine T.J., "CSF multianalyte profile distinguishes Alzheimer and Parkinson disease," *American Journal of Clinical Pathology*, 129, 526-529. 2008.
- [20] Gao D., Zhang Y.-X. and Zhao Y.-H., "Random forest algorithm for classification of multiwavelength data," *Research in Astronomy and Astrophysics*, 9, 220-226. 2009.
- [21] Stevens J., *Applied Multivariate Statistics for the Social Sciences*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1992.
- [22] Breiman L., "Random forests," *Machine Learning*, 45, 5-32. 2001.
- [23] Perdiguero-Alonso D., Montero F.E., Kostadinova A., Raga J.A. and Barrett J., "Random forests, a novel approach for discrimination of fish populations using parasites as biological tags," *International Journal for Parasitology*, 38, 1425-1434. 2008.
- [24] R Development Core Team, "R: a language and environment for statistical computing. R Foundation for Statistical Computing." Vienna, Austria <http://www.R-project.org>. 2009.